CLUSTERING OF FREQUENCY-BASED VIBRATION SIGNAL FOR BEARING FAULT DETECTION

CHIA MING XUAN

UNIVERSITI MALAYSIA PAHANG

CLUSTERING OF FREQUENCY-BASED VIBRATION SIGNAL FOR BEARING FAULT DETECTION

CHIA MING XUAN

Report submitted in partial fulfillment of the requirements
for the award of Bachelor of Mechanical Engineering

Faculty of Mechanical Engineering
UNIVERSITI MALAYSIA PAHANG

JUNE 2013

EXAMINER'S DECLARATION

I certify that the project entitled "Clustering of Frequency-Based Vibration Signal for Bearing Fault Detection" is written by Chia Ming Xuan. I have examined the final copy of this report and in my opinion, it is fully adequate in terms of language standard, and report formatting requirement for the award of the degree of Bachelor of Engineering. I herewith recommend that it be accepted in partial fulfillment of the requirements for the degree of Bachelor of Mechanical Engineering.

MR. MUHAMMAD HATIFI BIN MANSOR

Examiner   Signature

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this project and in my opinion, it is fully adequate in terms of language standard, and report formatting requirement for the award of the degree of Bachelor of Mechanical Engineering.

Signature	:
Name of supervisor	: MR. MOHD FADHLAN BIN MOHD YUSOF
Position	: Lecturer
Date	: 18 JUNE 2013

STUDENT'S DECLARATION

I hereby declare that the work in this report is my own except for quotations and summaries which have been duly acknowledged. The thesis has not been accepted for any degree and is not concurrently submitted for award of the other degree.

Signature :
Name : CHIA MING XUAN
ID Number : MA09042
Date : 18 JUNE 2013

I dedicate this report to my family, Chia Kee Woon, Koh Mui Guat, Chia Ming Yong, Chia Shan Shan and those who have guided me for this project

ACKNOWLEDGEMENTS

I am grateful and like to express my utmost sincere gratitude to my supervisor Mr. Mohd Fadhlan Bin Mohd Yusof for his profound ideas, invaluable guidance, continuous encouragement and constant support in making this research possible. He indefinitely has impressed me with his vision, outstanding professional conduct, strong conviction for engineering and his belief that a Degree program is only a start of a life-long learning experience. I greatly appreciate his consistent support from the first day I chose this title as my Final Year Project till these concluding moments. I salute his progressive vision about my work, tolerance in my naive mistake even though time and time again he has correct me to it and his commitment of my future career path. I would also to express very special thanks to Mr. Che Ku Eddy Nizwan Bin Che Ku Husin, for his teaching, guidance, suggestion and cooperation throughout the research. I also sincerely thanks for the time spent proofreading and correcting my tolerable mistakes.

My sincere thanks also go to my research-mate, members of the staff of the Mechanical Engineering Department, UMP, who helped me in various ways and made my stay at UMP pleasant and unforgettable. Infinite thanks also goes out to the member of the Advanced Structural Integrity and Vibration Research (ASiVR) group for their excellence cooperation, inspirations, guidance and support during my research.

I acknowledged my sincere indebtedness and gratitude to my parents, Chia Kee Woon and Koh Mui Guat for their love, dream, support and sacrifice from the moment I was born till now. I am also grateful to my brother and sister for their support and understanding. A big thanks to my friends who helped me, support me, push me, tolerate me and for their ideas to make this work possible. I cannot find any appropriate words to describe my appreciation and indebtedness to all of the people that have directly or indirectly involved in making this research successful. Lastly I would like to acknowledge all of their comments and suggestions, which was crucial for the successful completion of this study.

ABSTRACT

Bearing is one of the vital parts in any rotating machinery. Failure of this particular part can affect the machinery performance and in time will cause major failure to the machinery. Due to this crucial problem, on-line monitoring has become an alternative in prevention maintenance. The objective of this project is to study the trend of frequency spectrum from different bearing defects and to apply clustering approach using Principle Component Analysis, PCA on frequency domain signals. A set of good condition bearing is used along with four types of defective bearing which are inner race defect, corroded defect, contaminated defect and lastly roller defect. The signals are acquired using a PCB piezoelectric accelerometer and a National Instrument Data Acquisition System (NI-DAQ). The bearing will be run on three speed rotation which is 440, 1480 and 2672 RPM. The data is acquired by using DASYLab software, for both the time domain and frequency domain signals. The data then analyzed using PCA method through MATLAB software. Data is then plotted on scatter plot. After that, the data will be clustered using Agglomerative Hierarchical Clustering where a dendrogram is used to show a cluster of data in which the respective data for all types of bearing tested remain in their cluster. Finally, this method is suggested as an alternative in bearing fault detection, especially online monitoring.

ABSTRAK

Galas adalah salah satu komponen penting terdapat di dalam mesin. Kegagalan komponen ini boleh memberi kesan pada prestasi mesin dan apabila tiba masanya akan mengakibatkan kegagalan lebih besar kepada mesin berkenaan. Akibat daripada ini, pemantauan atas talian menjadi pilihan alternative dalam servis pencegahan. Objektif projek ini adalah untuk mengkaji isyarat getaran daripada galas yang berkeadaan baik dan juga galas rosak melalui senser accelerometer. Set galas berkeadaan baik digunakan bersama-sama dengan empat jenis galas rosak iaitu cela permukaan dalam, karat, galas tercemar dan cela pada pengguling . Data diperolehi menggunakan PCB piezoelectrik accelerometer dan Sistem Penerimaan Data National Instrument (NI-DAQ). Galas akan di uji menggunakan tiga jenis kelajuan iaitu 440, 1480 dan 2672 RPM. Kemudian data diperoleh dengan menggunakan perisian DASYLab. Data kemudian diplot membentuk taburan data plot. Selepas itu, data akan di guguskan menggunakan Agglomerative Hierarchical Clustering di mana dendrogram akan digunakan untuk menunjukkan gugusan data di mana semua data jenis-jenis galas ujian akan berada di dalam gugusan mereka sendiri. Akhir sekali, kaedah ini dicadangkan sebagai alternatif di dalam pengesanan galas rosak, terutama sekali di dalam pemantauan dalam talian.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Hz	hertz
kHz	kilo hertz
F	force
k	stiffness
c	damping coefficient
x	displacement
$\dot{x}$	velocity
$\ddot{x}$	acceleration
k	spring constant
m	mass
Δ	distance travel
w	weight
$\bar{x}$	mean of x
C^{n^2}	matrix with n rows and n columns
c, cov	covariance
Dim	dimension
eig	eigenvector
mV/g	mili volts per gravity

LIST OF ABBREVIATIONS

ASiVR	Advance Structural Integrity and Research
NI-DAQ	National Instrument Data Acquisition System
PCA	Principal Component Analysis
RPM	Revolution per minute
rev	revolution
min	minute
SOM	Strength of material
D. C.	Direct current
FFT	Fast Fourier Transform
FDI	Fault Detection and Isolation
PC	Principle component

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

A bearing is a mechanical device that supports the moving parts of a machine and to guide or to confine its rotational or linear motion, while preventing motion in the direction of applied load, i.e., reducing the friction and handling stress. Its primary purpose is to reduce friction with a minimum loss of energy between the surface of the bearing and the surface it's rolling over and at the same time controlling the rate of wear.

Generally, bearings are used in a wide, diverse range of general industrial applications, including electrical switch gears, solenoids, textile machinery, food processing equipment, traffic restrictor and pipe hangers, among many others. Due to their low friction and extended wear life, they are also used in an extensive variety of manufacturing equipment in applications such as bending presses, multi-ram press guides and machine tool slide-ways, as well as injection molding, and industrial solder machines, heat treating units and cable winding and braiding machines.

The faults or defects of bearing in rotating machineries can be extremely critical because these may lead to crucial machinery damage, production losses and serious personnel injury. Hence, it is a very important duty for the maintenance department to detect these faults right during in their initial stage.

1.2 PROBLEM STATEMENT

In industries, especially the manufacturing field industries, safety is the main concern. Safety here of course is mainly referring to the staff's safety and also the machinery tools. Furthermore, one of the main responsible of an engineer is to make sure the safety in certain production plant is well performed and assured.

On September 23, 2005, a motor coach incident probable caused by insufficient lubrication in the right-side tag axle wheel bearing assembly of the motor coach ended 23 passengers were fatally injured. An aviation disaster of LOT Polish Airlines Flight 5055, on 9 May 1987, the cause of the crash was the disintegration of an engine shaft due to faulty bearings inside engine No. 2, which seized, causing extensive heat and caused all 184 passenger perished. (Gero and David, 1997).

Based on crucial causes and incidents that previously revealed, it is obviously that the need of a proper bearing maintenance or condition monitoring method is really indeed required to look for and hence to avoid or minimize any unwanted incidents to be happen again.

Previously, to solve the bearing defect and bearing maintenance problems, normal practices in industries were mainly using manual inspections which are using visual inspection, statistical method (bearing life estimation), bearing condition monitoring by crack size detection, lubrication and simply changing the defect bearings in order to overcome bearing faults (Hirani, 2000). This seems that the bearing defects were detected only when the defects had shown obvious sound and vibration signals. Manual inspections are not only expensive, but also connected with a risk of accidentally causing damages when reassembling a machine. Thus there is a clear need for non-destructive methods for predicting bearing damages early enough to wait with bearing replacements until next scheduled stop for machine maintenance (Ericsson et al., 2005).

The most successful such methods in use today are all based on vibration analysis (Wowk, 1991). They do, however, require special competence from the user, whereas, as the industry optimizes there is less personnel and time available for condition monitoring. Moreover, to hire special competence from outsource involved huge cost. Thus important information to support decisions is lost and there is a demand for more automatized and supportive bearing monitoring software.

Bearing monitoring is then crucial in order to study the trend of defect bearing vibration signals. Classical bearing monitoring methods can usually be classified as either time domain methods (Braun, 1980; Dyer and Stewart, 1978; Logan and Mathew, 1996) or frequency domain methods (McFadden, 1984; Shiroishi, 1997; Taylor, 1980; Wang and Kootsookos, 1998). These methods look for periodically occurring high-frequency transients, which however is complicated by the fact that this periodicity may be suppressed. Moreover, classical Fourier methods tend to average out transient vibrations (such as those typical for defect bearings), thus making them more prone to drown in the background noise of harmless vibrations.

In order to help those incompetency users, clustering can be considered the most important unsupervised learning problem; so, as every other problem of this kind, it deals with finding a structure in a collection of unlabeled data. A loose definition of clustering could be the process of organizing objects into groups whose members are similar in some way. A cluster is therefore a collection of objects which are similar between them and are dissimilar to the objects belonging to other clusters.

The goal of clustering is to determine the intrinsic grouping in a set of unlabeled data. A good clustering can be shown that there is no absolute best criterion which would be independent of the final aim of the clustering. Consequently, it is the user which must supply this criterion, in such a way that the result of the clustering will suit their needs (Zaiane, 1999). Hence, with a reference done by a profession, the signal of faulty bearing can be grouped in a certain cluster and identified easily by incompetence users.

Referring to the problems mentioned above, it is found that the defect in bearings is hard to be detected by using direct observations. Hence, it is crucial for this research to identify and detecting major bearing faults as they are in the early stage. This can help to prolong bearing's service life span and also able to avoid any unwanted accidents that can make great losses. The need of clustering is that it is able to assist untrained or unprofessional users in order to identify the types of potential bearing defects based on the clustering data.

1.3 RESEARCH OBJECTIVE

The objectives of a research project summarize what is to be achieved by the study. Corresponds to the project background and problem statements, the objectives of this research are:

(i) To study the trend of frequency spectrum from different bearing faults.
(ii) To apply clustering approach using Principal Component Analysis, PCA on frequency domain signals.

1.4 SCOPE OF THE RESEARCH

The scope of research is the areas covered in the research. This part of the research paper will tell exactly what was done. The type of information that would be included in the scope of a research project would include facts and theories about this research. To achieve the above objectives mentioned in part 1.3, the scopes need to be considered as follow:

(i) Study defects on cylindrical roller bearings. Rolling element bearings are one of the most common components in rotating machinery including electric motors.
(ii) Types of bearing faults to be study are corrosion, contamination, inner race and roller defects.
(iii) Experimental RPM range used were 440 rpm (15%), 1480 rpm (50%) and 2672 rpm (90%).
(iv) Method of signal analysis will be used is Principle Component Analysis, PCA.

1.5 EXPECTED OUTCOME

In the end of this research, the expected outcomes are, firstly, able to perform signal clustering to ease bearing fault detection. Secondly, to contribute in develop an advanced signal processing technique for bearing defect monitoring system. Lastly is making bearing defect detection becomes easier for untrained or incompetence users.

1.6 SIGNIFICANCE OF THE RESEARCH

The significance of this research is the development of an advanced signal processing technique to recover the source signal corrupted during the transmission process, where detection of a small failure is made difficult due to the heavy background noise. Current techniques are inefficient to detect defective bearing signals in the presence of extraneous noise. By applying clustering method, it is believed that this able to help by to identify defects in bearing. The PCA method allows providing directly the redundancy relations between the variables without identifying the state representation matrix of the system. This task is often difficult to achieve (Ramahaleomiarantsoa et al., 2012).

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter covers the findings of information about the bearing types, bearing faults, the bearing fault detection techniques and follow-up by basic vibration theory. Chapter two also covers about vibration based condition monitoring, Principle Component Analysis (PCA) and its relative steps and lastly is the signal clustering for this research.

2.2 BEARINGS

Bearings are made to support radial loads, thrust loads, or combined radial-thrust loads. In order to serve all these functions, bearings make use of a relatively simple structure: a roller with internal and external smooth metal surfaces, to aid in rolling. The roller itself carries the weight of the load—the force of the load's weight is what drives the bearing's rotation. However, not all loads put force on a bearing in the same manner. There are two different kinds of loading: radial and thrust.

A radial load, as in a pulley, simply puts weight on the bearing in a manner that causes the bearing to roll or rotate as a result of tension. A thrust load is significantly different, and puts stress on the bearing in an entirely different way. If a bearing is flipped on its side and subject to complete force at that angle, this is called thrust load. A bearing that is used to support a bar stool is an example of a bearing that is subject only to thrust load.

2.3 TYPES OF BEARINGS

There are many types of bearings, each used for different purposes. These include ball bearings, roller bearings, ball thrust bearings, roller thrust bearings and tapered roller thrust bearings.

Ball bearings, as shown below, are one of the most common types of bearing. They are found in everything from inline skates to hard drives. These bearings can handle both radial and thrust loads, and are usually found in applications where the load is relatively small. In a ball bearing, the load is transmitted from the outer race to the ball and from the ball to the inner race. Since the ball is a sphere, it only contacts the inner and outer race at a very small point, which helps it spin very smoothly. But it also means that there is not very much contact area holding that load, so if the bearing is overloaded, the balls can deform or squish, ruining the bearing. Ball bearings are used in ways that are quite common, such as skateboards and roller blades, and some uses are not so common, such as military aircraft and oil drilling. Though ball bearings are useful, they are easily overloaded. By paying careful attention to the specifications of the type of ball bearing being used, this problem is minimized. Figure 2.1 shows the cutaway view of a ball bearing.

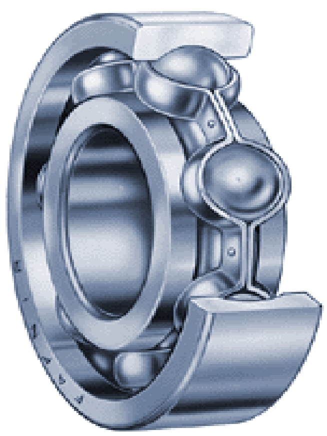


Figure 2.1: Cutaway view of a ball bearing

Source: The Timken Company, 2011

Roller bearings like the one illustrated below are used in applications like conveyer belt rollers, where they must hold heavy radial loads. In these bearings, the roller is a cylinder, so the contact between the inner and outer race is not a point but a line. This spreads the load out over a larger area, allowing the bearing to handle much greater loads than a ball bearing. However, this type of bearing is not designed to handle much thrust loading. A variation of this type of bearing, called a needle bearing, uses cylinders with a very small diameter. This allows the bearing to fit into tight places. Roller bearings are used in power generation, wind turbines, gear drives, rolling mills, machine tool spindles, gear reduction units etc. The cutaway view of a roller bearing is illustrated in Figure 2.2.

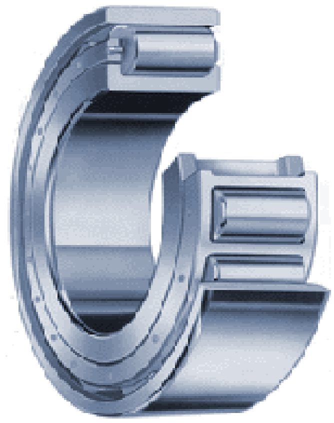


Figure 2.2: Cutaway view of a roller bearing

Source: The Timken Company, 2011

Figure 2.3 shows a ball thrust bearing. Ball thrust bearings like the one shown below are mostly used for low-speed applications and cannot handle much radial load. Barstools and Lazy Susan turntables use this type of bearing. They are mostly used for low speed appliances and can't handle heavy loads, for example they can be found in swivel bar stools and turn tables.

Figure 2.3: Cutaway view of a ball trust bearing

Source: The Timken Company, 2011

Roller thrust bearings like the one illustrated in Figure2.4 below can support large thrust loads. They are often found in gear sets like car transmissions between gears, and between the housing and the rotating shafts. The helical gears used in most transmissions have angled teeth -- this causes a thrust load that must be supported by a bearing. The applications of roller thrust bearings were blowout preventers, classifiers, extruders, gearboxes, metal mill work/back-up rolls, pre-heater fans, pumps and screw conveyors (The Timken Company, 2011).

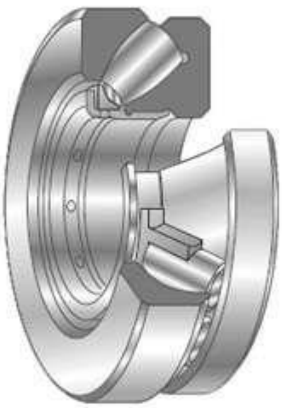


Figure 2.4: Cutaway view of a roller trust bearing

Source: The Timken Company, 2011

Tapered roller bearings can support large radial and large thrust loads. Tapered roller bearings are used in car hubs, where they are usually mounted in pairs facing opposite directions so that they can handle thrust in both directions. They are used in many car hubs. The disadvantages to taper roller bearings are that they are usually quite expensive and they add more friction than a ball bearing. Typical tapered roller bearing is illustrated in Figure 2.5 below.

Figure 2.5: Cutaway view of a tapered roller bearing

Source: The Timken Company, 2011

2.4 CYLINDRICAL ROLLER BEARING

Cylindrical roller bearings have exceptionally low friction torque characteristics that make them suitable for high speed operation. They also have high radial load carrying capacity. They are typically used in machine tools, transmissions, vibration machines and as wheel set bearings for rail vehicles. The surface finish of the tracks and rolling elements is critical to the running performance and noise characteristics of these bearings. Due to its ability to handle heavy radial loads and its important usage as discussed in Chapter 2.3, hence this type of bearing is then selected for this research.

There are numerous different kinds of bearings that are designed to handle radial load, thrust load, or some combination of the two. Because different applications require bearings that are designed to handle a specific kind of load and different amounts of weight, the differences between types of bearings concern load type and ability to handle weight.

Being the typical bearing found in electric motors, the rolling element (or anti friction) bearing is made up of outer race, inner race, cage and rolling elements which is cylindrical roller as illustrated in Figure 2.6.

The inner race fits inside a larger outer race. Metal balls or rollers rest between the two races, enabling the inner race to rotate within the stationary outer race. The outer race of the bearing will be stationary in a crankcase; the crankshaft ends will be inserted into the inner races of these bearings, and the crankshaft will spin with the inner race against the balls or rollers between the races. The cage's roles are to guide and drive the rolling body in the right rolling raceway; to separate the rolling body with equal distance, making it in the circle of raceway, so as to prevent mutual collision and friction when it works; to integrate the rolling body and ferrule together in the separation-bearing, so as to prevent rolling off (Schaeffler trading (China) Co. Ltd, 2012).

Figure 2.6 also shows the Pitch Diameter (Pd) which is the span between the centers of two opposite rolling elements (Felten, 2003). Together with this figure it shows and labels the basic components of the cylindrical roller bearing.

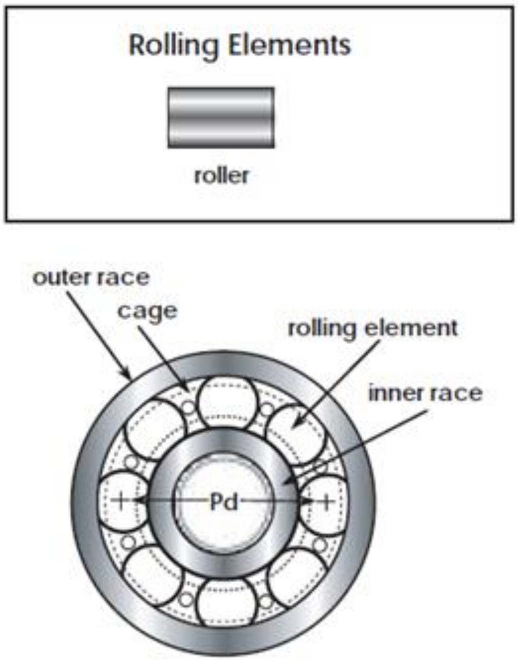


Figure 2.6: Rolling element bearing components

Source: Felten, 2003

Roller bearings are designed to carry heavy loads as the primary roller is a cylinder, which means the load is distributed over a larger area, enabling the bearing to handle larger amounts of weight. This structure, however, means the bearing can handle primarily radial loads. Rolling element bearings are found widespread domestic and industrial applications. Proper functioning of the bearings depends greatly on the smooth and quiet running of the appliances. In industrial applications, the bearings are one of the critical mechanical components and a solely minor defect, unless it is detected in time, might causes malfunction and may lead to catastrophic machinery failure.

Roller bearings can support heavier loads than ball bearings of the same size because of having line contact between the rollers and rings as opposed to point contact for ball bearings. Line contact has more area supporting the load and thus less stress than ball bearing point contact. This feature also makes roller bearings a stiffer support to a shaft.

There are many applications for cylindrical roller bearings. Examples include mining, petroleum production, power generation, power transmission, cement processing, aggregate crushing, and metal recycling. Some cylindrical roller bearings are used in briquetting machines, rubber mixing equipment, rolling mills, rotary dryers, or pulp and paper machinery. Others are used in construction equipment, crushers, electric motors, blowers and fans, gears and drives, plastics machinery, machine tools and traction motors and pumps. Ideal applications for cylindrical roller bearings are very large industrial machines such as presses, forges, conveyors, gear boxes and machine tool spindles. On vehicles, they are ideal axle bearings for dump trucks, cement mixers, bulldozers, load haulers, and lifts.

2.5 BEARING FAULTS

Generally, a rolling bearing cannot rotate forever. Unless operating conditions are ideal and the fatigue load limit is not reached, sooner or later material fatigue will occur. The period until the first sign of fatigue appears is a function of the number of revolutions performed by the bearing and the magnitude of the load. The major bearing faults shall discuss here are namely corrosion, contamination, races defect and rolling elements defect.

2.5.1 Corrosion

Rust is a film of oxide, hydroxide, or carbonate produced on a metallic surface by chemical action. Corrosion is the phenomena of oxidation or dissolution occurring on the surface and is produced by chemical action (electric chemical action including combination or cell restructuring) with acid or alkali (JTEKT Corporation, 2009).

Corrosion or etching is one of the most crucial problems encountered in anti-friction bearings. The high degree of surface finish on races and rolling elements makes them susceptible to corrosion damage from moisture and water if not adequately protected. Rust can be form if water or corrosive agents reaching to the inside of the bearing in such quantities that the lubricant unable to provide protection for the steel surfaces, as illustrated in Figure 2.7. This process will soon lead to deep seated rust. The general appearances are greyish black streaks across the raceways, mostly coinciding with the rolling element spacing. As for at a later stage, pitting of raceways and other surfaces of the bearing may occurred (The Timken Company, 2011).

Figure 2.7: Heavy corrosion on the race

Source: The Timken Company, 2011

Deep seated rust is referring to a thin protective oxide film forms on clean steel surfaces when exposed to the air. However, this film is not impenetrable and if water or corrosive elements make contact with the steel surfaces, patches of etching will form. This development soon leads to deep seated rust. Deep seated rust is a great danger to bearings since it can initiate the phenomena of flaking and cracks. Acid liquids corrode the steel quickly, while alkaline solutions are less dangerous. The salts that are present in fresh water constitute, together with the water, an electrolyte which causes galvanic corrosion, known as water etching. Salt water, such as sea water, is therefore highly dangerous to bearings (SKF Bearing Corporation, 1994).

Based on Ciprian Radu, 2010, his study shows that the main failure cause is the inappropriate lubrication of the bearing rolling elements (approximately 80% of the cases), which this lead to the formation of rust in such of corrosion or etching.

Lubricants are used between contact surfaces to keep the parts in continuous motion. The main purpose of rolling bearing lubrication is to avoid or reduce the metal-to-metal friction between the rolling and sliding contact surfaces. This is not the only function of rolling bearing lubrication. The supplementary functions are: heat dissipation from the bearing, removal of solid wear particles and contaminants from the rolling contact surfaces, corrosion protection, and increase of the sealing effect of the bearing seals. Lubrication is crucial for bearing life. In heavy duty applications, such as rolling mill machines, furnaces, ovens or high temperature fans, rolling bearings may be exposed to higher-than-normal temperatures. For these applications, appropriate selection of the lubricant and lubrication method is very important. In industrial applications there are two types of lubricants suitable for high temperature use: grease and oil. In special cases, bearings are lubricated with solid dry lubricants (FAG Bearing Corporation, 2002).

2.5.2 Contamination

Contaminate carry the meaning of to make a substance or place dirty or no longer pure by adding a substance that is dangerous or carries disease. It is also definite as to make a place or substance dirty or harmful by putting something such as chemicals or poison in it.

Contamination is one of the leading causes of bearing failure. Despite their study mechanical appearance, rolling element bearings are actually precision components with internal tolerances on the order of millionths of an inch – more precise than an expensive wristwatch. Contamination of the bearing shows up as scratches, pitting and scoring along the raceways, with corresponding marks on the ball and roller surfaces. Unlike brinelling, these small indentations are scattered, rather than centralized, on the bearing surface. Symptoms of contamination are particle denting of the rolling elements and raceways, resulting in high vibration and abrasive wear. The effect of contamination on bearing life depends on bearing type and size, relative lubricant film thickness and the size, hardness and distribution of solid contaminant particles. Although foreign matter can enter the bearing during mounting, the most direct and sustained area of entry is the housing seals and lubricant. Bearing manufacturers realize the damaging effect of dirt and take extreme precautions to assemble and deliver clean bearings. Some even assemble their bearings in air-conditioned clean rooms.

When seals or shields are defective it is not hard for abrasive particles, dirt, or dust to get into the bearing. There can even be dirt or dust in the lubricant. Wherever foreign material enters into the bearing by way of contaminated lubricant, the particles are pressed into the metal surface. Small dents and pits are formed which roughen the load-carrying surface. Severe roughness will result in flaking and premature bearing failure. Contamination can also be caused by improper cleaning of the housing or shaft, or by using dirty tools and hands during mounting and assembly (SKF Bearing Corporation, 1994). Typical contaminated defect on bearing is illustrated in Figure 2.8.

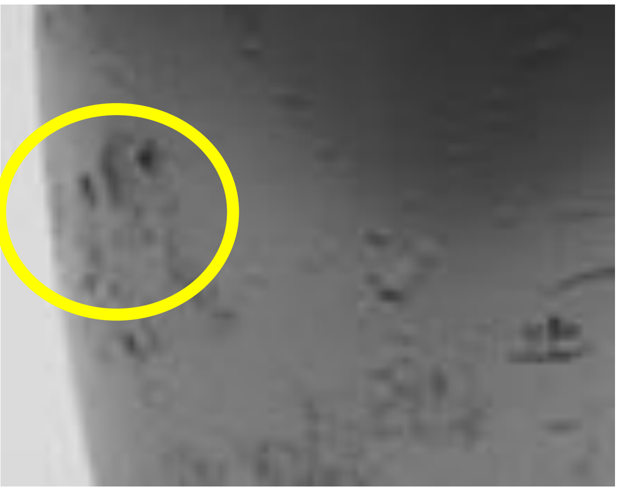


Figure 2.8: Hard particles caused contamination bruising on this roller bearing

Source: The Timken Company, 2011

2.5.3 Races Defect

Other common defects are the races defects, including inner race, as illustrated in Figure 2.9 and outer race, Figure 2.10, of the bearing. Generally, the races defects are mostly caused by wear. The appearances of such defect are small indentations and depression around the raceways. Wear may, however, occur as a result of the ingress of foreign particles into the bearing or when the lubrication is unsatisfactory. Vibration in bearings which are not running also gives rise to wear. Raceways may become dented if the mounting pressure is applied to the wrong ring, or if the bearing is subjected to abnormal loading while not running. Foreign particles in the bearing also cause indentations.

Apart from these, smearing is also another defect on races defect. When two inadequately lubricated surfaces slide against each other under load, material is transferred from one surface to the other. This is known as smearing and the surfaces concerned become scored, with a ‘torn’ appearance. Smearing happened in raceways is caused by roller rotation being retarded in the unloaded zone, where the rollers are not driven by the rings. Consequently their speed of rotation is lower than when they are in the loaded zone. The rollers are therefore subjected to rapid acceleration and the resultant sliding is so severe that it may produce smearing.

It can be said that most of the races defects are mainly caused by improper load and unsatisfactory lubrication. Other causes including housing misalignment or shaft bending the inner ring tilts as opposed to the outer ring and high moment loads result. In roller bearings this leads to a constraining force in the cage pockets and to more sliding in the raceways as well as the balls running on the shoulder edge. Hence, the bearing setup and its lubrication are extremely important in order to prolong the bearing service lifespan (SKF Bearing Corporation, 1994).

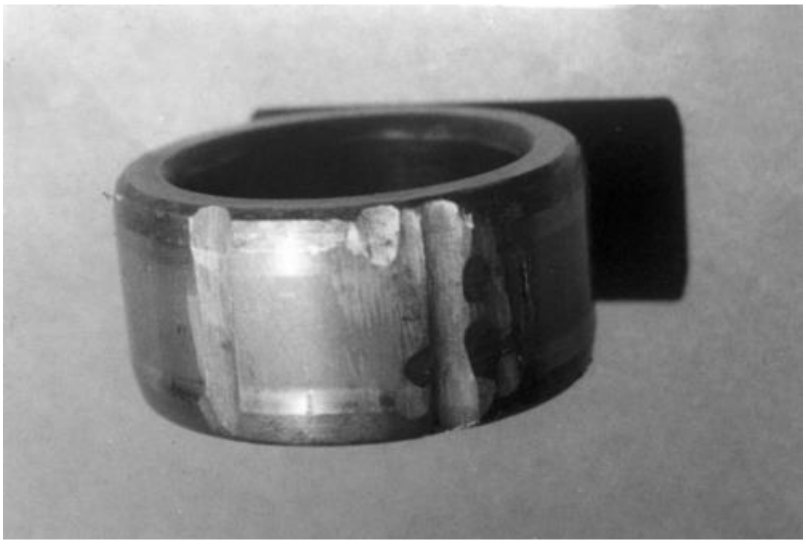


Figure 2.9: Defects on inner race of roller bearing

Source: Amarnath, 2004

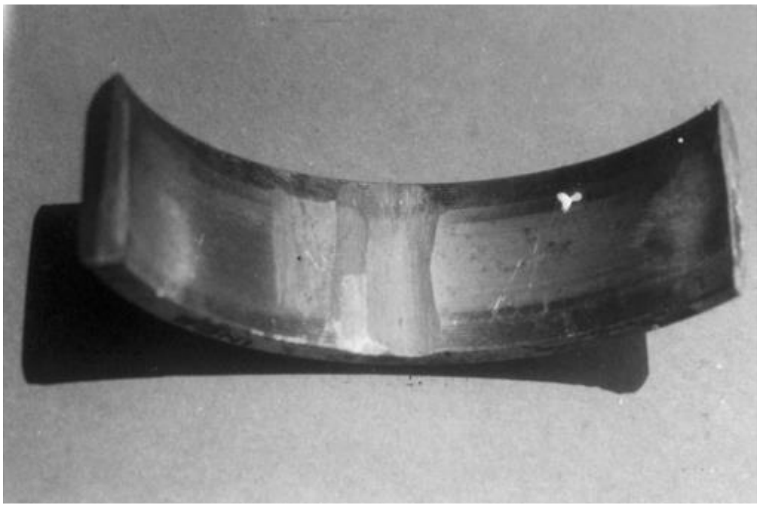


Figure 2.10: Defects on outer race of roller bearing

Source: Amarnath, 2004

2.5.4 Rolling Elements Defect

During cycling, the material of the raceways and rolling elements is subject to a continuous pulsating stress. This leads to failure patterns like those resulting from the fatigue of mating parts under bending stress: fatigue fractures develop. In rolling bearings these fractured areas run largely parallel to the surface and lead to material flaking and are referred to as fatigue damage, flaking, pitting, spalling, grey stippiness, micro pitting, steel pitting etc. As fatigue ball bearings react more sensitively to contamination than roller bearings, and bearings with small rolling elements more sensitively than those with large ones. The rolled-up material plays a very important role where the indentation of foreign particles is concerned. Like foreign particle indentations, rolling element indentations develop due to the bearing's high static overload and their rolled-up edges lead to failure. Figure 2.11 shows the defect found on the roller of a bearing.

The common symptoms are subsurface cracks of raceway and rolling elements, material flaking (relatively deep pitting), undamaged areas of the raceway indicate good lubrication in the early stage of damage, while more or less a lot of indentations by cycled fractured parts.

The major causes for such defects are static overload or shock impacts, mounting or dismounting forces, change in geometry of components in rolling contact due to wear in the case of contaminated lubricant applied via rolling elements (incorrect mounting order, unsuitable accessories) (Schaeffler Group Industrial, 2012).

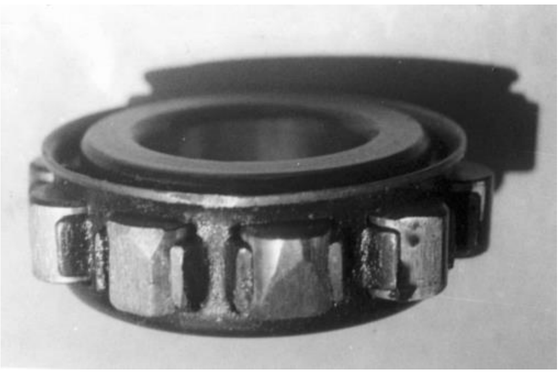


Figure 2.11: Defects on roller of roller bearing

Source: Amarnath, 2004

Hence, the defects mentioned in above and previous sections shall be study in this research. The variety of types of bearing defects shall include in this research is to vary the research outcome in order to cover much more bearing related defects.

2.6 FAULT DETECTION TECHNIQUES

Typically, the previous defects mostly arise during the operation of a bearing. Therefore, the detection of these defects at an early stage without machine disassembly is pivotal for condition monitoring, quality inspection, and predictive maintenance. Various methods are used for the diagnosis of bearing defects. The methods are broadly classified as acoustic measurements, current and temperature monitoring, wear debris detection, and vibration analysis.

The most effective acoustic-based bearing health monitoring is acoustic emission. It is a transient impulse generated by the rapid release of strain energy in solid material under mechanical or thermal stress. The detection of cracks is the prime application of acoustic emission; therefore, this technique can be used as a tool for condition monitoring of bearing faults and shaft cracks. The measurement of a machine's sound can also be employed for detecting defects in bearings. Typically, the accuracy of these methods depends on sound pressure and sound intensity data.

Bearing distributed defects generate excessive heat in the rotating components. Monitoring the temperature of a bearing housing or lubricant is the simplest method for fault detection in rotary machines.

The operating conditions of a machine can be monitored by analyzing the spectrum of the motor current. The changes in the electric background noise are associated with the changes in the mechanical components of the machine; therefore, fault signatures can be detected by motor current signal processing techniques.

In this method, the presence of metallic particles in the lubricant is detected by sensitive sensors. Furthermore, the spectrographic analysis of the different metallic elements in the lubricant can facilitate the location of the fault.

Since the abnormal vibration of rotary machines is the first sensory effect of rotary component failure, vibration analysis is widely employed in the industry. The fault vibration signal generated by the interaction between a damaged area and a rolling surface occurs regardless of the defect type. Consequently, a vibration analysis can be employed for the diagnosis of all types of faults, either localized or distributed. Furthermore, low-cost sensors, accurate results, simple setups, specific information on the damage location, and comparable rates of damage are other benefits of the vibration measurement method (Ghafari, 2007).

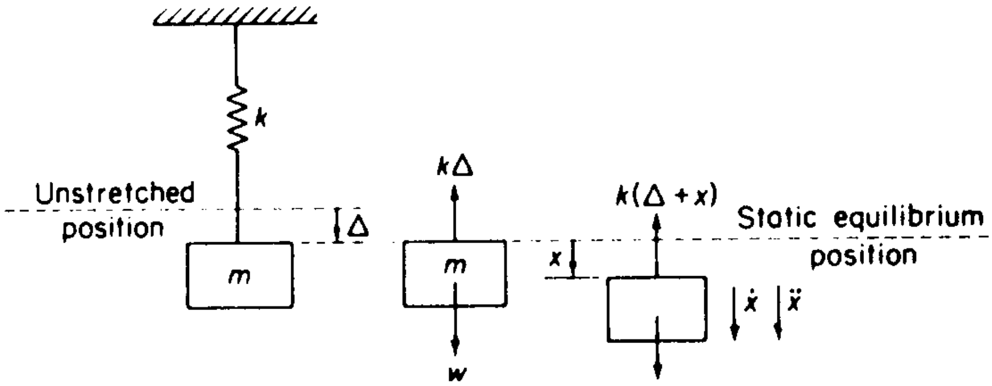
2.7 VIBRATION

Noise and vibration are constantly present in our high-tech society. Noise causes serious problems both at home and in the workplace, and the task of reducing community noise is a subject currently focused on by authorities in many countries. Similarly, manufacturers of mechanical products with vibrations causing acoustic noise, more and more find themselves forced to compete on the noise levels of their products. Such competition has so far occurred predominantly in the automotive industry, where the issues with sound and noise have long attracted attention.

2.7.1 Vibration Model

The basic vibration model of a simple oscillatory system consists of a mass, a massless spring, and a damper. The spring supporting the mass is assumed to be of negligible mass. Its force-deflection relationship is considered to be linear, following Hooke's law, $F=kx$ where displacement x is measured in meter and the stiffness k is measured in newton/meter.

The viscous damping, generally represented by a dashpot, is described by a force proportional to the velocity, or $F=cx$. The damping coefficient c is measured in newton/meter/second.



where x=displacement, $\dot{x}$=velocity, $\ddot{x}$=acceleration, k=spring constant, m=mass, Δ=distance travel and w=weight

Figure 2.12: Spring-mass system and free-body diagram

Source: Bhave, 2010

2.8 VIBRATION BASED CONDITION MONITORING

Under radial and misaligning loads bearing vibration is an inherent feature of rolling bearings even if the bearing is geometrically perfect and is not therefore indicative of poor quality. This type of vibration is often referred to as variable compliance and occurs because the external load is supported by a discrete number of rolling elements whose position with respect to the line of action of the load continually changes with time.

As the bearing rotates, individual ball loads, hence elastic deflections at the rolling element raceway contacts, change to produce relative movement between the inner and outer rings. The movement takes the form of a locus which under radial load is two dimensional and contained in a radial plane, whilst under misalignment it is three-dimensional. The movement is also periodic with base frequency equal to the rate at which the rolling elements pass through the load zone. Frequency analysis of the movement yields the base frequency and a series of harmonics. For a single row radial ball bearing with an inner ring speed of 1800 rev/min a typical ball pass rate is 100 Hz and significant harmonics to more than 500 Hz can be generated.

Variable compliance vibration is heavily dependent on the number of rolling elements supporting the externally applied load; the greater the number of loaded rolling elements, the less the vibration. For radially loaded or misaligned bearings running clearance determines the extent of the load region, and hence, in general, variable compliance increases with clearance. Running clearance should not be confused with radial internal clearance (RIC), the former normally being lower than the RIC due to interference fit of the rings and differential thermal expansion of the inner and outer rings during operation. Variable compliance vibration levels can be higher than those produced by roughness and waviness of the rolling surfaces. However, in applications where vibration is critical it can be reduced to a negligible level by using ball bearings with the correct level of axial pre-load (Lacey, 2008).

The majority of the research on the diagnosis and prognosis of bearings is based on signal processing techniques, independent of bearing vibration characteristics. In these works, first a localized or distributed defect is created on a bearing by means of grinding, acid etching, drilling, overloading, or over speeding to intentionally introduce defects in the bearing components. After a vibration signal is measured usually, by accelerometers, different signal processing techniques are employed to extract the fault sensitive features to serve as the monitoring indices. This procedure is quite similar among the published literature. The reported signal processing methods are categorized as time domain, frequency domain, and time-frequency domain (Ghafari, 2007).

2.8.1 Time Domain Analysis

Time-domain analysis is directly based on the time waveform itself. Traditional time-domain analysis calculates characteristic features from time waveform signals as descriptive statistics such as mean, peak, peak to peak interval, standard deviation, crest factor, high-order statistics: root mean square, skewness, kurtosis, etc. These features are usually called time-domain features.

Time domain analysis has been widely employed. Successful results of Root Mean Square (RMS) (Tandon, 1994; Downham, 1980), Kurtosis (Dyer and Stewart, 1978; Stronach et al., 1984; Williams et al., 2001; White, 1984), skewness, peak value (Mathew and Alfredson, 1984), Crest Factor (CF) (Ingarashi et al., 1980), and synchronous averaging (Hemmings et al., 1976; McFadden and Toozhy, 2000) have been reported in the low frequency range of <5 kHz. Band pass filtering has also been conducted in the time domain. It is based on the fact that the strike between the damage and the rotating component can excite high frequency resonances (10-100 kHz). The generated energy from this impact is not sufficient to excite the entire rotor's assembly, but is enough to excite vibration sensor resonance. Monitoring the vibration amplitude at the resonant band pass filtered frequency is the principle of the shock pulse method (Mathew and Alfredson, 1984; Tandon and Nakara, 1992; Butler, 1973; Smith, 1982). It is implemented in shock pulse meters which are the most accepted diagnostic instrument in the industry. Time domain analysis has the advantage of simple calculations, straightforward signal pre-processing, and speed independency. However, insensitivity to early stage faults and deeply distributed defects are drawbacks of this approach.

2.8.2 Frequency Domain Analysis

Frequency-domain analysis is based on the transformed signal in frequency domain. The advantage of frequency-domain analysis over time-domain analysis is its ability to easily identify and isolate certain frequency components of interest. The most widely used conventional analysis is the spectrum analysis by means of fast Fourier transform (FFT). The main idea of spectrum analysis is to either look at the whole spectrum or look closely at certain frequency components of interest and thus extract features from the signal.

Perhaps, frequency domain, also called spectral analysis, is the most reported signal processing method for bearing diagnosis. Each bearing component has a characteristic frequency, which is calculated from the kinematics of the rotating parts. Monitoring these frequencies or their harmonics at a low frequency range (<5 kHz) has been successful in bearing diagnosis (Dyer and Stewart, 1978; Igarashi and Hamada, 1989; Taylor, 1980). To decrease the effect of the noise level and frequency side bands, some researchers have adopted the amplitude demodulated or enveloped signal. The spectral analysis of a low and/or high frequency range enveloped signal is repeatedly reported as an efficient method for bearing diagnosis (Kadushin, 1991; McMohan and Scott, 1991; Milne et al., 1991; Martin and Thrope, 1992). A number of frequency domain features, based on simple or complex signal processing methods such as power cepstrum (Tandon, 1994), adaptive noise cancellation (Ho and Randall, 2000; Jianfang et al., 1989), and denoising (Bolaers et al., 2004), are also proposed for bearing diagnosis. The frequency domain approach is sensitive and robust to detect bearing defects and to identify the localized damage location. However, the accuracy of this method highly depends on the bearing dimensions and rotational speed. In addition, all the frequency domain methods require an intelligent selection of the frequency band in order to be effective.

The time domain vibration signal is typically processed into the frequency domain by the application of Fourier transform, usually in the form of fast Fourier transform (FFT) algorithm. The principal advantage of the method is that the repetitive nature of the vibration signals is clearly displaced as peaks in the frequency spectrum at the frequency where the repetition takes place. The FFT provides the fast and convenient means of computing the discrete Fourier Transform of vibration data and a window function can be used to force the vibration data to appear periodic, which reduces leakage from one frequency component to another. A smooth estimate of power spectrum computed from stationary random data can be obtained by averaging the spectrum obtained by equation over a number of windowed data record and is often referred to as the Welch method of power spectrum estimation. One of the main purposes in computing power spectra of vibrating data is to identify major frequency component which can be found in spectrum and then use these components and their amplitude for the trending purposes. The results taken from FFT and envelope analysis and conclude that FFT is more suitable for finding the fault whereas envelope analysis give excellent results in predicting the damage (Spyridon et al., 2009).

2.9 PRINCIPAL COMPONENT ANALYSIS (PCA)

Principle Component Analysis, PCA has been used for a number of applications including signal analysis and especially dynamic (acceleration) signal analysis (Moniz et al., 2005; Nichols et al., 2006; Choi and Lee, 2004; Trendafilova et al., 2007). In its dynamics applications it is indeed most frequently used for vibration-based damage and fault detection (Baydar et al., 2001; Moniz et al., 2005; Nichols et al., 2006; Trendafilova et al., 2007). A number of articles suggest its application for structural damage detection (Moniz et al., 2005; Nichols et al., 2006; Trendafilova et al., 2007). The case of structural damage and its detection from vibration measurements is somewhat similar to the detection of roller bearing faults. The presence of a fault can be seen to change some frequency ranges and the lines in these frequency ranges are further used as variables, which are subjected to PCA (Nichols et al., 2006; Trendafilova et al., 2007). In the present study a similar approach is taken by first applying the wavelet transform, which actually selects the frequency range containing the variability due to the presence of a fault, and then subjecting the already selected frequency lines to PCA to choose and weight properly only those that are responsible for the biggest part of the between-class variance.

PCA method is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. Since patterns in data can be hard to find in data of high dimension, where the luxury of graphical representation is not available, PCA is a powerful tool for analyzing data.

The other main advantage of PCA is that once have found these patterns in the data, and compress the data, i.e. by reducing the number of dimensions, without much loss of information.

2.9.1 STEPS IN PCA

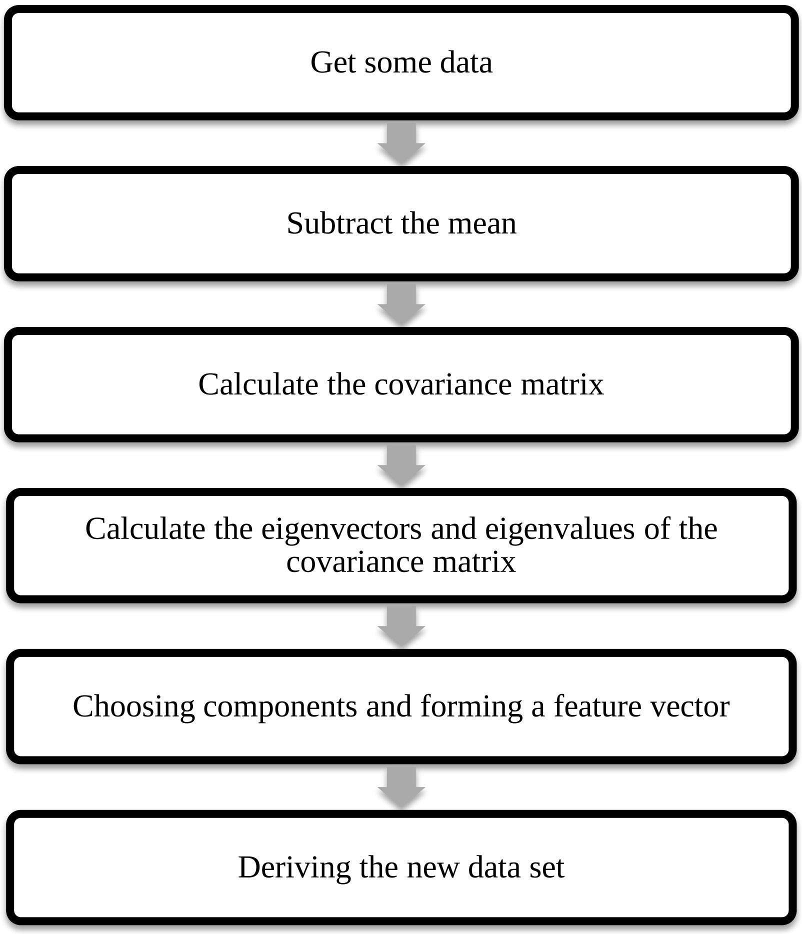


Figure 2.13: General steps in PCA analysis

According to Figure 2.13, vibration data, which is the frequency domain data, can be getting from data acquisition devices. For PCA to work properly, each of the data dimensions have to subtract the mean. The mean subtracted is the average across each dimension. So, all the data values, x have $\bar{x}$ (the mean of the x values of all the data points) subtracted, and all the y values have $\bar{y}$ subtracted from them (for a two dimensional as example). This produces a data set whose mean is zero.

A useful way to get all the possible covariance values between all the different dimensions is to calculate them all and put them in a matrix. So, the definition for the covariance matrix for a set of data with n dimensions is:

$$C^{n^2} = \left(c_{i,j}, c_{i,j} = cov\left(Dim_i, Dim_j\right)\right), \tag{2.1}$$

where C^{n^2} is a matrix with n rows and n columns, and Dim_x is the xth dimension. All that this formula says is that if have an n-dimensional data set, then the matrix has n rows and columns (so is square) and each entry in the matrix is the result of calculating the covariance between two separate dimensions. For example, the entry on row 2, column 3, is the covariance value calculated between the 2nd dimension and the 3rd dimension.

An example, the covariance matrix for an imaginary 3 dimensional data set, using the usual dimensions x, y and z. Then, the covariance matrix has 3 rows and 3 columns, and the values are:

$$C = \begin{pmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{pmatrix} \tag{2.2}$$

Some points to note: Down the main diagonal, it is noticed that the covariance value is between one of the dimensions and itself. These are the variances for that dimension. The other point is that since $cov(a,b) = cov(b,a)$, the matrix is symmetrical about the main diagonal.

Since the covariance matrix is square, the eigenvectors and eigenvalues for this matrix can be calculated. These are rather important, as they tell useful information about the data. For example as below:

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 5 \end{pmatrix} \tag{2.3}$$

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} \tag{2.4}$$

Theoretically, multiplication of two matrices together is possible, provided they are compatible sizes. Eigenvectors are a special case of this. Consider the two multiplications between a matrix and a vector above. In the first example (2.3), the resulting vector is not an integer multiple of the original vector, whereas in the second example (2.4), the example is exactly four times the vector that began with. The vector is a vector in two dimensional spaces. The vector $\binom{3}{2}$ (from the second example multiplication) represents an arrow pointing from the origin, $(0, 0)$ to the point$(3, 2)$. The other matrix, the square one, can be thought of as a transformation matrix. If multiply this matrix on the left of a vector, the answer is another vector that is transformed from its original position.

It is the nature of the transformation that the eigenvectors arise from. Imagine a transformation matrix that, when multiplied on the left, reflected vectors in the line $y = x$. Then if there were a vector that lay on the line $y = x$, it's reflection it itself. This vector (and all multiples of it, because it wouldn't matter how long the vector was), would be an eigenvector of that transformation matrix.

First know that the property of these eigenvectors is that eigenvectors can only be found for square matrices. And, not every square matrix has eigenvectors. And, given an n x n matrix that does have eigenvectors, there are n of them. Given a 3 x 3 matrix, there are 3 eigenvectors. Another property of eigenvectors is that even if scales the vector by some amount before multiply it, will still get the same multiple of it as a result, as below:

$$2 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix} \tag{2.5}$$

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 6 \\ 4 \end{pmatrix} = \begin{pmatrix} 24 \\ 16 \end{pmatrix} = 4 \times \begin{pmatrix} 6 \\ 4 \end{pmatrix} \tag{2.6}$$

This is because if scale a vector by some amount, all are doing is making it longer, not changing its direction. Lastly, all the eigenvectors of a matrix are perpendicular, i.e. at right angles to each other, no matter how many dimensions they have. By the way, another word for perpendicular, in math's talk, is orthogonal. This is important because it means that the data can express in terms of these perpendicular eigenvectors, instead of expressing them in terms of the x and y axes.

Another important thing to know is that when mathematicians find eigenvectors, they like to find the eigenvectors whose length is exactly one. This is because, the length of a vector doesn't affect whether it's an eigenvector or not, whereas the direction does. So, in order to keep eigenvectors standard, whenever getting an eigenvector, usually scale it to make it have a length of one, so that all eigenvectors have the same length. Here's a demonstration from previous example above. $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ is an eigenvector, and the length of that vector is

$$\sqrt{(3^2 + 2^2)} = \sqrt{13} \tag{2.7}$$

then divide the original vector by this much to make it have a length of 1:

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} \div \sqrt{13} = \begin{pmatrix} 3/\sqrt{13} \\ 2/\sqrt{13} \end{pmatrix} \tag{2.8}$$

Eigenvalues are closely related to eigenvectors, in fact, in examples mentioned previously. In both those examples, the amount by which the original vector was scaled after multiplication by the square matrix was the same. In that example, the value was 4. Hence, value 4 is the eigenvalue associated with that eigenvector. No matter what multiple of the eigenvector took before multiplied it by the square matrix, would always get 4 times the scaled vector as our result. So the eigenvectors and eigenvalues always come in pairs.

Here is where the notion of data compression and reduced dimensionality comes into it. If look at the eigenvectors and eigenvalues from the previous section, it is noticed that the eigenvalues are quite different values. In fact, it turns out that the eigenvector with the highest eigenvalue is the principle component of the data set. In our example, the eigenvector with the larges eigenvalue was the one that pointed down the middle of the data. It is the most significant relationship between the data dimensions.

In general, once eigenvectors are found from the covariance matrix, the next step is to order them by eigenvalue, highest to lowest. This gives the components in order of significance. The components of lesser significance can be ignoring. By doing so, some information will lose, but if the eigenvalues are small, information doesn't lose much. If some components being leave out, the final data set will have fewer dimensions than the original. To be precise, if originally have n dimensions in the data, and so calculate n eigenvectors and eigenvalues, and then choose only the first p eigenvectors, then the final data set has only p dimensions. What needs to be done now is to form a feature vector, which is just a fancy name for a matrix of vectors. This is constructed by taking the eigenvectors that want to keep from the list of eigenvectors, and forming a matrix with these eigenvectors in the columns.

$$Feature\ Vector = (eig_1\ eig_2\ eig_3\ \ldots\ eig_n) \tag{2.9}$$

This is the final step in PCA, and is also the easiest. Once chosen the components (eigenvectors) that wish to keep in the data and formed a feature vector, simply take the transpose of the vector and multiply it on the left of the original data set, transposed.

$$Final\ Data = Row\ Feature\ Vector \times Row\ Data\ Adjust \tag{2.10}$$

where $Row\ Feature\ Vector$ is the matrix with the eigenvectors in the columns transposed so that the eigenvectors are now in the rows, with the most significant eigenvector at the top, and $Row\ Data\ Adjust$ is the mean-adjusted data transposed, i.e. the data items are in each column, with each row holding a separate dimension. $Final\ Data$ is the final data set, with data items in columns, and dimensions along rows.

It will give the original data solely in terms of the vectors chose. The original data set had two axes, x and y, so data was in terms of them. It is possible to express data in terms of any two axes. If these axes are perpendicular, then the expression is the most efficient. This was why it was important that eigenvectors are always perpendicular to each other. After changed the data from being in terms of the axes x and y , and now they are in terms of 2 eigenvectors. In the case of when the new data set has reduced dimensionality, i.e. leaving some of the eigenvectors out, the new data is only in terms of the vectors that decided to keep.

In the case of keeping both eigenvectors for the transformation, the plot is basically the original data, rotated so that the eigenvectors are the axes. This is understandable since no information lost in this decomposition.

The other transformation is by taking only the eigenvector with the largest eigenvalue. The data resulting from that is found only has a single dimension. If compare this data set with the one resulting from using both eigenvectors, it is noticed that this data set is exactly the first column of the other. So, if plot this data, it would be 1 dimensional, and would be points on a line in exactly the x positions of the points in the plot, which have effectively thrown away the whole other axis, which is the other eigenvector.

Basically by transformed the data so that is expressed in terms of the patterns between them, where the patterns are the lines that most closely describe the relationships between the data. This is helpful because by classified the data point as a combination of the contributions from each of those lines. Initially had the simple x and y axes. This is fine, but the x and y values of each data point don't really tells exactly how that point relates to the rest of the data. Now, the values of the data points tell exactly where (i.e. above/below) the trend lines the data point sits. In the case of the transformation using both eigenvectors, the data was simply altered so that it is in terms of those eigenvectors instead of the usual axes. But the single-eigenvector decomposition has removed the contribution due to the smaller eigenvector and left with data that is only in terms of the other (Smith, 2002).

2.9.2 THE SIGNIFICANCE OF USING PCA METHOD

The aim of PCA is to reduce the quantity of spectral data, and thereby avoid over fitting problems, without discarding any useful information (Osborne et al., 1993). PCA uses projections to extract from a large number of variables, a much smaller number of new variables, which account for most of the variability between samples. Each of the new variables (principal components) is a linear combination of the original measurements and therefore contains information from the entire spectrum. PCA fits new axes (variables) in the data space. The first axis is chosen in the direction of maximum variability. This way the amount of information in the first new variable is maximized. The second axis is chosen to be orthogonal to the first, so the second new variable is uncorrelated with the first one. This operation is continued until a sufficient amount of variation is explained by the new variables. The higher a loading of a variable on a principal component, the more the variable has in common with this component. The loadings can be interpreted as correlations between the variables and the components. The score of the object is the value on the principal component axis, where the object is projected (Jorgensen, 2000).

PCA method is accurate and useful for vibration signal analysis (Ding et al., 2010; Bellino, 2010; Zimroz and Bartkowiak, 2012; Lu and Li, 2007). When the state information changes, the energy of non-stationary vibration signal are also change, this can be indicated accurately by the eigenvectors which is composed of PCA signals (Lu and Li, 2007).

2.10 SIGNAL CLUSTERING

A clustering is a type of classification imposed on a finite set of objects. The relationship between objects is represented in a proximity matrix in which rows and columns correspond to objects. There are two major categories of clustering types namely hierarchical and partitional clustering. The goal of clustering is to determine the intrinsic grouping in a set of unlabeled data. There is no absolute best criterion which would be independent of the final aim of the clustering. Consequently, it is the user which must supply this criterion, in such a way that the result of the clustering will suit their needs (Alexander and Daniel, 2000).

A hierarchical clustering is a sequence of partitions in which each partition is nested into the next partition in the sequence. An agglomerative algorithm for hierarchical clustering starts with the disjoint clustering, which places each of the n objects in an individual cluster. The clustering algorithm being employed dictates how the proximity matrix should be interpreted to merge two or more of these trivial clusters, thus nesting the trivial clustering into a second portion. The process is repeated to form a sequence of nested clustering in which the number of clusters decreases as the sequence progresses until a single cluster containing all n objects, called the conjoint clustering, remains. A divisive algorithm performs the task in the reverse order. Two specific hierarchical clustering methods are now defined called the single- link and the complete-link methods.

Partitional clustering, on the other hand, attempts to directly decompose the data set into a set of disjoint clusters. The criterion function that the clustering algorithm tries to minimize may emphasize the local structure of the data, as by assigning clusters to peaks in the probability density function, or the global structure. Typically the global criteria involve minimizing some measure of dissimilarity in the samples within each cluster, while maximizing the dissimilarity of different clusters. A commonly used partitional clustering method, K-means clustering, will be discussed in some detail since it is closely related to the SOM algorithm. In K-means clustering the criterion function is the average squared distance of the data items from their nearest cluster centroids.

Clustering can be used to reduce the amount of data and to induce a categorization. In exploratory data analysis, however, the categories have only limited value as such. The clusters should be illustrated somehow to aid in understanding what they are like. For example in the case of the K-means algorithm the centroids that represent the clusters are still high-dimensional, and some additional illustration methods are needed for visualizing them.

2.10.1 Distances and Similarities

After getting the eigenvectors from the PCA analysis, an important concept is to consider the distance between object points (data points) as a measure of similarity of the objects. Based on the distances, clusters of objects and outliers can be detected.

A fundamental idea in multivariate data analysis is to regard the distance between objects in the variable space as a measure of the similarity of the objects. Distance and similarity are both inverse: a large distance means a low similarity. Two objects are considered to belong to the same category or to have similar properties if their distance is small.

Feature vector from chapter 2.10.1 discussed previously shall be bring in into this section for clustering purposes. Let two objects be defined by the vectors x_x (with components=variables $eig_{x_1}\ eig_{x_2}\ eig_{x_3} \ldots eig_{x_n}$) and x_y (with components=variables $eig_{y_1}\ eig_{y_2}\ eig_{y_3} \ldots eig_{y_n}$) . Most used is the Euclidean distance, d(Euclid)—equivalent to distance in daily life. Applying Pythagoras' rule gives:

$$d(Euclid) = \left(\sum_{i=1}^{n} (eig_{y_i} - eig_{x_i})^2\right)^{1/2} = \left(\sum_{i=1}^{n} \Delta eig_i{}^2\right)^{1/2}$$
$$= \left[\left(eig_y - eig_x\right)^T (eig_y - eig_x)\right]^{1/2} \qquad (2.11)$$

The Euclidean distance is based on squared differences $\Delta eig_i = eig_{y_i} - eig_{x_i}$ of the variables. The City Block distance (or Manhattan distance) is the sum of the absolute differences of the variables.

$$d(city) = \sum_{i=1}^{n} \left|eig_{y_i} - eig_{x_i}\right| \qquad (2.12)$$

In general, the Minkowski distance is defined by

$$d(Minkowski) = \left(\sum_{i=1}^{n}(eig_{y_1} - eig_{x_1})^p\right)^{1/p} \tag{2.13}$$

The cosine of the angle a between the object vectors is a similarity measure; it is independent from the vector lengths and therefore only considers relative values of the variables.

$$\cos\alpha = \frac{eig_x^T\, eig_y}{\sqrt{(eig_x^T\, eig_x)(eig_y^T\, eig_y)}} = \frac{eig_x^T\, eig_y}{\|eig_x\|\cdot\|eig_y\|} \tag{2.14}$$

This measure is equivalent to the correlation coefficient between two sets of mean centered data—corresponding here to the vector components of eig_x and eig_y.

A similarity measure, s_{xy}, between objects x and y, based on any distance measure, d_{xy}, can be defined as

$$s_{xy} = 1 - d_{xy}/d_{max} \tag{2.15}$$

with d_{max} the maximum distance between objects in the used data set.

The Mahalanobis distance considers the distribution of the object points in the variable space (as characterized by the covariance matrix) and is independent from the scaling of the variables. The Mahalanobis distance between a pair of objects eig_x and eig_y is defined as

$$d(Mahalanobis) = \left[(eig_y - eig_x)^T \cdot C^{-1} \cdot (eig_y - eig_x)\right]^{0.5} \tag{2.16}$$

It is a distance measure that accounts for the covariance structure, here estimated by the sample covariance matrix C. Clearly, one could also take a robust covariance estimator. The Mahalanobis distance can also be computed from each observation to the data center, and the formula changes to

$$d(eig_j) = \left[\left(eig_j - \overline{eig}\right)^T C^{-1}\left(eig_j - \overline{eig}\right)\right]^{0.5} \quad for\ j = 1, \dots, n. \qquad (2.17)$$

Here, eig_j is an object vector, and the center is estimated by the arithmetic mean vector $\overline{eig}$, alternatively robust central values can be used.

Points with a constant Euclidean distance from a reference point (like the center) are located on a hyper sphere (in two dimensions on a circle); points with a constant Mahalanobis distance to the center are located on a hyper ellipsoid (in two dimensions on an ellipse) that envelops the cluster of object points. That means the Mahalanobis distance depends on the direction. Mahalanobis distances are used in classification methods, by measuring the distances of an unknown object to prototypes (centers, centroids) of object classes.

The most widely used distance measure for cluster analysis is the Euclidean distance. The Manhattan distance would be less dominated by far outlying objects since it is based on absolute rather than squared differences. The Minkowski distance is a generalization of both measures, and it allows adjusting the power of the distances along the coordinates. All these distance measures are not scale invariant. This means that variables with higher scale will have more influence to the distance measure than variables with smaller scale. If this effect is not wanted, the variables need to be scaled to equal variance.

The cosine of the angle between object vectors and the Mahalanobis distance are independent from the scaling of the variables. The latter accounts for the covariance structure of the data, but considering the overall covariance matrix of all objects might result in poor clustering. Thus usually the covariance for the objects in each cluster is taken into account (Varmuza and Filzmoser, 2009).

2.10.2 Linkage

The basic information for splitting or merging clusters is the similarity or distance between the clusters. Let us denote $x_1^{(j)}$ and $x_1^{(l)}$ for i=1,…,n as all observations that have been assigned to cluster j and l, respectively, with the cluster sizes n_j and n_l. Then the distance between the two clusters with index j and l can be determined by various methods, where the following are most frequently used:

$$\text{COMPLETE LINKAGE} = max_i \left\| x_i^{(j)} - x_i^{(l)} \right\| \tag{2.18}$$

$$\text{SINGLE LINKAGE} = min_i \left\| x_i^{(j)} - x_i^{(l)} \right\| \tag{2.19}$$

$$\text{AVERAGE LINKAGE} = average_i \left\| x_i^{(j)} - x_i^{(l)} \right\| \tag{2.20}$$

$$\text{CENTROID METHOD} = \left\| c_i - c_j \right\| \tag{2.21}$$

$$\text{WARD'S METHOD} = \left\| c_i - c_j \right\| \cdot \frac{\sqrt{2n_j n_l}}{\sqrt{n_j + n_l}} \tag{2.22}$$

Complete linkage takes the maximum of all pairwise distances between those objects that belong to cluster j and those that belong to cluster l. So, if even only two objects in both clusters are far away, the distance between both clusters will be large. In contrast, single linkage uses the minimum of the pairwise distances. Thus, the distance between two clusters will be small as soon as there exists one object in one cluster that is close to one or another object in the other cluster, independent on how far the remaining objects are. Average linkage computes the average distance between all pairs of objects in the two clusters, i.e., it computes the sum of all pairwise cluster distances and divides by the number of pairs $n_j \cdot n_l$. The centroid method uses the distance between the cluster centroids c_j and c_l. This, however, does not lead to strictly increasing distances within the clustering procedure, and thus a visualization of the results is difficult. This problem is corrected by the factor used in Ward's method. The other methods are based on averages or on distances between the cluster centroids, and are thus somewhere in between single and complete linkage. Complete linkage usually results in homogeneous clusters in the early stages of the agglomeration, but the resulting clusters will be small. Single linkage is known for the chaining effect, because even quite inhomogeneous clusters can be linked just by chance as soon as two objects are very similar. Hence, single linkage is chosen for this research.

The outline an algorithm for agglomerative clustering:

1. Define each object as a separate cluster and compute all pairwise distances.
2. Merge those clusters (or objects) with the smallest distance into a new cluster.
3. Compute the distances between all clusters using complete linkage, single linkage, average linkage, or other methods.
4. Merge those clusters with smallest distance from step 3.
5. Proceed with steps 3 and 4 until only one cluster remains.

The algorithm starts with linking pairs of objects (visualized by the connecting lines) into new clusters of size 2. Then single objects are merged to the clusters, and finally clusters are merged. This merging is indicated by the ellipses, where the increasing line thickness corresponds to higher levels in the hierarchy.

The results of hierarchical clustering are usually displayed by a Dendrogram. This tree-like diagram shows the objects as leaves at the bottom, and branches merge according to the order given by the algorithm. This presentation is appropriate because the objects in clusters of a lower level of the hierarchy are always a subset of the objects in clusters of a higher level of the hierarchy. The dendrogram shows the evolution of the cluster tree from bottom to top, in the scale of the cluster distance measure. It can be used for identifying the number of clusters that is most likely inherent in the data structure. If such a number, say k, exists, expect k clear branches of the complete tree, i.e., when merging the k clusters to $k - 1$ clusters, the cluster distance (height) will increase considerably. A given number of clusters can be achieved by cutting the tree at a certain height.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 RESEARCH DESIGN

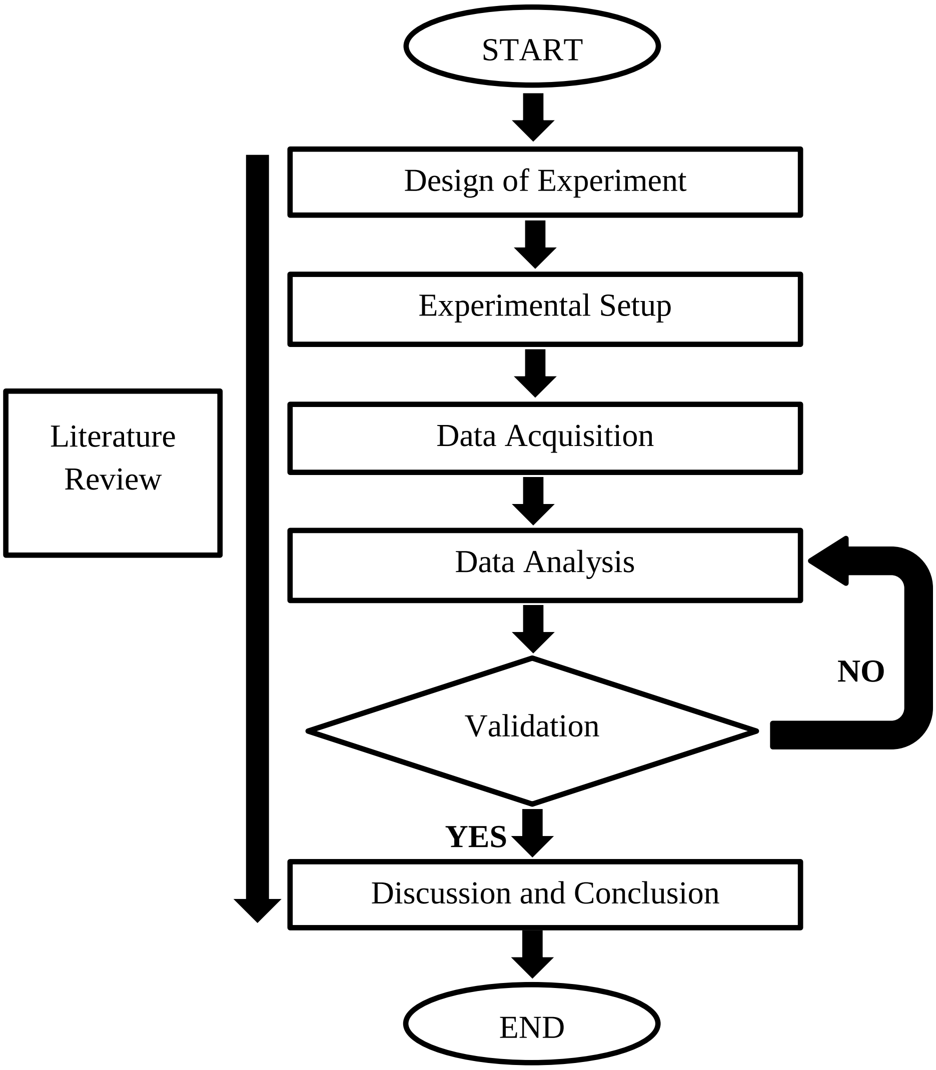
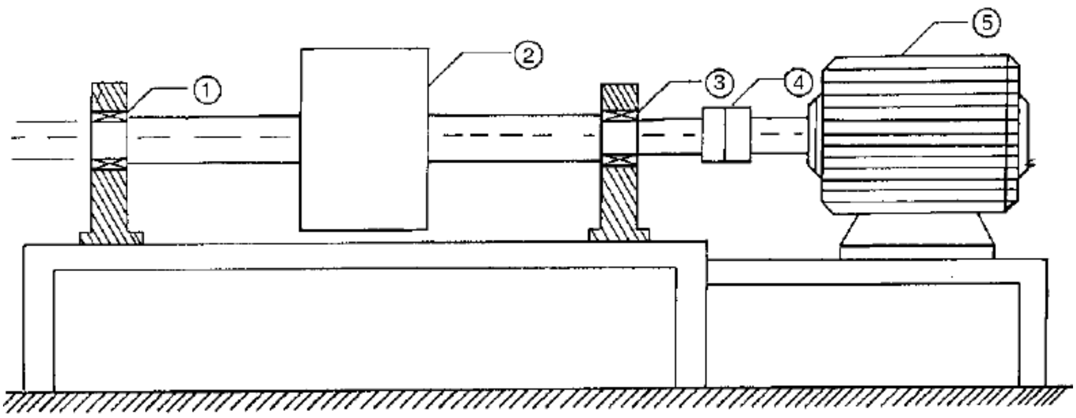


Figure 3.1: Research flow chart

Figure 3.1 shows the entire research flow chart for this study. Beginning with brainstorming for the design of experiment and experiment set up. Followed by data acquisition and data analysis which involving the software application such as DASYLab and MATLAB. Once the result validated and satisfied, the research flow continued for discussion and conclusion.

3.2 EXPERIMENT LAYOUT

An experimental test rig built to predict defects in antifriction bearings is shown in Figure 3.2 and Figure 3.3. The test rig consists of a shaft with central rotor, which is supported on two bearings. An induction motor coupled by a flexible coupling drives the shaft. Defect cylindrical roller bearing is mounted at driver end and healthy cylindrical roller bearing is mounted at free end. The cylindrical roller bearing is tested at certain constant speed of motor. Cylindrical roller bearing type FAG NU203-E-TVP2 used for analysis.



1-healthy cylindrical roller bearing; 2-rotor; 3-defect cylindrical roller; bearing; 4-flexible coupling and 5-motor.

Figure 3.2: Schematic diagram for experimental test rig

Source: Amarnath, 2004

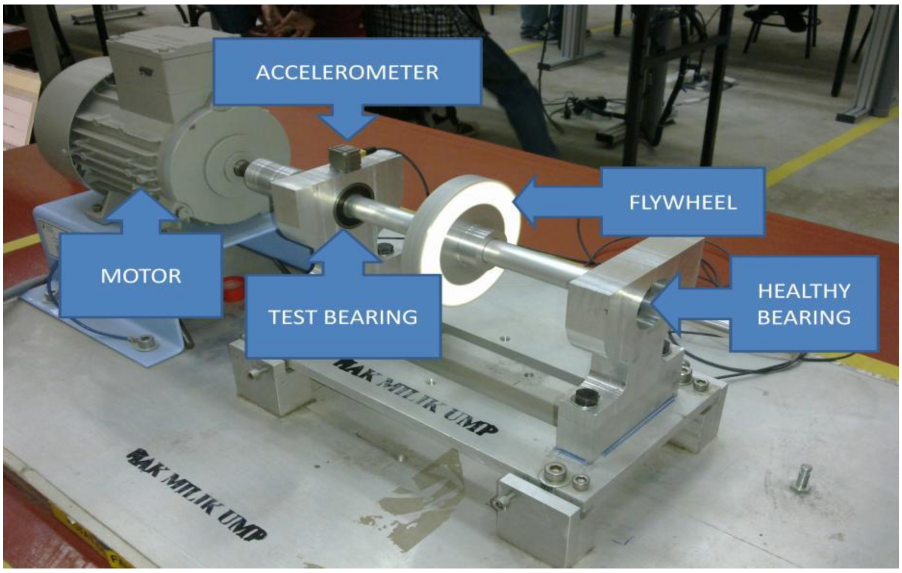


Figure 3.3: Experimental test rig

Table 3.1: Experimental speed states

Speed States	Shaft Speed (RPM)
S1	440 or 15%
S2	1480 or 50%
S3	2672 or 90%

The experiment conducted by a variable speed (0-3000 RPM) D.C. motor, where three different of speed is chosen at 440 RPM (or 15% of motor speed), 1480 RPM (or 50% of motor speed) and 2672 RPM (or 90% of motor speed) as listed in Table 3.1 above.

Four types of bearing defects, namely, inner race, contaminated, corroded and roller defects were studied. Experimental tests were carried out on four sets of bearings. Initially new bearing (good bearing) was fixed in the test rig and signals were recorded using FFT analyzer in DASYLab. The good bearing was replaced by defective bearing and signals were recorded for each one of the case separately under the same standard condition using frequency domain analysis were carried out.

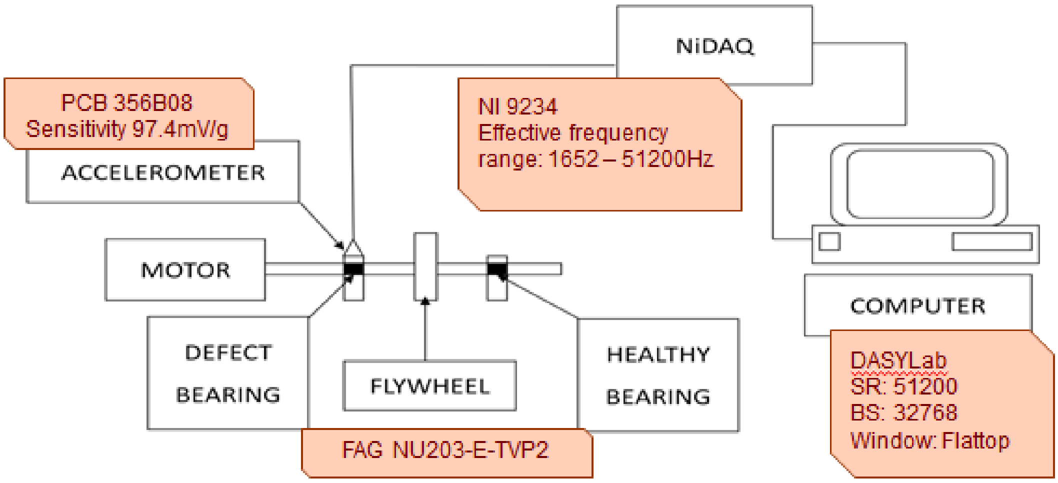


Figure 3.4: Standard condition setup

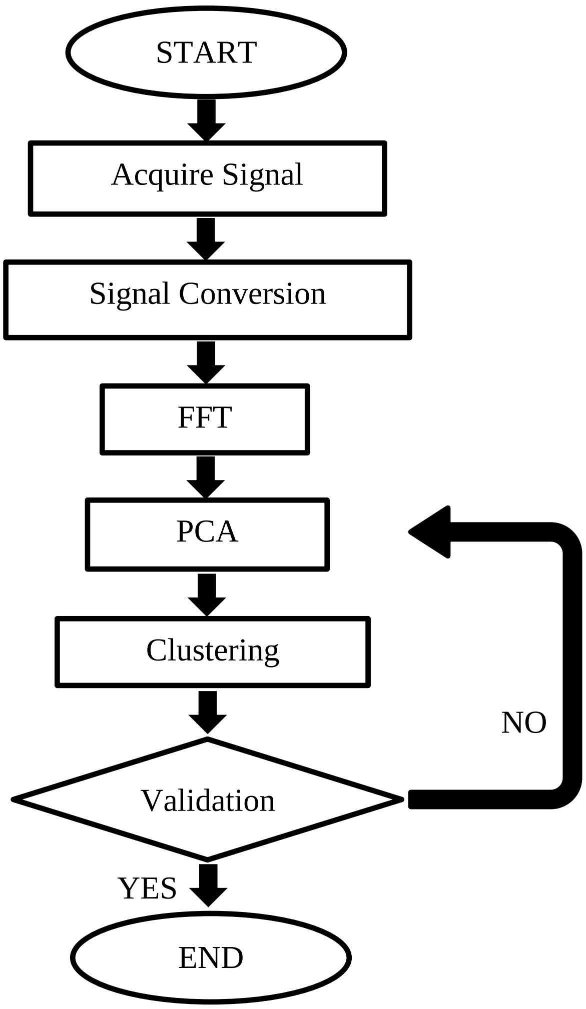


Figure 3.5: Signal analysis flow chart

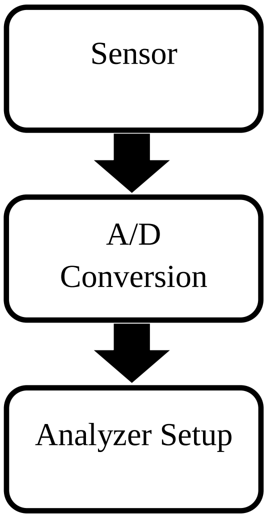


Figure 3.6: General flow for signal acquisition

Figure 3.4 shows standard condition setup for the experiment. The experiment and signal analysis involved is illustrated in the flow chart Figure 3.5. The sensor used is PCB piezoelectric sensor model 356B08 with the sensitivity of 95.5 mV/g for x-axis, 96.6 mV/g for y-axis and 97.4 mV/g for z-axis. Since cylindrical roller bearing handles heavier radial loads, hence, the z-axis signal is then studied.

As illustrated in Figure 3.4, the number of samples chosen is typically a number like 51200. This is also the maximum sampling rate for NI-DAQ 9234 which having effective frequency of 1652 Hz to 51200 Hz. The block size of 32768 data is chosen, which is also the maximum block size available in DASYLab software. By choosing the bigger block size enable the result of FFT having smaller resolution between each data, which a resolution of 1.56 Hz is obtained.

Additionally, the integration of signals (producing a velocity or displacement signal from an accelerometer or a displacement signal from a velocity pickup) will tend to lose low frequency information and introduce noise. Integration of the input signal is generally best accomplished in analog circuits due to the limited dynamic range of the analog-to-digital (A/D) conversion process. Digital circuits typically introduce more errors and if there is any jitter at low frequency, it becomes magnified upon integration.

Leakage (or smearing) is the result of the FFT algorithm trying to handle discontinuities in the sample. The FFT sees the discontinuities as a modulating frequency. This can be overcome by applying windowing function. The windowing function applied in this research is Flat-Top that gives broader band shape but maximum amplitude error of 1% compared to Hanning which gives narrow band shape but maximum amplitude errors of 16%. Figure 3.6 shows the general flow for signal acquisition.

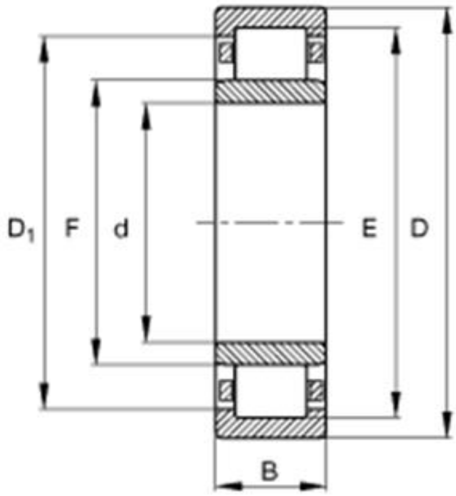
3.3 BEARINGS CHARACTERISTIC

The bearing that going to use in this research is the cylindrical roller bearings NU203-E-TVP2, manufactured by the Schaeffler Group's FAG brand. The details for the bearing mentioned are list as below in Table 3.2 and Figure 3.7:

Table 3.2: Cylindrical roller bearing details and dimensions (FAG NU203-E-TVP2)

Annotation	Details / Dimension
Inner diameter, $\boldsymbol{d}$	17mm
Outer diameter, $\boldsymbol{D}$	40mm
Roller diameter, $\boldsymbol{D_R}$	6.5mm
Pitch diameter, $\boldsymbol{D_P}$	10.4mm
Bearing’s width, $\boldsymbol{B}$	12mm
Number of rollers, $\boldsymbol{N_R}$	11pcs
Contact angle, $\boldsymbol{\alpha}$	0 º

Source: Schaeffler Group Industrial, 2012



d= 17mm

D= 40mm

B= 12mm

D_1= 32.5mm

F= 22.1mm

E= 35.1mm

Figure 3.7: Cylindrical roller bearing geometry (FAG NU203-E-TVP2)

Source: Schaeffler Group Industrial, 2012

The FAG brand started with an ingenious idea. In 1883, Friedrich Fischer designed a ball grinding machine in Schweinfurt, Germany that, for the first time, made it possible to produce absolutely round steel balls by grinding. This invention is regarded as the foundation for the entire rolling bearing industry. This is one of the reasons why FAG has long been considered to be a pioneer in rolling bearing technology.

3.4 PREPARATION OF TEST BEARINGS

Discrete testing required the preparation of test bearings with discrete faults of a known width in one of the raceways. For this research a number of bearing faults were considered, namely inner race defect, outer race defect, rolling element defect, cage defect, corrosion and contamination defects. In order to develop the required library of vibration records, a set of bearings with faults of differing faults was required. Schweinfurt FAG NU203-E-TVP2 bearings were used in this research: 6 were dismantled and used for discrete testing. About 40% of failures of electric motors are caused by bearings (Albrecht et al., 1986). In industrial applications, these bearings are considered as critical mechanical components and a defect in such a bearing, unless detected in time, causes malfunction and may even lead to catastrophic failure of the machinery (Tandon and Choudhury, 1999).

The common causes of these faults are impropriate lubrication (Timken Company, 2011). The main failure cause is the inappropriate lubrication of the bearing rolling elements, which is approximately 80% of the cases (Radu, 2010).

To achieve getting a corroded test bearing, it is rather simple to prepare for it simply by applying water onto the bearing and let it be oxidized for a few days and the outcome is illustrated in Figure 3.8. For a faster oxidize reaction, the test bearing can be expose to acidic solution simply by apply some acidic solution on it.

Figure 3.8: Corroded defect bearing specimen

For contamination fault, it can be achieve by applying some steel powder or burs together with lubricant in order to allow the steel powder to be evenly being applied throughout the bearing, as illustrated in Figure 3.9.

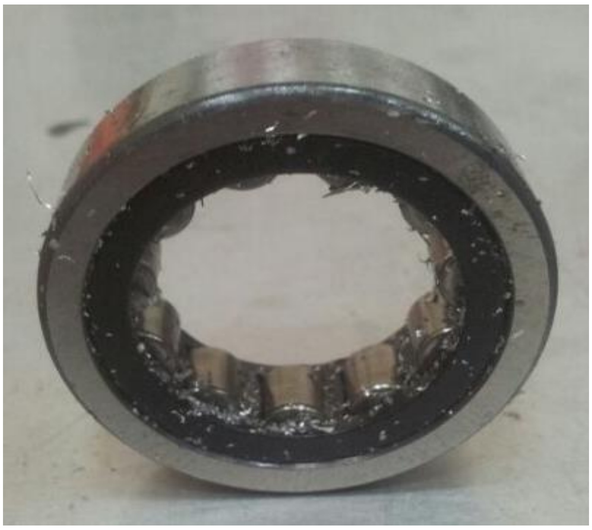


Figure 3.9: Contaminated defect bearing specimen

Figure 3.10 illustrate the inner race defect prepared for experiment purpose. To prepare an inner race defect bearing, simply by using a sharp triangular chisel tool to do filling on the inner race. Due to the reason that shape of inner race is cylindrical, the use of a G-clamp is needed in order to fix the inner race position while doing filing on it.

Figure 3.10: Inner race defect bearing specimen

As for preparing a rolling element defect, it can be achieve by applying the method that almost similar with the method applied on creating an inner race defect that discussed previously and is illustrated in Figure 3.11.

Figure 3.11: Roller defect bearing specimen

3.5 DATA ACQUISITION

Figure 3.12 shows that setup used for NI-DAQ software using NI-9234 device. The sensitivity of the sensor used must be key in correctly correspond to the axes. Figure 3.13 shows setup for DASYLab software, including the windowing function used is Flattop shown in the figure and is applied throughout the whole data acquisition processes.

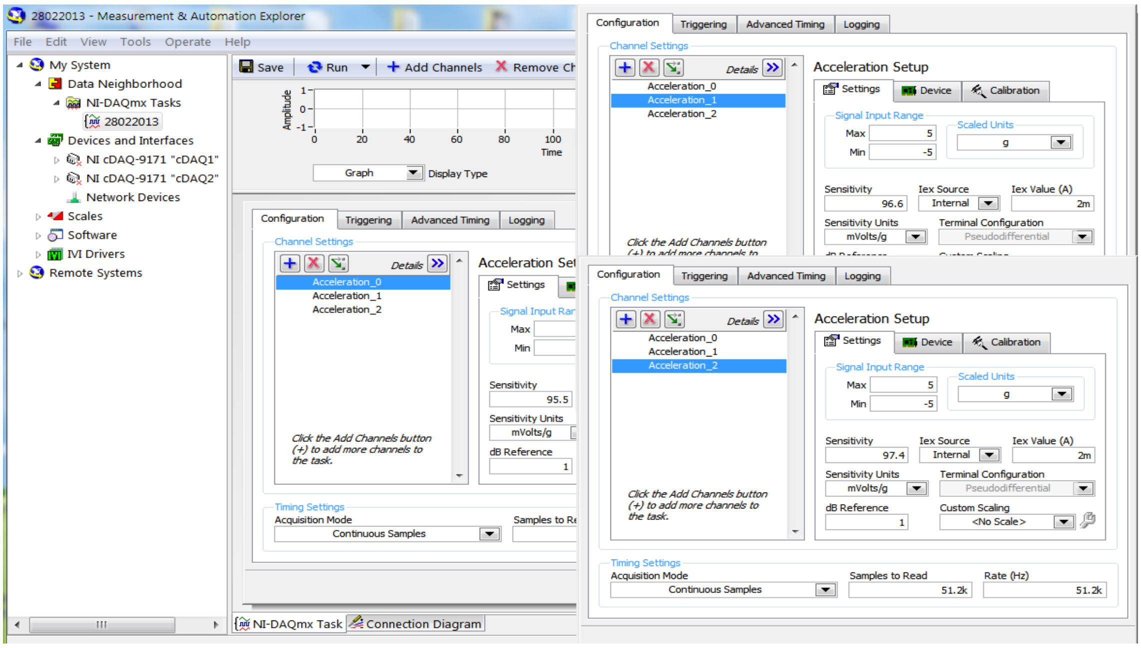


Figure 3.12: NI-DAQ software setup

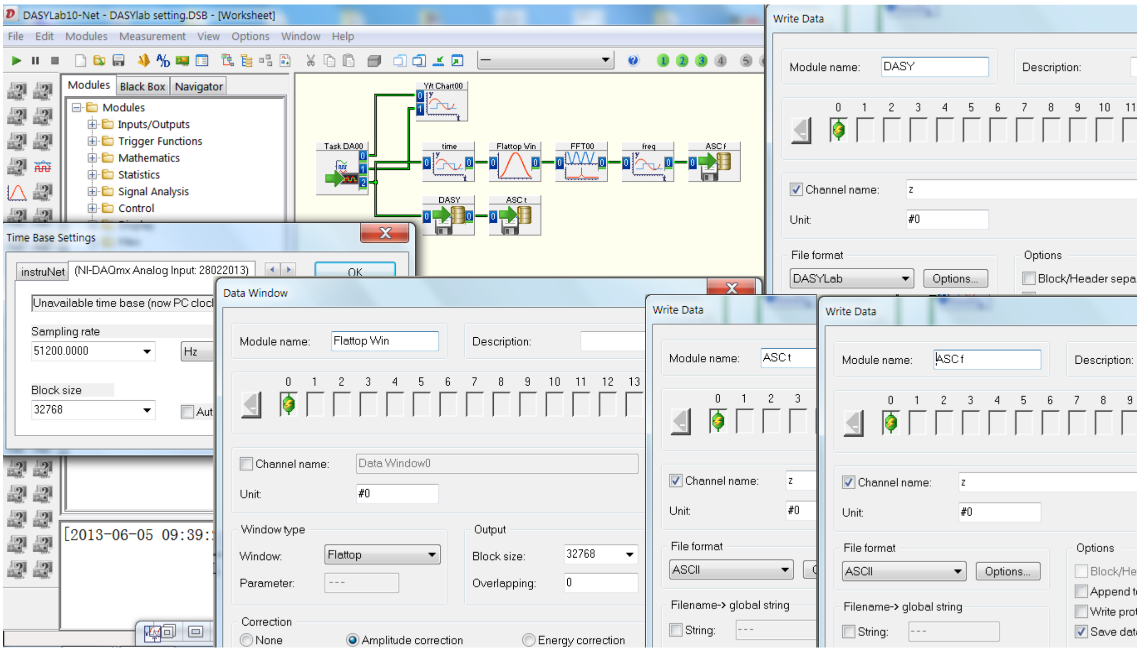


Figure 3.13: DASYLab software setup

3.6 SIGNAL ANALYSIS

Principle Component Analysis, PCA is the method used for signal clustering. The PCA method, which showed its effectiveness in the fault detection and isolation FDI, was implemented recently for the system diagnosis (Liu, 2006; Benaicha et al., 2010; Huang, 2001).

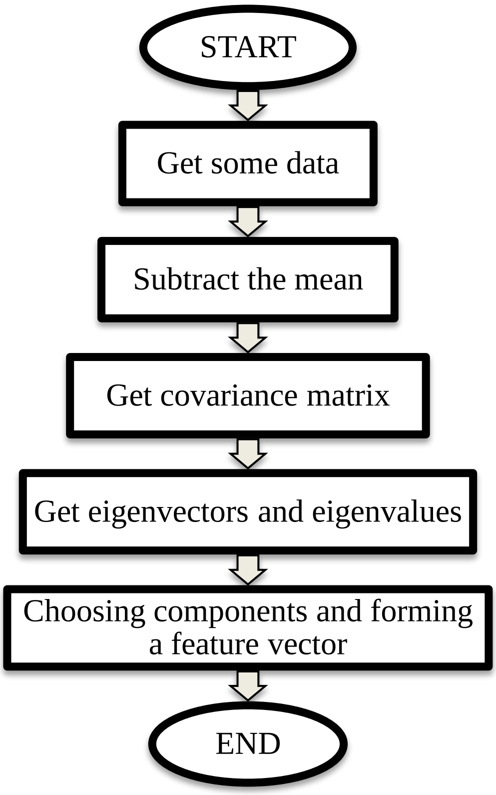
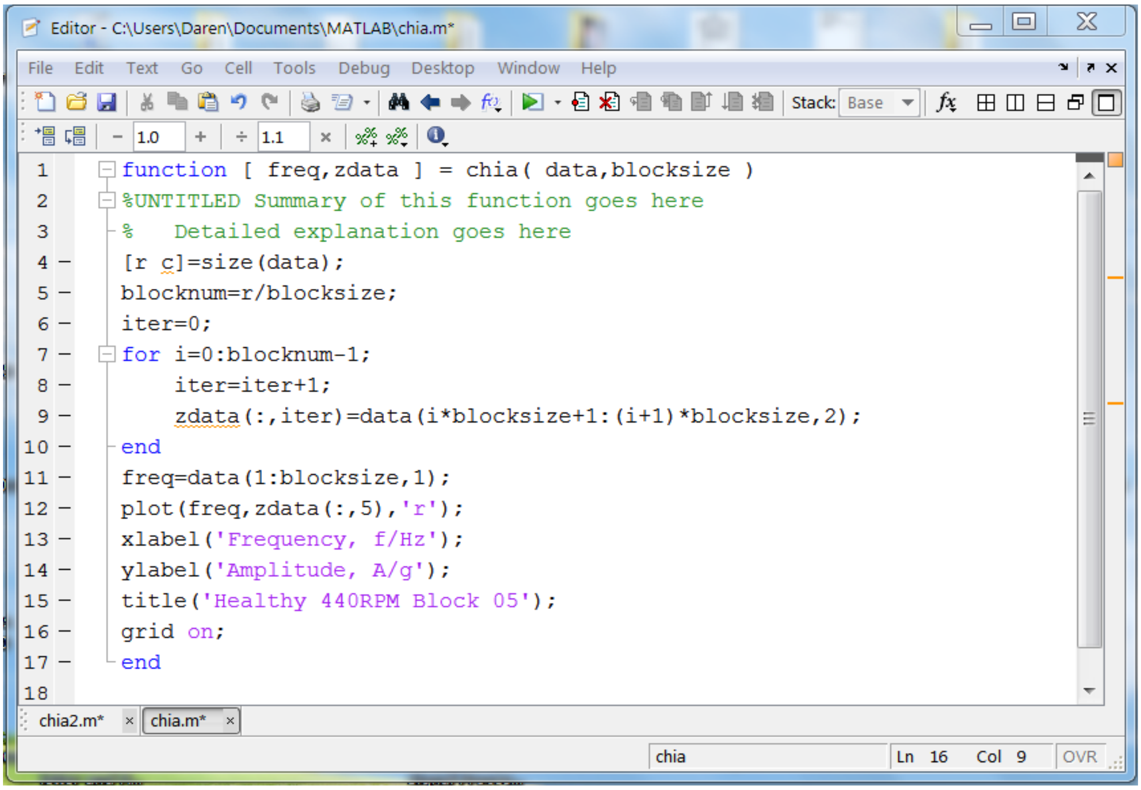


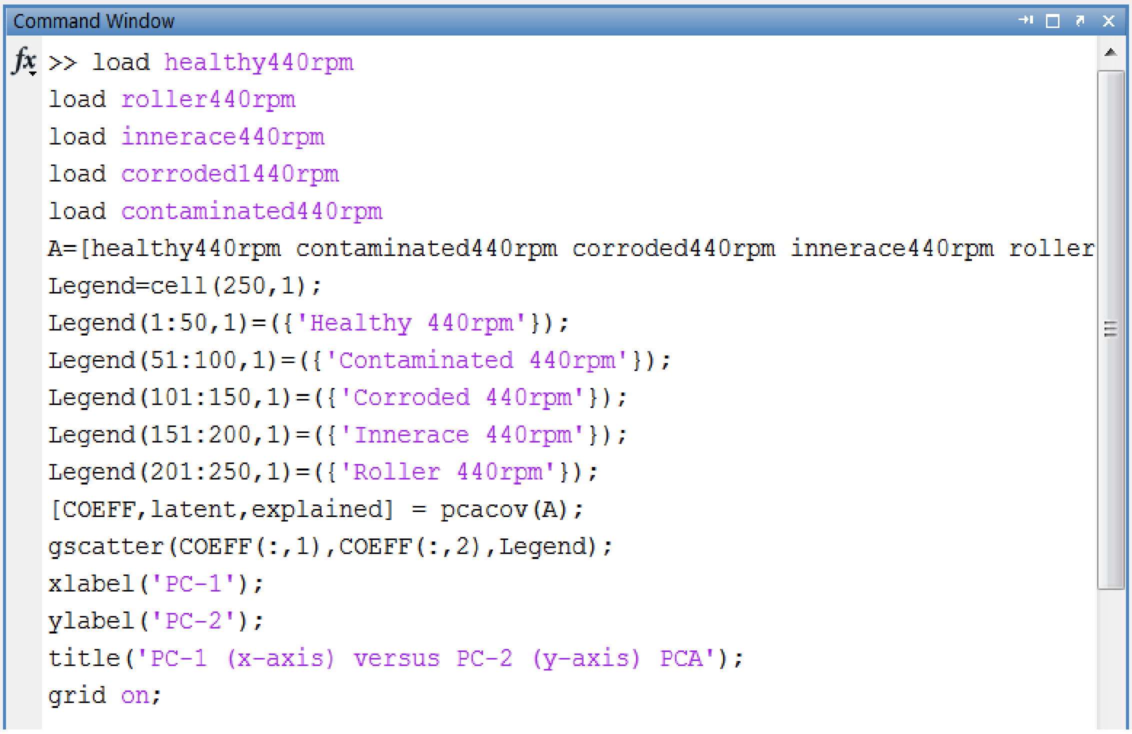
Figure 3.14: Applied PCA flow chart

Figure 3.14 shows the applied PCA flow chart. The software used to apply PCA is MATLAB. MATLAB coding in Figure 3.15 is the purpose to rearrange data into the correct time blocks while command in Figure 3.16 is to prepare the scatter plot.



```
function [ freq,zdata ] = chia( data,blocksize )
%UNTITLED Summary of this function goes here
%   Detailed explanation goes here
[r c]=size(data);
blocknum=r/blocksize;
iter=0;
for i=0:blocknum-1;
    iter=iter+1;
    zdata(:,iter)=data(i*blocksize+1:(i+1)*blocksize,2);
end
freq=data(1:blocksize,1);
plot(freq,zdata(:,5),'r');
xlabel('Frequency, f/Hz');
ylabel('Amplitude, A/g');
title('Healthy 440RPM Block 05');
grid on;
end
```

Figure 3.15: MATLAB coding to arrange data in time block



```
>> load healthy440rpm
load roller440rpm
load innerace440rpm
load corroded1440rpm
load contaminated440rpm
A=[healthy440rpm contaminated440rpm corroded440rpm innerace440rpm roller
Legend=cell(250,1);
Legend(1:50,1)=({'Healthy 440rpm'});
Legend(51:100,1)=({'Contaminated 440rpm'});
Legend(101:150,1)=({'Corroded 440rpm'});
Legend(151:200,1)=({'Innerace 440rpm'});
Legend(201:250,1)=({'Roller 440rpm'});
[COEFF,latent,explained] = pcacov(A);
gscatter(COEFF(:,1),COEFF(:,2),Legend);
xlabel('PC-1');
ylabel('PC-2');
title('PC-1 (x-axis) versus PC-2 (y-axis) PCA');
grid on;
```

Figure 3.16: MATLAB command for scatter diagram

3.7 SIGNAL CLUSTERING

Dendrogram is used for clustering display purposes. It is a tree-like diagram, the objects are as leaves at the right-hand side. The branches merge according to the linkage order given by the algorithm. Objects in clusters of a lower level of the hierarchy are always a subset of the objects in clusters of a higher level. Figure 3.17 illustrate the flow involved for signal clustering using dendrogram and the command involved in MATLAB software illustrated in Figure 3.18.

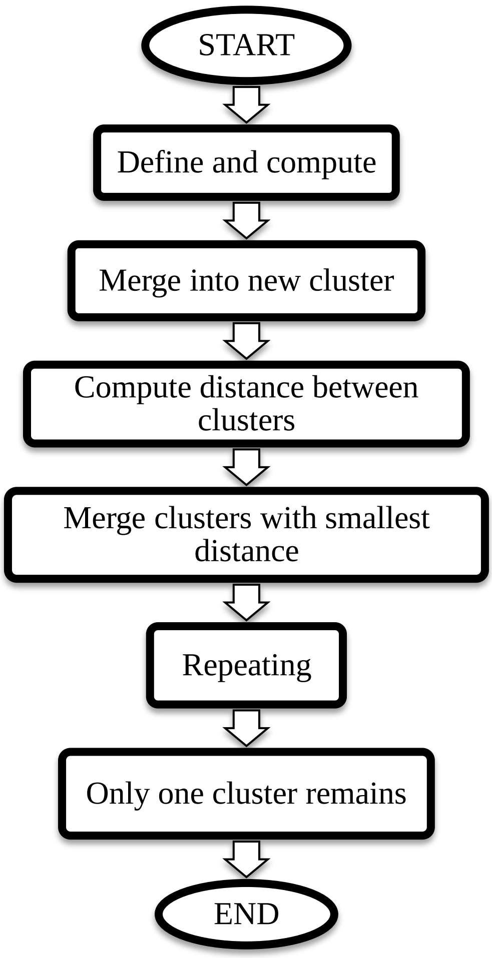


Figure 3.17: Flow chart for dendrogram clustering

```
B=COEFF(:,1:2);
C=pdist(B);
D=linkage(C);
dendrogram(D,0,'orientation','left','labels',Legend);
xlabel('Distance');
title('Dendrogram for 440RPM');
```

Figure 3.18: MATLAB command for dendrogram plotting

CHAPTER 4

RESULT AND DISCUSSION

4.1 INTRODUCTION

There are five groups of data which based on each type of bearing used for this project. Each of the group is known as healthy bearing, inner race defect bearing, contaminated bearing, corroded bearing, and lastly roller defect bearing. Each group of data contains three sets of data which was based on three speeds, which are 440 RPM, 1480 RPM and 2672 RPM. The data is obtained using sampling rate 51200, block size 32768, applying flattop data window for FFT purpose.

4.2 DATA ACQUISITION

Figure 4.1 to 4.3 shows the time domain graph obtained from five types of bearing with speed of 440 RPM, 1480 RPM and 2672 RPM. Healthy bearing (red) without defect has the lowest amplitude as observed. Contaminated (yellow) and roller (pink) defect bearings has a slightly higher amplitude range and has a characteristic spike on an each column of x-axis. This happen due to the rotating bearing hit the point defect on the bearing inner area (Qayyuum, 2012). Whereas for inner race (blue) and corroded (green) defect bearings, the spikes is much more visible than contaminated and roller defect bearings.

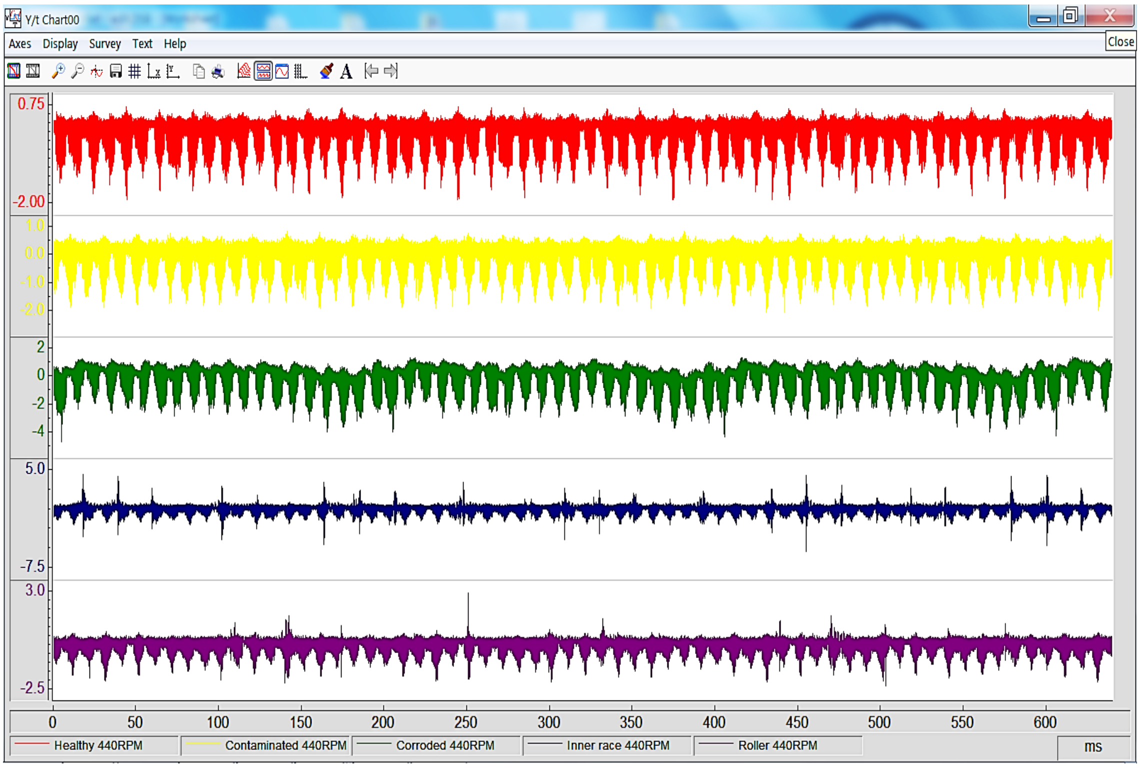


Figure 4.1: Time domain for speed 440 RPM

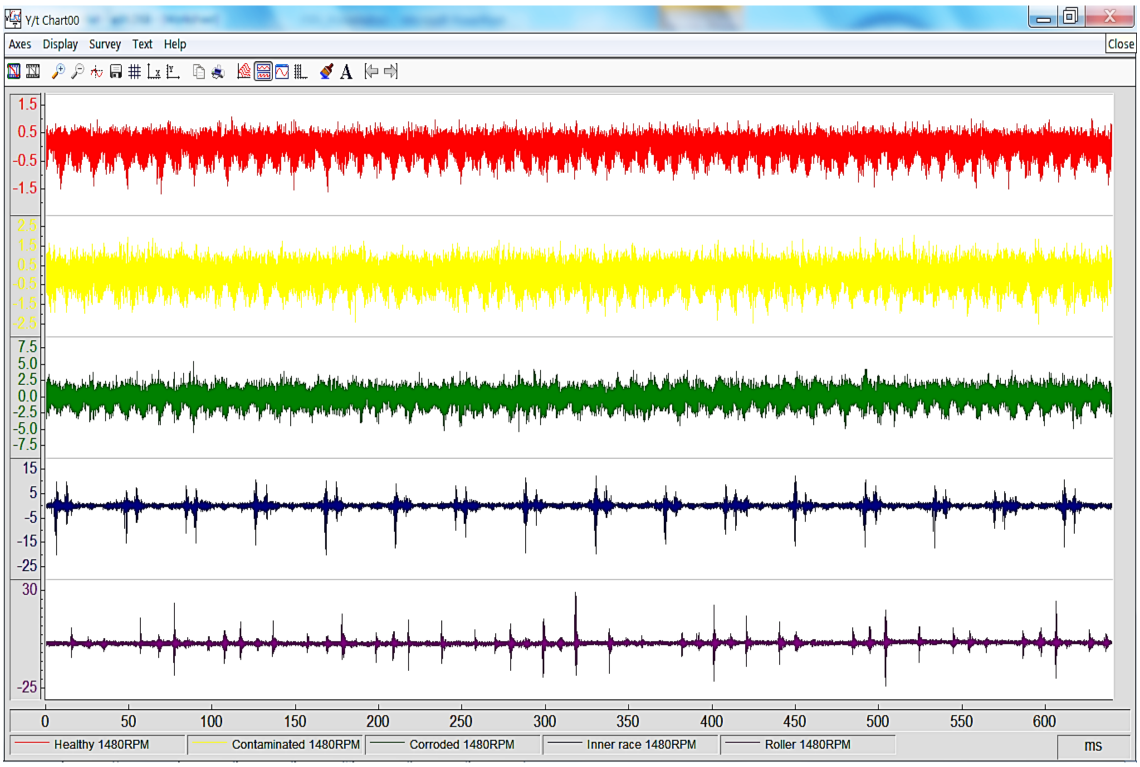


Figure 4.2: Time domain for speed 1480 RPM

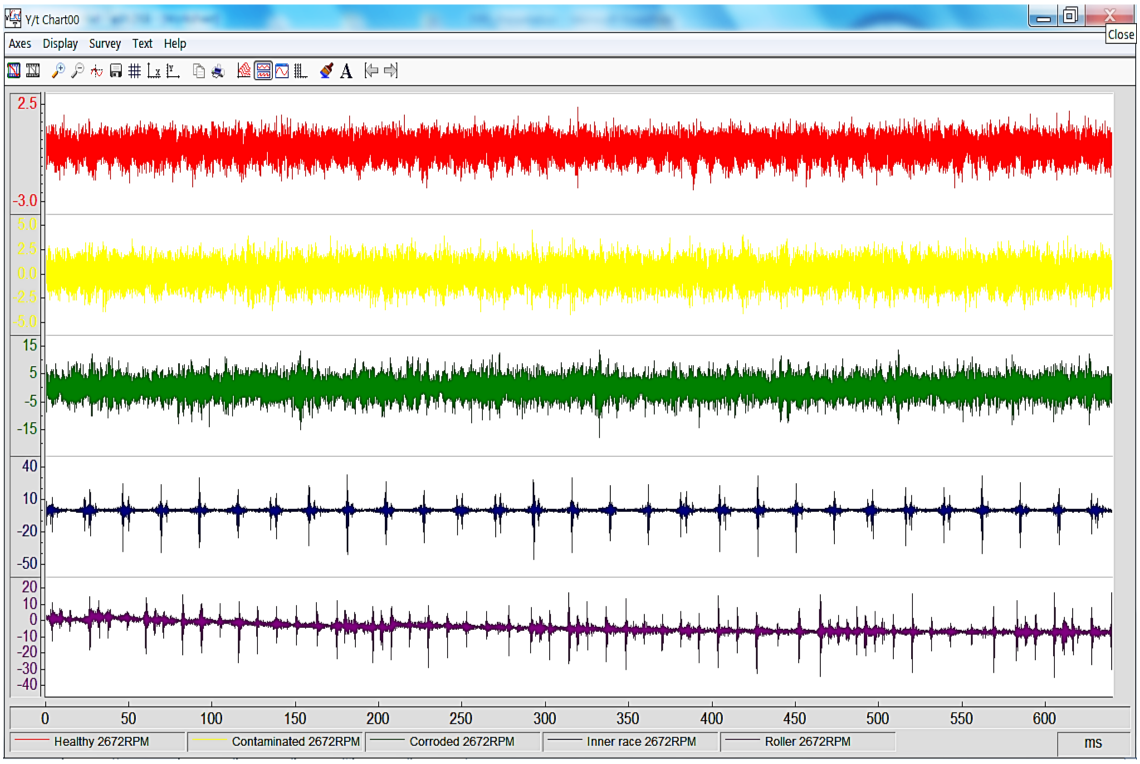


Figure 4.3: Time domain for speed 2672 RPM

Going through Figure 4.1 to 4.3, it is noticed that healthy, contaminated and corroded signals have spikes all over the graph. This could be explained as both contaminated and corroded detect were covered all over the bearing components. Inner race and roller signals show defect spikes in certain period as these two defects are considered as point defect which the faulty only occurred at specific one small part in the bearing component. The amplitude of spikes is noticed proportional to the operating RPM (Bhave, 2010).

Figure 4.4 to 4.6 are the corresponding frequency domain graph for the three different speeds in random block of signal. It is observed that the healthy bearing and roller defect bearing both having quite similar trend in terms of its amplitude values. Contaminated and inner race defect bearings both having higher amplitude trend compared to healthy and roller defect bearings. This is due to both contaminated and inner race defects are considered as heavy defects for bearings. Furthermore, the amplitude trend for corroded defect bearing is observed that having the highest value. This is due to corrosion all over the bearing, from inner race, roller to outer race.

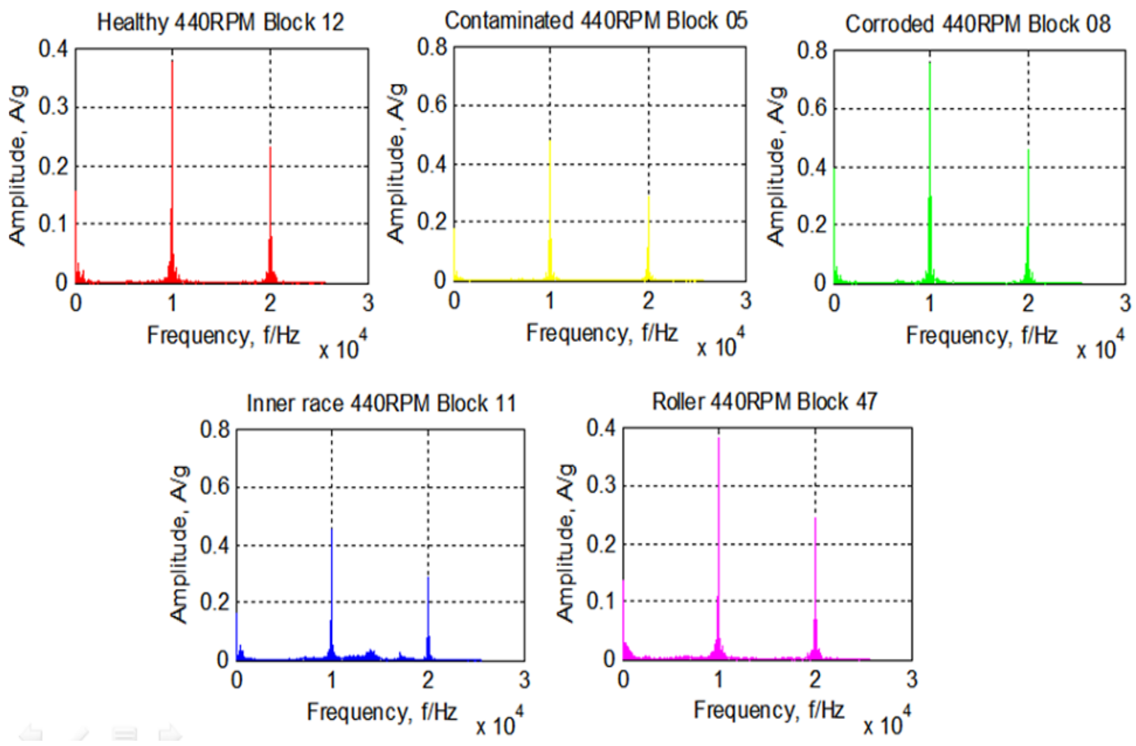


Figure 4.4: Frequency domain graph for five types of bearing on speed of 440 RPM

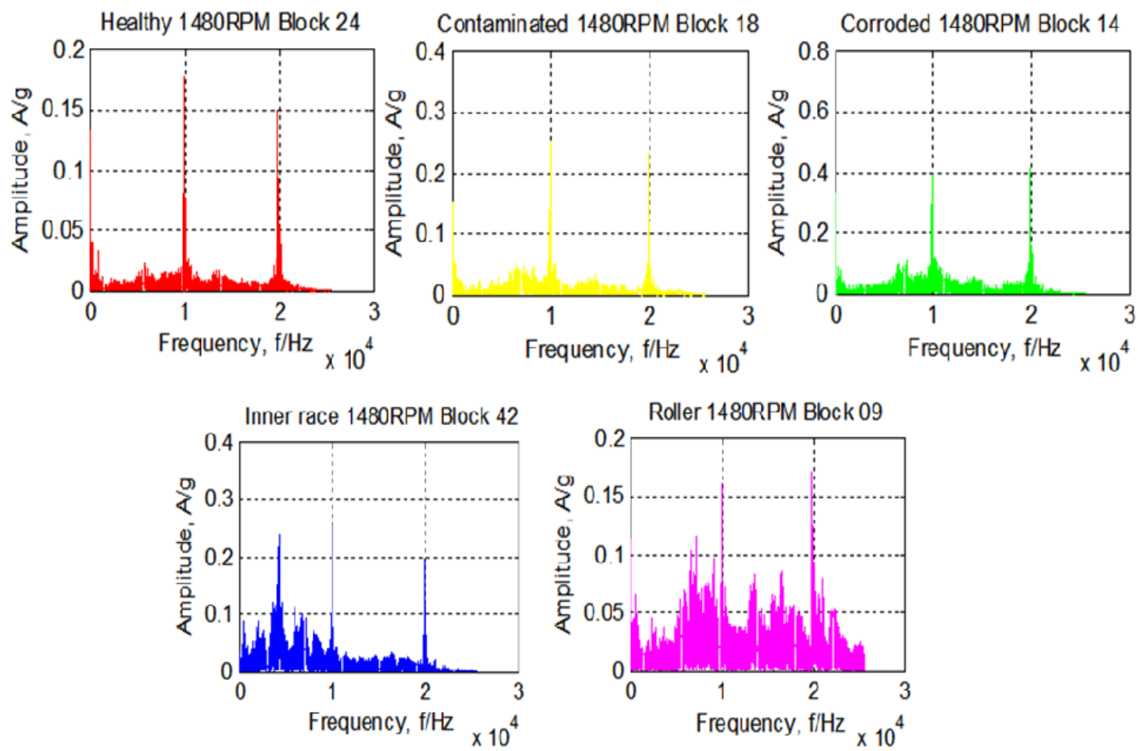


Figure 4.5: Frequency domain graph for five types of bearing on speed of 1480 RPM

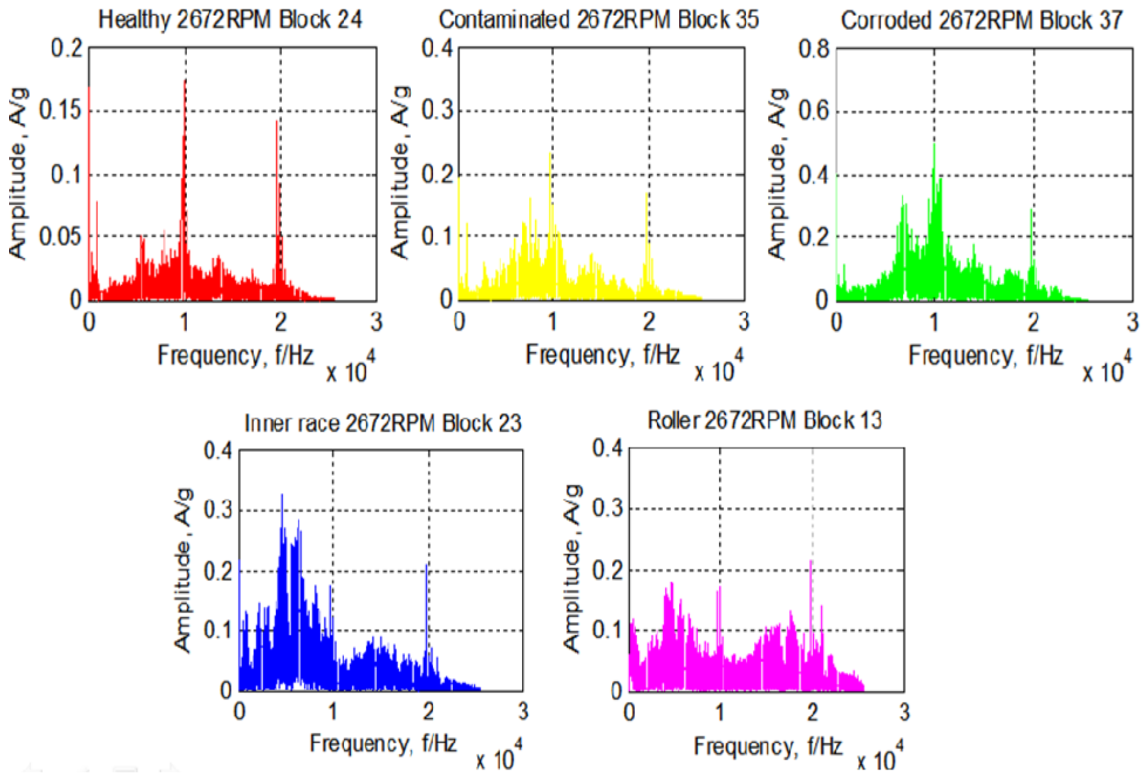


Figure 4.6: Frequency domain graph for five types of bearing on speed of 2672 RPM

Going through Figure 4.4 to 4.6, it is observed that inner race and roller defect both having lump shaped signals in higher RPM. The corroded defect bearing signals has the highest spikes of amplitude (Qayyuum, 2012). Signals are observed that mostly having increase in amplitude with respect to experimental RPM. Signals trend and pattern having similar lump shaped (Suri, 2010).

From all of this result, there the uses of different types of bearing faults where the most severe defect creates higher amplitude in signal and the good bearing creates less vibration (Martin and Honarvar, 1994).

4.3 PRINCIPAL COMPONENT ANALYSIS (PCA) RESULTS

The objective of PCA is to categorize the fifty data in their respective groups for each speed. Where, the plotted data must be separated from each other in the graph.

4.3.1 PCA Result for Speed of 440 RPM

Figure 4.7 shows the calculated principal component analysis (PCA) for five types of bearing that has been plotted based on two principal component analysis axis, PC-1 versus PC-2. This figure clearly shows the data are scattered in their own characteristic, where there is a thin separation of contaminated (yellow) data from the others. Inner race (blue) data share the same situation with contaminated data. As for corroded data, it scatters farthest away comparing to the rest of the defect data. While for healthy (red) and roller (purple) data, it is observed that both of the scatter within each group firmly and closed to each other. It also shows corroded data were not precisely distributed but the other four groups of data were noticed to be precisely distributed in their scatter group.

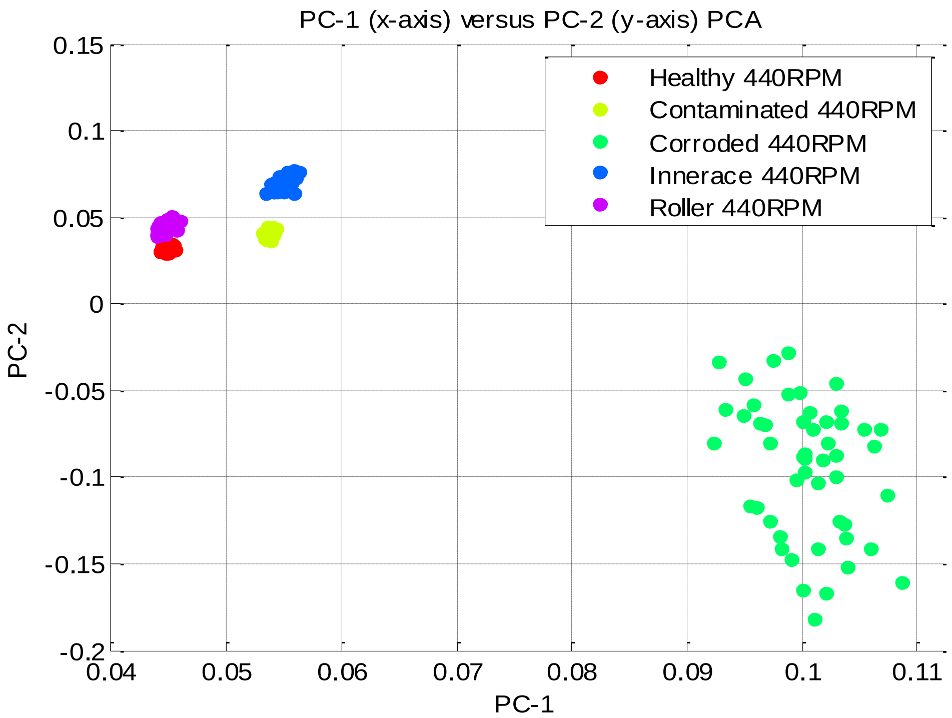


Figure 4.7: PC-1 versus PC-2 PCA scatter diagram for speed 440 RPM

4.3.2 PCA Result for Speed of 1480 RPM

Figure 4.8 shows clearly the separation and grouping for all bearing types. It is noticed that all the group of data is considered grouped firmly in their corresponding cluster. Noticing inner race in the situation where it is greatly isolated and remain within their group. It is noticed that both healthy and contaminated data scattered much more precise compared to the other three types of bearing fault data.

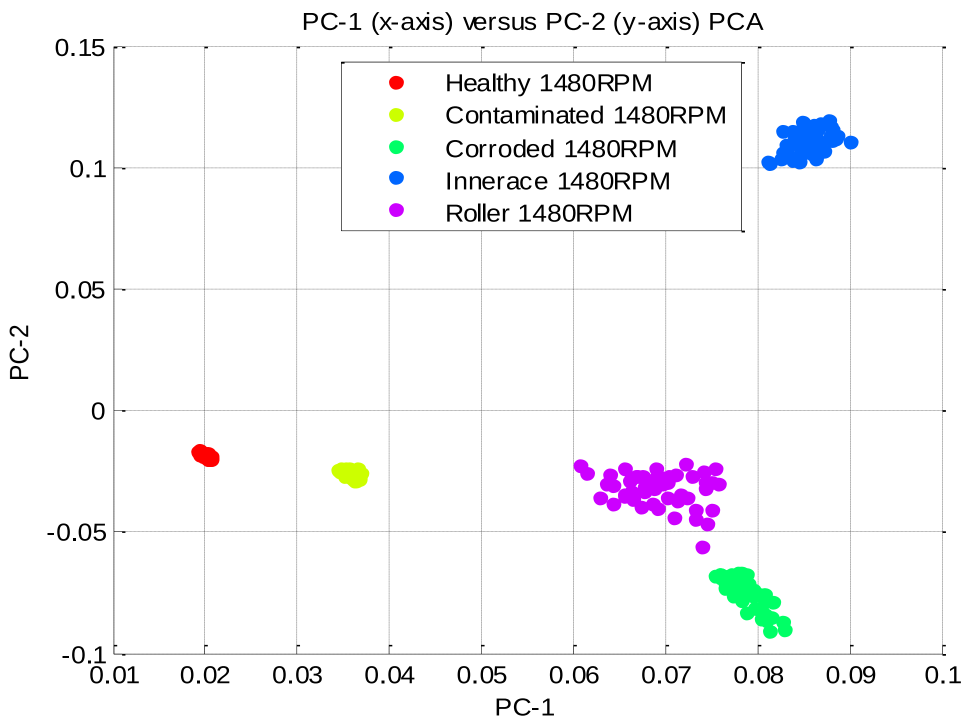


Figure 4.8: PC-1 versus PC-2 PCA scatter diagram for speed 1480 RPM

4.3.3 PCA Result for Speed of 2672RPM

Figure 4.9 shows a result that is also clear in separation on grouped data except roller (purple) having some data that might lead to outliers. But, for all other data, their separation is clear even to the naked eye. Same as previous speed analysis, corroded (green) and inner race (blue) are separated from each other and all other data by quite a fine margin. It is noticed that besides roller defect data, the rest of the data group scattered precisely within each of their defect group.

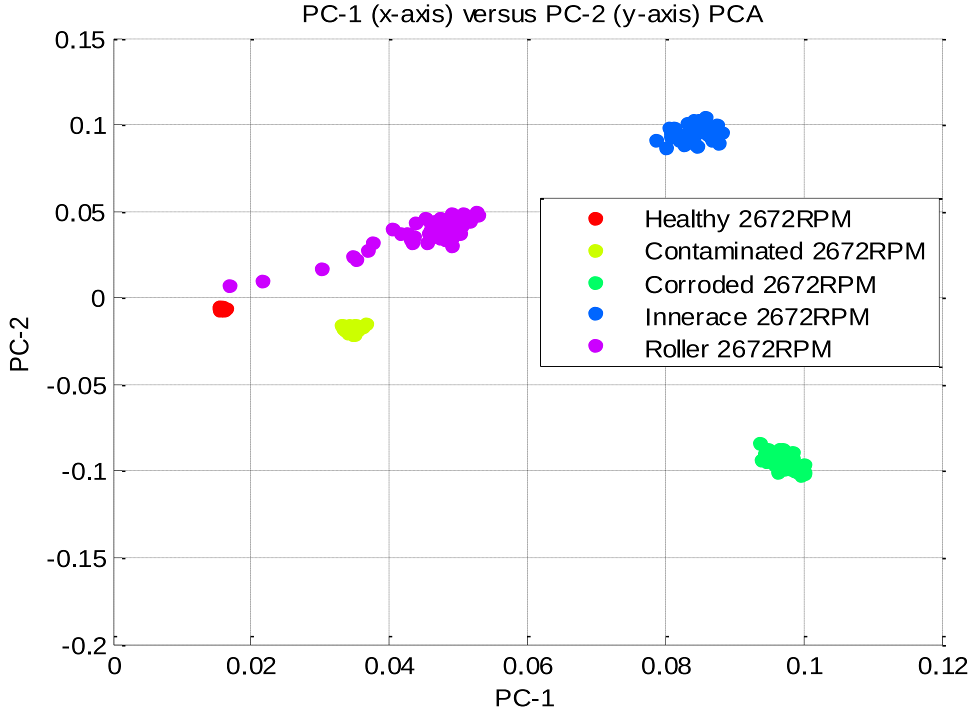


Figure 4.9: PC-1 versus PC-2 PCA scatter diagram for speed 2672 RPM

To conclude the result in PCA, by referring to the scatter diagrams plotted, it shows that all of the bearing type data are scattered in their own groups and none of them intrude and became outliers to other groups (Widodo et al. 2007).

4.4 AGGLOMEARATIVE HIERARCHICAL CLUSTERING

This is the final analysis in this research. For a clustered dendrogram to be successful, by observing for example, the data inner race defects lie in only a branch of the dendrogram. Meaning there is no other data in the branch and inner race defect data also do not intrudNie on other type data branches.

4.4.1 Dendrogram for speed 440 RPM

Since the data plotted in scatter diagram mostly scattered in circular cluster, the distance use for dendrogram plotting purposes is Euclidean distance the most suitable. While for the linkage involved, single linkage approach is used in order to get the minimum of the pairwise distances.

Figure 4.10 shows the complete view of dendrogram generated form MATLAB® software as a whole for speed of 440 RPM sorting from corroded, inner race, contaminated, healthy and lastly roller defect bearings data. By rough observation from the dendrogram, the grouped data is in their respective group or clustering by the type of bearing defects without any outliers occurred.

Overall, it is found that the healthy and the roller having the closest branch, followed by contaminated, then inner race and lastly corroded defect. This result is proven to be corresponding with scatter diagram Figure 4.7 where the corroded defect bearing data lies farthest in the plot. Followed by contaminated, then inner race and lastly corroded defect data.

Although healthy and roller data are observed clustered closed together but by applying dendrogram, all the data able to be separated with their corresponding group. Both defect data does not intrude in each other minimum and maximum range of data.

Figure 4.10: Dendrogram complete view for speed 440 RPM

4.4.2 Dendrogram Speed 1480 RPM

Figure 4.11 shows the full view of dendrogram generated from MATLAB® software as a whole for speed of 1480 RPM sorting from roller, corroded, contaminated, healthy and lastly inner race defect bearing data. Roughly observation from the figure it is observed that the grouped data is in their respective group or branches by the type of bearing defects.

Generally, it is observed that the branch of roller and corroded defect data are much nearer (or similar) to each other. While contaminated and healthy data considered sharing the almost similar condition. Lastly, the inner race defect data having the farthest distance compared to other groups of data. This result is true by referring to the scatter diagram Figure 4.8 where it is clearly that the scatter for inner race defect data lay farthest in the plot.

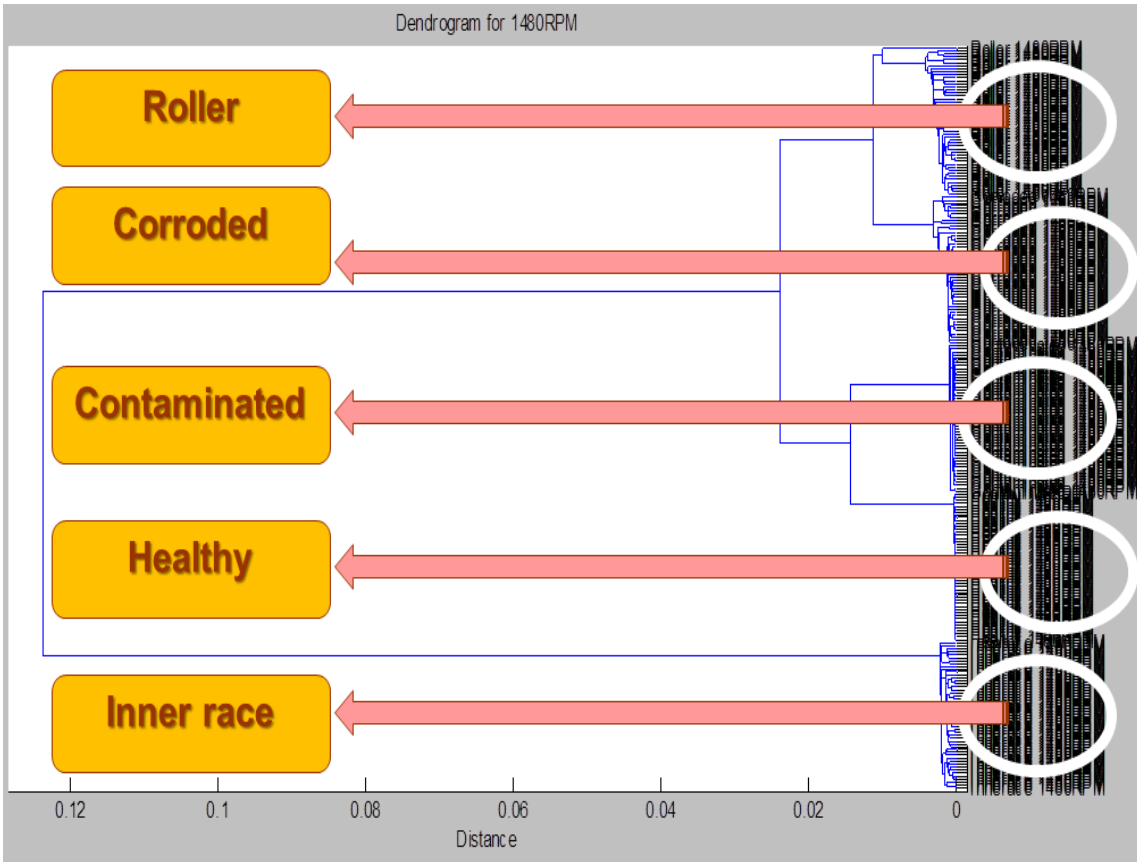


Figure 4.11: Dendrogram complete view for speed 1480 RPM

In scatter diagram of Figure 4.8 it is observed that roller and corroded data are observed clustered closed together. By applying dendrogram, both the data are proved to staying closed together but it is clearly that that stay at each of own group without having any outliers data that intrude other group. All these showing the PCA analysis able to do clustering for the vibration signal data accordingly with respect to their branches without any outliers.

4.4.3 Dendrogram Speed 2672 RPM

Figure 4.12 shows the complete view of dendrogram generated from MATLAB® software as a whole for speed of 2672 RPM sorting from top to the bottom are corroded, inner race, contaminated, roller and lastly healthy bearing data. Through rough observation, it is found that all the groups of data are able to stay in the same branch of each group of branch.

Going through the dendrogram, it is noticed that with the healthy and roller defect data having the closest distances apart, followed by contaminated, then inner race and lastly is corroded data. This condition is proven true by referring to scatter diagram Figure 4.9, where the data of healthy bearing and roller defect bearing group were observed to scatter nearest with each other and while corroded group of data were scatter farthest correspond to other groups of data.

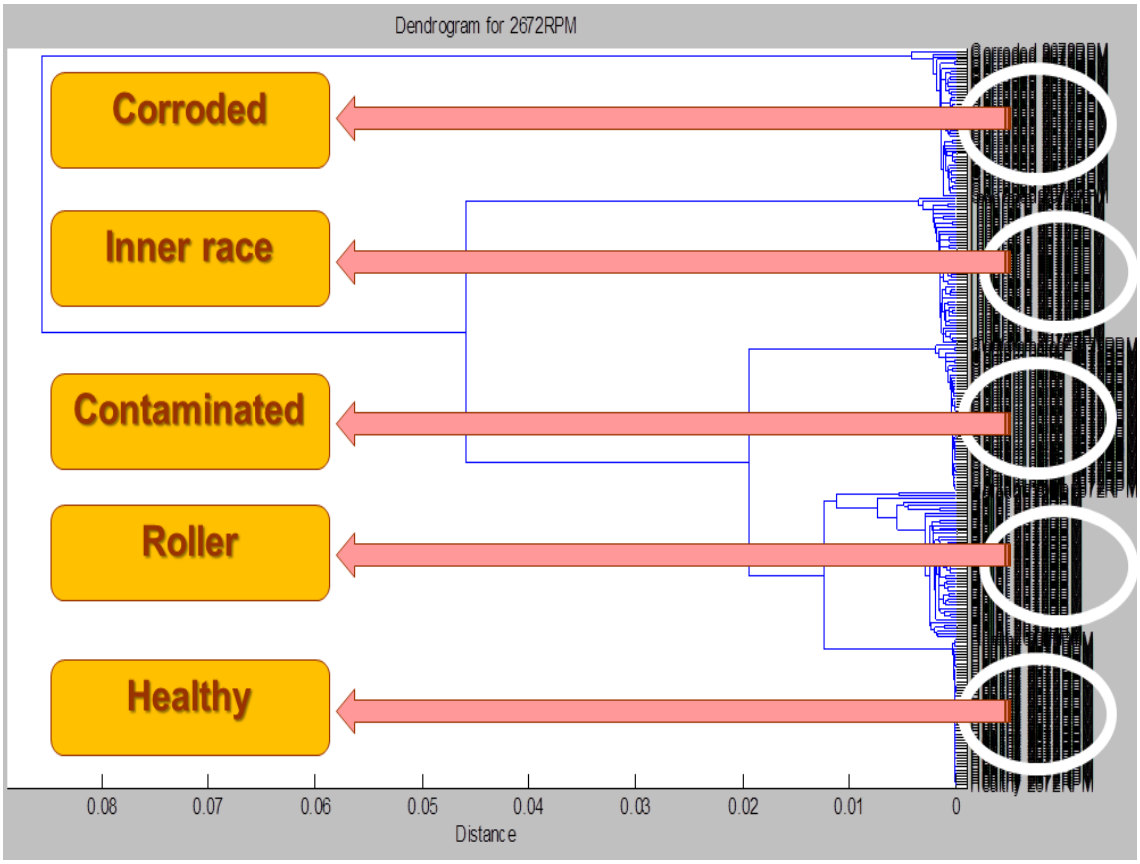


Figure 4.12: Dendrogram complete view for speed 2672 RPM

All in all, according to the results obtained, all of the data are clustered in their own clusters and situated in their respective branches without any outliers, assimilation or intrudes of branches for all the speed studied.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

Defect features was successfully detected on all of defective bearings and good condition bearing using frequency domain signal coupled with Principal Component Analysis to achieve signal clustering. Statistical parameters calculations are the basic of principal component analysis and finally the result are grouped in dendrogram of hierarchical clustering. Each of the defect bearings and healthy bearing has their own vibration trend when detail inspection is observed. By applying principal component analysis, the difference can be seen due to the presentation of the data using colors in a three dimensional plotting and finally clustering where the data remain in their own clusters without intertwining with other data. From the analysis, a monitored or tested system with all three variation of speed, results in a definite yet clear separation in principal component analysis and cluster of the defect bearings and good condition bearing all without any outliers. All of this shows that statistical analysis is effective in differentiating a multiple defect bearing with a good condition bearing at all three variation of speed of 440 RPM, 1480 RPM and 2672 RPM that tested in this research. Clear differentiation of the bearings in any running speed has made it possible for online monitoring without interrupting the machine for checking and maintenance purposes.

5.2 RECOMMENDATION

For further continuation of this project, many aspect of the experiment can be improved, such as making use of a one-stop instrumentation system for both data acquisition together with signal analysis. This is believed able to simplify the monitoring system by both time and cost aspects.

Another suggestion to make here is such technique can expand its usage to other types of bearings in order to verify its wide utility in condition monitoring field.

REFERENCES

Albrecht P.F., McCoy R.M. and Owen E.L. 1986. Assessment of the reliability of motors in utility application. *IEEE Transactions on Energy Conversation*. **EC-1**: 39-46.

Alexander, H. and Daniel, A.K. 2000. *Tutorial: Clustering Techniques for Larges Data Sets*.

Amarnath M., Shrinidhi R., Ramachandra A., Kandagal S.B. 2004. Prediction of Defects in Antifriction Bearings using Vibration Signal Analysis. *IE(I) Journal-MC*. pp. 88-92.

Baydar N., Chen Q., Ball A., Kruger U. 2001. Detection of incipient tooth defect in helical gears using multivariate statistics. *Mechanical Systems and Signal Processing*. **15**(2): 303–321.

Bellino A., Fasana A., Garibaldi L., Marchesiello S. 2010. PCA-based detection of damage in time-varying systems. *Mechanical Systems and Signal Processing*. **24**(7): 2250–2260.

Benaicha A., Guerfel M., Bouguila N., Benothman K. 2010. New PCA-based methodology for sensor fault detection and localization. *8th International Conference of Modeling and Simulation, MOSIM'10*. Tunisia.

Bhave, S. 2010. *Mechanical Vibration*. India: Pearson Education India.

Bolaers F., Cousinard O., Marconnet P., and Rasolofondraibe L. 2004. Advanced detection of rolling bearing spalling from de-noising vibratory signals. *Control Engineering Practice*. **12**: 181-190.

Braun, S. 1980. The signature analysis of sonic bearing vibration. *IEEE Transactions on Sonics and Ultrasonics*. **27**: 317-328.

Butler, D.E. 1973. The shock pulse method for detection of damaged rolling bearings. *NDT International*. pp. 92-95.

Choi, S.W. and Lee, I.B. 2004. Non-linear dynamic process monitoring based on dynamic kernel PCA. *Chemical Engineering Science*. **59**: 5897–5908.

Ding S., Zhang P., Ding E., Yin S., Naik A., Deng P., Gui W. 2010. On the Application of PCA Technique to Fault Diagnosis. *Tsinghua Science And Technology*. **15**(2): 138-144.

Downham, E. 1980. Vibration monitoring and wear prediction. *Proceeding of 2nd International conference on vibration in rotary machinery, IMechE*. pp. 29-33.

Dyer, D. and Stewart, R. 1978. Analysis of rolling element damage by statistical vibration analysis. *Journal of Mechanical Design*. **100**: 229-235.

Ericsson S., Niklas G., Elin J., Persson L., Ronny S., Stromberg J. 2005. Towards automatic detection of local bearing defects in rotating machines. *Mechanical Systems and Signal Processing*. **19**(3): 509– 535.

FAG Bearing Corporation. 2002. *Rolling Bearing Lubrication*. Publication number WL81115/4EA.

Felten, D. 2003. *Understanding Bearing Vibration Frequencies*, Mechanical Field Service Department, L&S Electric, Inc.

Gero and David. 1997. *Aviation Disasters*. Yeovil: Patrick Stephens Ltd (Haynes Publishing).

Ghafari, S.H. 2007. *A Fault Diagnosis System for Rotary Machinery Supported by Rolling Element Bearings*. University of Waterloo, Canada.

Hemmings R.C. and Smith, J.D. 1976. Information from bearing vibration. *IMechE*. pp.117-122.

Hirani, H. 2000. *Failure of Rolling Bearings: An Overview*. India: Indian Institute of Technology Bombay.

Ho, D. and Randall, R.B. 2000. Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mechanical Systems and Signal Processing*. **14**(5): 763-788.

Huang, B. 2001. Process identification based on last principal component analysis. *Journal of Process Control*, **11**: 19-33.

Igarashi, T. and Hamada, H. 1989. Studies on the vibration and sound of defective rolling bearings (first report: vibration of ball bearings with one defect). *JSME*. **25**: 248-253.

Ingarashi T., Noda B., and Matsushima E. 1980. A study on the prediction of abnormalities in rolling bearing. *JSLE*. **1**: 71-76.

Jianfang J., Yuemping S., and Qi Z. 1989. A method of bearing vibration signal extraction and fault detection. *Proceeding of 1st International Machinery Monitoring and Diagnostics Conference*. pp.149-154.

Jorgensen, A. 2000. *Clustering Excipient near Infrared Spectra Using Different Chemo metric Methods*. University of Helsinki.

JTEKT Corporation. 2009. *Koyo® Ball & Roller Bearings: Failures, Causes and Countermeasures*. Publication CAT.NO.B3001 E.

Kadushin, D. 1991. Roller element bearing fault analysis using envelope detection during an experimental case study. *Proceeding of 3rd International Machinery Monitoring and Diagnostics Conference, COMADEM 91*. pp. 132-141.

Lacey, S.J. 2008. *An Overview of Bearing Vibration Analysis, Maintenance & Asset Management*. **23**: 32-48.

Liu, L. 2006. Robust fault detection and diagnosis for permanent magnet synchronous motors, PhD dissertation, College of Engineering, The Florida State University, USA.

Logan, D. and Mathew, J. 1996. Using the correlation dimension for vibration fault diagnosis of rolling element bearing-i. Basic concepts. *Mechanical Systems and Signal Processing*. **10**: 241-250.

Lu, S. and Li, M. 2007. Bearing Fault Diagnosis Based on PCA and SVM. *Proceedings of the 2007 IEEE International Conference on Mechatronics and Automation*.

Martin, H.R. and Honarvar, F. 1995. Application of statistical moments to bearing failure detection. *Applied Acoustics*. **44**(1): 67-77.

Martin, K.F. and Thrope, P. 1992. Normalized spectra in monitoring of rolling bearing elements. *Wear*. **159**: 153-160.

Mathew J. and Alfredson R.J. 1984. The condition monitoring of rolling element bearing using vibration analysis. *Vibration and Acoustics, Stress, and Reliability in Design*. **106**: 447-453.

McFadden, P. and Smith, J. 1984. Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration*. **96**: 69-82.

McFadden, P. and Toozhy. M.M. 2000. Application of synchronous averaging to vibration monitoring of rolling element bearings. *Mechanical Systems and Signal Processing*. **14**: 891-906.

McMohan, S.W. and Scott, T. 1991. Condition monitoring of bearing using ESP and an expert system. *Proceeding of 3rd International Machinery Monitoring and Diagnostics Conference, COMADEM 91*. pp. 165-182.

Milne R., Aylett J., McMahon S., and Scott T. 1991. Portable bearing diagnostics using enveloping and expert systems. *Proceeding of 3rd international Machinery Monitoring and Diagnostics Conference, COMADEM 91*. pp. 75-79.

Moniz L., Nichols J.M., Nichols C.J., Seaver M., Trickey S.T., Todd M.D., Pecora L.M., Virgin L.N. 2005. A multivariate attractor-based approach to structural health monitoring. *Journal of Sound and Vibration*. **283**: 295–310.

Nichols J.M., Seaver M., Trickey S.T. 2006. Damage detection using multivariate recurrence quantification an analysis. *Mechanical Systems and Signal Processing*. **20**: 421–437.

Osborne B. G., Fearn T. and Hindle P. H. 1993. *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*. United Kingdom: Longman Group.

Qayyuum, M.R.M. 2012. Bearing Fault Detection using Vibration Responds: A Statistical Based Analysis. Degree Thesis. Universiti Malaysia Pahang, Malaysia.

Radu, C. 2010. *The Most Common Causes of Bearing Failure and the Importance of Bearing Lubrication*. RKB Technical Review.

Ramahaleomiarantsoa J.F., Nicolas H., Eric J., Razafimahenina J.M. 2012. *Performances of the PCA Method in Electrical Machines Diagnosis Using Matlab, MATLAB – A Fundamental Tool for Scientific Computing and Engineering Applications* . InTech Company.

Schaeffler Trading (China) Co.Ltd. 2012. *The Function of FAG Bearing Cage*.

Shiroishi J., Li Y., Liang S., Kurfess T., Danyluk S. 1997. Bearing condition diagnosis via vibration and acoustic emission measurements. *Mechanical Systems and Signal Processing*. **11**: 693-705.

SKF Bearing Corporation. 1994. *Bearing Failures and Their Causes*. Publication PI 401 E.

Smith, J.D. 1982. Vibration monitoring of bearings at low speeds. *Tribology International*. **15**: 139-144.

Smith, L.I. 2002. A tutorial on Principal Components Analysis.

Stronach A.F., Cudworth C.J., and Johnston A.B. 1984. Condition monitoring of rolling element bearing. *Proceeding of International Conference on Condition Monitoring*. pp. 162-177.

Spyridon, G.M. and Ioannis, C. 2009. Study and construction of an apparatus that automatically monitors vibration and wears in radial ball bearings which are loaded in radial direction. *2009 International Conference on Signal Processing Systems, IEEE Computer Society*. pp. 292-296.

Suri, G. 2010. Review of Techniques for Bearings and Gearbox Diagnostics, SpectraQuest, Inc.

Tandon, N. 1994. A comparison of some vibration parameters for the condition monitoring of rolling element bearings. *Measurement*. **12**: 285-289.

Tandon, N. and Choudhury, A. 1999. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*. **32**: 469–480.

Tandon, N. and Nakara, B.C. 1992. Comparison of vibration and acoustic measurement techniques for condition monitoring of rolling element bearings. *Tribology International*. **25**(3): 205-212.

Taylor, J. 1980. Identification of bearing defects be spectral analysis. *Journal of Mechanical Design*. **102**: 199-204.

The Timken Company. 2011. Timken Bearing Damage Analysis with Lubrication Reference Guide.

Trendafilova I., Cartmell M., Ostachowicz W. 2007. Vibration based damage detection in an air craft wing scaled model using principal component analysis and pattern recognition. *Journal of Sound and Vibration*. **313**: 560–566.

Varmuza, K. and Filzmoser, P. 2009. *Introduction to Multivariate Statistical Analysis in Chemometrics*. Austria: CRC Press.

Wang, Y. and Kootsookos, P. 1998. Modeling of low shaft speed bearing faults for condition monitoring. *Mechanical Systems and Signal Processing*. **12**: 415-426.

White M. F. 1984. Simulation and analysis of machinery fault signals. *Sound and Vibration*. **93**: 95-116.

Widodo A., Yang B.S., Han T. 2007. Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors. *Expert Sys Applications*. **32**: 299-312.

Williams T., Ribadeneira X., Billington S., and Kurfesss T. 2001. Rolling element bearing diagnostics in run-to-failure lifetime testing. *Mechanical Systems and Signal Processing*. **15**: 979-993.

Wowk, V. 1991. *Machinery Vibration: Measurement and Analysis*. New York: McGraw-Hill, Inc.

Zaiane, O. R. 1999. *Principles of Knowledge Discovery in Databases*. University of Alberta.

Zimroz, R. and Bartkowiak, A. 2012. Two simple multivariate procedures for monitoring planetary gearboxes in non-stationary operating conditions. *Mech. Syst. Signal Process*.