



2007

AUTOMATED CLASSIFICATION OF POWER QUALITY DISTURBANCES USING SIGNAL PROCESSING TECHNIQUES AND NEURAL NETWORKS

Praveen Settipalli

University of Kentucky, 2settipalli@gmail.com

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Settipalli, Praveen, "AUTOMATED CLASSIFICATION OF POWER QUALITY DISTURBANCES USING SIGNAL PROCESSING TECHNIQUES AND NEURAL NETWORKS" (2007). *University of Kentucky Master's Theses*. 430.

https://uknowledge.uky.edu/gradschool_theses/430

This Thesis is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Master's Theses by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

ABSTRACT OF THESIS

AUTOMATED CLASSIFICATION OF POWER QUALITY DISTURBANCES USING SIGNAL PROCESSING TECHNIQUES AND NEURAL NETWORKS

This thesis focuses on simulating, detecting, localizing and classifying the power quality disturbances using advanced signal processing techniques and neural networks. Primarily discrete wavelet and Fourier transforms are used for feature extraction, and classification is achieved by using neural network algorithms. The proposed feature vector consists of a combination of features computed using multi resolution analysis and discrete Fourier transform. The proposed feature vectors exploit the benefits of having both time and frequency domain information simultaneously. Two different classification algorithms based on Feed forward neural network and adaptive resonance theory neural networks are proposed for classification. This thesis demonstrates that the proposed methodology achieves a good computational and error classification efficiency rate.

KEY WORDS: Power Quality Classification, Frequency and Wavelet Domain, Multi-Resolution Analysis, Feed Forward Neural Networks, Adaptive-Resonance Theory Neural Networks.

Praveen Settipalli

(Author's Signature)

4th May, 2007

(Date)

AUTOMATED CLASSIFICATION OF POWER QUALITY
DISTURBANCES USING SIGNAL PROCESSING TECHNIQUES AND NEURAL
NETWORKS

By

Praveen Settipalli

Dr. Yuan Liao
Director of Thesis

Dr. Yu Ming Zhang
Director of Graduate Studies

15th June, 2007
Date

RULES FOR THE USE OF THESIS

Unpublished thesis submitted for the Master’s degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author with the usual scholarly acknowledgements.

Extensive copying or publication of the thesis in whole or in part also requires the consent of the Dean of the Graduate School at the University of Kentucky.

A library that borrows this thesis for use by its patrons is expected to secure the signature of each user.

Name

Date

Horizontal lines for recording signatures and dates.

THESIS

Praveen Settipalli

The Graduate School
University of Kentucky

2007

AUTOMATED CLASSIFICATION OF POWER QUALITY DISTURBANCES USING
SIGNAL PROCESSING TECHNIQUES AND NEURAL NETWORKS

THESIS

A thesis submitted in partial fulfillment of the
requirements for the degree of Masters of Science in the College of Engineering
at the University of Kentucky

By

Praveen Settipalli

Lexington, Kentucky

Director: Dr. Yuan Liao, Assistant Professor of Electrical Engineering

Lexington, Kentucky

2007

Dedicated to my Parents

ACKNOWLEDGEMENTS

Firstly, my deepest thanks to my advisor and thesis chair Dr. Yuan Liao for providing me with the support, valuable technical guidance and financial assistance through the span of the research. I would also like to thank Dr. Jimmie J. Cathey and Dr. Paul Dolloff for agreeing to be on my defense committee. Their critical reviews are very much acclaimed.

It is only the unparalleled love, support and vision of my parents, loved ones and friends that made this work a reality. Thank you one and all and lastly my whole hearted thanks to the Department of Electrical Engineering at the University of Kentucky for all the resources that helped me in successfully completing my degree requirements.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF FILES	ix
CHAPTER ONE: INTRODUCTION	1
1.1 INTRODUCTION	1
1.2 MOTIVATION	4
1.3 OUTLINE OF THE THESIS	5
CHAPTER TWO : LITERATURE REVIEW	7
2.1 POWER QUALITY STUDIES.....	7
2.2 DETECTION METHODS.....	8
2.3 CLASSIFICATION METHODS.....	10
CHAPTER THREE : POWER QUALITY DISTURBANCES AND THE DETECTION ALGORITHM.....	13
3.1 VARIOUS POWER QUALITY DISTURBANCES	13
3.1.1 HARMONICS	14
3.1.2 TRANSIENT.....	15
3.1.3 FLICKER	15
3.1.4 SAG	16
3.1.5 SWELL	16
3.1.6 OUTAGE	18
3.1.7 IMPULSE	18
3.1.8 NOTCH.....	19
3.2 DISCRETE WAVELET TRANSFORM	22
3.3 CHOICE OF THE WAVELET	24
3.4 MULTI RESOLUTION ANALYSIS	25
3.5 DETECTION ALGORITHM	32
CHAPTER FOUR : PROPOSED POWER QUALITY CLASSIFICATION ALGORITHM.....	34
4.1 PARSVEL'S THEOREM	34
4.2 STANDARD DEVIATION-MULTI RESOLUTION ANALYSIS CURVES.....	35
4.3 FEATURE VECTOR EXTRACTION ALGORITHM	38
4.4 FEED FORWARD NEURAL NETWOKS	43
4.5 TRAINING THE ARTIFICIAL NEURAL NETWORK	45
CHAPTER FIVE : CURRENT RESEARCH - CLASSIFICATION USING A NEW FEATURE VECTOR ALGORITHM AND ART NEURAL NETWORKS	48
5.1 FEATURE VECTOR EXTRACTION METHOD	49

5.2 ADAPTIVE RESONANCE THEORY NEURAL NETWORKS	53
5.2.1 ART-1 ALGORITHM	55
CHAPTER SIX : EXPERIMENTAL RESULTS	58
6.1 EXPERIMENTAL PROCEDURE	58
6.2 RESULTS	63
6.3 CONCLUSION	67
BIBLIOGRAPHY.....	68
VITA	72

LIST OF TABLES

Table 3.1: Spectral content, duration and magnitude of the various power quality disturbances.....	21
Table 3.2 : Frequency cut-off ranges for high pass coefficients at different MRA levels.....	31
Table 5.1 : A Possible result of grouping the clusters using the clustering algorithm.....	52
Table 6.1 : Various combinations of neural network architectures and their performance data.....	63
Table 6.2 : Error Classification during simulation for a 3 layer network and 90,000 dataset.....	66
Table 6.3 : Error Classification during simulation for a 4 layer network and 9,000 dataset.....	66
Table 6.4 : Error Classification during simulation for a 4 layer network and 90,000 dataset.....	67

LIST OF FIGURES

Figure 1.1 : Economic Implications of Power quality disturbances during a single year.....	2
Figure 1.2 : Generalized Block diagram of a disturbance classification algorithm.....	3
Figure 3.1 : Power Signal with even Harmonics.....	14
Figure 3.2 : Power Signal with Transients.....	15
Figure 3.3 : Power Signal with Flicker.....	16
Figure 3.4 : Power Signal with a Sag.....	17
Figure 3.5 : Power Signal with a Swell.....	17
Figure 3.6 : Power Signal with Outage.....	18
Figure 3.7 : Power Signal with a Impulse.....	19
Figure 3.8 : Power Signal with Notch.....	20
Figure 3.9 : Various forms of Daubechie mother wavelets generated using Matlab.....	24
Figure 3.10 : Frequency – Time Resolution in case of wavelet domain.....	25
Figure 3. 11 : Functional representation of Multi-Resolution Analysis of a signal $S(n)$	26
Figure 3.12 : MRA for a normal waveform, D1,D3,A4,D2,D4 in that order.....	27
Figure 3.13 : MRA for a Flicker waveform, D1,D3,A4,D2,D4 in that order.....	27
Figure 3.14 : MRA for a Harmonic waveform, D1,D3,A4,D2,D4 in that order.....	28
Figure 3.15 : MRA for a Notch waveform, D1,D3,A4,D2,D4 in that order.....	28
Figure 3.16 : MRA for a Outage waveform, D1,D3,A4,D2,D4 in that order.....	29
Figure 3.17 : MRA for a Sag waveform, D1,D3,A4,D2,D4 in that order.....	29
Figure 3.18 : MRA for a Swell waveform, D1,D3,A4,D2,D4 in that order.....	30
Figure 3.19 : Detection and localization algorithm, flow chart.....	33
Figure 4.1: Multi-Resolution Energy Distribution Curves for a Harmonic Disturbance & a Impulse.....	36
Figure 4.2 : Multi-Resolution Energy Distribution Curves for a Flicker and a Notch.....	37
Figure 4.3 : Multi-Resolution Energy Distribution Curves for a Sag & a Swell.....	37
Figure 4.4 : Multi-Resolution Energy Distribution Curves for an Outage and a Transient.....	38
Figure 4.5 : Feature Vector matrix for ‘m’ signals - Typical Input matrix to the classifier.....	42
Figure 4.6 : A 3 layer neural network with ‘n’ inputs and ‘m’ outputs and 1 hidden layer.....	43

Figure 4.7 : A 4 layer neural network with 2 hidden layers, n inputs and z outputs.....	44
Figure 5.1: Transforming the time domain signal into 12 scale MRA with zero padding.....	49
Figure 5.2 : Possible way of clustering the threshold matrix Gb.....	51
Figure 5.3 : Architecture of a ART-1 neural network with two layers, F1 and F2.....	56
Figure 5.4 : Adaptive Resonance Theory1 functional algorithm.....	57
Figure 6.1 : Performance function for a 3 layer neural network using ‘Trainbfg’ function.....	59
Figure 6.2 : Performance function for a 4 layer neural network using ‘Trainbfg’ function.....	60
Figure 6.3 : Performance function for a 5 layer neural network using ‘Trainbfg’ function.....	61

LIST OF FILES

Filename: Praveen_Thesis.pdf

Type : PDF

Size : 2.36 MB

CHAPTER ONE: INTRODUCTION

1.1 INTRODUCTION

Any disturbance in the voltage, current or frequency of the power signal that can adversely affect the customer's equipment can be termed as a power quality problem. The deregulation of the power industry and the proliferation of sensitive semiconductor equipment into almost all kinds of industrial machinery and consumer electronics generated the demand for power quality and techniques for the reduction in power quality problems. [1] refers to power quality as the combination of voltage quality and current quality, where voltage quality is concerned with deviations of the actual voltage from the ideal value and current quality is the equivalent definition for current. However, most often a disturbance in voltage also causes a disturbance in the current and hence the term Power quality is used when referring to both voltage quality or current quality. These disturbances even though last only a fraction of a second can cause huge losses and hours of manufacturing downtime in case of industrial applications. The figures given below taken from [2] and [3] show the monetary loss caused due to power quality disturbances for a given year in the United States. Hence studies related to classification of power quality disturbances and the corresponding equipment sensitivity studies and equipment modeling are in demand.

A unified frame work for integrating the data processing and analysis and modeling and simulation of the system and equipment was proposed in [4]. These studies can help us to understand about the power quality disturbances and take corresponding actions to avoid the economic losses caused because of it. In this process, an efficient and yet simple classification algorithm which can accurately classify the disturbances is

needed. This thesis work hence focus on developing an algorithm for feature extraction and then classification of those disturbances with a high degree of error classification efficiently.

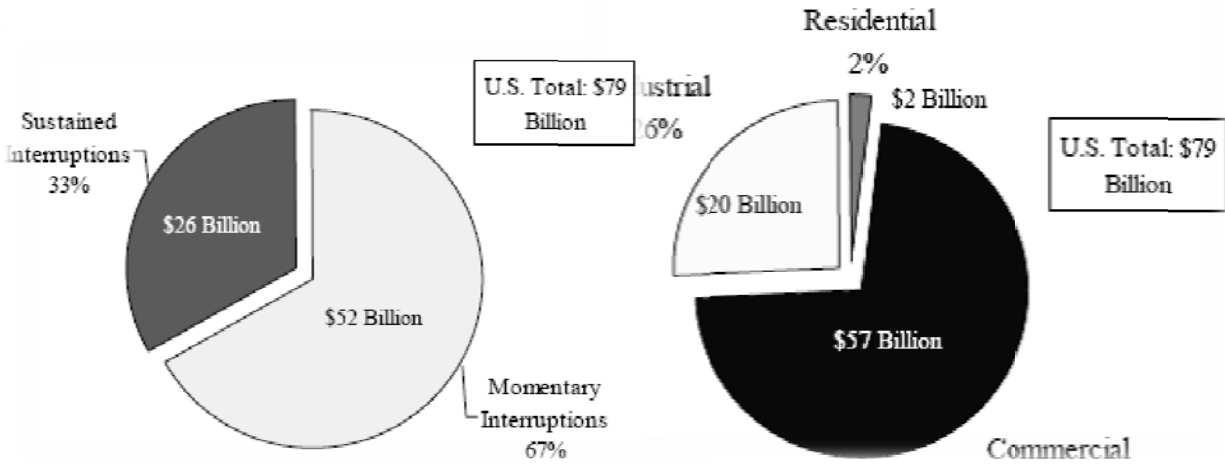


Figure 1.1 Economic Implications of Power quality disturbances during a single year[2]

The Literature review regarding the various such other methods present in the literature is given in chapter 2. Fig 1.2, presents the over view about the various blocks present in a power quality classification system/algorithm. The basic block diagram of the structure of any classification algorithm using Artificial Intelligence techniques will be similar to one shown in fig 1.2, except that in case of a neural network other AI tools like Expert Systems, Hidden Markov Models etc [5] were used.

In real time applications, instead of the signal generation block in fig 1.2, real time data is fed into the system. Real time data can be live data from transmission lines or stored data from a database. [6] presents a J2EE framework for a online power quality classification system. The paper focus on the 'n'- tier J2EE architecture and this thesis in

its feature work will focus on implementing such framework using the proposed feature extraction and classification algorithms.

The feature extraction block contains the algorithms to convert the signal from a large R space to a limited value F space. Discrete wavelet transforms and Fourier transforms are using to extract those features. The classification block generally consists of AI tools to classify the feature vector into its corresponding disturbances. Decision making will then depend on the output of the classifier. Various methods like voting schemes, Dempster-Shafter theory of evidence, rule based approaches are sometimes used for decision making.

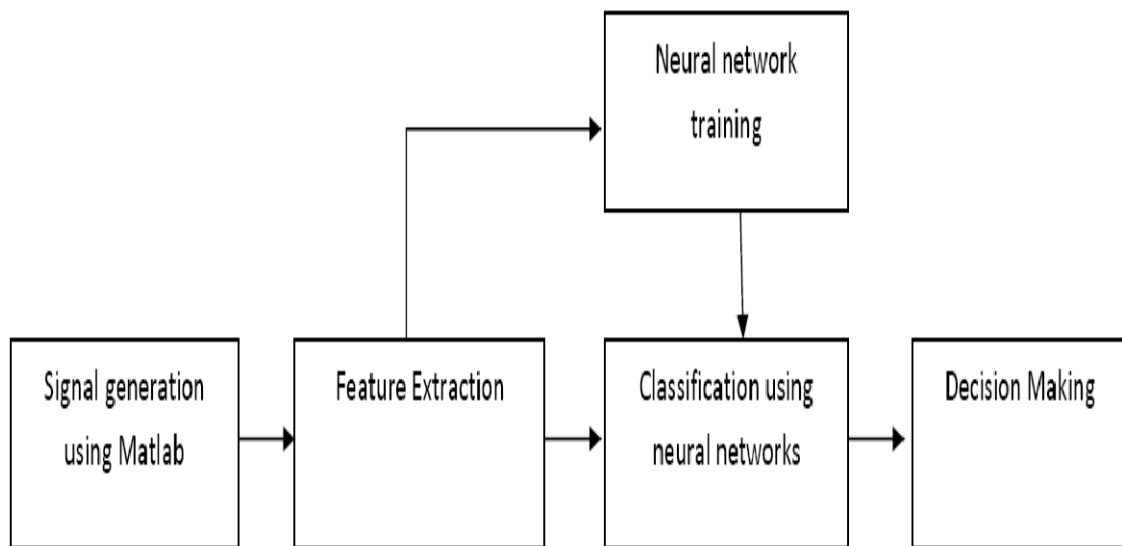


Figure 1.2 : Generalized Block diagram of a disturbance classification algorithm

1.2 MOTIVATION

Advanced signal processing techniques like wavelet transforms and artificial intelligence tools like neural networks, expert systems were continuously being used to achieve significant success in the areas of finger print recognition, finger print database compression, speech recognition etc. So power quality problems which are similar to any of those above areas in terms of design can also take advantage of these techniques and researchers have started looking at this problem from the above perspective.

This thesis drew its initial motivation from [4] which proposes a classification methodology by using fuzzy logic. However, neural networks when trained in a significantly very large amount will be able to produce very similar accurate results. Moreover, the use of fuzzy expert system always needs a person who has expert knowledge in power systems whereas feature extraction and classification using neural networks can be done by anyone with signal processing knowledge. So this entire work focus on a classification algorithm from a non-power systems perspective.

Given the huge size of power quality data and the enormous loss a nano-second of a power quality disturbance can cause, a very efficient and yet simple power quality classification algorithm is the need of the day. Moreover, the current state of VLSI enables us to fabricate the algorithm onto a Field Programmable Gate Arrays (FPGA). [7] and [8] proposed hardware implementations for their classification algorithm. This thesis also focus in its future work, in the direction towards implementing a VHDL based FPGA realization for the proposed classification algorithms.

1.3 OUTLINE OF THE THESIS

The second chapter deals with the various power quality disturbances that are considered in this thesis work and then briefly reviews the state of art and literature in power quality classification. Various feature extraction techniques and classification techniques using fuzzy logic, Neuro-fuzzy techniques, expert systems, hidden markov models etc are stated.

The third chapter deals with explaining about how Discrete Wavelet and Fourier transforms can be used to identify the detecting and then localizing the disturbance present in the waveform. Various disturbance scenarios and the corresponding discrete wavelet transform coefficients at various levels using Multi Resolution Analysis are shown and then the algorithm used to detect and localize the disturbances is explained.

The fourth chapter presents the proposed feature extraction and classification algorithm. Feature extraction comprises of wavelet multi resolution analysis(MRA) and parameters extracted from discrete Fourier transform. Then a 4 layer and 5 layer feed forward neural network is presented and the classification of disturbances by first training the neural network with the feature vector and then testing the network on different data sets are presented.

The fifth chapter deals with a new algorithm for feature extraction. This method is based on extracting more features from the resolution levels which contain significantly more information than from those that contain less information. A method to do this, which was initially proposed by [43] was explained and then a classification algorithm which uses Adaptive Resonance theory neural network is proposed.

The sixth chapter explains the experimental results. This chapter details about the application development using MATLAB 7.01 and summarizes about the results obtained for various unique datasets. Details regarding feature extraction and training of neural networks are presented.

The seventh chapter is basically a summary of the entire work and also deals with inferences regarding the results obtained and it also provides insight into the feature work and improvements that can be done to this work.

CHAPTER TWO : LITERATURE REVIEW

This section is a discussion about the state of art literature in areas related to power quality disturbance classification. The first few paragraphs deal with history related to power quality and the various standards present in literature regarding it, followed by literature about various tools for feature classification starting with papers which introduce the Wavelet transforms, multi resolution analysis and related approaches. Then literature on various kinds classification techniques which use the state of art Artificial Intelligence tools is provided, concluded with literature which deals with the hardware implementation of the power quality classification algorithms.

2.1 POWER QUALITY STUDIES

There are many references [10] – [13] which dealt with the various guidelines regarding monitoring power quality disturbances. [12] Provides the basic introduction to all the various power quality disturbances possible in power distribution scenario. [10] provides a survey of various distribution sites and concluded various interesting observations about the various disturbance occurrence statistics which includes statistics that the majority of the voltage sags have a magnitude of around 80% and a duration of around 4 to 10 cycles and that the total harmonic distortion on harmonic disturbances is around 1.5 times the normal value. These surveys provide basic introductory information about the occurrence and cause of the disturbances.

2.2 DETECTION METHODS

Wavelet transform is a recent signal processing tool which is widely being used for disturbance detection in PQ. Wavelets can provide accurate frequency resolution and poor time localization at low frequencies and the vice versa at high frequencies. The property that the wavelets integrate to zero shows the ability of the standard deviation of different resolution levels to represent the distribution of the distorted signals. This capacity is used to classify and quantify the short duration variations within the power signals. Papers [14] – [17] presents the properties of wavelet transforms and their use to scenarios similar to power quality disturbance classification.

Poisson, P. Rioual and M.Meunier [18] discussed the possibility of using Continuous Wavelet transform, Multi-resolution analysis and Quadratic Transform for detection of the disturbances. They concluded from their analysis that multi-resolution analysis and quadratic transforms are adequate and reliable tools for the detection of sharp changes in the signals and frequency transients. The continuous wavelet along with all the above features can directly compute the magnitude of the 50 Hz signal. We can observe that none of these methods uniquely or jointly are able to detect all kinds of disturbances. However their paper could successfully analyze the strengths of each of these above techniques.

Olivier Poisson, Pascal Rioual, and M.Meunier [19] proposed a method of using continuous wavelet transforms to detect and analyze voltage sags and transients. The characteristics of the analyzed signals are measured on a time-frequency plane and are compared with the standard benchmark values. Any inconstancy will imply that there is a disturbance in the signal. This algorithm enabled accurate time localization, magnitude

measurement of voltage sags and transient identification. Many more artificial intelligence based automated detection techniques followed this paper.

Another interesting transform called S-transform which is an extension of wavelet transform is proposed by P. K. Dash, B. K. Panigrahi, and G. Panda [20]. The S-transform had many promisingly impressive time-frequency resolution characteristics. The S-transform is obtained by multiplying the CWT with a phase factor as

$$S(\tau, \rho) = e^{i2\pi\tau\rho} * \mathcal{F}(\tau, \alpha) \quad (1)$$

$$\text{where } \mathcal{F}(\tau, \alpha) = \int_{-\infty}^{+\infty} h(t)f(t - \tau, \alpha)dt \quad (2)$$

$\mathcal{F}(\tau, \alpha)$ being the mother wavelet. The more defining properties of the S-transform are due to the fact that the modulating sinusoids are fixed with respect to the time axis while the localizing scalable Gaussian window dilates and translates. There are many other papers [21-25] which dealt with applying wavelet transforms for power quality analysis. Potential investigation in this area deals with searching for a better feature extraction tool. Approaches like combining FT with various WT functions and similar methods depending on the type of disturbances [4], can be investigated.

2.3 CLASSIFICATION METHODS

Classification of disturbance signals requires the use of pattern recognition techniques. Pattern recognition is a process of perceiving a pattern of a given object based on the knowledge already possessed [26]. So automated pattern recognition uses various artificial intelligence techniques like fuzzy logic (FL), artificial neural networks (ANN) and adaptive fuzzy logic (AFL) for the classification of disturbance signals. Recently techniques based on probabilistic models like Hidden Markov models, Dynamic time wrapping, Dempster-Shafer theory of evidence are also proposed. A survey of various AI tools relevant to power quality research can be found at [27].

Artificial neural networks are the oldest among the pattern recognition tools [28]. They are defined as software algorithms that can be trained to learn the relationships that exist between input and output data. The disadvantage of using artificial neural networks is that they require a lot of time to train them, before they are fully functional. The advantage of using neural networks is that they do not make any assumptions regarding the underlying distribution. They recognize the patterns by experience acquired during the training session. The network adjusts its internal parameters by prescribed rules during the training session.

Fuzzy logic is the next approach in pattern recognition [29]. It was developed from the fact that human brain doesn't make decisions based on sharp decision boundaries. Fuzzy logic uses exactly the same concept. Unlike the classical digital logic which uses either a 0 or 1, fuzzy logic uses a decision boundary which smoothly transitions between stages. The membership function sets this smooth transition between the decision boundaries. Classification of signals is made by using a fixed set of fuzzy

rules which consists of fuzzification, inference, composition and defuzzification. Approaches which combine both neural networks and fuzzy logic are recently being published [30 – 34].

Liao Y and Lee J.B [4], [35] presented a novel approach of using a fuzzy-expert system for automated detection and classification of power quality disturbances. Fuzzy logic is used for the classification of disturbance signals. A combination of Fourier and wavelet based techniques is used to the detection of signals. They compared the classification results with those of using ANN and proved that fuzzy logic is an efficient tool for the classification of power signal disturbances in terms of computational efficiency and accuracy.

P. K. Dash, S.Mishra, M.M.A.Salama, and A.C.Liew [36] proposed the use of a Fourier linear combiner and a fuzzy expert system for the classification of signals. A Fourier linear combiner estimates the normalized peak amplitude of the voltage signal and its rate of change. These values are given as input to the fuzzy expert system which classifies the disturbances based on the rules formulated. Even though this system seems to be computationally simple compared with using wavelet transforms and ANN or FL, the authors have not provided the computational error efficiency or other comparison strategies which could prove its efficiency over the existing methods.

M. Gaouda, S. H. Kanoun, M. M. A. Salama, and A. Y. Chikhani [23] proposed the use of computationally simple pattern recognition techniques. They used the wavelet multi-resolution transform for feature extraction and proposed a minimum Euclidean distance classifier, K-nearest neighbor classifier and neural network classifier for pattern recognition. But in scenarios where building a rule based classifier is tedious, statistical

models like Dynamic Time Wrapping (DTW), Hidden Markov models or Dempster-Shafer theory of evidence are more efficient. The DTW is a template matching algorithm derived from dynamic programming [37]. Conceptually, template matching is based on the comparison of the test signal against all of the stored templates in the dictionary. A measure of similarity is calculated and then used to achieve a recognition decision. But DTW requires a huge computation time. HMM is conceptually defined as a doubly stochastic process, comprised of an underlying stochastic process that is not directly observable, but can only be visualized through another set of stochastic processes that produce a sequence of observations [38]. Dempster-Shafer theory of evidence is a similar theory of probable reasoning and combining evidence. It pools several pieces of evidence bearing on a hypothesis under consideration to assess the truth of the hypothesis. This theory provides a partial belief of the accepted hypothesis. The literature which deals with these approach can be found in [39], [40].

CHAPTER THREE: POWER QUALITY DISTURBANCES AND THE DETECTION ALGORITHM

This chapter introduces the various power quality disturbances that are being considered in this chapter. Their characteristics were detailed and then the signal processing tools used for extracting the features from these waveforms are introduced. This follows the proposed algorithm which used wavelet multi resolution analysis to detect and hence localize the disturbances.

3.1 VARIOUS POWER QUALITY DISTURBANCES

The duration and the frequency of the interruption influence the loss due to the interruptions. The losses due to interruptions can be categories as direct losses – loss in production environments, damages to equipment etc, indirect losses – delay in the delivery of the product etc and non-economic inconveniences. Chapter 1 already explained about the economic aspects of power quality disturbances. The first step towards classification of power quality disturbances is to know about the characteristics of various disturbances. In a general sense power quality disturbances can be broadly classified as transients, long duration voltage variations, short-duration, voltage variations, voltage unbalances, voltage fluctuations, power frequency variations, voltage variations. However, its tedious to consider all variants of these disturbances and hence the disturbances which occur the most are considered in this thesis. Based on it, Power quality disturbances can be classified as below [12].

3.1.1 HARMONICS

Harmonics can be defined as sinusoidal waveforms with frequencies that are multiples of the frequency at which the supply voltage is designed to be delivered. It is generally produced due to the non-linear characteristics of the load and devices. A parameter used to compute the

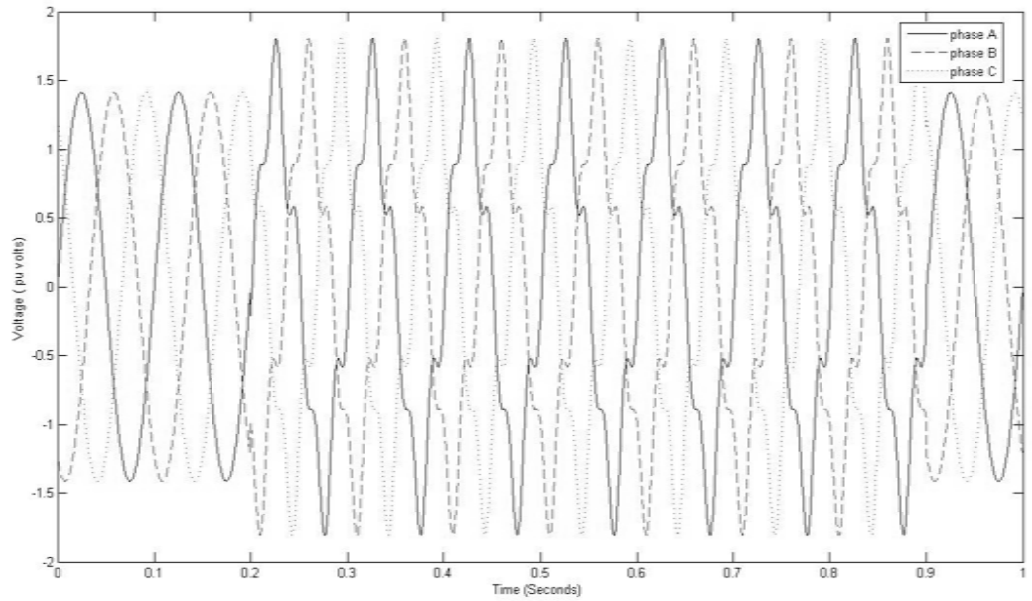


Figure 3.3 Power Signal with even Harmonics

harmonics is the Total Harmonic Distortion(THD). The THD of a signal in time domain can be calculated as given below

$$\text{THD} = \frac{\sqrt{V_2^2 + V_4^2 + V_6^2 + \dots}}{V_1} \quad (3)$$

where V_n is the harmonic component of the voltage signal V . Fig 3.1 below shows a power signal with even harmonics.

3.1.2 TRANSIENT

A transient is a signal with a disturbance that dies to zero in a finite time. Transients can be again divided as Impulsive transients and Oscillatory transients. Impulsive transients are a sudden, non-power frequency change in the steady-state condition of power signal, that is generally unidirectional in polarity, where as an Oscillatory transients are sudden frequency change in the steady state condition of the power signal and this generally includes both positive and negative polarity values. Fig 3.2 shows a power signal with a transient disturbance.

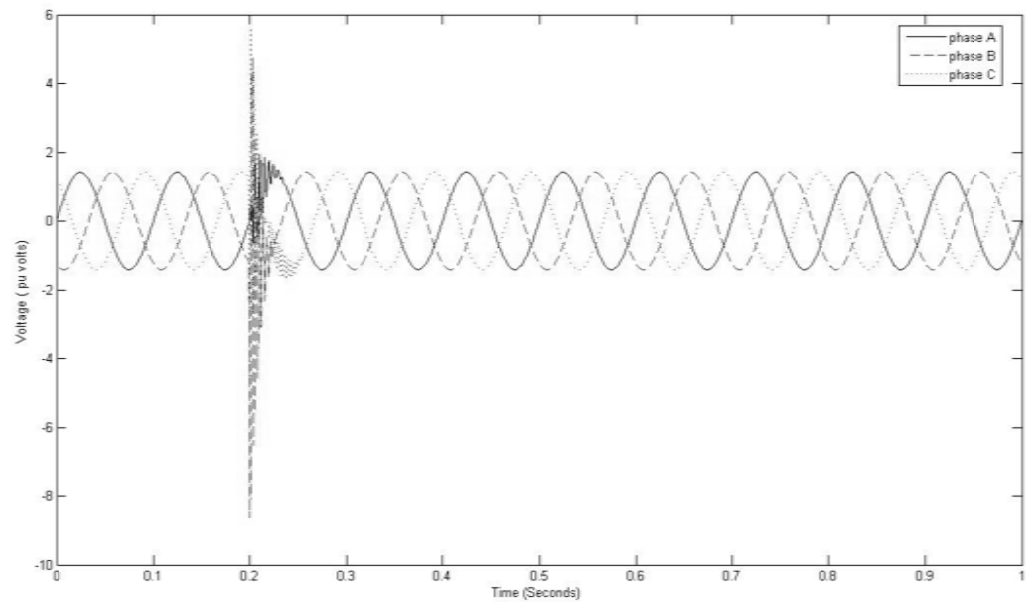


Figure 3.2 Power Signal with Transients

3.1.3 FLICKER

Voltage fluctuations are series of random voltage changes or spikes. Flicker is defined by its RMS magnitude expressed as a percent of the fundamental frequency magnitude. Their magnitude generally will be in the range of 0.9 to 1.1 pu. The main

source of voltage fluctuations are the continuous rapid variations of load. Arc furnace is one of the common cause for voltage flickers. Fig 3.3 shows the voltage flickers in a signal.

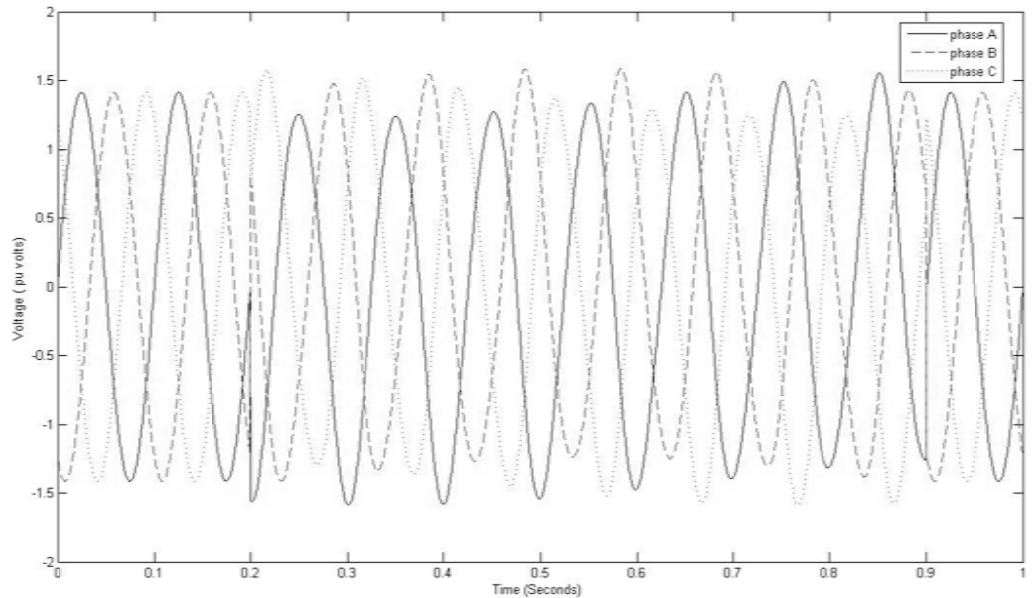


Figure 3.3 Power Signal with Flicker

3.1.4 SAG

A sag is a decrease in RMS voltage or currents to about 0.1 to 0.9 pu at normal supply frequency for a duration of $\frac{1}{2}$ cycles to 60 seconds. Voltage sags are usually associated with system disturbances but can also be caused by connection of heavy loads or starting of large motors[12]. Fig 3.4 shows a voltage sag in a power signal

3.1.5 SWELL

A swell is a inverse of sag, characterized by an increase in RMS voltage or current to between 1.1 and 1.8 p.u for a duration of $\frac{1}{2}$ cycles to 60 seconds. These are

mostly associated with system disturbance conditions, but they are not very common like Sags. Fig 3.5 shows a voltage swell in a power signal.

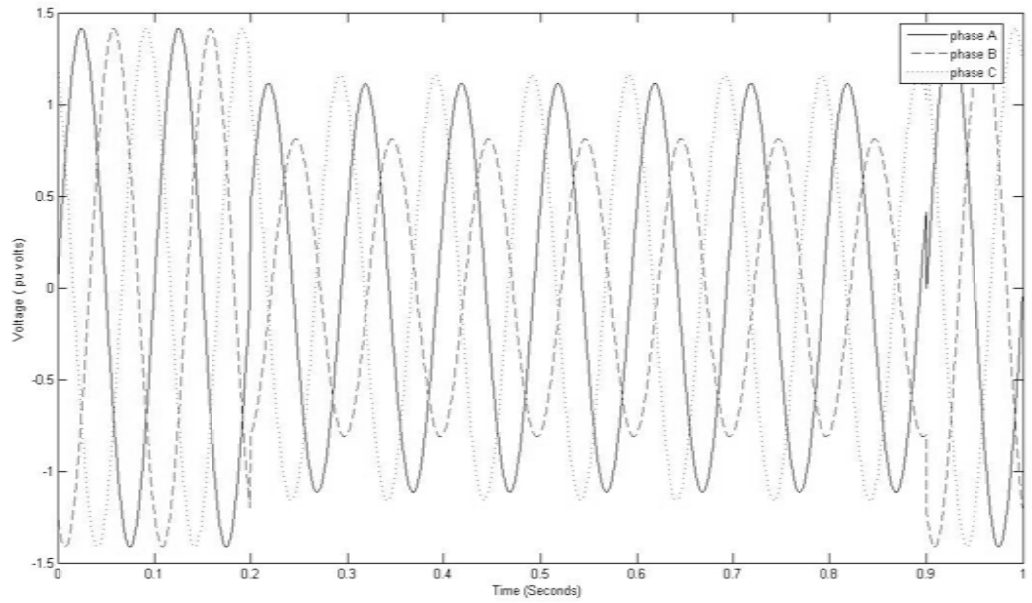


Figure 3.4 Power Signal with a Sag

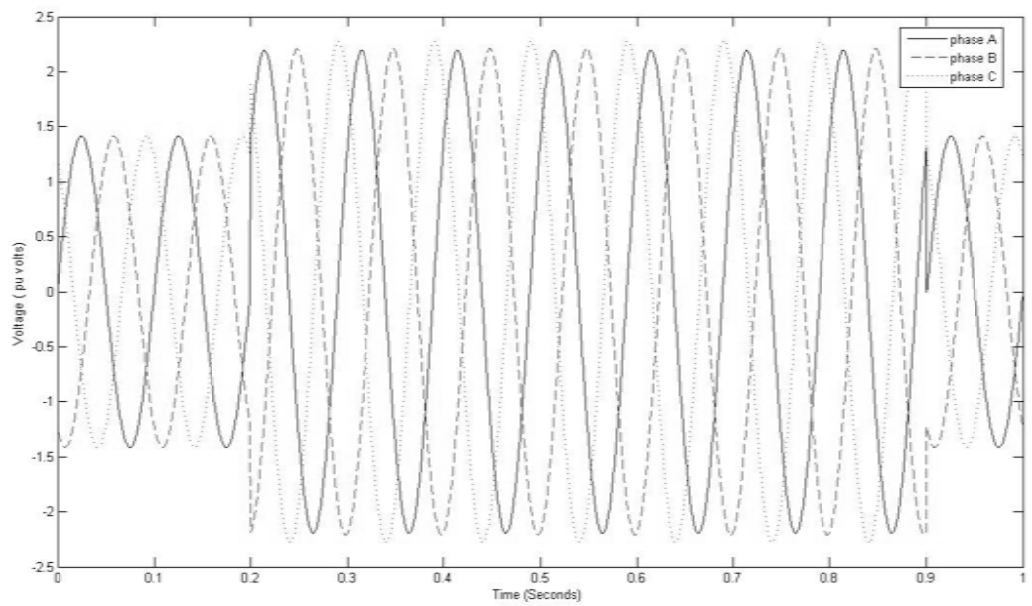


Figure 3.5 Power Signal with a Swell

3.1.6 OUTAGE

An outage can be defined as the reduction in the supply voltage or load current to less than 0.1 pu for a period which sometimes may exceed sixty seconds. These interruptions are measured by their duration since the voltage magnitude is almost 0-10% of its normal value. Fig 3.6 shows a outage in a power signal.

3.1.7 IMPULSE

Impulses are generally spikes that occur due to lightning effects. They can be defined as a momentary, non-power frequency change in the steady state voltage waveform. These are generally unidirectional. They have a very high frequency component and magnitude.

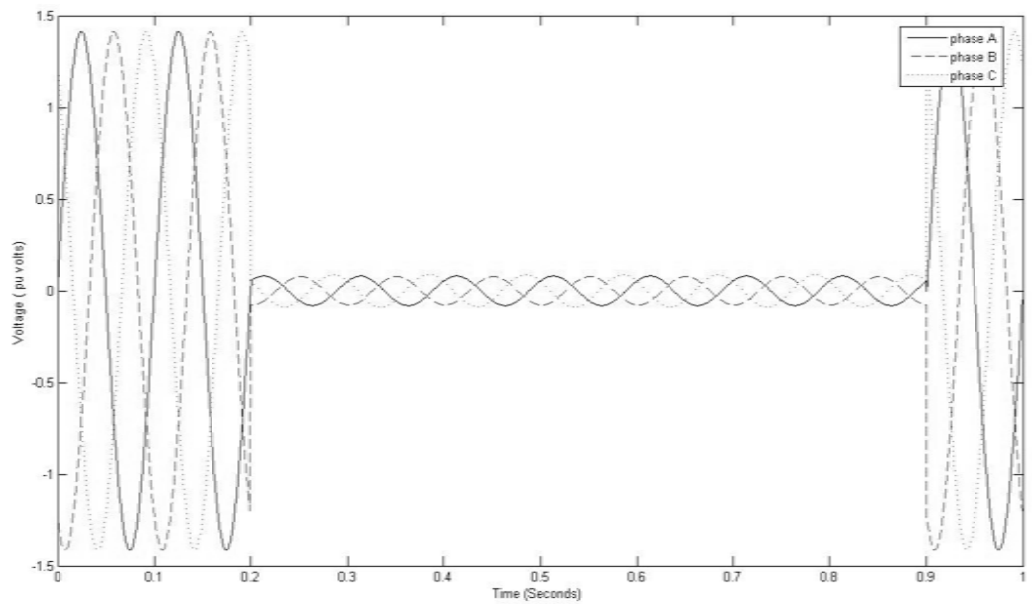


Figure 3.6 Power Signal with Outage

3.1.8 NOTCH

Notches are steady state power disturbances with high frequencies present in this. These contain sudden spikes which naturally give rise to the high frequency content. Notches generally occur due to the switching of inductive circuits using solid state switches and they generally cause adjustable speed drives. Figure 3.1 shows a notch in a power signal.

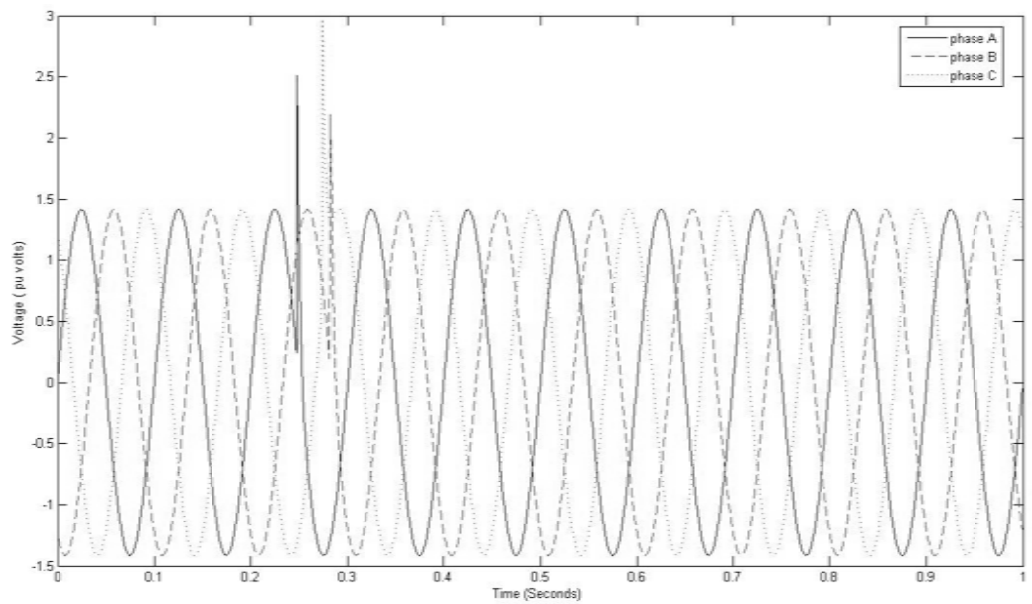


Figure 3.7 Power Signal with a Impulse

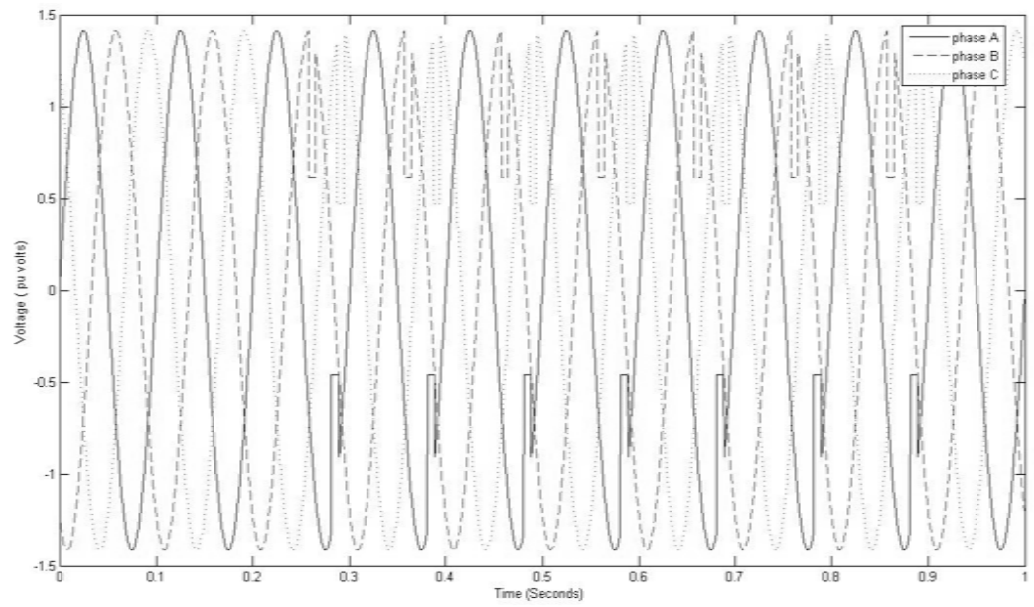


Figure 3.8 Power Signal with Notch

Table 3.1 Spectral content, duration and magnitude of the various power quality disturbances[12]

<i>Categories</i>	<i>Typical spectral content</i>	<i>Typical duration</i>	<i>Typical magnitude</i>
Transients	0.5 – 5 MHz	5 μ sec	0 – 4 pu
Sag	Supply frequency	½ cycles - 1 min	0.1 – 0.9 pu
Swell	Supply frequency	½ cycles - 1 min	1.1 – 1.9 pu
Outage	none	½ cycles – 1 min	< 0.1 pu
Harmonics	Nth multiple of the supply frequency	Steady state	0 – 20 % of supply voltage
Notch	> 5 MHz	Steady State	60 – 80 % of supply voltage
Impulse	> 5 MHz	1 cycle to steady state	Sharp increase in the supply voltage.
Flicker	Within the range of supply frequency	1 cycle to steady state	50 % – 150 % of supply voltage.

Table 3.1 summarizes the data from [12] presented in the above paragraphs. It presents the frequency duration and the magnitude of each of the power quality disturbance signal that is considered in this thesis.

3.2 DISCRETE WAVELET TRANSFORM

Wavelets provide efficient and fast algorithms to represent a signal split in its distinct frequency bands using multi resolution analysis. In core signal processing applications like audio and video compression etc, the properties like orthogonality, symmetry help in reconstruction of the signal with minimal error. Chapter 2, already dealt with the various application of wavelets in different branches of signal processing. This chapter provides with fundamentals of wavelets followed by multi resolution analysis and the detection algorithm for power quality disturbance detection.

A time representation of a signal $f(t)$ can be represented as in (4) and its frequency representation or the Fourier domain representation will be as in (5). We can observe that (4) gives information of maximum time resolution and no frequency information conversely (5) provides the frequency information and no time localization is available as proposed in [14].

$$f(t) = \int_{-\infty}^{\infty} f(u)\alpha(t - u)du \quad (4)$$

$$f(w) = \int_{-\infty}^{\infty} f(t)e^{iwt} dw \quad (5)$$

Where the range is from minus infinity to plus infinity, so whenever the component with frequency w appears in time, it will affect the result of the integration equally as well. The lack of time information in Fourier transform gives rise to a Windowed Fourier Transform. In this case, the signal is divided into small segments, where these segments can be assumed to be stationary. But to make the signal stationary we might need to narrow the window and this results in a poorer frequency resolution.

Wavelets can be defined as a class of functions used to localize a given signal in both time and frequency domains. This set of wavelets would be constructed from a mother wavelet, which dilated or expanded to change the size of the window. This implies that a dilated wavelet gives more of the time information and an expanded version of it looks into the frequency information. Thus wavelets adapt to both high frequency and low frequency components automatically by using various window sizes.

Wavelets as defined above, are generated from a mother wavelet Ψ , using dilations and expansions as proposed in [15].

$$\Psi_{a,b}(x) = \frac{1}{|a|} \Psi\left(\frac{x-b}{a}\right) \quad (6)$$

$$\text{where } \Psi \text{ must satisfy } \int \Psi(x) dx = 0 \quad (7)$$

3.3 CHOICE OF THE WAVELET

The choice of the mother wavelet plays an important role in the extraction of the required features. Several wavelets have been considered for the decomposition of the power disturbance signals. They are, family of Daubechie wavelets (db4, db6, db8, db10), Symlets, Coiflets and Bi-orthogonal wavelets [14], [15], [17]. Usually the choice of the mother wavelet depends on the type of the disturbance signal to be analyzed. At low level decompositions, i.e. highest frequency decomposition the mother wavelet is most localized in time and oscillates very rapidly in a very

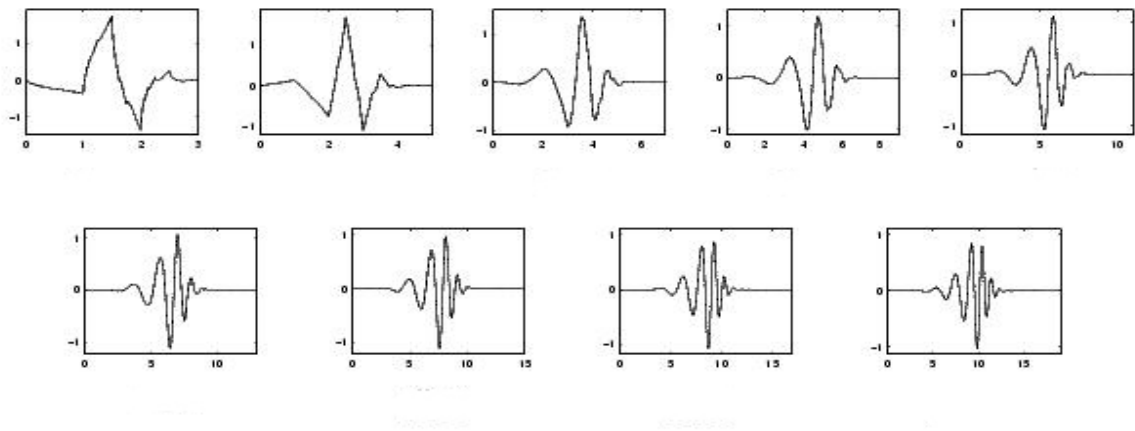


Figure 3.9 Various forms of Daubechie mother wavelets generated using Matlab[45]

short interval of time, and at higher levels of decomposition the wavelet becomes less localized in time and oscillate less due to the dilation nature of the DWT. As a result, fast and shorter disturbances will be detected at lower levels and slow and long durations variations will be detected at higher levels. It is observed that Daubechie4 wavelets are better suited for both short – fast transients as well as slow and steady state disturbances. Hence Daubechie4 is chosen as the mother wavelet in this thesis.

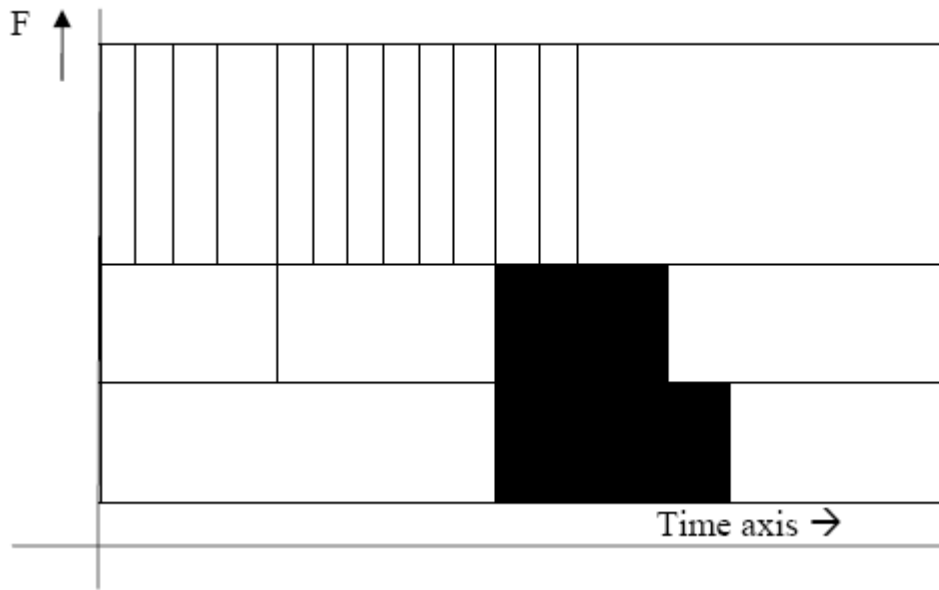


Figure 3.10 Frequency – Time Resolution in case of wavelet domain

3.4 MULTI RESOLUTION ANALYSIS

Multi-resolution Analysis is used to decompose any signal and represent it at different other resolutions. The concept of wavelet multi-resolution analysis is that the power signal $s(n)$, can be approximated step by step in such a way that $s(n)$ is passed through a low pass filter $l(n)$ and a high pass filter $h(n)$ and with the output from the low-pass filter the process repeats for the given number of decomposition levels. The goal is to develop representations of a complicated signal $s(n)$ in terms of several simpler ones and analyze them. This helps in achieving two important properties. The first is the localization property in time of the disturbance signal and the other one is the presence of specific frequencies at different resolution levels. Therefore $s(n)$ will be approximated in different precision in different steps. The number of decompositions is chosen by trial

and error. If the number is too small the decomposed sub-series can't show obvious regularities, and if the number is too big the result might get aggravated.

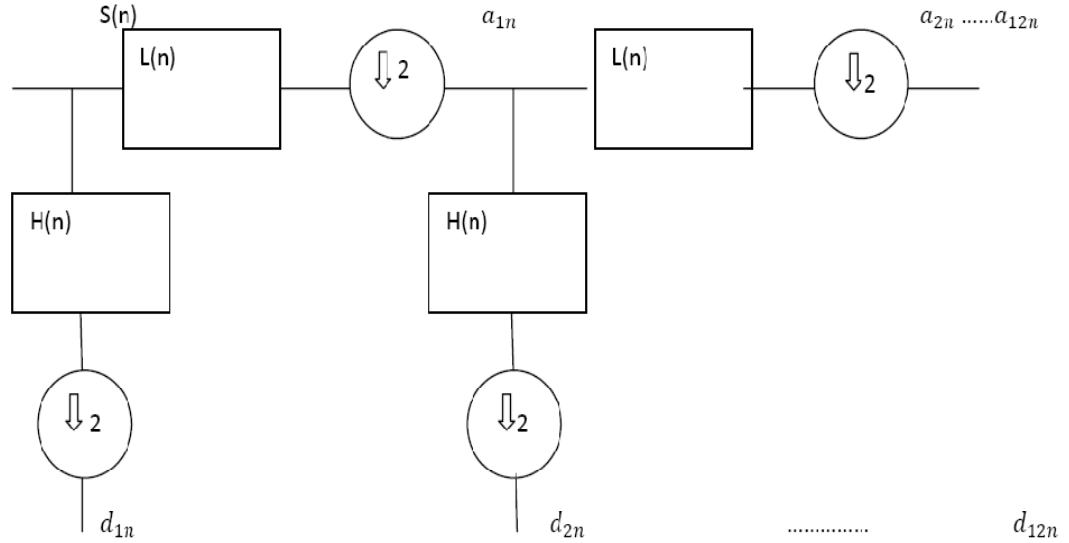


Figure 3. 11 Functional representation of Multi-Resolution Analysis of a signal $S(n)$

With the use of scaling and wavelet functions the signal is mapped into the wavelet domain and analyzed into an approximate and detail coefficients A_j and D_j respectively, where

$$A_j(k) = \sum_m l(m - 2k)s_{j+1}(m) \quad (8)$$

$$D_j(k) = \sum_m h(m - 2k)s_{j+1}(m) \quad (9)$$

The reconstruction process uses the approximation and detail coefficients, $A_j(k)$ and $D_j(k)$ at resolution j to reconstruct the coefficient $s_{j+1}(m)$ at the next resolution level, $j+1$

$$s_{j+1}(n) = \sum_k l(n - 2k)s_j(k) + \sum_k h(n - 2k)s_j(k) \quad (10)$$

The figures 3.11 to 3.17 shows the wavelet domain multi-resolution analysis for the various power quality disturbance signals considered in this thesis work. We can observe that the detail level decomposition acts as a high pass filter with cut-off proportional to the sampling frequency of the signal waveform.

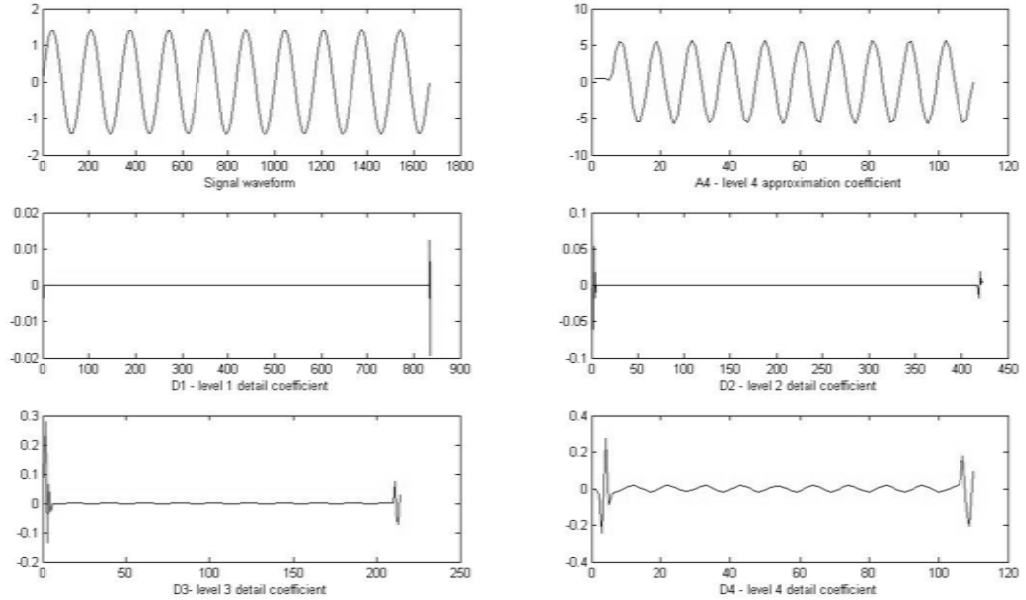


Figure 3.12 MRA for a normal waveform, D1,D3,A4,D2,D4 in that order

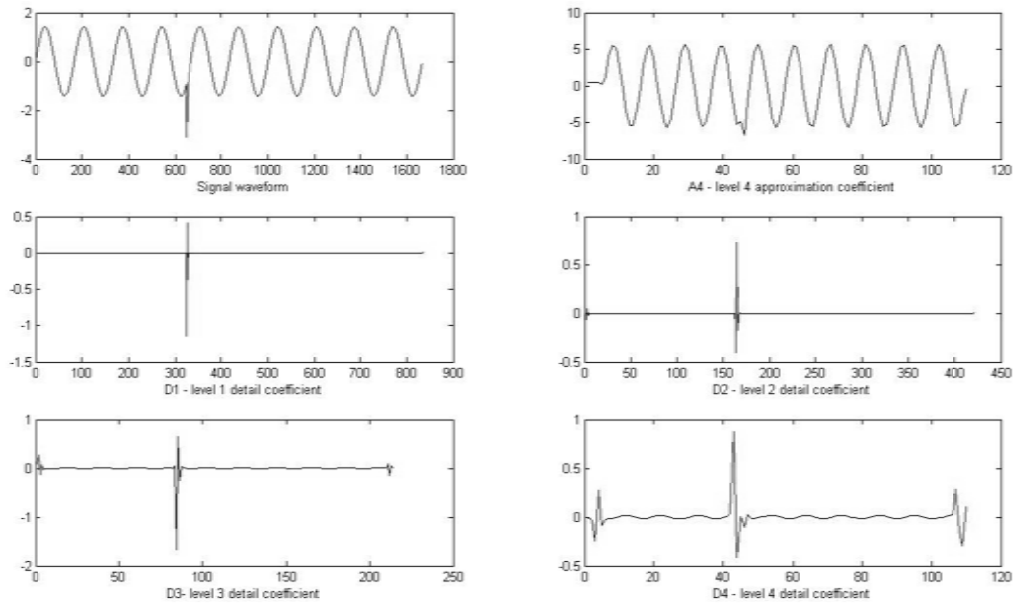


Figure 3.13 MRA for a Flicker waveform, D1,D3,A4,D2,D4 in that order

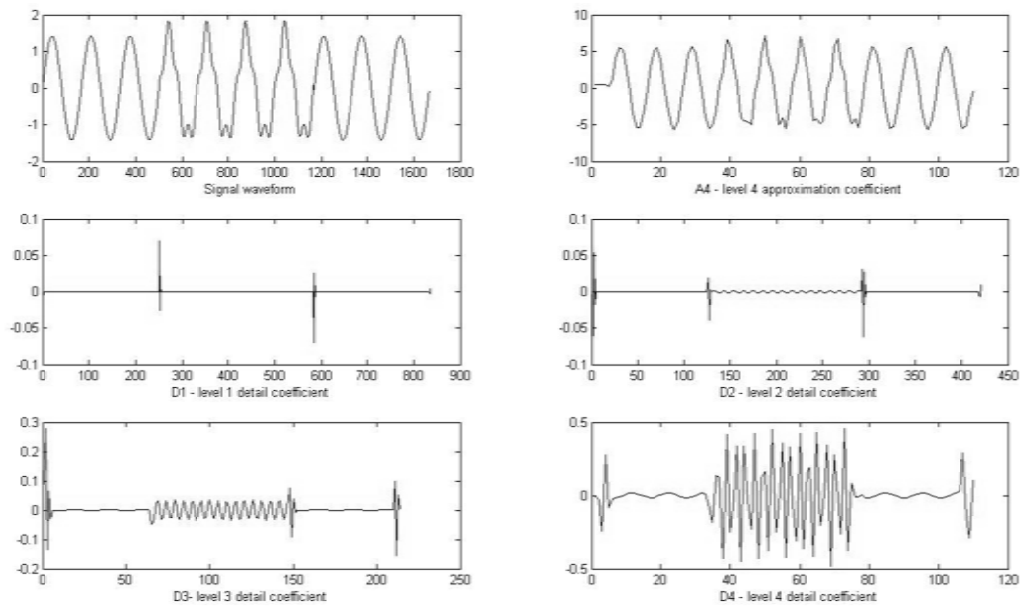


Figure 3.14 MRA for a Harmonic waveform, $D1, D3, A4, D2, D4$ in that order

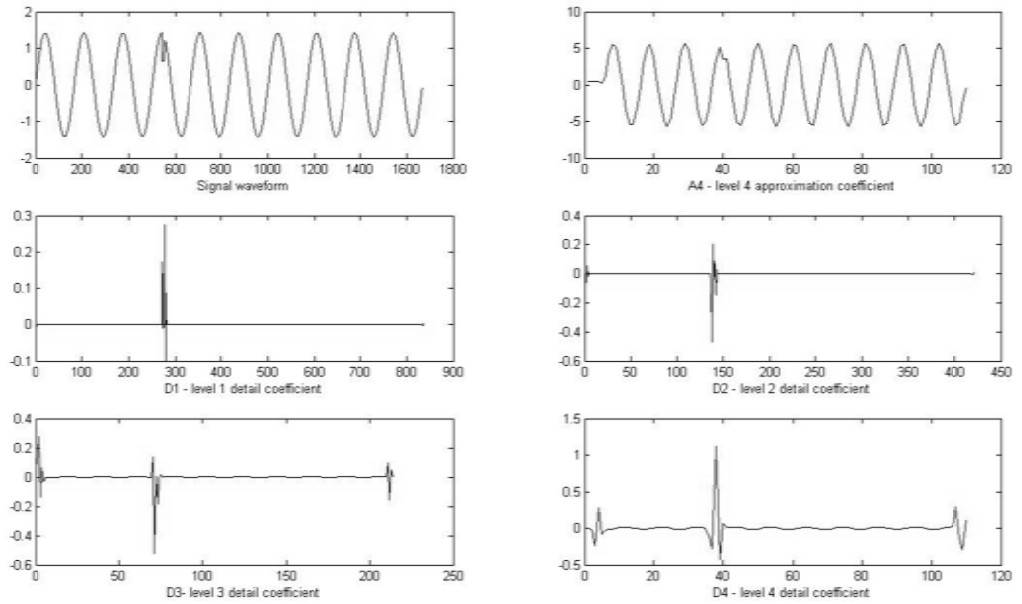


Figure 3.15 MRA for a Notch waveform, $D1, D3, A4, D2, D4$ in that order

Table 3.2 shows the various cut-off frequencies for the resolution levels considered in the above MRA. The spectral components of the various disturbance signals can be observed in table 3.1.

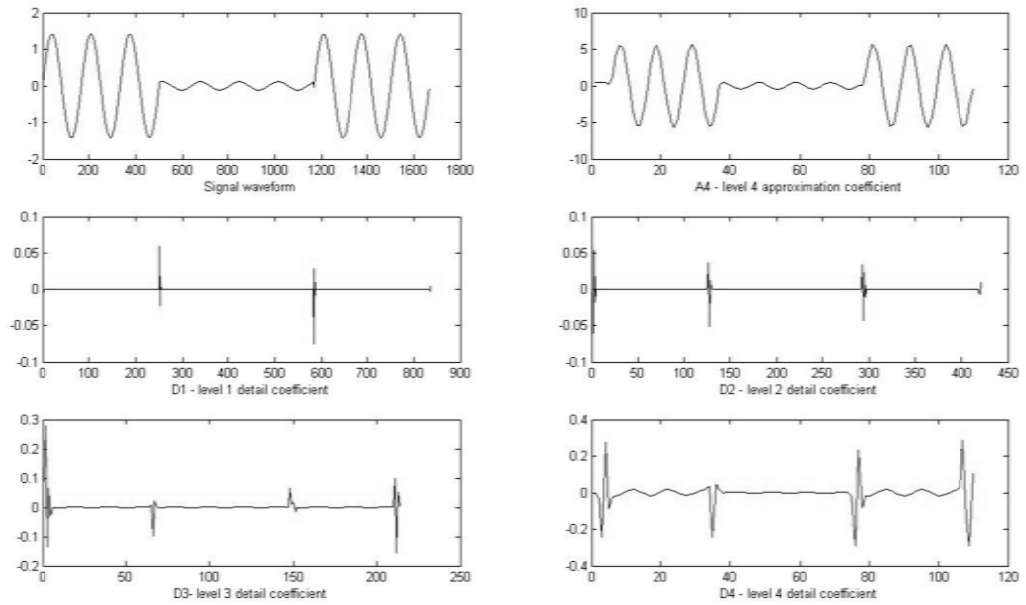


Figure 3.16 MRA for a Outage waveform, $D1, D3, A4, D2, D4$ in that order

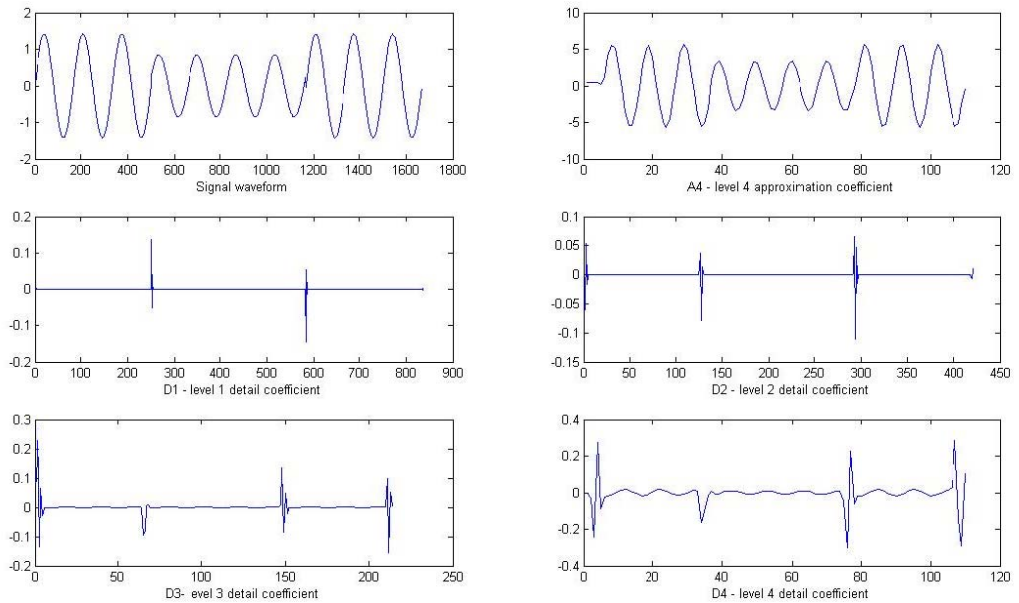


Figure 3.17 MRA for a Sag waveform, $D1, D3, A4, D2, D4$ in that order

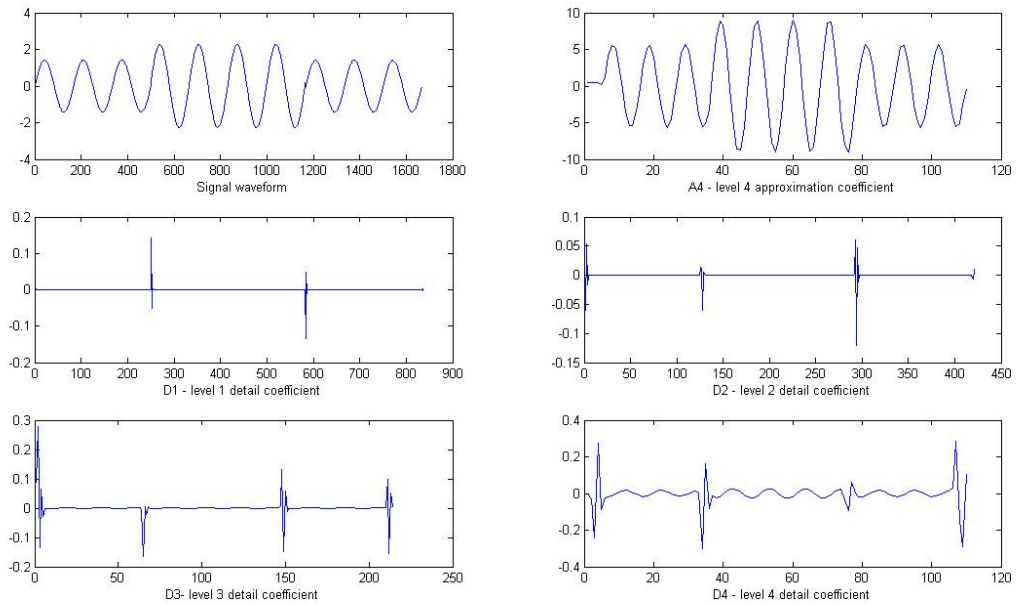


Figure 3.18 MRA for a Swell waveform, $D1, D3, A4, D2, D4$ in that order

Hence we can conclude that with an appropriate resolution level, we can be able to detect and localize the power quality disturbance present in the signal. However, there are a few steady state disturbances like harmonics which are stationary and hence be better detected using the Fourier domain analysis. In Section 3.4 a classification algorithm which uses both Wavelet and Fourier domain for detection and localization of the disturbances is proposed.

As discussed above, the detail coefficients for a particular level represents the output of the high pass filter with cut off frequency equal to the sampling frequency / $2^{level-1}$. Table 3.1 tabulates the various frequency ranges that are covered by the various levels of wavelet multi-resolution analysis. For a sampling frequency of 2^{14} , the various cut-off frequencies are tabulated below.

Table 3.2 Frequency cut-off ranges for high pass coefficients at different MRA levels.

FREQUENCY CUT OFF RANGES FOR HIGH PASS COEFFICIENTS AT DIFFERENT MRA levels				
Resolution level	Detail Coefficient range		Resolution level	Detail coefficient range
Level 1	8192 Hz		Level 5	512 Hz
Level 2	4096 Hz		Level 6	256 Hz
Level 3	2048 Hz		Level 7	128 Hz
Level 4	1024 Hz		Level 8	64 Hz

3.5 DETECTION ALGORITHM

In this proposed algorithm we will explore the various level multi-resolution analysis features to find the detail coefficient's threshold value. Generally as a single frequency window may not be able to cover all the spectral content of the disturbance signal, a 4 level MRA coefficients is considered. However, as already discussed wavelet domain is most efficient for non-stationary signals, if we have a steady state disturbance Fourier domain resolution is much efficient compared with wavelet domain. Hence initially the fundamental frequency and the phase angle shift for each window are also computed.

In Fourier domain, the FFT provides information about the frequency content of a signal by resolving into n bins, where the number of bins determines the accuracy, or resolution. Using FFT, compute the Fundamental frequency and the phase angle shift of that particular window and then a search algorithm will find the coefficient which are not in the acceptable range of magnitude and the phase angle shift. Then the disturbance can be localized by mathematically computing which window that particular coefficient belongs to.

Equations (8), (9), (12) and (14) presents the mathematical formulas for computing the 4 level detail coefficients for wavelet multi-resolution analysis and the fundamental frequency and the phase angle shift in Fourier domain. The flow chart in Fig 3.18 shows the detection algorithm. Chapter 6 tabulates the efficiency of the algorithm.

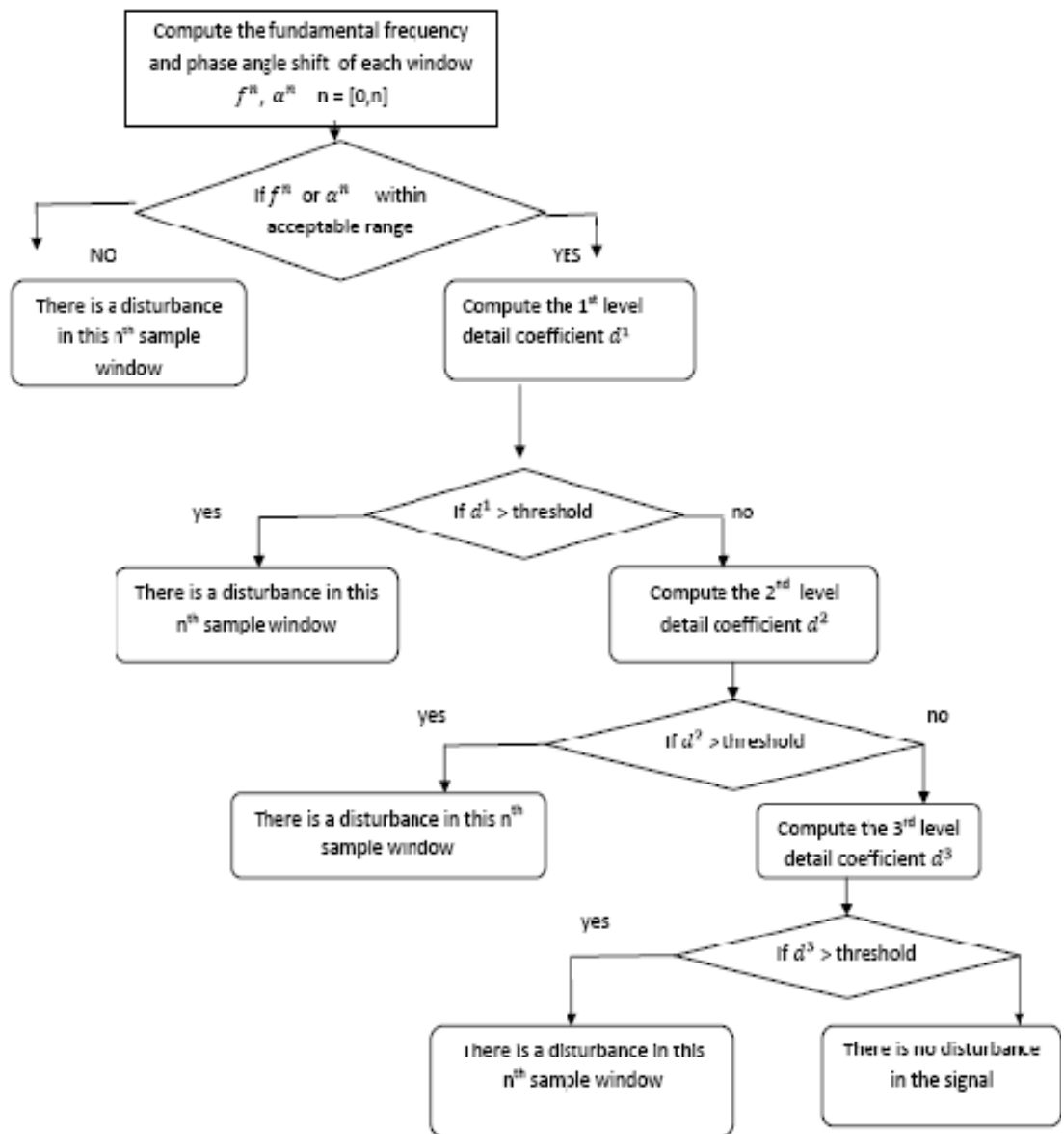


Figure 3.19 Detection and localization algorithm, flow chart

CHAPTER FOUR : PROPOSED POWER QUALITY CLASSIFICATION ALGORITHM

The detection and localization of the disturbance signal is explained in chapter 3. The main objective of this thesis is the automated classification of the power quality disturbances in such a way that measures for eventual improvement in the quality of the power can be incorporated. This chapter proposes a new set of feature vectors which can be used as an input to a proposed classifier. After examining the various combinations, a Feed forward neural network is used as a classifier in this chapter.

4.1 PARSVEL'S THEOREM

Parsvel's theorem states that if the used scaling function and the wavelets form an orthonormal basis, then the energy of the distorted signal is related to the energy of each expansion components and their wavelet coefficients. This implies that the energy of the signal can be partitioned in terms of its expansion coefficients and still be wholly represented as its unique original form. (11) represents the mathematical representation of Parsvel's theorem where $f(t)$ is the signal under consideration and $A_m(k)$ is the m^{th} level approximation coefficient where 'm' is the maximum resolution level considered, and $D_j(k)$ is the detail level coefficient for all values of j in the range [0, m]

$$\int |f(t)|^2 dt = \sum_{k=0}^{length(a_m(k))} |A_m(k)|^2 + \sum_{j=0}^m \sum_{k=0}^{length(d_j(k))} |D_j(k)|^2 \quad (11)$$

[23] proposed a method of using multi-resolution analysis curves for the classification of power quality disturbances. Energy of the distorted signal will be partitioned at different resolution levels based on the power quality problem being

considered. Assuming a zero mean, the standard deviation can be considered as a measure of the energy of the considered signal. So in this thesis along with the other features STD-MRA curves explained in the next paragraph are used. Hence this thesis exploits the concept that the energy of the entire signal can be instead represented as the sum of the energies of all the resolution coefficients.

4.2 STANDARD DEVIATION-MULTI RESOLUTION ANALYSIS CURVES

The energy of the power quality disturbance at various levels varies depending on the type of the disturbance. For a normal sinusoidal waveform, the standard deviation value is equal to the energy of the signal (as the mean is zero). Hence, the standard deviation value of the various levels of multi-resolution analysis can give us a comparative indication of any disturbance present within power signal as shown in equation (12).

$$d_j = \sqrt{\frac{\sum_{k=0}^{length(D_j(k))} |D_j(k) - \mu(D_j)|^2}{length(D_j)}} \quad (12)$$

Where $length(D_j)$ is the total number of samples present in D_j and $\mu(D_j)$ is the average value of the signal. As seen in the figures 4.1 to 4.4, the curves change with respect to the variation in the magnitude and duration of the disturbance. The parts to the left and the right of the peak will change according to the changes in the frequency content of the disturbance signal. The left most part is the resolution level with the highest frequency and the right most part is the resolution level with the lowest frequency. So observing from figures 4.1 to 4.4, it is clear that STD-MRA curve values

can be a distinct parameter that can be used in the feature vector matrix, as an input to the neural network.

In figures 4.1 to 4.4 the comparison of the standard deviation values of all the 12 levels of resolution levels with the same values for pure sine waves is shown. It can be observed that when there is a disturbance of high frequency the energy level at the corresponding resolution level has a spike. In case of steady state disturbances like Sag, Swell and Outage the energy corresponding to the resolution level at supply frequency has a dip and hike respectively. These figures show that standard deviation value curves obtained through multi-resolution analysis proves to be an efficient way of feature extraction. Along with these values other frequency domain feature extraction parameters proposed by [4] are also used in this thesis.

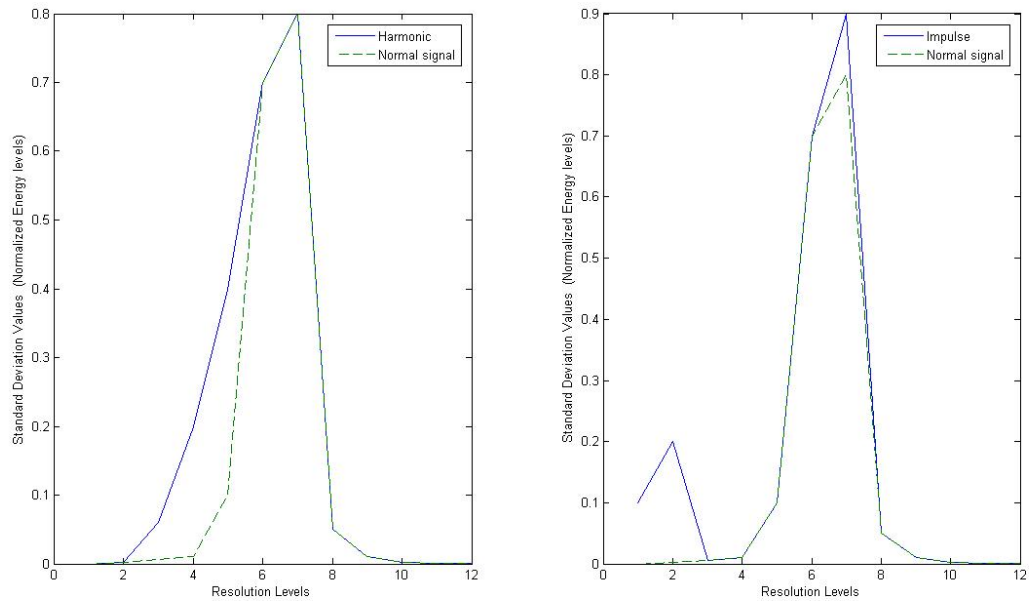


Figure 4.1 Multi-Resolution Energy Distribution Curves for a Harmonic Disturbance & a Impulse

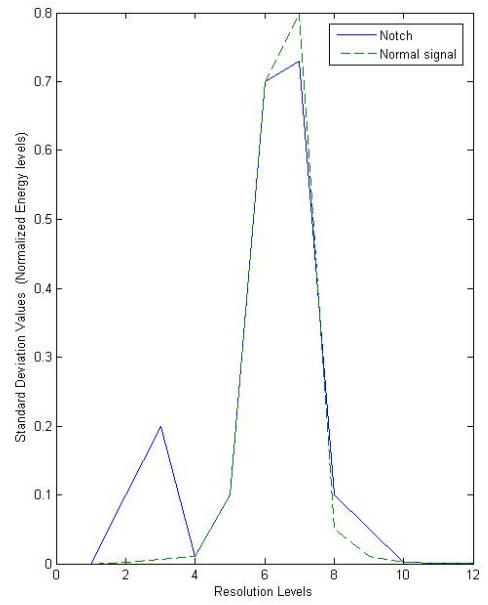
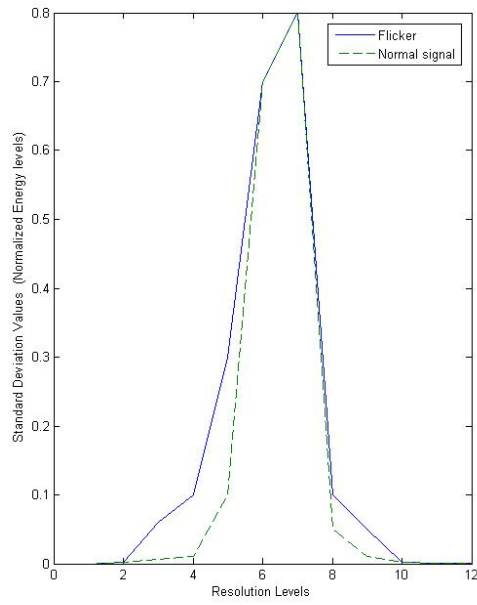


Figure 4.2 Multi-Resolution Energy Distribution Curves for a Flicker and a Notch

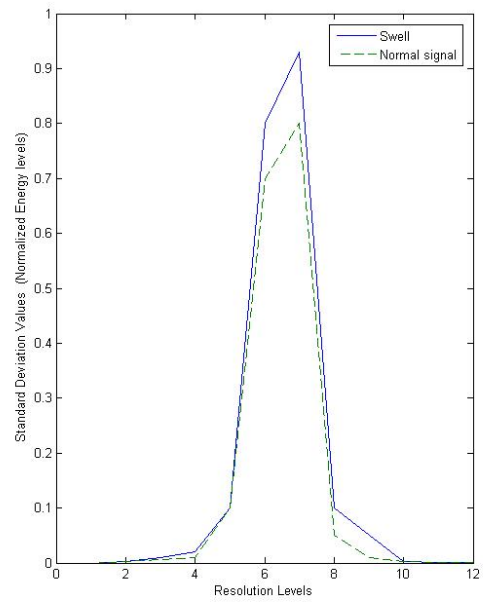
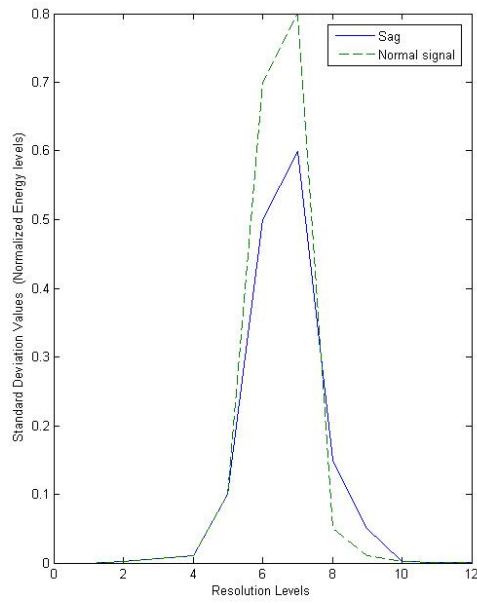


Figure 4.3 Multi-Resolution Energy Distribution Curves for a Sag & a Swell

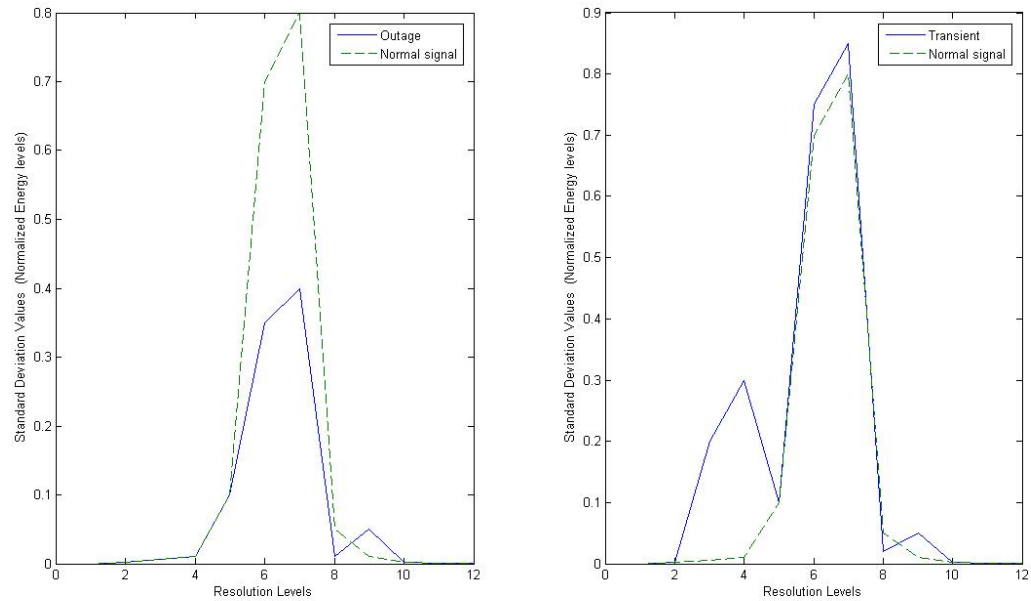


Figure 4.4 Multi-Resolution Energy Distribution Curves for an Outage and a Transient

4.3 FEATURE VECTOR EXTRACTION ALGORITHM

From [4] and [35] and based on various test simulations run on Matlab the best feature extraction parameters are proposed. They are

- 12 values of Standard deviation for the 12 level MRA detail coefficients.
- Fundamental component.
- Phase angle shift.
- Total harmonic distortion.
- Oscillation number of the missing voltage.
- Lower harmonic distortion.
- Variation of the RMS voltage.

All the frequency domain feature extraction parameters, Fundamental component, Phase angle shift, Total harmonic distortion, Oscillation number of the missing voltage, Lower

harmonic distortion, Oscillation number of the RMS variation are initially proposed by [4]. The efficiency of using these parameters for the power quality disturbances considered in this thesis are well documented in [4].

The fundamental component of any frequency component can be defined as the magnitude of the DC component in the Fourier domain. Hence it can be obtained as

$$V_n = 1.414 * abs(V^n[1])/N \quad (13)$$

Where $V^n[1]$ is the DC component or the first value in the discrete Fourier transform series. The windowed DFT V^n of a signal $v(t)$ for a window 'n' of 'L' samples can be computed as

$$V^n = \sum_{i=0}^{N-1} v(i + (n - 1) * N) e^{-j(\frac{2\pi ki}{N})} \quad (14)$$

Phase angle shift can be computed as

$$\alpha_n = angle(V^n[1]) - angle(V^1[1]) \quad (15)$$

In frequency domain total harmonic distortion can be computed as

$$THD_n = \frac{\sqrt{\sum_{k=2}^{floor(\frac{t}{2})} \{abs(V^n[k])\}^2}}{V^1[1]} \quad (16)$$

Lower harmonic distortion can be computed as

$$LHD_n = \frac{\sqrt{\sum_{k=2}^{10} \{abs(V^n[k])\}^2}}{V^1[1]} \quad (17)$$

Oscillation number of the missing voltage is computed as

$$OSM_n = root(v_{missing}^s) \quad (18)$$

where $root(\cdot)$ is the function which returns the number of zero crossings of the function.

Variation of the RMS voltage is computed as

$$OSR_n = \sqrt{(V_{rms}^s - avg(V_{rms}^s))} \quad (19)$$

Where $v_{missing}^s$ and V_{rms}^s are computed from equations as proposed in [4] as

$$v_{missing}^s(i) = v(i) - \frac{2}{N} * abs(V^1[1]) * \cos\left(\text{angle}(V^1[1]) + 2 * \frac{\pi(i-1)}{N}\right) \quad (20)$$

$$V_{rms}^s = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} v^2 [i + (n-1) * N]} \quad (21)$$

Where $v(i)$ is the voltage signal for $i = [0, N]$ where N is the length of the signal and $V^1[1]$ is the same as that used in (12).

In the feature vector algorithm, for each of these parameters, the parameters of the window just before the occurrence of the disturbance is also considered. This way a distinct set of features can be obtained. The feature vector hence consists of the twelve level detail coefficients and the six fourier domain parameters taken from [4] and the another set of six fourier domain parameters which belong to the window just before the occurrence of the disturbance. Hence for each power quality signal considered the extracted matrix looks like as in (21), where $d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9, d_{10}, d_{12}, a_{12}$ are the values computed as shown in section 4.1 and $V_n, \alpha_n, THD_n, LHD_n, OSM_n, OSR_n, V_{n-1}, \alpha_{n-1}, THD_{n-1}, LHD_{n-1}$ are the computed from equations (12), (13), (14), (15), (16), (17), (18) respectively.

$$\text{Input} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \\ d_7 \\ d_8 \\ d_9 \\ d_{10} \\ d_{11} \\ d_{12} \\ a_{12} \\ V_n \\ \alpha_n \\ THD_n \\ LHD_n \\ OSM_n \\ OSR_n \\ V_{n-1} \\ \alpha_{n-1} \\ THD_{n-1} \\ LHD_{n-1} \end{bmatrix}$$

(22)

Where j is any given signal and ' n ' is the detected window.

$$NN_{input} = \left(\begin{array}{cccccccc} d_{1,1} & d_{2,1} & d_{3,1} & d_{4,1} & & & & d_{m,1} \\ d_{1,2} & d_{2,2} & d_{3,2} & d_{4,2} & & & & d_{m,2} \\ d_{1,3} & d_{2,3} & d_{3,3} & d_{4,3} & & & & d_{m,3} \\ d_{1,4} & d_{2,4} & d_{3,4} & d_{4,4} & & & & d_{m,4} \\ d_{1,5} & d_{2,5} & d_{3,5} & d_{4,5} & & & & d_{m,5} \\ d_{1,6} & d_{2,6} & d_{3,6} & d_{4,6} & & & & d_{m,6} \\ d_{1,7} & d_{2,7} & d_{3,7} & d_{4,7} & & & & d_{m,7} \\ d_{1,8} & d_{2,8} & d_{3,8} & d_{4,8} & & & & d_{m,8} \\ d_{1,9} & d_{2,9} & d_{3,9} & d_{4,9} & & & & d_{m,9} \\ d_{1,10} & d_{2,10} & d_{3,10} & d_{4,10} & & & & d_{m,10} \\ d_{1,11} & d_{2,11} & d_{3,11} & d_{4,11} & & & & d_{m,11} \\ d_{1,12} & d_{2,12} & d_{3,12} & d_{4,12} & & & & d_{m,12} \\ a_{1,12} & a_{2,12} & a_{3,12} & a_{4,12} & \dots & \dots & \dots & a_{m,12} \\ V_{1,n} & V_{2,n} & V_{3,n} & V_{4,n} & & & & V_{m,n} \\ \alpha_{1,n} & \alpha_{2,n} & \alpha_{3,n} & \alpha_{4,n} & & & & \alpha_{m,n} \\ THD_{1,n} & THD_{2,n} & THD_{3,n} & THD_{4,n} & & & & THD_{m,n} \\ LHD_{1,n} & LHD_{2,n} & LHD_{3,n} & LHD_{4,n} & & & & LHD_{m,n} \\ OSM_{1,n} & OSM_{2,n} & OSM_{3,n} & OSM_{4,n} & & & & OSM_{m,n} \\ OSR_{1,n} & OSR_{2,n} & OSR_{3,n} & OSR_{4,n} & & & & OSR_{m,n} \\ V_{1,n-1} & V_{2,n-1} & V_{3,n-1} & V_{4,n-1} & & & & V_{m,n-1} \\ \alpha_{1,n-1} & \alpha_{2,n-1} & \alpha_{3,n-1} & \alpha_{4,n-1} & & & & \alpha_{m,n-1} \\ THD_{1,n-1} & THD_{2,n-1} & THD_{3,n-1} & THD_{4,n-1} & & & & THD_{m,n-1} \\ LHD_{1,n-1} & LHD_{2,n-1} & LHD_{3,n-1} & LHD_{4,n-1} & & & & LHD_{m,n-1} \end{array} \right)$$

Figure 4.5 Feature Vector matrix for ‘m’ signals - Typical Input matrix to the classifier

Fig 4.5 represents the entire feature vector matrix for ‘m’ signals. Each disturbance signal has 23 features extracted and hence the input matrix when training ‘m’ signals simultaneously would be $23 \times 'm'$.

4.4 FEED FORWARD NEURAL NETWORKS

An artificial neural network can be defined as an interconnection of a group of artificial neurons which are based on a mathematical model used for the sake of information processing. An ANN is implemented as a different layer hierarchy where each neuron in a layer receives the same information at the same point of time and has the same activation function (in most practical cases). It is worth mentioning here that in neural network literature, even the input and the outputs are considered as 2 independent layers and hence any network with just the input and the output layer can be termed as a 2 layer neural network. So it could be understood that any network which is termed as a 3 layer neural network will have a hidden layer apart from the input and the output layers. This hidden layer as implied by the term is not accessible to any outside parameters.

A feed forward neural network as the name implies consists of sets of neurons to which the information is passed in the forward direction (from the input to the output or left to right). A typical feed forward neural network is shown below. We can observe that there are no interconnections between neurons of the same layer.

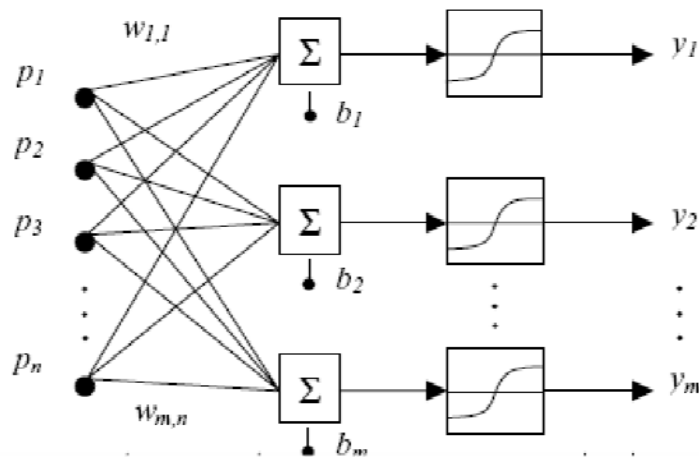


Figure 4.6 A 3 layer neural network with 'n' inputs and 'm' outputs and 1 hidden layer

The computation process involves the scaling of the inputs $p_1, p_2, p_3 \dots p_n$ through the individual channels and weighed by weights $w_{1,1}, w_{1,2} \dots w_{m,n}$. The resultant signals are passed to all the neurons in the network where they are summed together. We can observe that there are bias values $b_1, b_2 \dots b_i$. These values enable the converge of the network during the computation of the error. These signal values are now sent to a transfer function which is also called the activation function and hence the output of the activation function is the output of the neural network.

From the description it could be inferred that if there are ‘n’ neurons in the network there would be ‘n’ output functions. However the number of inputs is always independent of the number of neurons.

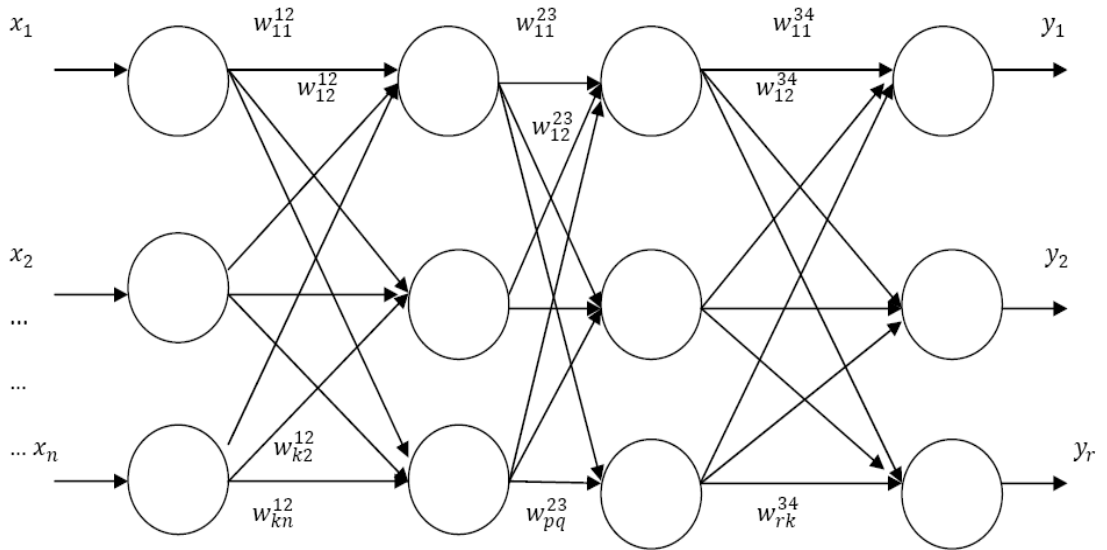


Figure 4.7 A 4 layer neural network with 2 hidden layers, n inputs and z outputs

In case of a neural network which contains multiple layers, the outputs from the first layer is connected to the inputs of the second network and to generalize the outputs of the $(n-1)^{\text{th}}$ layer are connected to the inputs of the n^{th} network. The figure below shows a fully interconnected feed forward neural network which consists of three layers and an arbitrary number of layers. A standard convention of representing a neural network of this type is m-r-z, which refers to the number of neurons in each network.

4.5 TRAINING THE ARTIFICIAL NEURAL NETWORK

The most important thing to make the neural network work is to train it. So a neural network can be used to classify the disturbance patterns only when it had been trained with sufficient number of pre-classified data. This training changes the biases and the weights of the neural network, so that the output meets the required classification values.

Training a neural network can be accomplished in two different ways, either using a supervised learning or an unsupervised learning. Supervised training requires the correct output for a given set of input values. The given set of output values are compared with the resultant output from the neuron and some sort of algorithm is used to change the values of the biases and the weights so that they match the target output values. Unsupervised training on the other hand, requires the network to learn the process on its own. It doesn't involve the training based on a target set of outputs. This sort of ANN are mostly using for data clustering applications, where the data is to be grouped as clusters based on given set of parameters. The first paper implemented in the thesis uses the feed forward neural network with error-back propagation training algorithm. This is a

supervised mode of training algorithm. A brief treatment about this algorithm follows in the next paragraph.

Back Propagation is the most widely used ANN training algorithm for the reason that it is very powerful and at the same time one of the most simplest training algorithm. Back propagation algorithm examines the output of the network and compares it to the target output provided, calculates the error and then the error is once again back – propagated to the network so that the weights and biases are adjusted in a way that the network during the next epoch produces an output which is much closer to the target value. This process is repeated until the desired target is reached (or the desired error or epoch parameters are reached). The back propagation algorithm only deals with propagating the error back to the network. The way the information is processed through the network depends on the learning function used. Hence there are quite a large number of back propagation algorithms proposed, which basically defer only in the way the learning function is implemented.

A critical aspect of classification using neural networks is choosing an appropriate network size for the given application. The network size involves both the number of layers in the network and the number of nodes per layer. Network size greatly effects the ability of the network to produce accurate classification results on signals outside its training set. Hence this thesis implemented various FFNN architectures and came up with the below two architectures which have the optimized error rates and training time for the power quality disturbance signals in the problem

The feed forward neural network with back propagation (FFN-BP) employs a supervised training mode, such that for a input vector, X and a desired output response S ,

must also be presented to the network. The difference in the error between the simulated output Y and the desired output response S is back propagated during the training., and hence based on the number of epochs(and other network parameters like the performance goal etc) the weights are adjusted to achieve the performance goal. Chapter 6, documents the various kinds of FFN-BP structures considered in the classification of these power quality disturbances. The most optimal neural network architecture is obtained from the results and is used for the classification task at hand.

CHAPTER FIVE : CURRENT RESEARCH - CLASSIFICATION USING A NEW FEATURE VECTOR ALGORITHM AND ART NEURAL NETWORKS

The use of STD MRA curves in chapter 4, is a good way to extract the features from power signal disturbances. However computing the energy of the multi-resolution analysis provides a uniformly distributed energy vector, i.e. the amount of data extracted is the same irrespective of the frequency resolution. In case of power signals, some frequencies which have the disturbance signals are more important and hence more values should be extracted from it than from frequencies which have no disturbances associated with it. There is an un-proved and un-objected assumption that all the frequency levels comprise of equal weights. So this thesis work proposed using a new method of Multi-resolution analysis for power quality signals which was initially proposed by Pittner and Kamathi [41] for Flank wear estimation and lung signal sound classification. This method focus on extracting more information from frequency ranges of more importance and less information from frequency ranges of significantly less importance. Thus most of the intrinsic information of the power signal disturbance is retained and hence it is assumed that with a good classification tool the error classification efficiency can be improved. The next few sub-sections explains the method and its relevance to feature extraction in the power quality disturbance scenario. Then the use of ART neural networks [42] – [45] as a classification tool is explored. Adaptive resonance Theory is a well established neural network framework, developed by Carpenter, Grossberg and their collaborators at Boston University [42]. There is an extensive variety of ART architectures. The next sections consists of explanation regarding the proposed feature extraction and the classification tool.

5.1 FEATURE VECTOR EXTRACTION METHOD

In this method, the 12 scale multi-resolution matrix of the computed wavelet coefficients is divided into disjoint clusters equal to row vectors and then calculated the

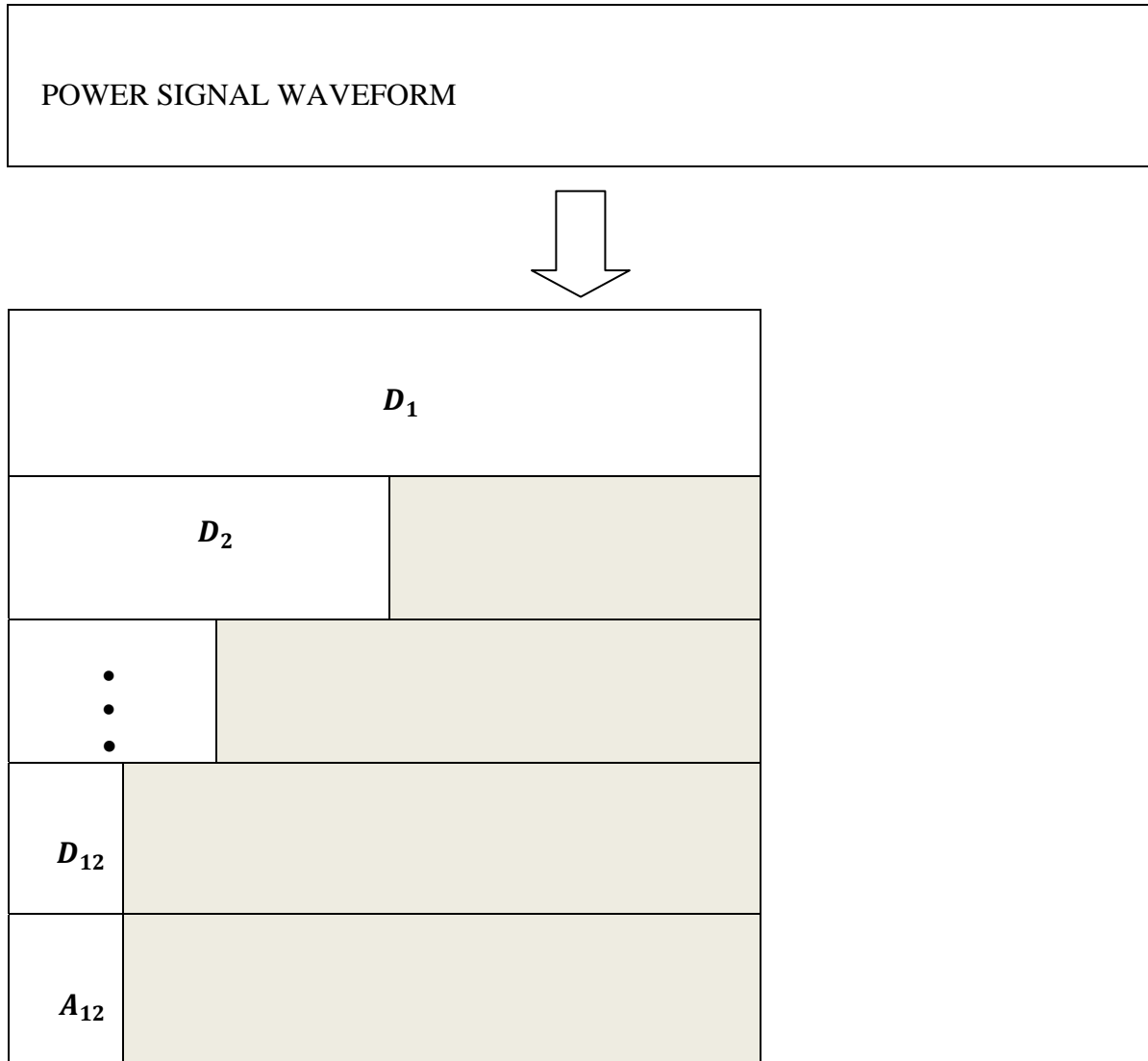


Figure 5.1 Transforming the time domain signal into 12 scale MRA with zero padding
standard deviation value(energy of that cluster) for each of the clusters. Thus, in this way the coefficients at more important scales produce a larger number of clusters and thus can be able to contribute significantly to the feature vector. The algorithm used to cluster the

multi-resolution wavelet coefficient matrix is presented below in 6 steps. Table 5.1 presents a possible result of grouping the clusters using the clustering algorithm.

Step 1

Compute the multi-resolution coefficients for up to 12 levels. The wavelet used is a Daubechie4 and the obtained coefficient values are arranged as in figure 5.1. Let us call this matrix A, where A is a 2 dimensional matrix with 13 rows and columns equals the size of the largest coefficients for any given scale.

Step 2:

Generate matrices A_k for $k = [1,9]$ which implies that a 2 dimensional matrix is to be generated for all the types of signal waveforms.

Step 3:

For the matrix A_k generated in Step 2, compute

$$A_S = \sum_{k=1}^9 A_k \quad (23)$$

Step 4:

Create a Identity Matrix H for A_S computed in Step 3.

Step 5:

After computing the Identity matrix H, now compute

$$G = \frac{1}{\sigma(R(A_S))} (A_S - \mu(A_S) * (H)) \quad (24)$$

Step 6

Compute a binary matrix G_B from G, with a threshold value 'thrhd' using (22) where γ is an arbitrary value adjusted to achieve the desired clustering.

$$thrhd = \sqrt{2(\ln(n) - \ln(\gamma))} \quad (25)$$

Here the two dimensional multi-resolution wavelet coefficient matrix is arranged for each signal as shown in figure 5.1. Here the creation of clusters from the coefficient matrix is guided by the assumption that the regions of two dimensional matrix 'A' containing large-value wavelet coefficients allow a much better replication of the original signal than other regions of the matrix. After the 6 steps, we obtain the binary threshold matrix . As seen in figure 5.2, the clusters of each row vector in is determined in such a way that each cluster contains a 1,

$$\dots \underbrace{0\dots 010\dots 0}_{cluster(i-1)} \underbrace{0\dots 010\dots 0}_{cluster(i)} \underbrace{0\dots 010\dots 0}_{cluster(i+1)} \dots$$

Figure 5.2 Possible way of clustering the threshold matrix

located near the midpoint of the cluster. If the entire row of the matrix contains no 1's, it is treated as a single cluster. This cluster pattern is then superimposed on the main multi-resolution matrix and the standard deviation values for each cluster are determined and is provided as the input to the classification scheme.

Table 5.1 shows a possible result of grouping the clusters using the clustering algorithm written in Matlab. As proposed in chapter 4, a 12 level wavelet multi-level resolution matrix is considered and the clustering mechanism explained above is imposed on the matrix.

Table 5.1 A Possible result of grouping the clusters using the clustering algorithm

Resolution Level	Wavelet Coefficient Cluster
d_1	$\{ d_{1,1}, d_{1,2}, d_{1,3}, \dots, d_{1,1505} \}$
d_2	$\{ d_{2,1}, d_{2,2}, d_{2,3}, \dots, d_{2,752} \}$
d_3	$\{ d_{3,1}, d_{3,2}, d_{3,3}, \dots, d_{3,52} \}, \{ d_{3,53}, d_{3,54}, d_{3,55}, \dots, d_{3,374} \}$
d_4	$\{ d_{4,1}, d_{4,2}, d_{4,3}, \dots, d_{4,22} \}, \{ d_{4,23}, d_{4,24}, d_{4,25}, \dots, d_{4,86} \}, \{ d_{4,87}, d_{4,88}, d_{4,89}, \dots, d_{4,120} \}, \{ d_{4,121}, d_{4,122}, d_{4,123}, \dots, d_{4,182} \}$
d_5	$\{ d_{5,1}, d_{5,2}, d_{5,3}, \dots, d_{5,12} \}, \{ d_{5,13}, d_{5,14}, d_{5,15}, \dots, d_{5,42} \}, \{ d_{5,43}, d_{5,44}, d_{5,45}, \dots, d_{5,52} \}, \{ d_{5,53}, d_{5,54}, d_{5,55}, \dots, d_{5,72} \}, \{ d_{5,73}, d_{5,74}, d_{5,75}, \dots, d_{5,91} \}$
d_6	$\{ d_{6,1}, d_{6,2}, d_{6,3}, \dots, d_{6,12} \}, \{ d_{6,13}, d_{6,14}, d_{6,15}, \dots, d_{6,22} \}, \{ d_{6,23}, d_{6,24}, d_{6,25}, \dots, d_{6,45} \}$
d_7	$\{ d_{7,1}, d_{7,2}, d_{7,3}, \dots, d_{7,12} \}, \{ d_{7,13}, d_{7,14}, d_{7,15}, \dots, d_{7,22} \},$
d_8	$\{ d_{8,1}, d_{8,2}, d_{8,3}, \dots, d_{8,7} \}, \{ d_{8,8}, d_{8,9}, d_{8,10}, d_{8,11} \}$
d_9	$\{ d_{9,1}, d_{9,2}, d_{9,3}, \dots, d_{9,6} \}$
d_{10}	$\{ d_{10,1}, d_{10,2}, d_{10,3} \}$
d_{11}	$\{ d_{11,1}, d_{11,2} \}$
d_{12}	$\{ d_{12,1} \}$

There are around 3010 samples present in a signal waveform that is considered for clustering and hence the first decomposition contains half the number of samples, i.e. 1505. This pattern continues with all the 12 levels of decomposition and based on the threshold values used a possible clustering matrix could look as shown in table 5.1.

After the clusters are obtained the feature extraction mechanism is the same as proposed in chapter 4. However, as the frequency domain components are not considered, the feature vector will be the standard deviation values for each of these clusters. Hence the number of inputs to the classifier will be equal to the number of clusters which in turn depend on the threshold values used.

5.2 ADAPTIVE RESONANCE THEORY NEURAL NETWORKS

Adaptive Resonance Theory was introduced by Grossberg [42] in 1976. It evolved into various flavors, ART-1, ARTMAP, Fuzzy-ART, Fuzzy-ARTMAP [43]. Fuzzy-ART neural network is capable of implementing both analog and digital data and has a much faster implementation compared with other architectures. Historically for neural networks, the algorithms for the binary value input – networks are developed first and they are later extended to the case of analog input patterns. In case of ART neural networks, ART-1 is the first evolved network. ART-1 is an unsupervised learning network which outputs the clustering data sets for binary-valued inputs. ARTMAP evolved from ART-1, it is a supervised learning network for classifying binary-valued input patterns. Fuzzy-ART is a hybrid version of ART-1 which implements unsupervised learning architecture but however, can input analog valued data sets. Fuzzy-ARTMAP as the name indicates is a hybrid of ARTMAP and Fuzzy-ART. It has a supervised learning classification architecture and can input analog valued data sets too.

In architectures like back propagation networks or perceptrons, an input training set is presented sequentially until the network finishes learning the entire training dataset. So when a new pattern is presented to the network, the network has to once again retrain with all the new training patterns. In the process of making the network learn the new

training patterns, if the network forgets the old training pattern, it can't any more classify the old data sets. The ability of a network to learn a new training set pattern is called plasticity, and its ability to retain the previous learning along with the new learning is called stability. The quest for the presence of both stability and plasticity in neural networks led to the development of ART networks[42].

ART-1 is the most computationally simple prototype among the ART networks and hence this thesis investigates the possibility of using ART-1 for the sake of classification of power quality disturbances. [46] proposed the use of ART-1 and Fuzzy-ART algorithms for the classification of power system faults. Hence the feasibility of using Adaptive Resonance Theory neural networks for the classification of power quality disturbances is worth investigating and this thesis is currently focusing on implementing a classification algorithm using ART-1 and Fuzzy-ART architectures. However, such implementation is still under current research and hence this chapter deals with only the insights in using ART neural networks for this classification problem. The use of feature vector extraction method proposed in 5.2 matches well with the resonance property of the ART networks.

Fuzzy-ART neural network is an architecture evolved from ART-1. The main differences between the ART-1 and Fuzzy-ART architectures are that, the input vectors are analog valued in case of a Fuzzy ART. The second difference being, there would only be one set of analog weight vector Z_j in case of Fuzzy ART where as ART-1 consists of two sets of binary weights. The other difference is that fuzzy operators are used instead of the crisp relational operators (like greater-than, less-than) that are used in ART-1. ARTMAP and Fuzzy-ARTMAP are the other hybrid varieties of ART where ARTMAP

is a supervised form of learning for binary input datasets and Fuzzy-ARTMAP is a structure identical to ARTMAP except that the ART1 module has been substituted by Fuzzy-ART structures.

5.2.1 ART-1 ALGORITHM

The ART-1 architecture is a massively parallel neural network structure which self organizes recognition codes in response to a sequence of binary valued input patterns [42]. This network is fed with a data set of binary valued input patterns and the data set gets clustered into a set of most feasible categories in an unsupervised manner. Fig 5.3 shows the architecture of an ART1 neural network as proposed in [43]. F1 and F2 are the two layers of neurons. The F1 layer is provided with 'n' number of inputs, which also contain 'n' nodes. The F2 layer consists of 'm' nodes. All the neuron in F1 layer are connected to all the neurons in the F2 layer through the weights ' $Z_{ji}^{bottom_up}$ ', and all the neurons in F2 are again have a new connection to F1 through ' $Z_{ji}^{top_down}$ '. Index 'i' indicates that connection goes from the *i-th* neuron in F1 layer to index 'j' which indicate that till the *j-th* neuron in F2.

The clustering unit with the largest sum input becomes the eligible candidate to learn the input vector. The algorithm shown in figure 5.4 proposed in [43] explains the process of adjusting the weights till they resonate and hence get the ability to cluster the input data sets. The activations of all the F2 vector are initially set to zero. The decision whether this particular cluster unit is allowed to learn the input vector depends

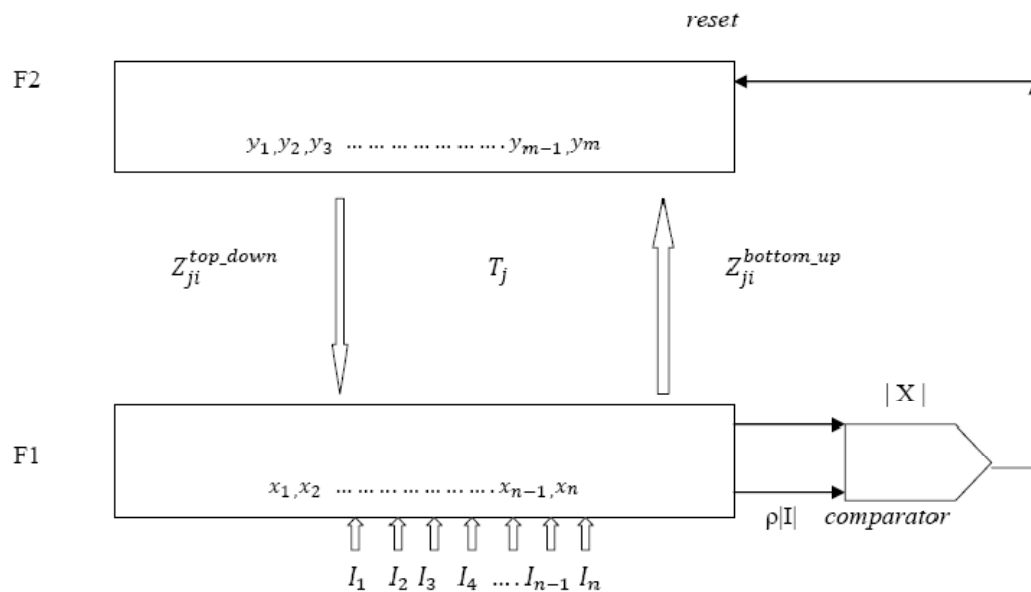


Figure 5.3 Architecture of a ART-1 neural network with two layers, F1 and F2

on the similarity of its top-down weight vector to its input vector. This decision is made by the reset unit depending on the value it receives from the input and the comparator. If the particular cluster is not allowed to learn, it is temporarily isolated and a new cluster unit is selected as the candidate.

The clustering depends on the vigilance parameter ρ . The criteria for an adequate match between an input pattern and the obtained cluster is determined by this vigilance parameter. This always ranges between 0 and 1. A higher vigilance parameter imposes a strict matching criteria, and this in turn partitions the input set into finer and more number of clusters. The reverse is true in case of a low vigilance parameter. The ART-1 algorithm tested in this thesis work didn't yet give efficient classification clusters. Repeated testing of various vigilance parameter values might ' ρ ' bring a better clustering

output but the exact reasons for not being able to achieve the desired efficiency are still under investigation.

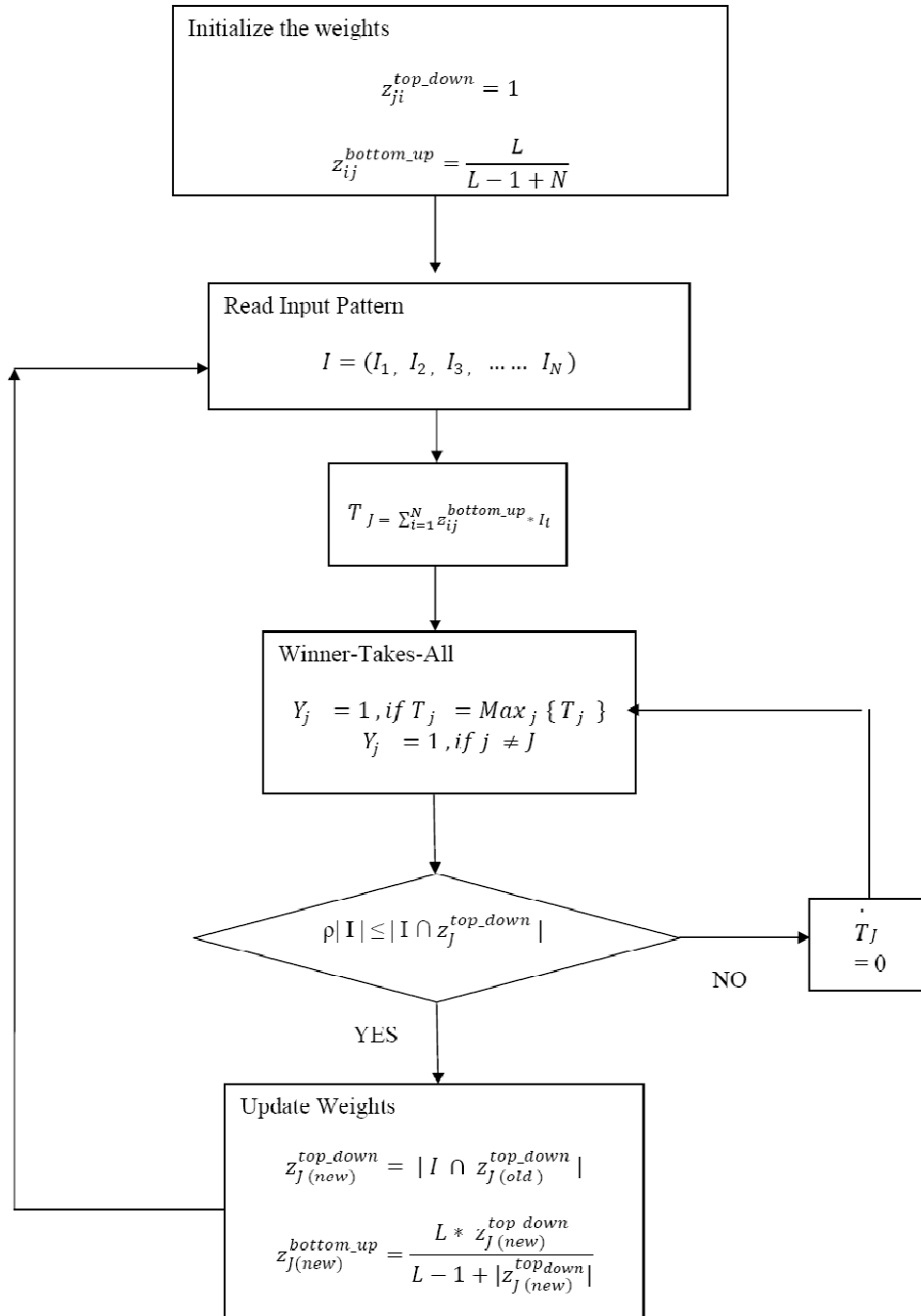


Figure 5.4 Adaptive Resonance Theory 1 functional algorithm [43]

CHAPTER SIX : EXPERIMENTAL RESULTS

This chapter describes the experimental / simulation studies carried out on the proposed algorithms. The creation of the feature vector matrix which is the input to the neural networks is demonstrated and the various kinds of neural networks used and the classification results achieved are tabulated in this chapter.

6.1 EXPERIMENTAL PROCEDURE

The various kinds of disturbances that are discussed in 3.1 are generated using Matlab. This simulation assumes the knowledge about the detection and localization of the disturbance . Hence a 10 cycle window where the disturbance starts at cycle 3 and ends after the 6th cycle are considered. The supply frequency is assumed to be 60Hz and the supply test voltage is assumed as 1 Volt. The signals have a sampling frequency of 10KHz. The hardware used is a Intel Processor 1.7 GHz and a 1 GB RAM. Matlab 7.1 is used for all the simulation studies.

A set of 1000 iterations (which is 9000 different waveforms) and other set of 10,000 iterations (90,000) are generated. The various combinations of neural network activation, training functions and node size are used and the optimal combination is discovered.

Various toolbox functions such as different types of feed forward neural networks, training functions, activation functions, learning functions, initialization and performance functions available in the Matlab toolbox are tested and the most efficient combination is presented. It is also observed that for the given size of the feature vector, a 4 layer network is the best optimal solution for this classification problem. The Table 6.1

tabulates the various combinations of transfer functions, node sizes and training functions used for achieving the most optimal network. It is observed that a 4 layer network is the best optimal solution, as in case of a an 5 layer network even though it provides a much faster convergence towards the performance goal, it requires a lot of memory and hence a 4 layer network is used. It is also observed that the training algorithm is a very important factor in achiving the desired performance and accuracy of the network. Fig 6.1 represents a 3 layer FFNN with Sigmund activation functions('Tansig') and BFGS quasi-Newton back propagation training algorithm. Fig 6.2 represents a 4 layer FFNN with Sigmund activation functions('Tansig') and BFGS quasi-Newton back propagation training

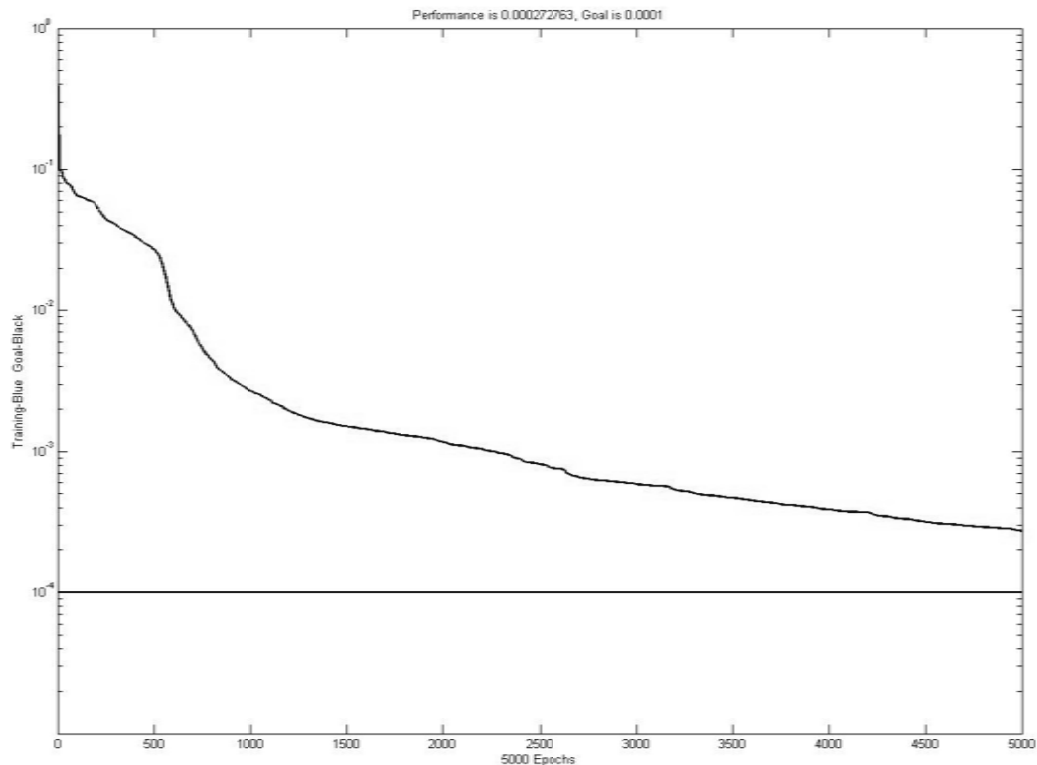


Figure 6.1 Performance function for a 3 layer neural network using 'Trainbfg' function

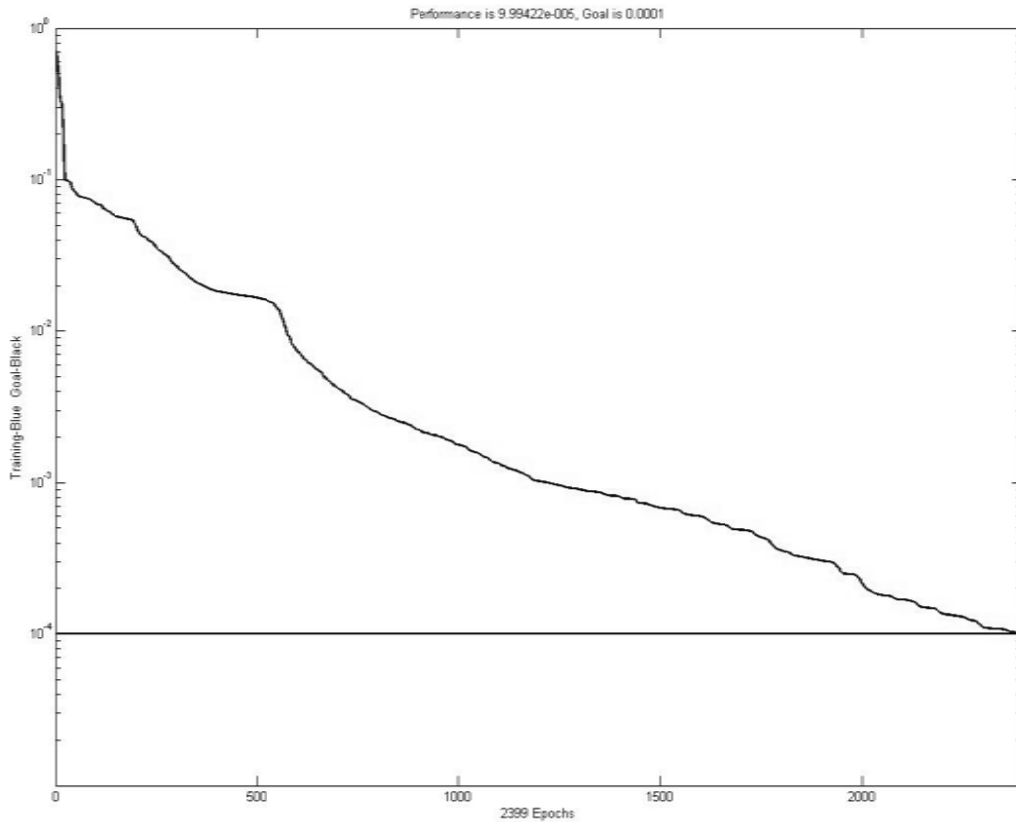


Figure 6.2 Performance function for a 4 layer neural network using 'Trainbfg' function

algorithm. Fig 6.3 represents a 3 layer FFNN with Sigmoid activation functions('Tansig') and BFGS quasi-Newton back propagation training algorithm. From table 6.1, it is concluded that Sigmoid activation function and BFGS quasi-Newton back propagation training algorithm are best suited for our classification purpose and from fig 6.1 – 6.3 we can conclude that the more the size of the network the faster the performance convergence. However, a larger network requires a lot of memory and hence based on this tradeoff a 4 layer network is chosen to be the most optimal solution for this power quality disturbance classification.

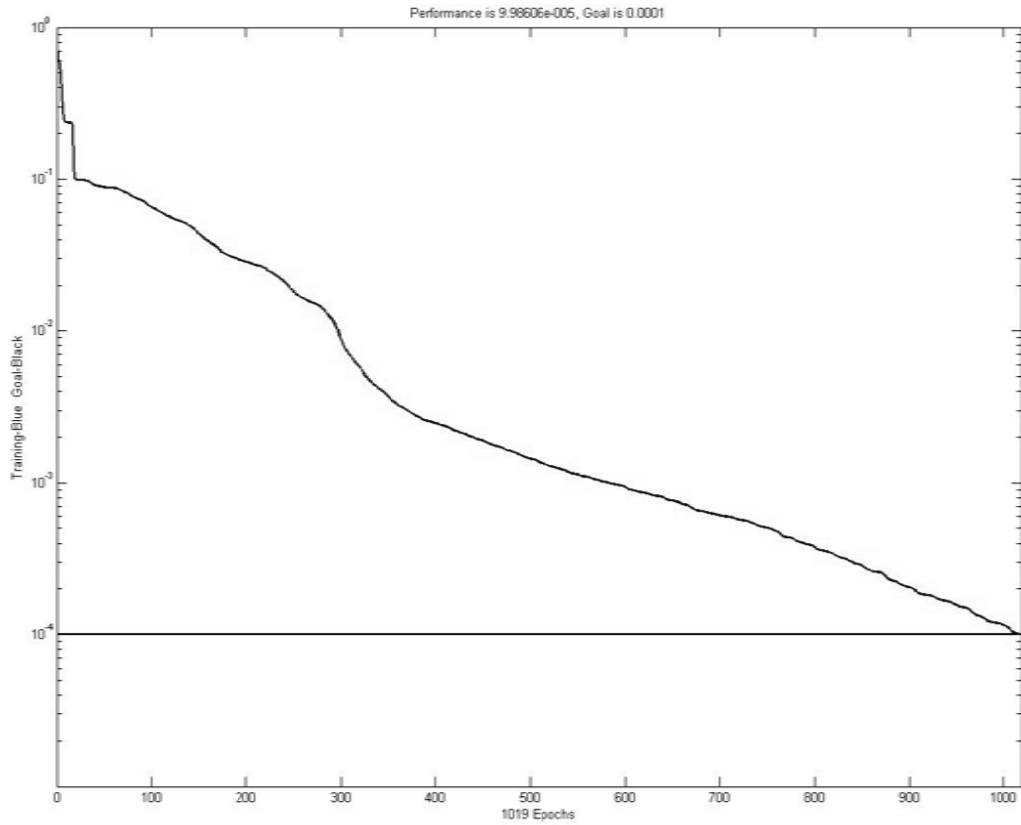


Figure 6.3 Performance function for a 5 layer neural network using 'Trainbfg' function

Various training algorithms [44] that include BFGS-Quasi-Newton (BFG), Resilient Back propagation (RP), Scaled Conjugate Gradient (SCG), Conjugate Gradient with Powell/Beale Restarts (CGB), Fletcher-Powell Conjugate Gradient (CGF), Polak-Ribière Conjugate Gradient (CGP), One-Step Secant (OSS) and Variable Learning Rate Back propagation (GDX), which are the most famous training algorithms, are explored. The training process consists of various steps which are given below,

Step 1

The data set containing the feature vector is saved as a matfile and loaded into the workspace during the training phase. This file contains both the Input vector and the Target vector.

Step 2

Find the maximum and minimum values for the input vector after removing the redundancies and normalize (if required) to suit the selected the feed forward neural network

Step 3

Select the set of parameters that are initialized to the network to train it. The parameters that are initialized are maximum number of epochs, performance goal with MSE, maximum and minimum gradient.

Step 4

Train the network for the given transfer function and training function and the given set of activation functions and number of neurons. The number of nodes in each layer is initially varied and an optimum combination is figured out based on the training times and performance errors.

Step 5

Change the training function keeping same transfer function and the number of nodes in each layer.

Step 6

Find the most efficient network based on the training time and the least possible mean square error.

6.2 RESULTS

Thus from these steps the various simulations tabulated in Table 6.1 are obtained and the FFNN architecture that is best optimized for this problem is selected by defining the number of nodes in the layer, the training and the activation functions.

Table 6.1 Various combinations of neural network architectures and their performance data

Transfer Function combination used	Training Function used	Performance Goal (Mean Square Error)	Number of Epochs used
Feed Forward network with			
(10,20,9) Tansig, Tansig, Tansig	Trainbfg	0.0001	2234
(10,10,20,9) Tansig, Tansig, Tansig, Tansig	Trainbfg	0.0001	1045
(10,9) Tansig, Tansig	Trainbfg	0.0001	4352
(10,20,9)Tansig, Tansig, Purelin.	Trainlm	Performance goal (0.0001) not met	10,000
(10,20,9)Tansig,Logsig,Purelin	Traindx	Performance goal (0.0001) not met	10,000
(10,20,9) Tansig, Tansig, Tansig	Trainrp	0.0001	6,345
(10,20,9)Tansig, Tansig, Tansig	Traincgp	Performance goal(0.0001) not met	10,000

Then a different set of power signal data is loaded and the neural network is simulated for the new input data. The results of various such input sets and the corresponding classification / misclassification is tabulated in tables 6.2 to 6.4.

The first row in the table belongs to a 4 layer neural network with 10 and 20 nodes in the first and the second hidden layers respectively. The output layer contains 9 nodes. The training function used is 'Trainbfg' and the activation functions are 'Tansig', 'Tansig' and 'Tansig' for the two hidden layers and the output layer respectively. For a targeted mean square error (Performance goal of 0.0001) the network took 2234 epochs to achieve it.

The second row in the table belongs to a 5 layer neural network with 10, 10 and 20 nodes in the first, second and the third hidden layers respectively. The output layer contains 9 nodes. The training function used is 'Trainbfg' and the activation functions are 'Tansig', 'Tansig', 'Tansig' and 'Tansig' for the three hidden layers and the output layer respectively. For a targeted mean square error (Performance goal of 0.0001) the network took 1045 epochs to achieve it. However, as already discussed there is a huge memory trade-off for using a five layer network.

The third row in the table belongs to a 3 layer neural network with 10 and 9 nodes in hidden and the output layers. The training function used is 'Trainbfg' and the activation functions are 'Tansig' and 'Tansig'. For a targeted mean square error (Performance goal of 0.0001) the network took 4352 epochs to achieve it.

The fourth row in the table belongs to a 4 layer neural network with 10 and 20 nodes in the first and the second hidden layers respectively. The output layer contains 9 nodes. The training function used is 'Trainblm' and the activation functions are 'Tansig',

‘Tansig’ and ‘Purelin’ for the two hidden layers and the output layer respectively. For a targeted mean square error (Performance goal of 0.0001) the network couldn’t achieve it even after 10,000 epochs.

The fifth row in the table belongs to a 4 layer neural network with 10 and 20 nodes in the first and the second hidden layers respectively. The output layer contains 9 nodes. The training function used is ‘Traindx and the activation functions are ‘Tansig’, ‘Logsig’ and ‘Purelin’ for the two hidden layers and the output layer respectively. For a targeted mean square error (Performance goal of 0.0001) the network took 6,345 epochs.

The sixth row in the table belongs to a 4 layer neural network with 10 and 20 nodes in the first and the second hidden layers respectively. The output layer contains 9 nodes. The training function used is ‘Traincgp’ and the activation functions are ‘Tansig’, ‘Tansig’ and ‘Purelin’ for the two hidden layers and the output layer respectively. For a targeted mean square error (Performance goal of 0.0001) the network couldn’t achieve it even after 10,000 epochs.

Tables 6.2 to 6.4 presents the error classification rate for all the nine different signal sets considered and with a 4 layer feed-forward neural network structure finalized in the previous sections. Table 6.2 belongs to a 3 layer neural network and consists of a 90,000 dataset. The error classification rates are tabulated against their misclassification category. Table 6.3 belongs to a 4 layer network with the same topology and consists of 9,000 dataset. Its error classification rates are tabulated against their misclassification category. Similarly, table 6.4 represents the same 4-layer network topology against a 90,000 dataset. Its error classification rates are tabulated against their misclassification category.

Table 6.2 Error Classification during simulation for a 3 layer network and 90,000 dataset

Disturbance Type	Normal	Harmonic	Impulse	Flicker	Notch	Sag	Swell	Outage	Transient	Error %
Normal	—	0	0	0	0	0	0	0	0	0
Harmonic	0	—	0	0	0	0	0	0	0	0
Impulse	0	95	—	0	0	0	0	0	0	0.95%
Flicker	0	51	0	—	142	0	0	0	0	1.93 %
Notch	0	12	0	3	—	0	0	0	0	0.15%
Sag	0	12	0	0	0	—	0	0	0	0.12%
Swell	0	0	0	0	0	0	—	0	0	0%
Outage	0	0	54	0	0	0	0	—	0	0.54%
Transient	0	0	0	0	0	0	0	0	—	0%

Table 6.3 Error Classification during simulation for a 4 layer network and 9,000 dataset

Disturbance Type	Normal	Harmonic	Impulse	Flicker	Notch	Sag	Swell	Outage	Transient	Error %
Normal	—	0	0	0	0	0	0	0	0	0
Harmonic	0	—	0	0	0	0	0	0	0	0
Impulse	0	0	—	0	0	0	0	0	0	0
Flicker	0	0	0	—	0	0	0	0	0	0
Notch	0	0	0	0	—	0	0	0	0	0
Sag	0	0	0	0	0	—	0	0	0	0
Swell	0	0	0	0	0	0	—	0	0	0
Outage	0	0	0	0	0	0	0	—	0	0
Transient	0	0	0	0	0	0	0	0	—	0

Table 6.4 Error Classification during simulation for a 4 layer network and 90,000 dataset

Disturbance Type	Normal	Harmonic	Impulse	Flicker	Notch	Sag	Swell	Outage	Transient	Error %
Normal	—	0	0	0	0	0	0	0	0	0
Harmonic	0	—	0	0	0	0	0	0	0	0
Impulse	0	0	—	0	0	0	0	0	0	0
Flicker	0	0	0	—	0	0	0	0	0	0
Notch	0	0	0	0	—	0	0	0	0	0
Sag	0	0	0	0	0	—	0	0	0	0
Swell	0	0	0	0	0	0	—	0	0	0
Outage	0	0	0	0	0	0	0	—	0	0
Transient	0	0	0	0	0	0	1	0	—	0.0001

6.3 CONCLUSION

This thesis proposed two kinds of classification algorithms for the disturbance types considered and could successfully use one of those algorithms to achieve a very high error classification rate. The use of wavelet multi-resolution analysis standard deviation curves along with fourier domain parameters with the help of a 4 layer feed forward neural network proved to achieve almost a zero error rate for the considered datasets. Also, the insights provided in this thesis on using clustering based wavelet MRA energy parameters along with Adaptive Resonance theory classification algorithms is worth further investigation.

BIBLIOGRAPHY

1. H. J. Bollen and Irene Gu, *Signal Processing of Power Quality Disturbances*, Wiley-IEEE Press, Hoboken, USA, 2006.
2. Hamachi LaCommare, J. Eto, *Understanding the Cost of Power Interruptions to U.S. Electricity Consumers*, Consortium for Electric Reliability Technology Solutions, USA, August 2004.
3. F.C. Pereira, O.C.N. Souto, J.C. De Oliveira, A.L.A Vilaca, P.F. Ribeiro, “An analysis of costs related to the loss of power quality”, *Proceedings of the International Conference on Harmonics And Quality of Power*, vol. 2, no. 1, Oct 1998, pp. 777 – 782.
4. Yuan Liao, *Automated Analysis of Power Quality Data and Transmission Line Fault Location*, Dissertation of Texas A&M University, College Station, Texas, 2000.
5. M. M. Morcos and W. R. Anis Ibrahim, “Electric power quality and artificial intelligence: Overview and applicability,” *IEEE Power Engineering Review*, vol. 19, no. 3, April 1999, pp. 5-10.
6. Rong-Ceng Leou, Wen-Ruei Tsai and Yong-Nong Chang, “A power quality monitoring system based on J2EE architecture”, *TENCON 2004, IEEE Region 10 Conference*, Melbourne, Australia, Nov. 2004, pp. 13-17.
7. D. Piombo, R. Zunino, “FPGA realization of power quality disturbance detection: an approach with wavelet”, *ANN and fuzzy logic, IJCNN '05 - IEEE International Joint Conference on Neural Networks*, 30 July-4 Aug. 2005, pp. 132-138.
8. F. Choon, M. B. I. Reaz and F. Mohd-Yasin, “Power Quality Disturbance Detection Using Artificial Intelligence: A Hardware Approach”, *19th IEEE International Parallel and Distributed Processing Symposium*, 04-08 April 2005, pp. 146a - 149a.
9. S. Pittner and S.V. Kamarthi, “Feature extraction from wavelet coefficients for pattern recognition tasks”, *IEEE Transactions on Pattern Analysis and Machine*, vol. 21, no. 1, January 1999, pp. 83-89.
10. E. W. Gunther and H. Mehta, “A survey of distribution system power quality— preliminary results,” *IEEE Transactions on Power Delivery*, vol. 10, no. 1, Jan. 1995, pp. 322-329.
11. M. B. Hughes, J. S. Chan, and D. O. Koval, “Distribution customer power quality experience,” *IEEE Transactions on Industrial Applications*, vol. 29, no. 1, Nov./Dec. 1993, pp. 1204-1211.

12. IEEE Recommended Practice for Monitoring Electric Power Quality, IEEE Inc., New York, USA, 1995.
13. J. J. Burke, D. C. Griffith, and J. Ward, "Power quality—Two different perspectives," IEEE Transactions on Power Delivery, vol. 5, no.3, June 1990, pp. 1501-1513.
14. G. Beylkin, R. Coifman, I. Daubechies, S. G. Mallat, Y. Meyer, L. Raphael and M. B. Ruskai, Introduction to Wavelets, Jones and Bartlett, Boston, 1991.
15. I. Daubechies, "Ten Lectures on Wavelets", CBMS-NSF Regional Conference Series in Applied Mathematics for the Society for Industrial and Applied Mathematics, Philadelphia, 1992.
16. S. G. Mallat, "Multiresolution approximations and wavelet orthonormal bases" Transactions of American Mathematical Society, vol. 315, no. 1, 1989, pp. 69-87.
17. M. P. Collins, W. G. Hurley, and E. Jones, "The application of wavelet theory in an expert system for power quality diagnostics," 30th Universal Power Engineering Conference, 1995.
18. Oliver Poisson, Pascal Rioual and Michel Meunier, "New Signal processing tools applied to power quality analysis", IEEE transactions on Power Delivery, vol. 14, no. 2, July 1999, pp. 324-327.
19. Oliver Poisson, Pascal Rioual and Michel Meunier, "Detection and Measurement of Power quality disturbances using Wavelet transform", IEEE transactions on Power Delivery, vol. 15, no. 3, July 2000, pp. 214-219.
20. P K Dash, B K Panigrahi and G Panda, "Power quality analysis using S transform", IEEE transactions on power delivery, vol. 18, no. 2, April 2003, pp. 23-29.
21. M. P. Collins, W. G. Hurley, and E. Jones, "The application of wavelet theory to power quality diagnostics," 29th Universal Power Engineering Conference, 1994.
22. Y. Xu, X. Xiao, Y. Yang, and X. Chen, "Application of wavelet transform in power quality analysis," Automated Electric Power Systems, vol. 23, no. 23, 1999, pp.87-93.
23. A. M. Gouda, M. M. A. Salama, M. R. Sultan, and A. Y. Chikhani, "Application of multi resolution signal decomposition for monitoring short-duration variations in distribution systems," IEEE Transactions on Power Delivery, vol. 15, no. 1, Apr. 2000, pp. 137-145.
24. A. M. Gaouda, M. M. A. Salama, A. Y. Chikhani, and M. R. Sultan, "Application of wavelet analysis for monitoring dynamic performance in industrial plants," North American Power Symposium, Laramie, WY, 1997.

25. S. Santoso, E. J. Powers, W. M. Grady, and P. Hoffman, "Power quality assessment via wavelet transform analysis," *IEEE Transactions on Power Delivery*, vol. 11, no.1, Apr. 1996, pp.56-62.
26. S.Madan and K. E.Bollinger, "Applications of Artificial Intelligence in Power Systems," *Electric Power Systems*, vol. 41, 1997, pp. 117-131.
27. R.C. Bansal, "Bibliography on the fuzzy set theory applications in power systems", *IEEE Transactions on Power Systems*, vol. 18, no.4, Nov. 2003, pp. 1291- 1299.
28. R. L. King, "Artificial neural networks and computational intelligence," *IEEE Power Computing Applications*, vol. 11, 1998, pp. 14-25.
29. M. E. El-Hawary, *Electric Power Applications of Fuzzy Systems* IEEE Press, New York, USA, 1998.
30. J. S. R. Jang and C.T. Sun, "Neuro-fuzzy modeling and control", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 83, Mar. 1995, pp. 378-406.
31. W. R. Anis Ibrahim, M. M.Morcos, and D. G.Kreiss, "An adaptive neuro-fuzzy intelligent tool and expert system for power quality analysis—Part I: An introduction," *IEEE Power Engineering Society Summer Meeting*, Edmonton, Canada, 1999, pp. 21-29.
32. W. R.Anis Ibrahim and M. M.Morcos, "Preliminary application of an adaptive fuzzy system for power quality diagnostics," *IEEE transactions on Power Engineering*, vol. 20, no. 1, 2000, pp. 55-58.
33. G. P. Damarla, A. Chandrasekaran, and A.Sundaram, "Classification of power system disturbances through fuzzy neural network," *Canadian Conference on Electrical and Computer Engineering*, 1994.
34. P. K. Dash, S. Mishra, M. M. A.Salama, and A. C.Liew, "Classification of power system disturbance using a fuzzy expert system and a Fourier linear combiner," *IEEE Transactions on Power Delivery*, vol. 15, Apr. 2000, pp. 472-477.
35. Yuan Liao and J B Lee, "A fuzzy-expert system for classifying power quality disturbances", *International Journal of Electrical Power and Energy Systems*, vol. 26, no. 3, March 2004, pp. 199-205.
36. P K Dash, B K Panigrahi and G Panda, "Power quality analysis using S transform", *IEEE transactions on power delivery*, vol. 18, no 2, April 2003, pp. 231-236.
37. Abdel Galil, E F Saadany, A M Youssef and M M A Sallama, "Disturbance classification using dynamic time warping classifier", *IEEE transactions on power delivery*, vol. 19, no. 1 January 2004, pp. 117 – 122.

38. Abdel Galil, E F Saadany, A M Youssef and M M A Sallama, "Disturbance classification using hidden markov models and vector quantization", IEEE transactions on power delivery, vol. 20, no. 3, July 2005, pp. 71-77.
39. G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton USA, 1976.
40. J. A. Barnett, "Computational methods for a mathematical theory of evidence," Seventh International Joint Conference on Artificial Intelligence, Vancouver, B.C., 1981.
41. S.Pittner and S.V.Kamathi, "Feature extraction from wavelet coefficients for pattern recognition tasks", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 1, Jan. 1999, pp. 83-88.
42. G.A. Carpenter, Stephen Grossberg and David B. Rosen, "Fuzzy ART : Fast Stable learning and categorization of analog patterns by an Adaptive Resonance System", Neural Networks, Vol.4, no.1, 1991, pp. 759-771.
43. Serrano-Gotarredona Teresa, Linares-Barranco Bernabé, Andreou Andreas G, "Adaptive Resonance Theory Microchips", Springer, USA, 1998.
44. The Mathworks Inc., Using Matlab, Natick MA, July 2002.
45. C.G.Looney, Pattern Recognition Using Neural Networks: Theory and Algorithms for Engineers and Scientists, Oxford University Press, New York, 1997.
46. K. Kasinathan, Power System Fault Detection and Classification using wavelet transforms and Adaptive Resonance Theory neural networks, Master's Thesis, University of Kentucky, 2006.

VITA

Praveen Settipalli was born on May 13th, 1983 in Andhra Pradesh, India. He received his under-graduate degree in Electronics and Communications Engineering from VR Siddhartha Engineering College, affiliated to Acharya Nagarjuna University in June 2004. In Fall 2004, he enrolled as a MSEE student in the Department of Electrical & Computer Engineering at the University of Kentucky. He is currently with the Lexmark Manufacturing Systems Application Development team in Lexington, KY.