Development of Fault Progression Model for Utilization in Prognostic Techniques

Submitted by:

Ayesha Siddiqui

2010-NUST-MS PhD-Elec (Comm-N) 03

Supervised by

Cdr Dr Syed Sajjad Haider Zaidi

Thesis
Submitted to
Department of Electronics and Power Engineering
Pakistan Navy Engineering College
National University of Science and Technology, Islamabad

In requirements for the award of the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING
With Specialization in Communications

March 2013


بسم الله الرحمن الرحيم

Dedication

Dedicated to my father Mr. Shamim Ahmed Siddiqui.

Abstract

Machine reliable operation is of key importance especially when they are used for critical applications. In order to guarantee machine reliability and life estimation, machine fault diagnosis and prognosis is of great importance. This work aims to develop an understanding towards machine fault diagnosis algorithms and to develop a fault progression model that can be used for estimation of remaining machine useful life. Time frequency methods are used in analysing the transient machine faults. Further faults classification is done. In this work healthy machine and faulty machines data is obtained. On the data a non intrusive analysis is done to identify the electrical fault with the help of the field oriented currents analysis in the permanent magnet synchronous motors.

Acknowledgements

I really am thankful to Almighty Allah who provided me with the opportunity to complete my MS degree from this prestigious institute. I am grateful to my supervisor Cdr Dr Syed Sajjad Haider Zaidi for his continuous guidance and encouragement from initial to final level that enabled me to develop the understanding of this subject.

I would like to show my gratitude to my committee members who supported me through out this research phase and guided me.

I am a proud daughter of my parents. The courage, support, trust, encouragement, duas and love that my parents showed through out my life made me believe that nothing is out of my reach. You are the best parents in the world. I would also like to thank my in laws especially my father in law for his support through out this research phase.Regards to my husband for his words of encouragement at times.

Contents

List of Figures

List of Algorithms

List of Tables

Chapter 1

Introduction

In the industrial world, electrical machines fault diagnosis and prognosis field is gaining a lot of attention because of the need of reliability, performance efficiency and cost effectiveness as they act as critical component in several process. Machines are more susceptible to failures as they get mature making their maintenance activity difficult and costly. The tasks related to their maintenance can be curative or preventive [26]. Prognosis as defined in some research works as Remaining Useful Life (RUL) [21].

- Curative: In this, components are replaced when they are not able to complete the task for which they are designed. The demerit of this solution is occurrence of fault i.e happening of unwanted state. In order to avoid this, significant parameter monitoring is needed. Component replacement is done when parameters exceeded the threshold value. This is also called Condition Based Monitoring (CBM).

- Preventive: As fault arises non availability of the resources or spare parts can delay the maintenance so predictive one is a better choice in comparison to curative one. In this health state of the system is predicted and suitable plans are made for its maintenance. It can be seen as a prognosis activity.

For machine maintenance diagnosis of faults and prognosis for RUL of a machine can be done. For diagnosis of faults several techniques are available whereas prognosis is a relatively new field. Diagnosis refer to as the detection of early faults whereas prognosis which is generally performed after the diagnosis estimates the remaining machine life. Prognosis requires a large amount of historical data. It is difficult to do the prognosis of system when a system model is not present. In this work, data of healthy and faulty machine is sampled and then analysis is done on that data. Fault diagnosis and failure

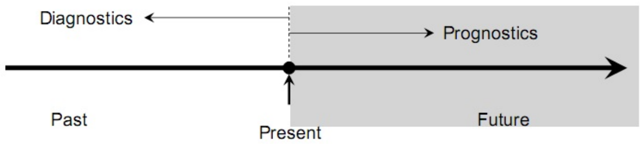

Figure 1.1: Diagnosis vs Prognosis [26]

prognosis has a wide variety of applications such as health assessment systems, electro-mechanical systems, computer software fault detection and prediction methods, and manufacturing systems etc. If the system is complex, noisy, non-linear and contains different subsystem then this field of diagnosis and prognosis becomes difficult to realize. Modelling a system is important for prognosis.

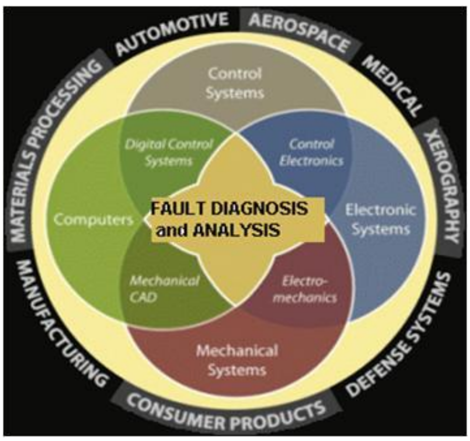

Figure 1.2: Fault diagnosis and analysis areas

Objectives of this work includes non intrusive algorithm development for efficient diagnosis of complex systems using supervised learning approach and to forecast the upcoming fault. Also to develop a fault progression model for RUL of system. Permanent Magnet AC Machine (PMAC) was the target system in this work. Data samples were taken from the machine in the healthy condition and with the occurrence of faults and this is provided by my supervisor.

1.1 Thesis Organization

This thesis report is organized as follows:

- **Chapter 1** contains the importance and objectives of this research along with the details of thesis organization.
- **Chapter 2** provides a comprehensive details about techniques that have been in practise for diagnosis of faults and prognosis of machine failures.
- **Chapter 3** presents the problem statement of the research work and the proposed approach adopted as a solution.
- **Chapter 4** describes the details of parameters involve in experiment setup along with the details of explored fault.
- **Chapter 5** provides details of the issues involved in implementing adopted approach along with obtained results.
- **Chapter 6** aims at concluding this work and proposing future work that can be done in this field.

Chapter 2

Literature Review

2.1 Chapter Scope

Before failure there must be some sudden changes in the behaviour of the electrical drive [2] during its normal operation. So a developing fault will ultimately lead to a failure. In order to protect machine from failure some earlier fault mitigation techniques can be applied. For the diagnosis of fault and prognosis of failure various algorithms are available in literature. So for understanding of those methods a review of them is presented here. This will allow to develop an understanding towards merits and demerits of them. The available methods comprised of the different concepts related to time frequency analysis, clustering, classification and estimation.

2.2 Fault Diagnosis

For the fault diagnosis there are two methods such as intrusive and non intrusive methods. Intrusive method requires addition of external item that increases the cost and it is classical method whereas non intrusive method does not require any additional installations and sensors. In this kind of method current or voltage of the motor is taken as a dataset and analysis is done on that dataset.This method has gained attention over the recent years. Non intrusive methods are divided in to three categories model based, signal based and data based[31].

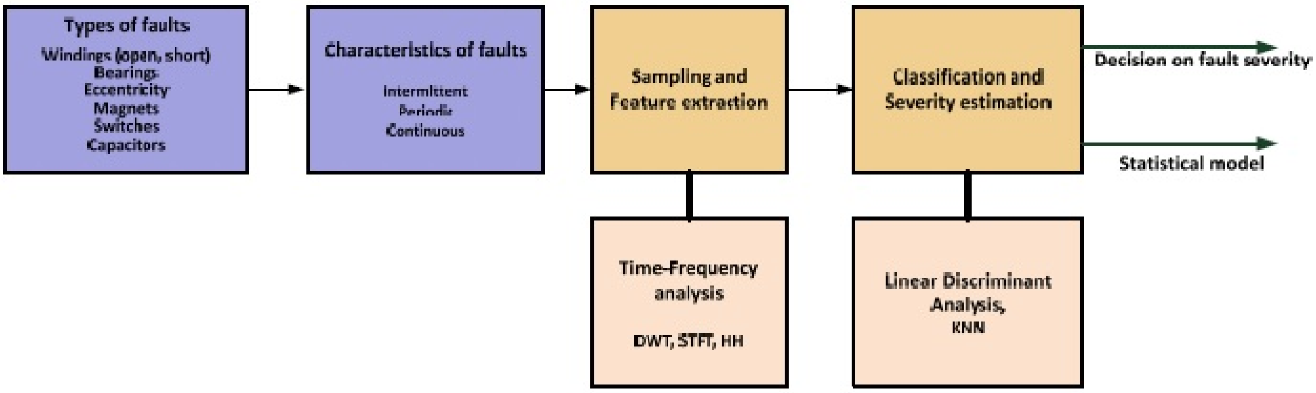

Figure 2.1: Fault diagnosis

2.2.1 Model Based

The diagnosis based on machine model which further can be used for fault signature prediction. Model-based diagnosis relies on theoretical analysis of machine whose model will be used to predict fault signatures [25]. There are different types of fault diagnosis:

2.2.2 Data Based

In this type of diagnosis no knowledge of machine parameters are required. It is comprised on signal processing and clustering techniques .

2.2.3 Signal Based

In this type of diagnosis, sampled quantities from actual machine are taken for identification of known fault signatures. Using suitable signal processing techniques these signatures are monitored. Frequency, Time or time frequency analysis techniques can be of interest. In the recent years this has gained a lot of attention. For signal to noise ratio enhancement and data normalization it plays an important role.

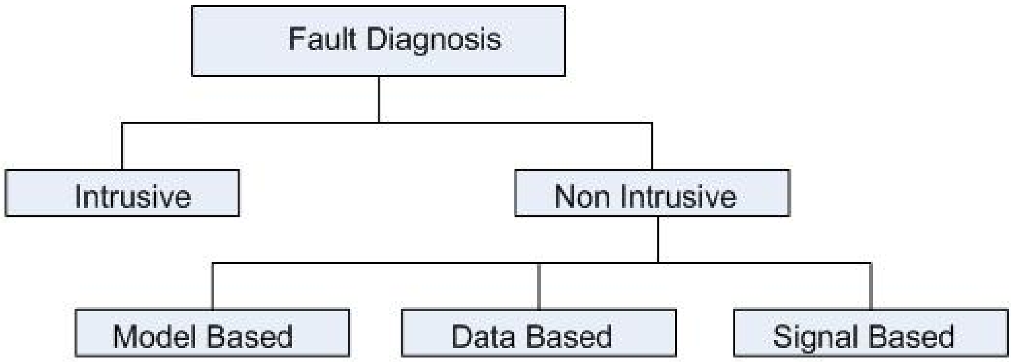

Figure 2.2: Fault diagnosis types

2.3 Failure Prognosis

Estimation of the remaining usable life and prediction of the future state is machine failure prognosis. This is an important field as it helps in timely machine maintenance and also to protect it from disastrous failures. Following are the techniques of machine failure prognosis.

- Model Based (Physics of failure)
- Data Based
- Hybrid

This classification is according to four criteria: cost, precision, applicability and complexity [9]. The fault is considered continuously variable, whose evolution is defined by a deterministic or stochastic law. In model based method a set of mathematical laws are used to represent the physical component and its degradation phenomenon which will be further used for RUL. The data driven approach transforms sensors monitored data in to reliable degradations behavioural model. The data driven models are suitable for systems where to monitor data and transformation in to degraded behavioural model is easy.

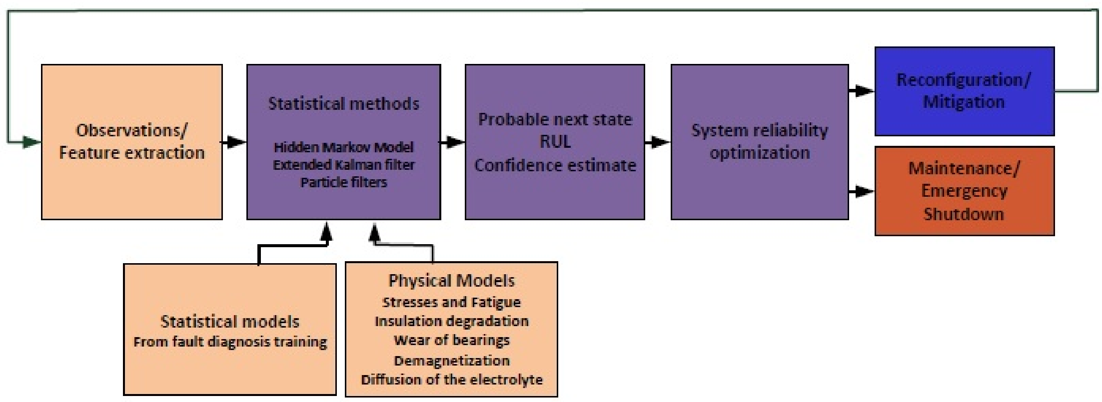

Figure 2.3: Prognosis

2.4 Types of Faults Occur in Electrical Drives

Following are the types of faults that are generally expected to occur in the electrical drives[25].

- Insulation: over-voltage, initial manufacturing quality, temperature.
- Bearing faults: environment, affected by wear, temperature, loading.
- Connections: corrosion, welding, crimping.
- Rotor bar breakage in induction motors: starting cycles, manufacturing problems.
- Rotor eccentricity: manufacturing, loading, wear.

- Power electronics components: capacitors, switches, gate drivers.
- Permanent magnet demagnetization: load, temperature, controller error, noise.
- Gears.
- Sensor failure (e.g. current sensor,rotor position sensor,).

2.5 Motor Current Signature Analysis

For fault diagnosis in electrical machines, there are various condition monitoring techniques which includes vibration monitoring, acoustic monitoring, chemical and thermal monitoring etc but as they require specialized sensors and other tools so the most suitable is to monitor current [10]. This method uses current spectrum for the location of fault frequencies. Motor Current Signature Analysis (MCSA) has been extensively used in the domain of electrical machines for diagnosis and prognosis [12, 17, 30, 27, 29, 6]. Even some of the research works provide a comparison as to which method is more preferable for condition monitoring using comparative approach [24]. MCSA is a maintenance tool

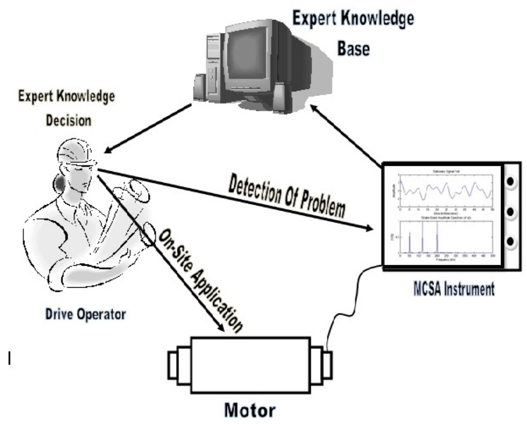

Figure 2.4: MCSA

which is used for valuable prediction of machine behaviour. It has gained much acceptance in nowadays world. By this approach mechanical faults can also be found. In this approach the current is recorded and analysis is done in time frequency domain. It has been in practice since 1985 and used over the years. In time domain format the current samples are taken. Reference [10] provides a detailed background of how MCSA has been applied in order to find out wide variety of machine faults.

2.6 Methodology

This sections presents a comprehensive review of the available techniques that are adopted for feature extraction, classification and for prognosis [30]. Among the available tools and techniques, relevant techniques have been adopted.

2.6.1 Feature Extraction

For the extraction of features a transformation is needed since observing the signals with respect to different domains is of vital importance as different domain representation reveals different aspects of signal. While time domain representation seems enough but it is actually incomplete. For complete analysis different domain representation is required. Further knowledge can be obtained if it is represented in frequency domain. So time frequency analysis together aims at reconciling spectral and temporal analysis within a single distribution. Applications involve physiological origin, radar and sonar signals, acoustic signals, astrophysical data, etc. So time frequency transformation is applied.

Our lives are generally measured with respect to time so we are more comfortable to represent quantities w.r.t to time. When representing signals in time domain we get time-amplitude representation but applications of signal processing mostly requires the frequency content in the signal and hence via frequency representation properties such as spectrum, bandwidth and roll off can be obtained easily. The time frequency transformation had been effectively used for the fault analysis of electrical machines [8]. For frequency representation the most common time to frequency transform is Fourier Transform whose continuous form can be applied to continuous time signal and discrete one can be applied to discrete time sequences. Discrete Fourier transform is seen as first choice with the advent of its fast computation method namely Fast Fourier

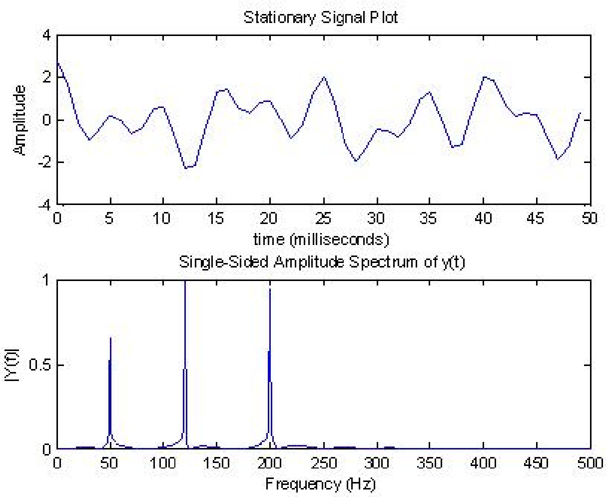

Figure 2.5: Stationary Signal

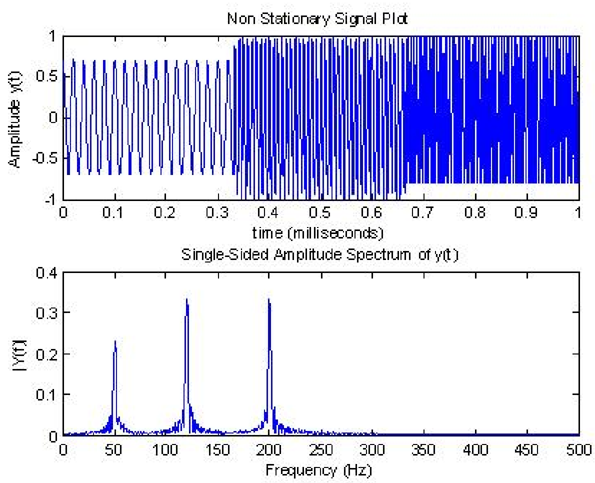

Figure 2.6: Nonstationary Signal

Transform (FFT). Talking about the Fourier Transform and transform similar to it. They just give information that how much of each frequency is present in the signal but does not tell at which time it is existing. Information regarding frequency content of a signal locally in time is of viable importance if the signal is non stationary. Signal whose frequency content does not change with time is called stationary signals. In such case same frequency content exists through out the signal duration and hence time information is not that important. The signals encountered in real life are mostly non-stationary examples of such signals are music signals, sinusoidal or linear FM signals, chirp signals etc.

Figure 2.7: FFT Tiling

Consider the signal plotted in Figure 2.5 the frequency content as identified by its FFT is existed throughout. Hence its time information is not really needed.

Now consider the Figure 2.6 both stationary and non stationary signals have the same FFT. Thus the signals in both the figures are unable to differentiate whether the frequency components are lying throughout the spectrum or are localized in time. There are various time to frequency transformation techniques that are commonly used. Some of these techniques are as under. If we consider time frequency plane tiling of FFT then it indicates no time axis information is available whereas frequency plane is divided in to several bands as in Figure 2.7

2.6.1.1 Short Time Fourier Transform

Previously from Figure 2.5 and Figure 2.6 it was seen that FFT does not work for non stationary signals. Here the main consideration is some portions of the non stationary signals can be assumed as if they are stationary. The size of the portion is known as the window size. Window length must satisfy the criteria that the signal segment must be stationary where the window is applied [18]. The 'Uncertainty principle' is the problem associated with the Short time Fourier Transform (STFT) which says that exact time and frequency location of a signal is not known. What maximum one can know is which frequency band exists in which time interval and hence due to this, resolution problem exists in STFT.

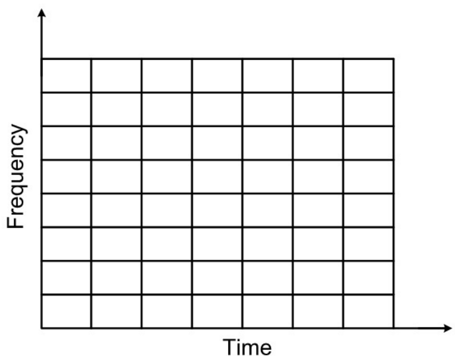

Figure 2.8: STFT Tiling

The performance of the STFT are affected due to the window size. If the window size is narrow then it gives poor resolution in time. Furthermore, wide windows can violate the condition of stationarity. The problem with the STFT is resolved with the Wavelet Transform (WT). In order to see the phenomenon of how spectrum w.r.t time changes see Figure 2.8. The mathematical expression for STFT is as below.

$$STFT(t,f) = \int_{-\infty}^{\infty} h(t-\tau)x(\tau)e^{-j2\pi f\tau}d\tau \tag{2.1}$$

where h is the window function.

A design tradeoff in implementing the STFT is taken in to consideration between resolution of time and frequency. A limitation due to uncertainty principle is made that limits time bandwidth product lower bound.

$$TB \geq \frac{1}{2} \tag{2.2}$$

Here B is bandwidth and T is duration.

2.6.1.2 Wavelet Transform

The wavelet transform allows multi-resolution analysis . As its name implies, it analyses the signal at different frequency with variable resolutions. The purpose of this is to give for shorter durations high frequency components and for longer durations low frequency components[32]. Wavelets unique property is of finite energy. The basis function, scaling and wavelet function all have energy concentrated around a point. There are variety of basis functions that can be used in the Wavelet analysis in comparison to Fourier transform that uses fixed sinusoid. For desired application the best suitable wavelet can be chosen [35].

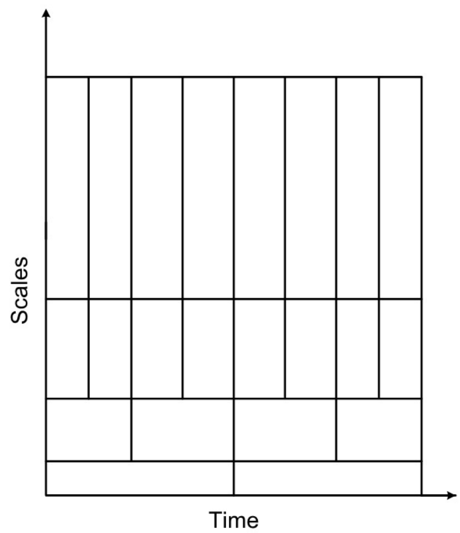

Figure 2.9: DWT Tiling

Considering the tiling in Figure 2.9 which is variable and has the property of good resolution in frequency at the lower frequency and at higher frequency it has good resolution in time. It also has a demerit that this transformation is not shift invariant. Various types of wavelets can be applied in building the detection and classification algorithm for machines [7]. A comparison of results by applying various mother wavelets for prognosis of failures in electric motors is presented in [34].

2.6.1.3 Wigner Ville Distribution

The distribution of energy as time and frequency function is given by this transform. The resolution problem that we have encountered in STFT, it doesn't have that problem of resolution. Signal s(t) Wigner Ville Distribution (WVD) [5] is given as

$$W(t,\omega) = \frac{1}{2\pi}\int s^*(t-\frac{\tau}{2})s(t+\frac{\tau}{2})e^{-j\tau\omega}d\tau \tag{2.3}$$

WVD is a quasi probability distribution. Eugene Wigner in 1932 introduces this in classical mechanics to study quantum corrections [28]. Then in 1948 J.Ville derived it again independently as a representative of local time frequency energy of signal. The important property of this distribution is the correlation of signal with time and frequency version of itself. It doesn't contain a windowing function and this one reason frees it from smearing effect. WVD provides the highest time and frequency plane resolution. The properties a WVD [4] has are the reality i.e its transformation is always real even if the signal is complex, symmetric in nature i.e if the frequency spectra is symmetric then the transformation is also symmetric whereas for the time symmetric real signal the distribution will be symmetric in time.

$$W(t,\omega) = W(t,-\omega)\ For\ Real\ Signals \equiv symmetrical\ spectraS(\omega) = S(-\omega) \tag{2.4}$$

$$W(t,\omega) = W(-t,\omega)\ For\ Real\ Signals \equiv symmetrical\ spectraS(t) = S(-t) \tag{2.5}$$

Marginals in time and frequency are satisfied in this distribution [3]

$$\int W(t,\omega)d\omega = |s(t)|^2 \tag{2.6}$$

$$\int W(t,\omega)dt = |S(t)|^2 \tag{2.7}$$

Time shift and the frequency shift i.e if the signal is shifted in time or in frequency the distribution is shifted accordingly [31].
if $s(t) \rightarrow s(t-t_0)$ then $W(t,\omega) \rightarrow W(t-t_0,\omega)$
if $s(t) \rightarrow e^{j\omega_0 t}s(t)$ then $W(t,\omega) \rightarrow W(t,\omega-\omega_0)$

The demerit of this transformation are the cross terms which comes when WVD applied to multicomponent signals and makes WVD difficult to understand. The unique characteristic of this generated interference term is of its highly oscillatory nature in comparison to the the auto terms which are representative of the original signal. Due to this nature of the interference term, idea can be inferred that these cross terms can be attenuated using domain kernel functions without any significant effect on the signal.

2.6.1.4 Choi Williams Distribution

In approach to suppress interference terms encountered in the WVD a kernel in WVD is placed. This added kernel is a two dimensional Low Pass Filter (LPF) and has the capability of suppressing interference term in time and frequency plane. There are wide range of kernels constructed for this purpose each of them has their own advantages but the most popular of them is Choi Williams Distribution (CWD) [4]. Mathematical expression for the Choi Williams distribution of a signal is given by [5]

$$C(t,f) = \int\int\int \phi(\theta,\tau)s(u-\frac{\tau}{2})s^*(u-\frac{\tau}{2})e^{j(\theta u-\theta\tau-\tau\omega)}dud\theta d\tau \tag{2.8}$$

where $\phi(\theta,\tau) = e(-\frac{(\theta,\tau)^2}{\sigma})$ is a kernel function. It acts as signal autocorrelation function filter. The amount of smoothing is controlled by σ. Due to this smoothing phenomenon the cross terms are removed and resolution is reduced.

- If kernel satisfies the identical conditions $\phi(\theta,\tau) = \phi*(-\theta,-\tau)$ then its distribution will be a real distribution.

- This distribution satisfies the time frequency marginals

$$\int C(t,\omega)d\omega = |s(t)|^2 \tag{2.9}$$

$$\int C(t,\omega)dt = |S(t)|^2 \tag{2.10}$$

if $\phi(\theta,0) = 1$ and $\phi(0,\tau) = 1$ respectively. Total energy is preserved if kernel function satisfies $\phi(0,0) = 1$.

- Distribution is time shift invariant as kernel is independent of time. The distribution will be shifted according to the signal shifted in time as if $s(t) \to s(t-t_0)$ then $C(t,\omega) \to C(t-t_0,\omega)$

- The distribution will be shifted as according to the signal shifted in spectrum. The kernel is independent of frequency so the distribution is frequency shift invariant. if $s(t) \to e^{j\omega_0 t}s(t)$ then $W(t,\omega) \to W(t,\omega-\omega_0)$

2.6.2 Principal Component Analysis

The Principal Component Analysis (PCA) is a statistical data analysis technique. The aim of PCA is to find out from a given multivariate data to a less redundant smaller variable data set that in the same time is the most efficient representative of available data. In PCA the merit of redundancy is correlation between elements of data [13]. PCA application ranges from neuroscience to computer graphics.
Consider a vector x with n elements. A sample x(1),.....,x(n) available from this random variable. It is important that vector are correlated and there is some redundancy making compression possible. If elements are independent, PCA cannot achieve anything. In PCA transform first mean is subtracted from x:

$$x \leftarrow x - E\{x\} \tag{2.11}$$

After centering x is transformed linearly to vector y having n elements, so $m<n$ and hence redundancy due to correlations is removed. This is achieved by finding rotated orthogonal coordinate system such that elements of x becomes uncorrelated in new coordinates. The variances of x projections on the new coordinates at the same time are maximized, due to this the maximal variance is along the first axis, second axis

corresponds to maximal variance in the first axis orthogonal direction and so on.
PCA can be employed for selecting features. It can be used in selecting significant individuals from feature vector [22]. Thus dimensionality is reduced without affecting the accuracy.

2.6.3 Pattern Recognition

Once signal features are available the next step is to determine the type and severity of fault which is done with the help of classifiers. The process of creating groups of objects in such a way that the objects in the same group have same properties and the distinct objects are in other groups is called clustering. Groups are called clusters. The clustering phenomenon can be seen from Figure 2.10. Then comes the classification phase that refers to assign these data points to pre-defined classes. There are two approaches for clustering process

- Supervised Learning Approach:
  Data used for training specifies what classes we try to learn. Difficulty lies with in the implementation if data classification is unknown.
- Unsupervised Learning Approach:
  The task is to learn classification from data. For clustering no predefined classification is required.

In order to categorize data there are two ways dynamic or static. For the dynamic one the criteria is usually the intra-cluster distance i.e compactness and for inter-cluster distance is separation. For one dimensional dataset X and division in j clusters i.e C_1, $C_2, \ldots, C_m$ Let z_i be the cluster center for cluster C_i. It can be defined intra-cluster and inter-cluster distance as [19].

$$D_{intra} = \frac{1}{n} \sum_{i=1}^{m} \sum_{x \in C_i} \|x - z_i\|^2 \tag{2.12}$$

$$D_{inter} = \min_{1 \leq i < j \leq m} \|z_i - z_j\|^2 \tag{2.13}$$

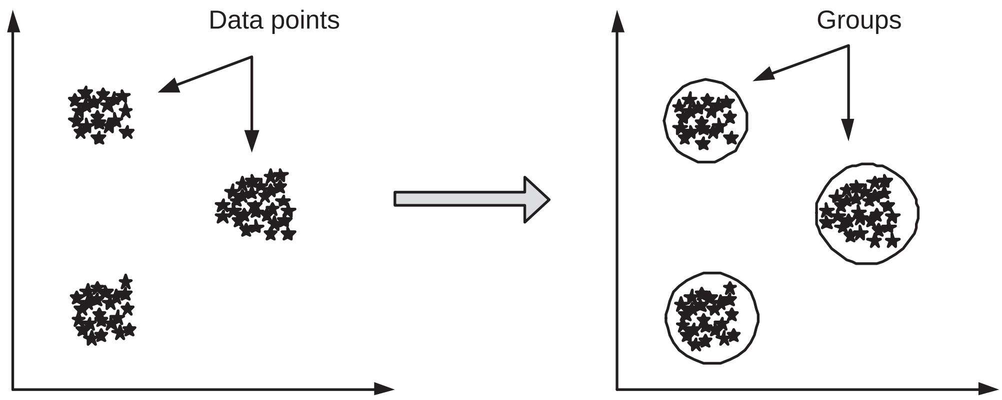

Figure 2.10: Three clusters are shown; the similarity measure is "distance"

Mean of cluster can be defined as,

$$m_j = \frac{1}{n_j} \sum_{\forall z_p \in c_j} z_p \tag{2.14}$$

where n_j is the number of data points that belong to cluster j and c_j is the subset of data vectors that form cluster j. where n is the number of data points of X. For good clusteringD_{inter} should be a large value and D_{intra} should be a small value .

2.6.3.1 Similarity Measures

The association between two data points are indicated by the similarity measure . Suppose $\mathbf{v}=(v_1, v_2, \ldots, v_d)$ and $\mathbf{u}=(u_1, u_2, \ldots, u_d)$ be two d-dimensional data points. The measure of similarity between $\mathbf{v}$ and $\mathbf{u}$ will be their attribute values function,

$$f(\mathbf{u}, \mathbf{v}) = f(u_1, u_2, \ldots, u_d, v_1, v_2, \ldots, v_d) \tag{2.15}$$

f is a metric and it is a distance function whose definition is in a set E having following properties:

1. reflexivity: $f(\mathbf{u}, \mathbf{v}) = 0 \Longleftrightarrow \mathbf{u} = \mathbf{v}$;

2. non-negativity: $f(\mathbf{u}, \mathbf{v}) \geq 0$;

3. triangle inequality: $f(\mathbf{u}, \mathbf{v}) \leq f(\mathbf{u}, \mathbf{w}) + f(\mathbf{v}, \mathbf{w})$,

4. commutativity: $f(\mathbf{u}, \mathbf{v}) = f(\mathbf{v}, \mathbf{u})$;

where $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are arbitrary data points.

For applications, distance selection is important and the best one can be attained with experience as well as knowledge. Some of the common measures are:

2.6.3.1.1 Euclidean Distance Euclidean distance between two data points $\mathbf{p}$=$(p_1, p_2, \ldots, p_n)$ and $\mathbf{q}$=$(q_1, q_2, \ldots, q_3)$ in n-dimension is given as,

$$D_{euc}(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \tag{2.16}$$

where p_i and q_i are the values of i^{th} attribute of $\mathbf{p}$ and $\mathbf{q}$, respectively.

2.6.3.1.2 Tchebychev Distance Tchebychev or Chebyshev distance between two points is defined as,

$$D_{cheb}(\mathbf{p}, \mathbf{q}) = \max_i |p_i - q_i| \tag{2.17}$$

where p_i and q_i are the values of i^{th} attribute of $\mathbf{p}$ and $\mathbf{q}$, respectively. It is also known as "maximum distance".

2.6.3.1.3 Manhattan Distance Manhattan distance between two points is given as

$$D_{man}(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{n} |p_i - q_i| \tag{2.18}$$

where p_i and q_i are the values of i^{th} attribute of $\mathbf{p}$ and $\mathbf{q}$, respectively. Manhattan distance is also called "city block distance".

2.6.3.1.4 Minkowski Distance The generalization of Euclidean,Manhattan and Tchebychev is the Minkowski distance . The Minkowski distance can be defined between two data points as,

$$D_{mink}(\mathbf{p}, \mathbf{q}) = \left(\sum_{i=1}^{n} |p_i - q_i|^r \right)^{\frac{1}{r}} \tag{2.19}$$

where p_i and q_i are the values of i^{th} attribute of **p** and **q**, respectively. If we get r =1 and 2 then we get Euclidean as well as Manhattan distance, respectively. In the limiting case of r reaching infinity, we get Tchebychev distance:

$$\lim_{r\to\infty}\left(\sum_{i=1}^{n}|p_i - q_i|^r\right)^{\frac{1}{r}} = \max_i |p_i - q_i| \tag{2.20}$$

The other distances to measure numerical data are Mahalanobis distance, Average distance and so on. According to application one can choose the suitable one.

2.6.3.2 Supervised Learning Approach

The classifiers that are used by convention in machine fault diagnosis and prognosis includes:

2.6.3.2.1 Linear Discriminant Classifier Linear Discriminant Classifier (LDC) is trained on input vectors comprised on known faults set. Feature space is divided in C sub-regions, where C is number of known fault classes, each of different severity. The weighting coefficients are computed for each class. These coefficients are such that they maximize the corresponding input vector linear discriminant function. It is defined as

$$D_c(x) = x_1\alpha_{1c} + x_2\alpha_{2c} + \cdots + x_k\alpha_{kc} + \alpha_{k+1,c}, wherec = 1, 2, \cdots C \tag{2.21}$$

where x is is k dimensional feature vector and α is normalized weighting coefficients for Cth class. A sample will be said to belong a particular class if the discriminant function for that class is greater than any other. x belongs to class p if

$$D_p(x) > D_k(x) \tag{2.22}$$

Through a training procedure the weighing coefficients are adjusted from initial guess. This procedure algorithm makes adjustments to weighting coefficients till correct classification of training sample vector.

2.6.3.2.2 Multiple Discriminant Classifier Multiple Discriminant Analysis (MDA) analyses fault signals. Fault analysis with different time frequency transforms indicates time frequency features have unnecessary information. In order to choose optimal transformation suitable data reduction techniques are required. The basic concept of MDA is to transform feature space from high dimension to lower dimension and to maximize different classes discrimination. MDA can be seen as an extended version of Fisher Discriminant Ratio which uses intra-class to inter-class scatter ratio. If time frequency feature vectors x and class labels are given, inter-class and intra class scatter matrices can be calculated. Inter-class scatter matrix is given as

$$\sum_{b} = \sum_{i=1}^{C} K_i(m_i - M)(m_i - M)^T \tag{2.23}$$

Intra-class scatter matrix is given as

$$\sum_{w} = \sum_{i=1}^{C} \sum_{x \epsilon C} (x - m_i)(x - m_i)^T \tag{2.24}$$

Here m_i is average of class, Total average is indicated by M and the number of samples in class i is given as K_i. An optimal linear transformation is selected which project C classes data to C-1 space along with the consideration that lower dimensional space gives maximum separation between different clusters.

2.6.3.2.3 Support Vector Machine Classifier For regression analysis and various classifiers, machine learning algorithms utilize supervised learning models that analyze data for certain patterns or statistical parameters, called Support Vector Machine (SVM) or Support Vector Networks. An SVM works by taking input a training data set, pre-assigned with given categories, to build a statistical model based on which new inputs can be assigned to one of the given categories. An SVM model can be viewed as the representation of the training data set as points in space mapped to different categories based on localization, so that new points falling into vicinity of a particular group can be assigned to the respective category. A simplest SVM assigns each given input to two possible classes or categories, called binary linear classifier. Such classification is non-probabilistic or linear; however non-linear classification can be achieved by employing high-dimensional feature spaces.

2.6.3.2.4 The k-Nearest Neighbor Algorithm The k-Nearest Neighbor is one of the simple classification method. It may also be known as Single-link method. It was first introduced by Florek in 1951 and then independently by McQuitty and Sneath in 1957. The k-NN algorithm is summarized in Algorithm 3.2 . It simply treats each data point as single cluster, tries to connect each data point on the bases of closeness. Let c_i, c_j, c_k be the centroid of three groups of data points. Then the distance between c_k and $c_i \cup c_j$ can be defined, using Lance-Williams Equation as follow:

$$D(c_k, c_i \cup c_j) = \min(D(c_k, c_i), D(c_k, c_j)) \tag{2.25}$$

where $D(.,.)$ is distance between two clusters can be calculated using Equation 2.16.

In the k-NN the choice of how many k neighbors have to be merged depends upon the data set. Generally large value of k reduce the effect of noise but make boundaries between the classes less distinct. The selection of a good k can be done by heuristically. Its computational complexity is $O(n^2 \log n)$ but it can be improved to $O(n \log n)$ using the graph theory.

2.6.3.3 Unsupervised Learning Approach

There are many techniques that are used for unsupervised clustering analysis. The most common of which is and it is generally employed in machine fault diagnosis and prognosis [23]. The conventional k-means algorithm is first described by James MacQueen in 1967. The k-means clustering algorithm groups data vectors into pre-defined number of clusters, based on the Euclidean distance as similarity measure. Each data point is assigned to the cluster based on the closeness, i.e. the minimum distance between a data point and a centroid, and is associated with one centroid which represents the "mid-point" of that cluster. The k-means algorithm is summarized in Algorithm 2.1.

The k-means clustering algorithm can be stopped if any one of the following criterion are satisfied:

- Maximum number of iterations has been reached
- When the centroid value does not change over a number of iterations
- Required tolerance value of centroid is achieved after some iterations

Require: Data set , Number of cluster k

Initialize:

Randomly select the k cluster means

repeat:

(**a**) for each data vector, assign the vector to the class with the closest mean, using Equation(2.16)

(**b**) recalculate the k cluster means, using Equation(2.14)

until *stopping criteria are met*

Algorithm 2.1: Conventional k-means clustering algorithm

One of the drawbacks of the k-means algorithm is trapped into local optima. The k-means algorithm is also depending on initialization of randomly selected cluster centroid.

The k-means is efficient in clustering of large data sets. The conventional k-means has several variations, which are used to minimize the problem of local optima, such as fuzzy-k-means, PSO-k-means, etc.

2.6.4 Curve Fitting

Curve fitting is used to find the best fit line or curve for a series of data points. It is also known as regression analysis. Curve fitting usually helps in the visualization of data, to deduce values of a function where no data is available, and to find out the relationships among two or more variables.

Curve fitting can involve either interpolation or extrapolation.

2.6.4.1 Interpolation

Interpolation new data points are constructed within the range of a discrete set of known data points. We usually have a number of data points, obtained by sampling or experimentation, which represent the values of a function for a limited number of values of the independent variable. In some cases we need to estimate the value of that function for an intermediate value of the independent variable. This may be achieved by curve fitting. Interpolation errors are usually present during data estimation depending on the method used.

2.6.4.2 Extrapolation

If data points are estimated, beyond the original observation interval, on the basis of its relationship with another variable then this process is called extrapolation. It is similar to interpolation, which produces estimates between known observations, but in extrapolation there is greater uncertainty and higher risk of producing meaningless results. Extrapolation is basically the extension of a method, assuming similar methods can be applicable to data set beyond the observed data interval.

Basically in curve fitting, the trend among the data points is captured by assigning a single function across the entire range. Different techniques for curve fitting are available involving both linear and non linear methods. Some of them are described briefly.

2.6.4.3 Least Square Curve Fits

Least Square curve fitting methods minimize the square of the error between the original data and the values predicted by the equation. This technique of curve fitting may not be statistically robust method of fitting yet it is simple and easy to understand. The major drawback of this technique of curve fitting is its sensitivity to data points that are wide apart. If a data point is widely different from the majority of the data, it can skew the results of the regression. Therefore the data points must be examined sensibly before using this type of fitting. Different least square curve fits are:

1. Linear
   The following function try to fit a straight line through data.

   $$y = a1 + a2 * x \tag{2.26}$$

   No restrictions on data is associated with this type of curve.

2. Polynomial
   The following is the polynomial curve fitting equation.

   $$y = a1 + a2 * x + a3 * x^2 + a4 * x^3 + ... + a10 * x^9 \tag{2.27}$$

As the complexity of data curvature increases the polynomial order required for fitting also increases. No restrictions on data is associated with this type of curve.

3. Exponential
Following is the governing equation for exponential curve fit

$$y = a1 * e^{(a2*x)} \tag{2.28}$$

Generally this type of curve is used where the data decreases or increases at a higher rate. Data equal to zero or negative data cannot be fitted with this curve.

4. Logarithmic
The governing equation for this curve is

$$y = a1 + a2 * log(x) \tag{2.29}$$

The data which is fitted with type of curve is one that spans decades $(10^0, 10^1, 10^2...$ and so on) Data equal to zero or negative data cannot be fitted with this curve.

5. Power
The governing equation for this type of curve is

$$y = a1 * x^{a2} \tag{2.30}$$

Data equal to zero or negative data cannot be fitted with this curve.

2.6.5 Prognosis Methods

2.6.5.1 Kalman Filtering

It is an estimator which is optimal in nature. The parameters in which one is interested in can be obtained via inaccurate, uncertain and indirect measurements. It is recursive. Optimality refers to if all noise is Guassian then Kalman filter mean square error of estimated parameter is minimized. The popularity of Kalman filter is because of practical good results attainment due to structure. It is convenient for online real time processing. If given a basic understanding its formulation and implementation is easy. Inversion of measurement equations is not needed. It has been employed in various applications which includes: From limited earth observations, determination of planet orbit parameters. To track targets aircraft, missiles using RADAR, Localization of Robot and Building map from range sensors. The Kalman filtering approach consists of first a recursive Kalman estimator and then a Kalman predictor. Here the representation of pattern vector component is done by recurrent linear first order model. With this a modal state space representation is obtained

$$x_{k+1} = F_k x_k + G_k u_k + w_k \tag{2.31}$$

$$y_k = C_k x_k + v_k \tag{2.32}$$

The table 2.6.5.1 from [15] presents the correspondence between a Kalman Filter from system view in correspondence with view. They have applied Kalman Filter in operation modes tracking.

Table 2.1: Kalman Filter variables dimension and it correspondence with PR variables

	System point of view	**PR point of view**	**Dimension**
$\mathbf{x}_k$	State Vector	Pattern Vector	(d x 1)
$\mathbf{u}_k$	Control Input	Severity Degree	(1 x 1)
$\mathbf{y}_k$	Measurement	Polynomial Evolution	(d x 1)
$\mathbf{w}_k$ $\mathbf{v}_k$	Noises	-	(1 x 1)
$\mathbf{F}_k$	State Transition Matrix	Evolution Matrix	(d x d)
$\mathbf{G}_k$	Control Matrix	Severity Degree Weighting Matrix	(d x 1)
$\mathbf{C}_k$	Output Matrix	Identity Matrix	(d x d)
$\mathbf{L}_k$	Observation Matrix	Identity Matrix	(d x d)
$\mathbf{K}$	Evolution Variable	Severity Degree Modification	-

2.6.5.2 Particle Filters

It is a Bayesian method, recursive estimation is used for filtering problem solution. Estimation of the first two state vector moments is the filtering problem. Dynamic state space model along with noisy observation is what the space vector is governed. Following are the equations for particle filtering.

$$x_k = f_k(x_{k-1}, w_{k-1}) \tag{2.33}$$

measurement equation $y \epsilon \Re^k$ is given as

$$y_k = h_k(x_k, v_k) \tag{2.34}$$

The equation 2.33 is the state equation where as equation 2.34 is called the output equation. Instead of future state vector the probability of future state vector is estimated in Bayesian form. Using the pattern update equation the prior Probability Density Function (pdf) is calculated and from measurement equation posterior pdf is calculated. Particle Filters are analogus to Markov chain Monte Carlo (MCMC) batch methods. Sometimes they are even similar to importance sampling methods. It can be used as alternative to Extended Kalman Filter (EKF) or Unscented Kalman Filter (UKF) and with added advantage that with the availability of sufficient samples, they will lead to Bayesian optimal estimate. In that case they can be more accurate than EKF or UKF.

2.6.5.3 Hidden Markov Modelling

For signal monitoring Hidden Markov Model (HMM) is a stochastic technique through finite states. An HMM assumes states as hidden and system to be a Markov system. Main aim is to characterize state with the observation given. At time k state S_k is a state which is hidden, O_k is the observation sequence with the assumption of C possible states. From the observable parameters hidden parameters are to be computed.The main issues to be solved by HMM are as under:

- When observation sequence $y = y_1, y_2, \cdots, y_k$ and model parameter set $\theta = \pi, A, B$ is given then when model is given how to efficiently compute p(y) which is observation sequence probability.

- Observation sequence $y = y_1, y_2, \cdots, y_k$ and model parameter set $\theta = \pi, A, B$ is given then for optimal observation sequence generation how to choose state sequence $x = x_1, x_2, \cdots, x_k$. Maximum likelihood can be optimal measure.

- $\theta = \pi, A, B$ parameter adjustment for maximization of observation sequence likelihood.

Chapter 3

Problem Statement and Solution

3.1 Chapter Scope

Aim of this chapter is to develop a problem statement and to discuss the approach adopted as a solution.It also identify the inputs and data required for processing.The functional block diagram of overall algorithm is also given in this chapter.

3.2 Problem Statement

In this work, fault diagnosis of electromechanical systems with a focus on PMAC drive is addressed. The faults that occur in machine is of different natures. The problem here is to propose a cost effective diagnosis technique in order to deal the non-catastrophic electrical faults occurring in PMAC drive. Then forecasting of future machine state in order to detect upcoming fault behaviour falls in category of prognosis. Major problem with prognosis is system non-linearity and most important is non availability of system model.

3.3 Approach Adopted

. For the cost effective technique non intrusive analysis is preferred as recommended in [31]. As the machine fault is apparent in the current of the machine, MCSA is the

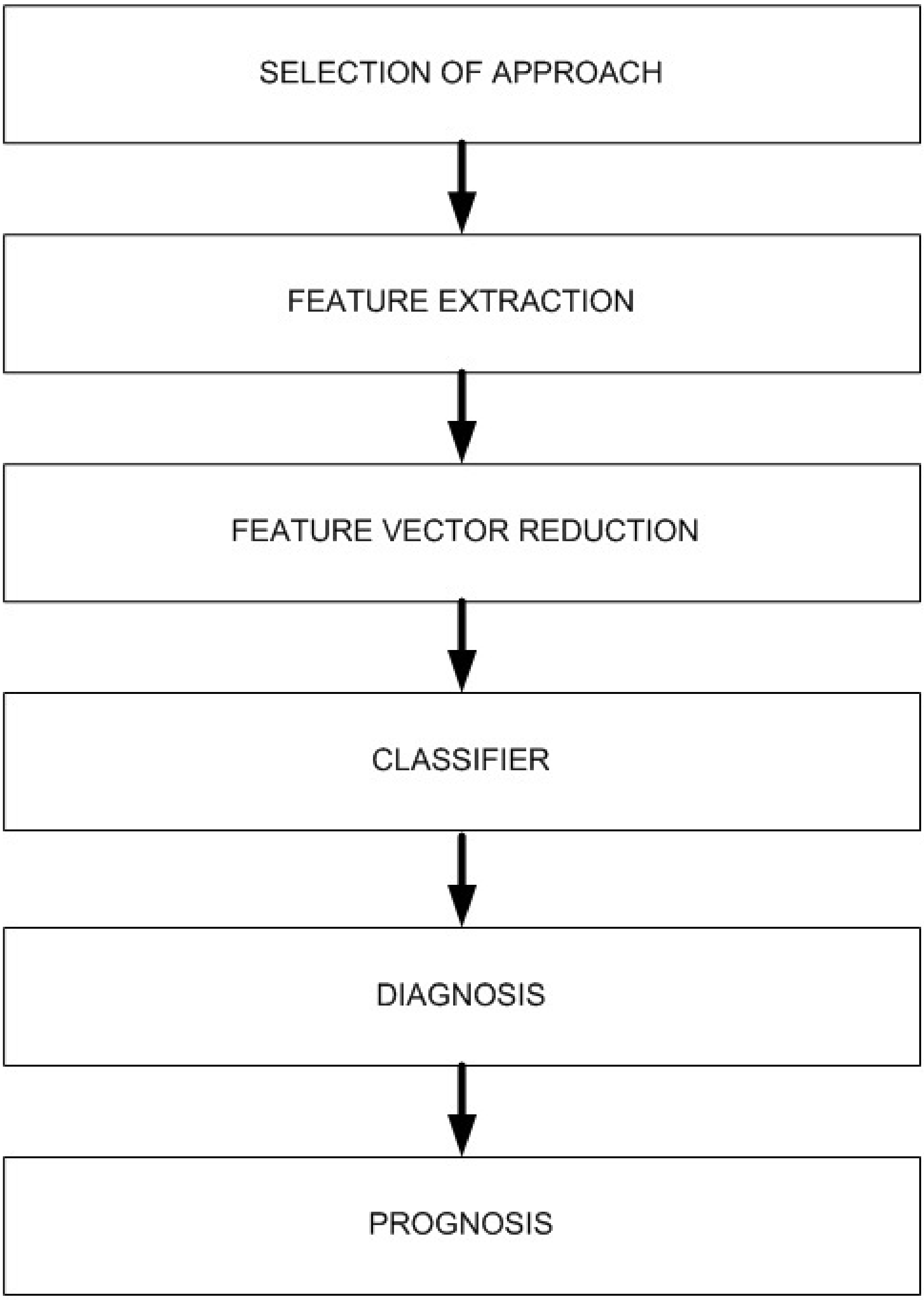

Figure 3.1: Methodology Adopted

best non-intrusive way to deal with faults as proposed in[17]. Among three different approaches that are discussed in section 2.2 and section 2.3 the approach adopted here is signal based approach and the two other approach which includes the data based and the model based is left. The model based approach requires a model which is impractical as electromechanical systems are complex and extensive efforts and approximations are required. In data based approach huge amount of data along with patterns of faults are required which is not practical. Using the supervised learning approach signal based method is adopted.

3.3.1 Feature Extraction Methods

The signals can be analysed in time, frequency or time frequency domains. As fault will arises from intermittent increased contact resistance so the fault will be transient one. Either time or frequency domain information separately is not enough for the extraction of features. Hence time frequency distribution can efficiently extract features from the signal. The joint representation of signal in time and frequency domains have three axis, one for time, one for frequency and one to indicate the amplitude of the signal. Here for the extraction of features Wigner Ville is applied. It is a bilinear transform and the computations it require is more than $(N^2 logN)$. This is adopted because bilinear transform produces better results [37]. Fault inception is detected by setting a threshold on the energy of the analysis coefficients.

3.3.2 Principal Component Analysis

Principal Component Analysis by variance maximization has been implemented in order to obtain the most discriminant features from the signal. The algorithm implemented for this is given in Algorithm 3.1. The algorithm works on Eigen Values and Eigen Vector calculation. For the observed data space, analysis of Principal component for correlation matrix results in orthonormal eigen-basis. The base largest eigen value refers to principal-components associated to most of co-variability in number of observed data.

Require: Arranged Data set.

Procedure:

Calculate size of the matrix.

Subtract mean of the input signal from each input sample.

Taking covariance of input samples.

Obtain Eigen Values and Eigen Vectors from covariance matrix.

Select the highest Eigen value and obtain the corresponding Eigen Vector.

Multiply the selected Eigen Vector to the zero mean shifted signal.

Each class samples are discriminated.

Calculate mean of each class sample that will act as each class representative point.

Algorithm 3.1: Principal Component Analysis

3.3.3 KNN Classifier

This was also adopted in [23] before applying clustering. K Nearest Neighbour (KNN) is selected as a classification technique because of its simplicity and efficiency. The distance metric used for KNN is Euclidean and Mahalanobis. The classification aims to classify different severity faults.

Require: Data set , Number of neighbors k which is set to 1.

Initialize:

construct distance matrix of data set using Equation(2.16)

repeat:

(**a**) find the closest k neighbor using Equation (2.16)

(**b**) merge the pair c_i, c_j where $i \neq j$, $\longrightarrow c_{ij}$

(**c**) find the neighbor of c_{ij} and update the distance matrix

until *stopping criteria are met; i.e. desired numbers of cluster of k neighbors*

Algorithm 3.2: k-Nearest Neighbor clustering algorithm

Some of the data is taken as training data whereas some of the data is taken as test data. The class representative points obtained from PCA is taken as training. Processes are applied on the same lines on the test data. Single representation is obtained from each test vector and distance of each training sample is calculated from the training means in order to classify from which class the each test point belongs to.

3.3.4 Forecasting the Future State

The phase of prognosis starts after the diagnosis phase. It refers to track the behaviour of machine's operating mode in order to estimate the machine useful life. Diagnosis and prognosis phase in terms of implementation can be seen from figure below showing algorithm flow. In order to track the evolution of fault, polynomial evolution approach is adopted [14] as it is relatively new and interesting. Using each fault representation a polynomial is fitted on those points to obtain a fault model also with this fault evolution can be extrapolated to know the future values. The problems with the Kalman Filtering [1] is that fault progression model is needed. The fault model obtain from polynomial approach will be used as seed model for Kalman Filtering.

3.3.5 Polynomial Curve Fitting

In order to obtain the fault progression model curve fitting tool in from MATLAB is applied. The parameter selected for curve fitting is polynomial one among the other methods. Coefficients are calculated on the basis of least squares method. It is an excellent tool that provides complete analysis of the fitted curve. It calculates residuals, show equations and offer variety of different methods for fitting of curve.

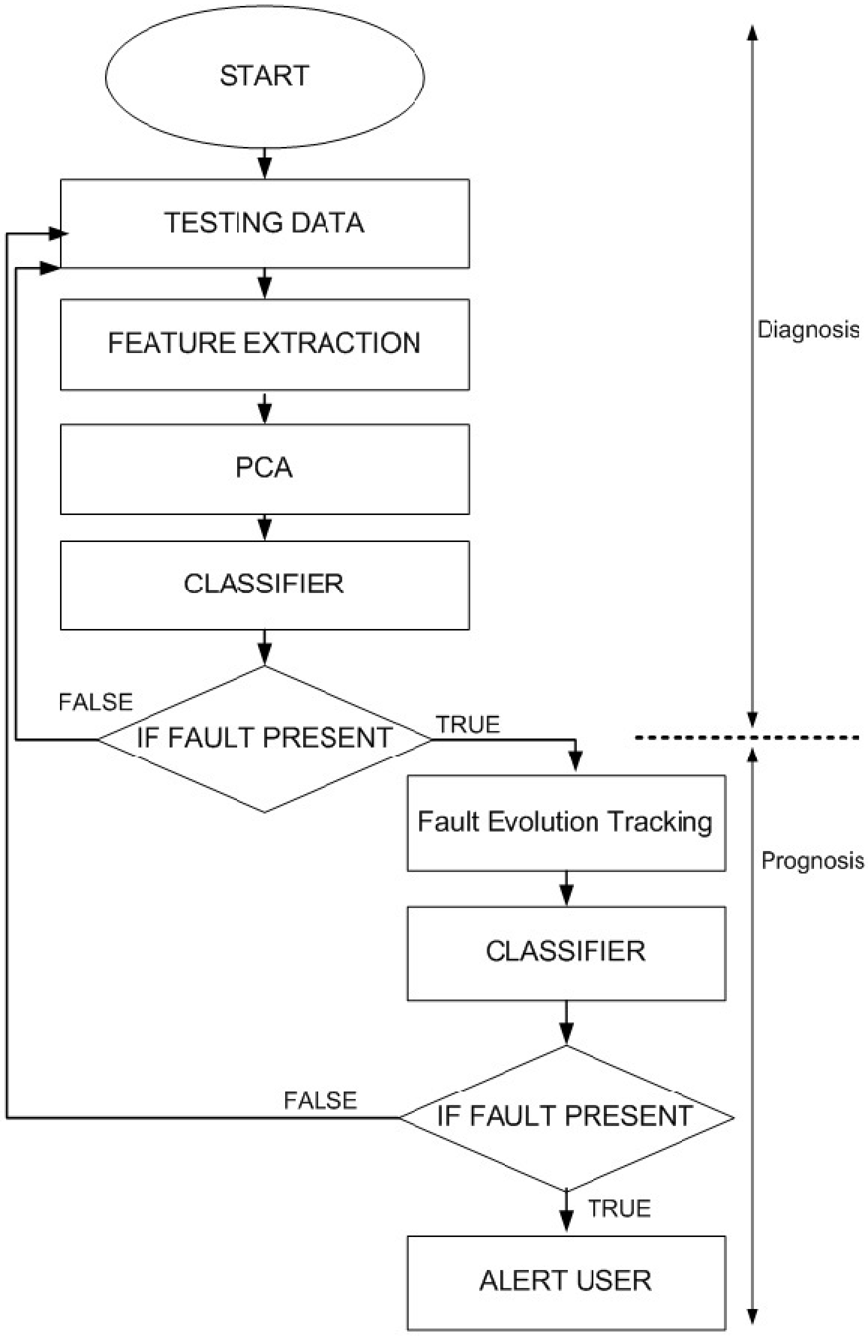

Figure 3.2: Algorithm Flow

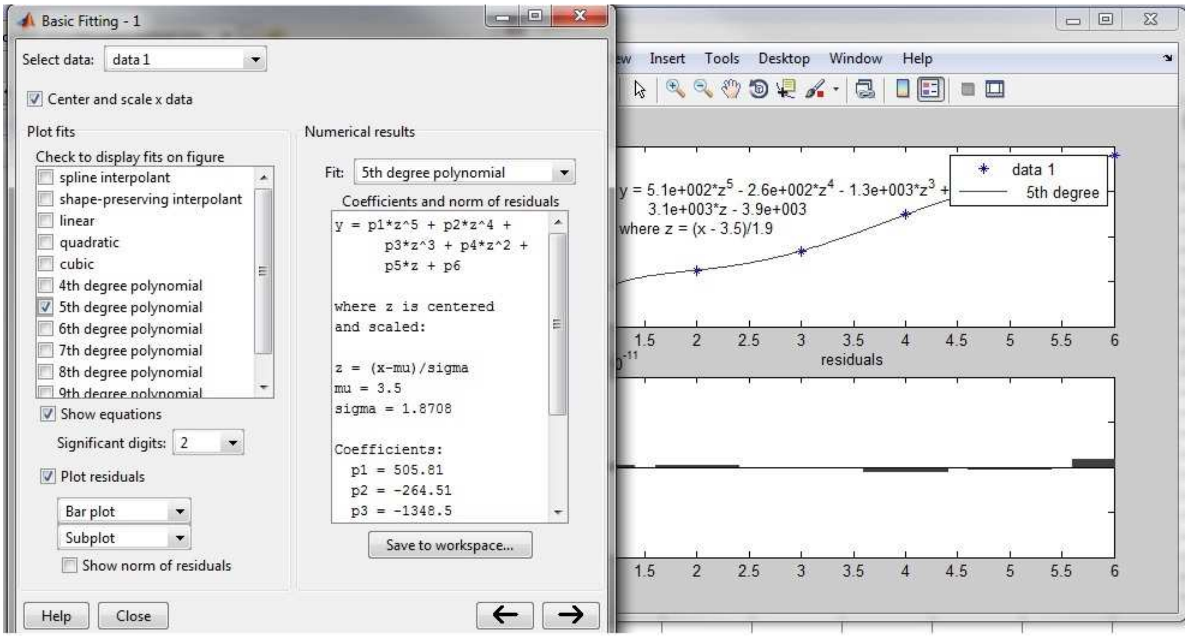

Figure 3.3: Curve Fitting Tool of MATLAB

Chapter 4

Experimental Setup and Explored Fault

The electrical machine in focus is Permanent Magnet Synchronous Machine (PMSM) of 12 V used in automotive application. The power rated is 1 hp and the no load speed is 3000 r/min approximately. Machine operated using constant control of torque-angle in a drive vector with set torque angle of $pi/2$ [20]. By this mode machine losses are minimized and this is suitable to speed of operation up to base speed. To determine fault type and severity a collection of data from both healthy and faulty drives are obtained. The synchronous machine rotor is equipped with one or more damper windings and field winding and in general all windings of rotor have different electrical characteristics. Variable change for rotor variables offers no benefit as a result of rotor asymmetries. However for stator variables, change of variables is advantageous [16]. In order to detect and classify faults machine stator currents are analysed. Mostly, stator variables is transformed to reference frame fixed in rotor as in Park‘s Equation [11]. Rather than analysing three phase machine current independent to each other field oriented currents i_{qs} and i_{ds} are used. By this the benefit is non presence of fundamental electrical frequency. The complete representation of stator currents can be obtained with i_{qs} and i_{ds}. However it is seen that with only i_{qs} achievement of accurate fault detection and classification can be obtained. Test condition are as follows:

1) The torque producing component of stator current command was iqs=0.3pu

2) Flux is held weakening component ids=0
3) Dynamometer speed is held constant to 400rpm
The recommended practice for data collection along with fault introduction and detection is experimental. Fault considered here is non catastrophic that implies continuous motor operation with the likelihood of increased failure. The fault explored in this work is intermittently increased contact resistance between motor and controller. For fault creation series resistance is added with normal closed switch in parallel to one of the phases of motor [38, 36]. By opening of switch fault is initiated for short time interval which causes current flow through resistance [33]. Different severity faults were intro-

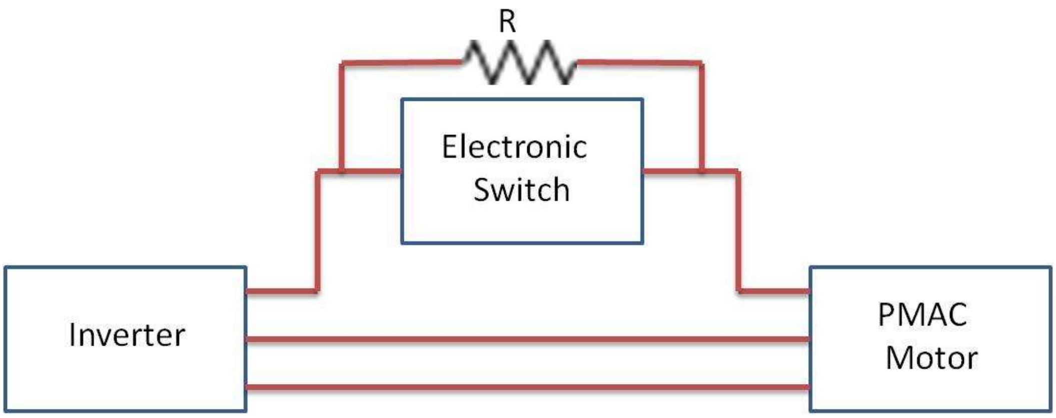

Figure 4.1: Series Resistance Fault

duced in the form of intermittent increased contact resistance. In comparison to the stator resistance the series resistance is almost ten times the value of stator resistance. As soon as fault is initiated the phase current rises to 95 percent of peak amplitude. The fault durations are 5ms and 10ms in order to know result invariances with respect to this parameter. Table 4.1 show the different fault severities in terms of pu.

Table 4.1: Fault Severity

Type of Fault severity	Increased Contact Resistance
Fault Severity 1	2.14pu
Fault Severity 2	2.80 pu
Fault Severity 3	4.03 pu
Fault Severity 4	6.33 pu
Fault Severity 5	15.84 pu

A computer running on Linux used as project controller. The computer running on Linux was used a controller for the project. Computer is superior to a Digital Signal Processor in terms of cost, central processing unit power and capacity of memory. The limitation in input and output capability of a computer is biggest hindrance. As a remedy custom Xilinx FPGA based input output board was developed. The resources of this board were twelve analog channels and quadrature encoder counter. Outputs were twelve digital and four analog channels. Parallel port is used for communication between input output board and computer.

Using current transducers with the rated accuracy of 0.45 percent two phases currents were measured having bandwidth of zero to two hundred kilohertz. For single line calibration single line-line was also measured. With the quadrature encoder count value of 1024 per revolution (4096 for quadrature) and with a pulse of index rotor position can be measured. A conducting electronic switch bidirection in nature was designed for the initiation of fault in stator.

This experimental setup was arranged with the permanent magnet synchronous motor along with the other circuitry to measure current. The fault phenomenon that will be addressed here will be transient and non catastrophic. The acceptable method for collection of data is that the fault development and detection strategy remains experimental. The experiments are conducted in order to mimic the electrical faults. From continous sampled signals data from these faults were extracted. There will be some issues that needs to be addressed in order to deal with the transient fault detection one is to make detection invariant to duration of fault and with the start of fault when fault is incepted. Algorithms are compared based on torque producing component analysis of field oriented stator currents in alternating current PMSM.

Chapter 5

Implementation and Obtained Results

The objective of this chapter is to present the results that are obtained by the implementation of different algorithms. Issues related to implementation is also discussed. MATLAB is the implementation platform. The results obtained are presented in the relevant sections. There are twelve samples from each class i.e healthy as well as faulty

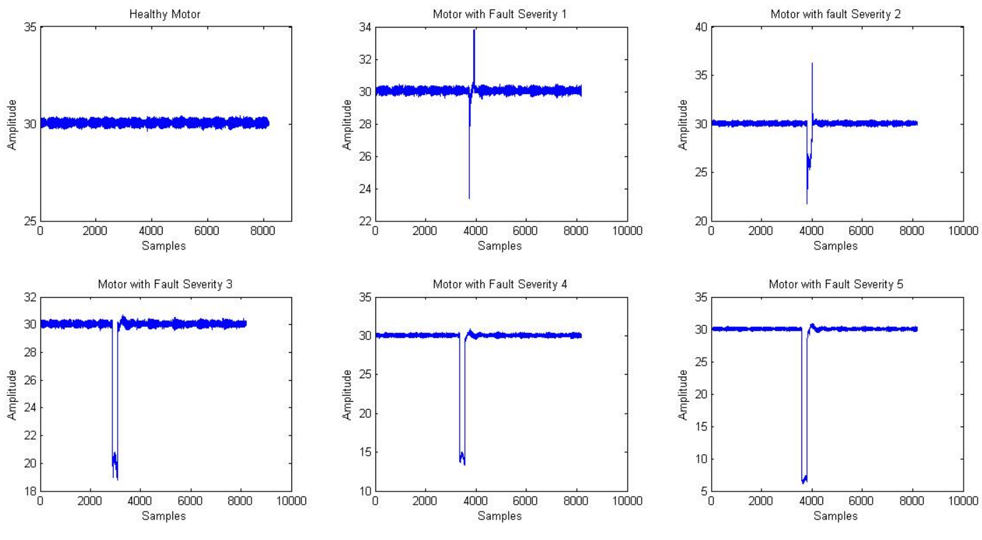

Figure 5.1: Healthy and Faulty Motor Current

machines. Each sample has 8192 observations. Some of the samples are taken as training and some are taken as test data. The figure 5.1 shows the healthy and faulty motor behaviour in time domain. Seeing the behaviour it can be concluded that motor faults are intermittent and non catastrophic.

5.1 Wigner Ville Distribution Results

As nature of signals is transient Wigner Ville time frquency transformation is applied up to 16 levels. Wigner Ville time frequency transformation was applied because it provides highest time frequency plane resolution. The time frequency features resulted will be taken as machine health indicators. The Wigner Ville Transform of the healthy motor in comparison to faulty motors can be seen as smooth in graph. This can be seen by a series of figures below. The fault was detected by thesholding energy to be 25 percent greater than the largest which was observed in all samples from healthy machine data.

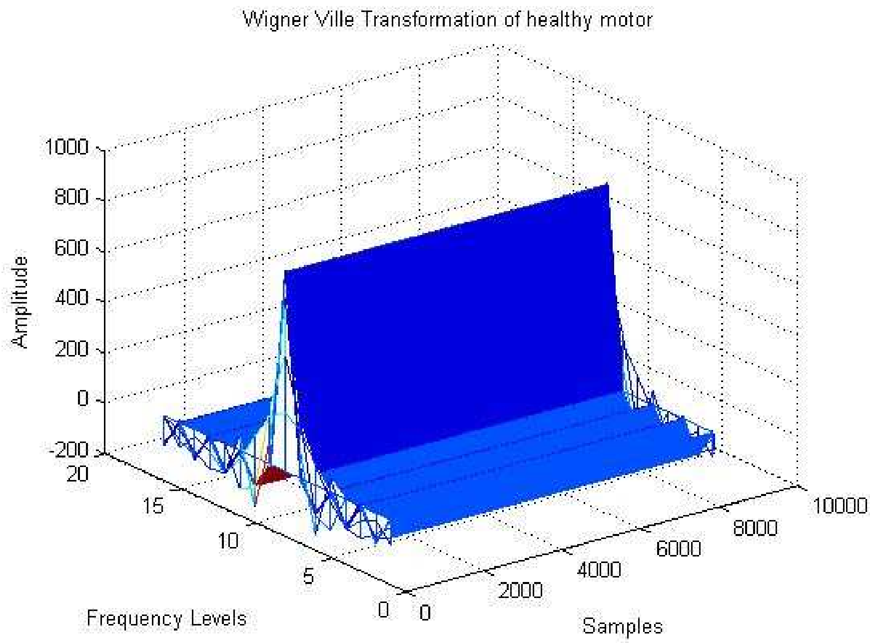

Figure 5.2: Wigner Ville Transformation of Healthy Motor

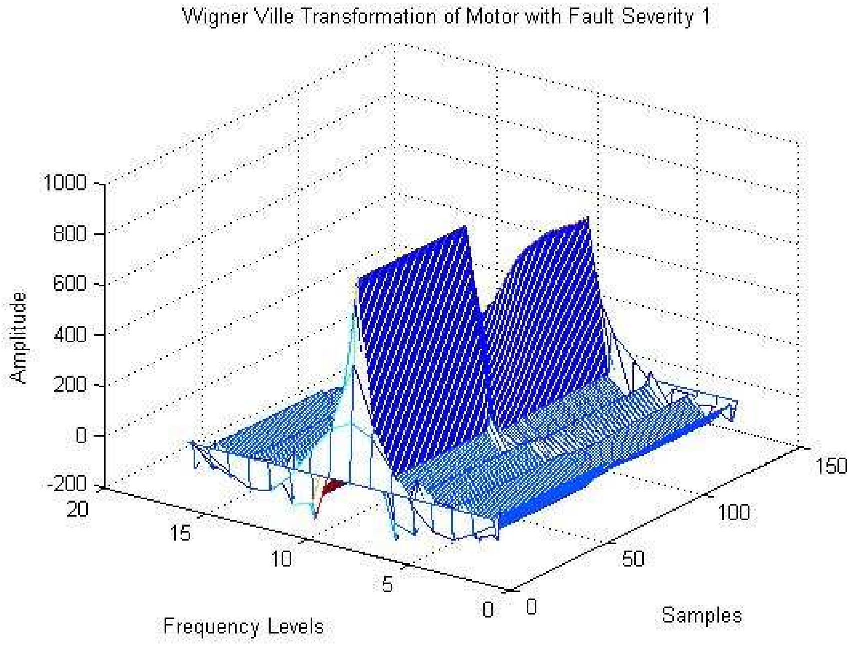

Figure 5.3: Wigner Ville Transformation of Fault Severity 1

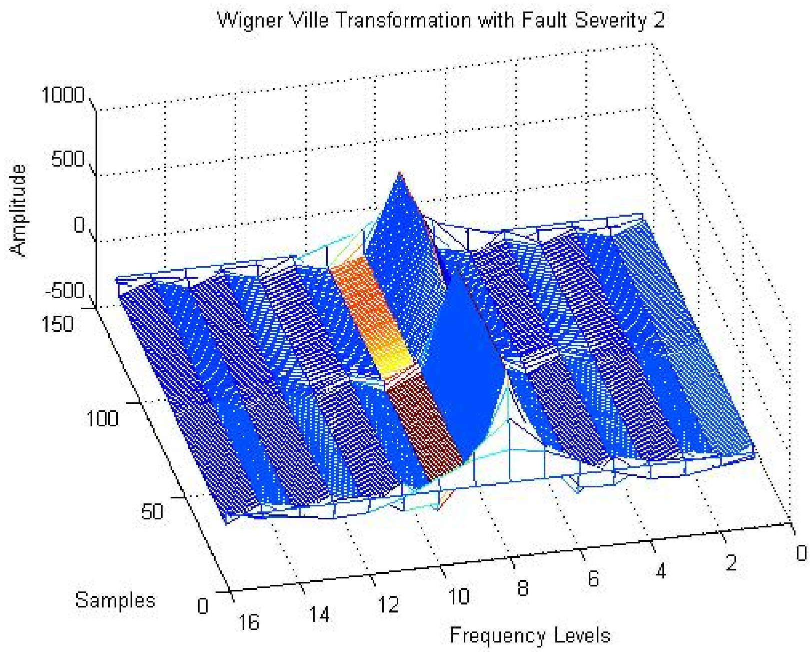

Figure 5.4: Wigner Ville Transformation of Fault Severity 2

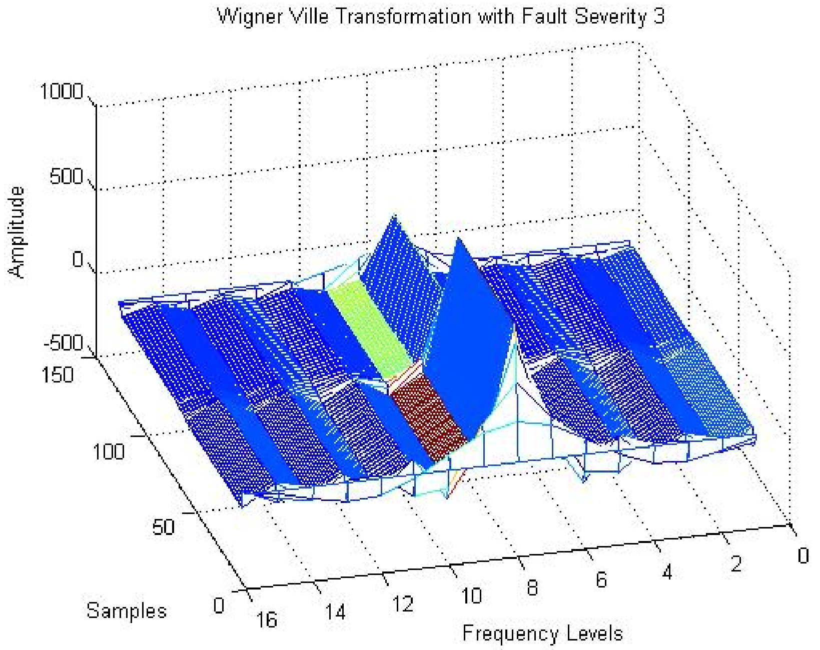

Figure 5.5: Wigner Ville Transformation of Fault Severity 3

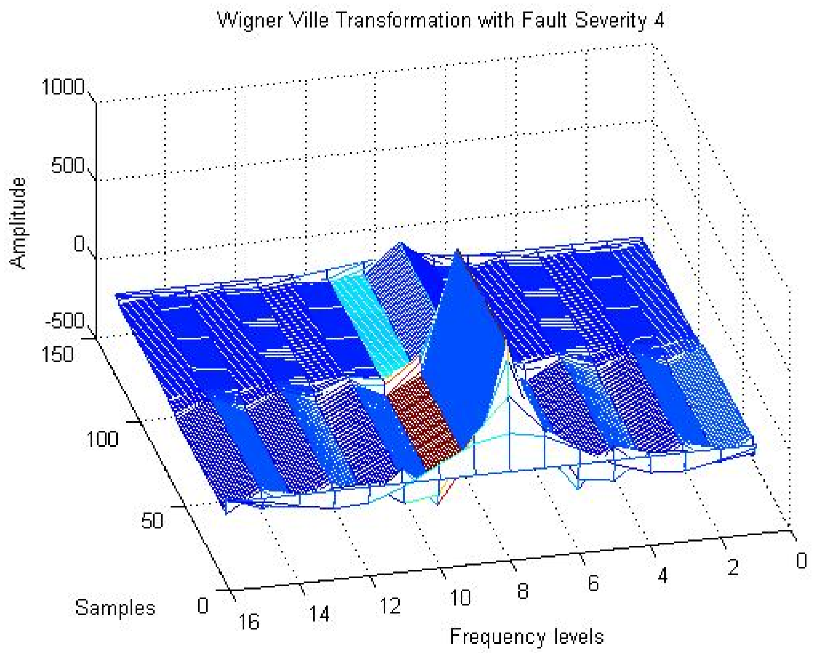

Figure 5.6: Wigner Ville Transformation of Fault Severity 4

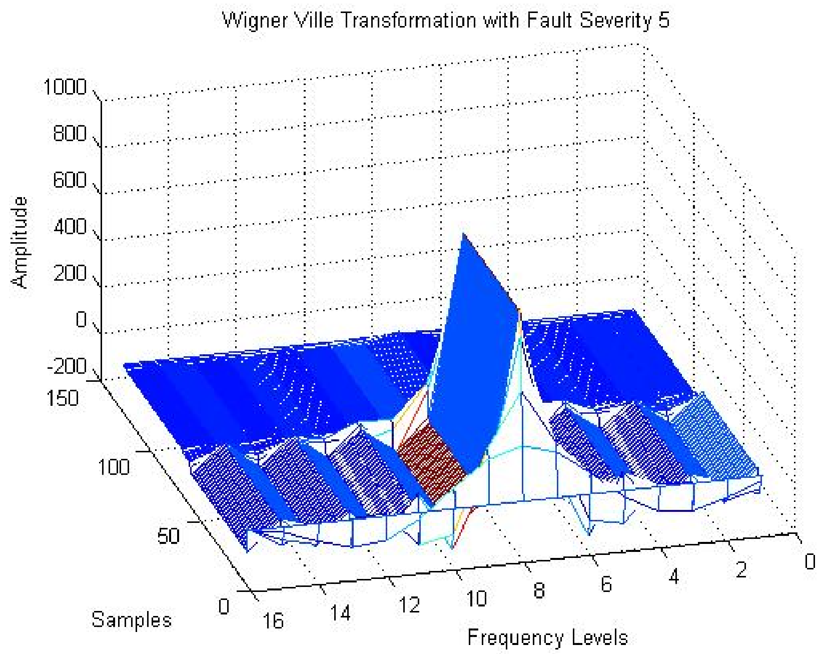

Figure 5.7: Wigner Ville Transformation of Fault Severity 5

5.2 Principal Component Analysis Results

All the data obtained from time frequency analysis is then formed in to a matrix. After applying time frequency analysis coefficients at each level of frequency is aligned to make a 1D matrix of that samples. All such coefficients are concatenated in matrix for application of PCA. Resulted Eigen values showed that the first Eigen value is very large in comparison to others. Data is projected according to first Eigen value. Plotting the resulted data shows each class of data discriminatingly. Each class data mean is taken as a representative of that class. The figure 5.8 shows how each class is represented discriminatingly.

5.3 KNN Classifier Results

The test data according to the same lines was classified with the help of the training data using the KNN algorithm with two different distance metric Euclidean and Mahalanobis.

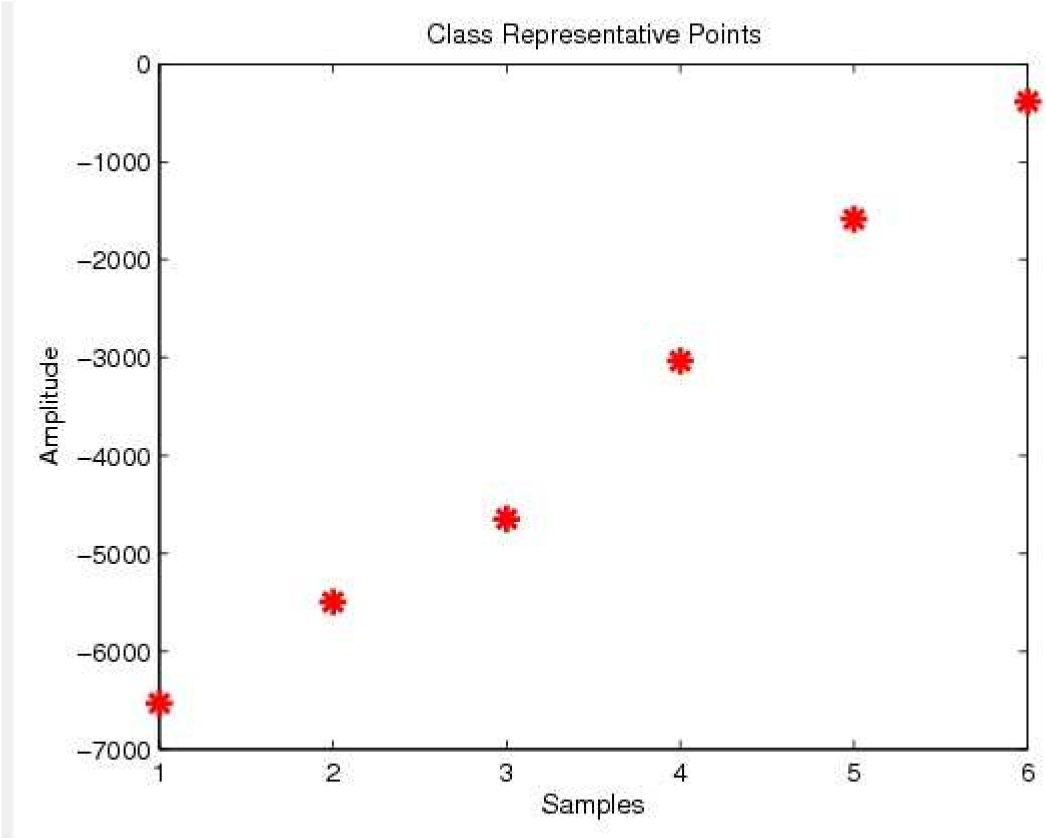

Figure 5.8: Data Class Represented with Points

When the training data set is 60 percent of the data and the test data set was forty percent of the test data accurate results were observed as indicated in table 5.1 and table 5.2. But when the training and test dataset was fifty percent of the data accurate results were not found with this implemented table 5.3. The value of K in KNN classifier is taken to be equal to 1.

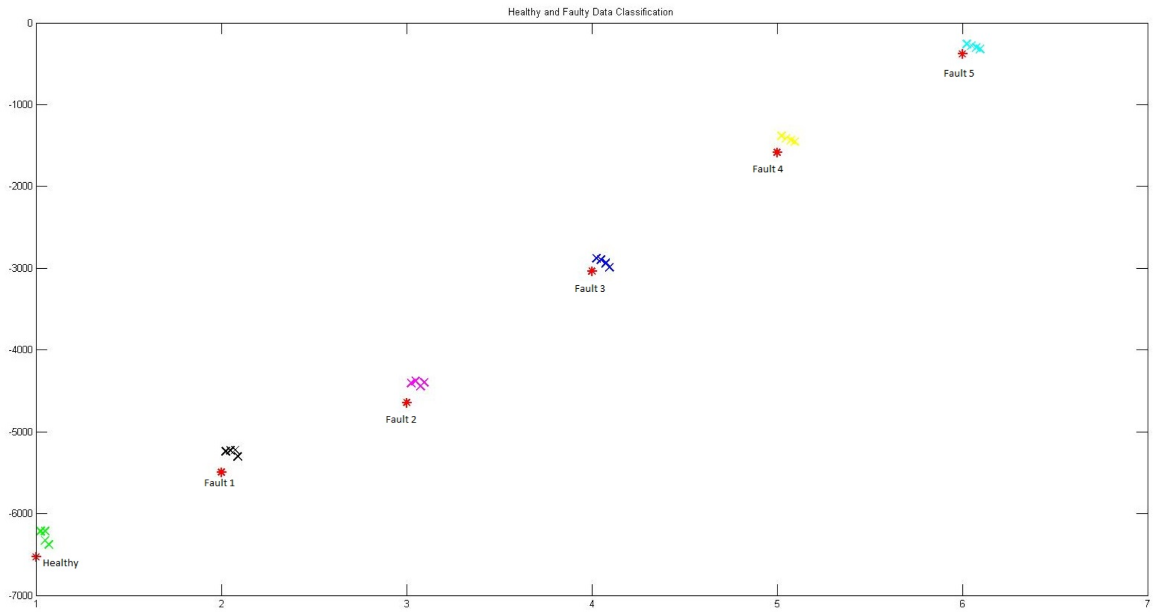

Figure 5.9: Fault Classification

Table 5.1: KNN Algorithm Performance Using Euclidean and Mahalanobis Distance(Training 50 percent of data Test 50 percent of data)

S.No	Test Description	False Detection	Total Samples/Detected/Classified
1	Healthy	0	4/0/0
2	Contact Resistance 2.14 pu	0	4/4/0
3	Contact Resistance 2.80 pu	0	4/4/4
4	Contact Resistance 4.03 pu	0	4/4/4
5	Contact Resistance 6.33 pu	0	4/4/4
6	Contact Resistance 15.84 pu	0	4/4/4

Table 5.2: KNN Algorithm Performance Using Euclidean Distance(Training 60 percent of data Test 40 percent of data)

S.No	Test Description	False Detection	Total Samples/Detected/Classified
1	Healthy	0	4/0/4
2	Contact Resistance 2.14 pu	0	4/4/4
3	Contact Resistance 2.80 pu	0	4/4/4
4	Contact Resistance 4.03 pu	0	4/4/4
5	Contact Resistance 6.33 pu	0	4/4/4
6	Contact Resistance 15.84 pu	0	4/4/4

Table 5.3: KNN Algorithm Performance Using Mahalanobis Distance(Training 60 percent of data Test 40 percent of data)

S.No	Test Description	False Detection	Total Samples/Detected/Classified
1	Healthy	0	4/0/4
2	Contact Resistance 2.14 pu	0	4/4/4
3	Contact Resistance 2.80 pu	0	4/4/4
4	Contact Resistance 4.03 pu	0	4/4/4
5	Contact Resistance 6.33 pu	0	4/4/4
6	Contact Resistance 15.84 pu	0	4/4/4

5.4 Polynomial Approach for Fault Model Results

A polynomial approach can be employed for fault modelling. If there are C classes then when a fault is present $C - 1$ represent the system evolution when failure is present as indicated in [14],[15].As this progression is slow in nature polynomial function with the following expression will be used.

$$x(i) = a(i) + \sum_{j=1}^{C-1}(b_j \mu^j) \tag{5.1}$$

where μ is the degree of fault severity where i is the number of dimension. So in order to see the trend each class representation is taken and polynomial is obtained through this. The fault progression polynomial will have one healthy and other faulty class. On the same lines using the techniques for the curve fitting techniques a fault progression polynomial is obtained. Figure 5.10 shows when there are six different classes the polynomial of degree five will yield the results with lowest residuals. The fault model obtained is

$$x = 13000 + 12000\mu - 8300\mu^2 + 2700\mu^3 - 390\mu^4 + 21\mu^5 \tag{5.2}$$

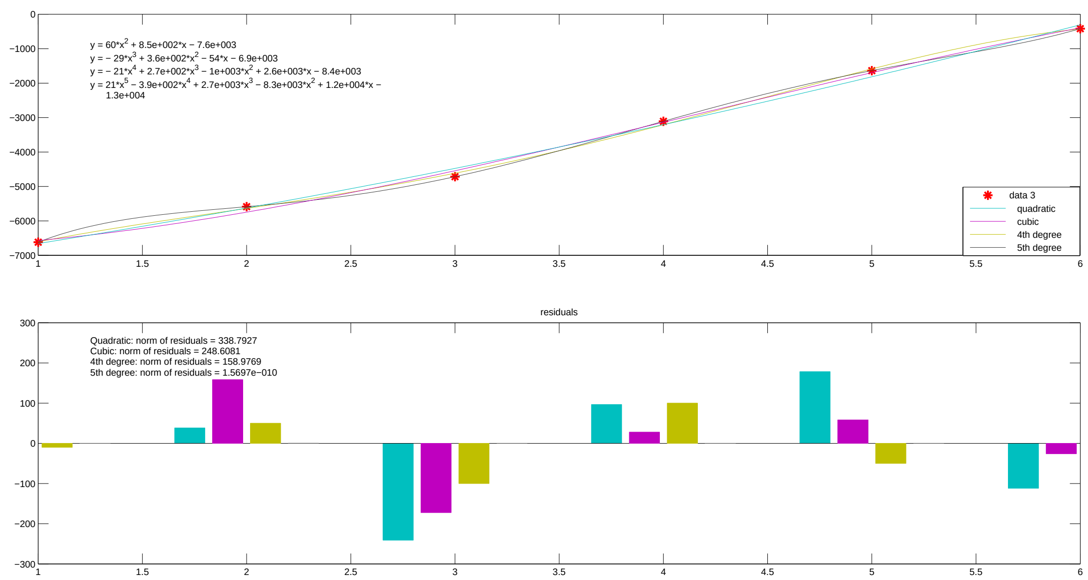

Figure 5.10: Fault Progression Model

Now, in order to go for prediction, polynomial function is extrapolated from the fault fault 4 severities to further severe faults in order to see the future trend. Seeing the forecast of the polynomial approach it is seen that it does not yield very good results. This fault model will be used as a seed fault model in Kalman Filtering system.

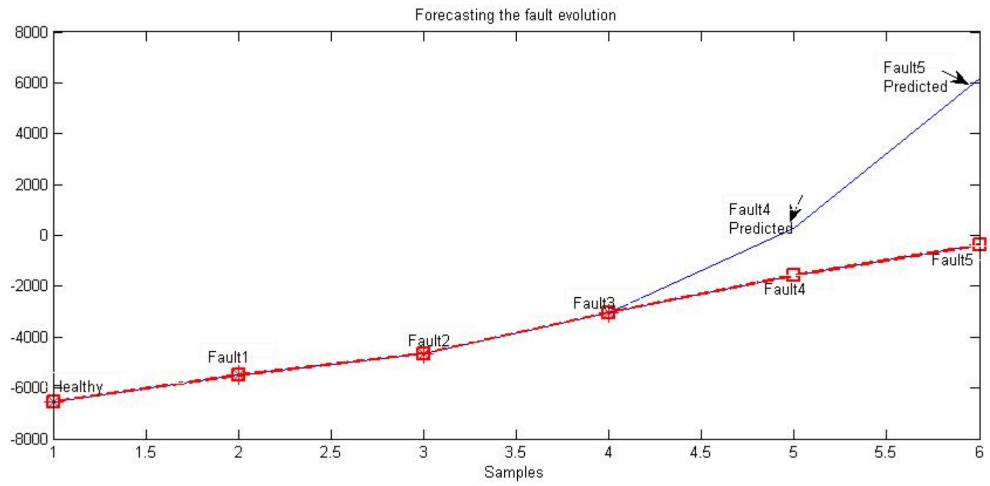

Figure 5.11: Forecasting Fault Evolution through Polynomial Approach

Chapter 6

Conclusion and Future Recommendations

6.1 Conclusion

This research work is composed of two phases one is diagnosis in which a technique of principal component analysis is introduced leaving the traditional ones. All the analysis has been done using the torque producing component of the stator current. The developed algorithm for diagnosis was tested on the test data. In the other phase of prognosis a fault model is developed using the polynomial approach. But this approach is unable to give appreciable results for forecasting. As it cannot model the dynamics involved from one fault to another. So the developed model is proposed to be used as a seed model for the Kalman Filtering approach with the expectation to give better results. The approach can be adopted on industrial scale in order to schedule maintenance of machines.

6.2 Future Recommendations

The technique developed for diagnosing and prognosis of faults focusing a specific electrical machine. These technique can be applied to other machines. The developed technique can be used at system level maintenance with the addition of monitoring. The algorithms developed so far for mitigating these faults and to protect machine from failures includes signal processing techniques, statistical technique and prediction techniques so there is a room of selecting better approaches in these domains instead of the adopted algorithms. The main outlines for the future recommendations are:

- The algorithms developed has been tested on the offline machine data. This can also be used for online machine fault diagnosis and prognosis.

- For online analysis this work can also be implemented on reconfigurable platform such as Field Programmable Gate Array (FPGA). Other than FPGA, Digital Signal Processor (DSP) is also a good choice.

- The developed technique can also be implemented on other machine such as DC machines and induction machines.

- The modelling with the Kalman Filter can give more accurate results with the better estimated parameters.

- For the failure prognosis several other techniques can also be implemented other than HMM and Kalman Filtering. Particle Filtering and Neuro fuzzy techniques can be used for prognosis.

- A comparative analysis of various prognosis techniques in terms of accuracy, implementation complexity and cost can also be done to find out the best suitable one.

Appendix A

Acronyms

CBM	Condition Based Monitoring
CWD	Choi Williams Distribution
DSP	Digital Signal Processor
DWT	Discrete Wavelet Transform
EKF	Extended Kalman Filter
FFT	Fast Fourier Transform
FPGA	Field Programmable Gate Array
HMM	Hidden Markov Model
LDC	Linear Discriminant Classifier
LPF	Low Pass Filter
MCMC	Markov chain Monte Carlo
MCSA	Motor Current Signature Analysis
MDA	Multiple Discriminant Analysis
OSH	Optimal Separating Hyperplane

PCA	Principal Component Analysis
pdf	Probability Density Function
PMAC	Permanent Magnet AC Machine
PMSM	Permanent Magnet Synchronous Machine
RUL	Remaining Useful Life
STFT	Short time Fourier Transform
SVM	Support Vector Machine
UKF	Unscented Kalman Filter
WT	Wavelet Transform
WVD	Wigner Ville Distribution

Bibliography

[1] I.B.A. Bazine, S. Tnani, T. Poinot, G. Champenois, and K. Jelassi. On-line detection of stator and rotor faults occurring in induction machine diagnosis by parameters estimation. In *Diagnostics for Electric Machines, Power Electronics Drives (SDEMPED), 2011 IEEE International Symposium on*, pages 105 –112, sept. 2011.

[2] B.L.Theraja and A.K.Theraja. *A textbook of Electrical Technology (In S.I Units)*, volume 2. S.Chand, 1997.

[3] Boualem Boashash. *Time-Frequency Signal Analysis and Processing A Comprehensive Reference.*

[4] P. J. Chi and C. T. Russell. Use of the wigner ville distribution in interpreting and identifying ulf waves in triaxial magnetic records. *Journal of Geophysical Research: Space Physics*, 113, January 2008.

[5] L. Cohen. *Time-Frequency Analysis.* Prentice Hall, 1995.

[6] Yao Da, Xiaodong Shi, and M. Krishnamurthy. Health monitoring, fault diagnosis and failure prognosis techniques for brushless permanent magnet machines. In *Vehicle Power and Propulsion Conference (VPPC), 2011 IEEE*, pages 1–7, 2011.

[7] H. Douglas, P. Pillay, and A.K. Ziarani. A new algorithm for transient motor current signature analysis using wavelets. In *38th IAS Annual Meeting. Conference Record of the Industry Applications Conference, 2003.*, volume 2, pages 911–916 vol.2, 2003.

[8] S. Aviyente E. Strangas and S. Zaidi. Time frequency analysis for efficient fault diagnosis and failure prognosis for interior permanent magnet ac motors. In *IEEE Transactions on Industrial Electronics*, volume 55, pages 4191–4199, December 2008.

[9] K. Medjaher, D.A. Tobon-Mejia, and N. Zerhouni. Remaining useful life estimation of critical components with application to bearings. *Reliability, IEEE Transactions on*, 61(2):292–302, 2012.

[10] Neelam Mehala. *Condition Monitoring and Fault Diagnosis of Induction Motor Using Mo Current Signature Analysis*. PhD thesis, National Institute of Technology,Kurukshetra,India, October 2010.

[11] M.Riaz. *Simulation of Electrical Machine and Drive Systems using MATLAB and SIMULINK*. Department of Electrical and Computer Engineering University of Minnesota, http://www.ece.umn.edu/users/riaz/, 2013.

[12] Ratna Dahiya Neelam Mehala. Motor current signature analysis and its applications in induction motor fault diagnosis. In *7th WSEAS International conference on Signal Processing, Robotics and Automation (ISPRA-08),*, volume 2, 2007.

[13] Aapo Hyvtirinen Juha Karhunen Erkki Oja. *Independent Component Analysis*. JOHN WILEY & SONS, INC., 2001.

[14] O. Ondel, E. Blanco, and G. Clerc. Beyond the diagnosis: the forecast of state system application in an induction machine. In *Diagnostics for Electric Machines, Power Electronics and Drives, 2007. SDEMPED 2007. IEEE International Symposium on*, pages 491–496, 2007.

[15] O. Ondel, E. Boutleux, Eric Blanco, and G. Clerc. Coupling pattern recognition with state estimation using kalman filter for fault diagnosis. *IEEE Transactions on Industrial Electronics*, 59(11):4293–4300, 2012.

[16] Scott D. Sudhoff Paul C. Krause, Oleg Wasynczuk. *Analysis of Electric Machines and Drive Systems*. John Wiley & Sons, second edition, 2002.

[17] P. Pillay and Z. Xu. Motor current signature analysis. In *Thirty-First Annual Meeting on Industry Applications Conference, 1996.*, volume 1, pages 587–594 vol.1, 1996.

[18] Robi Polikar. The engineer‘s ultimate guide to wavelet tutorial. Web, 2006.

[19] S. Ray and R Turi. Determination of number of clusters in k-means clustering and application in colour image segmentation. In *4th international conference on advances in pattern recognition and digital techniques*, pages pages 137 –143. Narosa Publishiing House, 1999.

[20] R.Krishnen. *Electric Motor Drives: Modeling, Analysis and Control.* Englewood Cliffs,NJ:prentice-Hall, 2001.

[21] K. Bongort S. Engel, B. Gilmartin and A. Hess. Prognostics, the real issues involved with predicting life remaining. In *Aerospace Conference Proceedings*, volume 6, pages 457–469, 2000.

[22] Fengxi Song, Zhongwei Guo, and Dayong Mei. Feature selection using principal component analysis. In *System Science, Engineering Design and Manufacturing Informatization (ICSEM), 2010 International Conference on*, volume 1, pages 27–30, 2010.

[23] A. Soualhi, G. Clerc, H. Razik, and O. Ondel. Detection of induction motor faults by an improved artificial ant clustering. In *IECON 2011 - 37th Annual Conference on IEEE Industrial Electronics Society*, pages 3446–3451, 2011.

[24] Thanis Sribovornmongkol. Evaluation of motor online diagnosis by fem simlations. Master's thesis, School of Electrical Engineering Royal Institute of Technology Stockholm, 2006.

[25] Elias Strangas and Selin Aviyente. Fault diagnosis, prognosis and reliability of electrical drives. Web.

[26] D.A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot. A data-driven failure prognostics method based on mixture of gaussians hidden markov models. *IEEE Transactions on Reliability,*, 61(2):491–503, 2012.

[27] H. K. Khalil W. G. Zanardelli, E. G. Strangas and J. M. Miller. Wavelet- based methods for the prognosis of mechanical and electrical failures in electric motors. In *Mechanical Systems and Signal Processing*, pages 411–426, 2005.

[28] Wikipedia. Wigner quasiprobability distribution. Web, Feb 2013.

[29] Ronald J. Gilmore William T. Thomson. Motor current signature analysis to detect faults in induction motor drives–fundamentals, data interpretation, and industrial case histories. In *Proceedings of the Thirty Second Turbomachinery*, pages 145–156, 2003.

[30] S.S.H. Zaidi, W.G. Zanardelli, S. Aviyente, and E.G. Strangas. Prognosis of electrical faults in permanent magnet ac machines using the hidden markov model. In *IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society*, pages 2634–2640, 2010.

[31] Syed Sajjad Haider Zaidi. *Fault Diagnosis and Faliure Prognosis of Electrical Machines*. PhD thesis, Michigan State University, 2010.

[32] W. G. Zanardelli. The use of wavelet analysis for the prognosis of failures in electric motors. Master's thesis, Michigan State University, 2000.

[33] Wesley G. Zanardelli. *Methods to Identify Intermittent Electrical and Mechanical Faults in Permanent Magnet AC Drives Based on Time-Frequency Analysis*. PhD thesis, Michigan State University, 2005.

[34] W.G. Zanardelli, E.C. Strangas, H.K. Khalil, and J.M. Miller. Comparison of wavelet based methods for the prognosis of failures in electric motors. In *Power Electronics in Transportation, 2002*, pages 61 –67, oct. 2002.

[35] W.G. Zanardelli, E.G. Strangas, and S. Aviyente. Failure prognosis for permanent magnet ac drives based on wavelet analysis. In *Electric Machines and Drives, 2005 IEEE International Conference on*, pages 64–70, 2005.

[36] W.G. Zanardelli, E.G. Strangas, and S. Aviyente. Intermittent fault identification for permanent magnet ac drives based on the short-time fourier transform. In *Diagnostics for Electric Machines, Power Electronics and Drives, 2005. SDEMPED 2005. 5th IEEE International Symposium on*, pages 1 –6, sept. 2005.

[37] W.G. Zanardelli, E.G. Strangas, and S. Aviyente. Comparative study of time frequency methods for the detection and categorization of intermittent faults in electrical drives. In *IEEE Transactions on Industry Applications*, number 4, pages 971 –980, july-aug. 2007.

[38] W.G. Zanardelli, E.G. Strangas, and S. Aviyente. Identification of intermittent electrical and mechanical faults in permanent-magnet ac drives based on time frequency analysis. *IEEE Transactions on Industry Applications,*, 43(4):971–980, 2007.

Index