CHAPTER 8

FAULT DIAGNOSIS USING VIBRATION ANALYSIS

8.1 INTRODUCTION

There are several methods available to monitor the condition of the machines. Experts having more knowledge in the field by using sight and sound perform the machine condition monitoring. Vibration monitoring and analysis is most widely used techniques in condition monitoring applications. Intent of vibration analysis is to capture the pertinent vibration frequencies and amplitudes in order to create the trend for each machine over an extended usage of the period.

The condition of bearing needs to be monitored regularly to avoid unexpected catastrophic breakdown of the machinery. In order to determine whether the machinery is in good or bad condition, the various parameters like vibration, temperature, oil pressure, noise and performance is measured. If the condition of the machine is not good, then the monitoring makes it possible to determine the cause for the problem. The condition monitoring techniques are used in conjunction with predictive maintenance.

During the past few years, vibration-based techniques have been proved an effective method of detecting the faults in rotating machinery. The traditional condition diagnosis techniques used on general rotating machinery often fail when applied to reciprocating machinery, such as reciprocating compressors and diesel engines. This is because the signal measured in
reciprocating machinery, contains a strong noise component and its vibration levels are higher even during normal conditions. All these factors make the application of vibration-based techniques for reciprocating compressor is very difficult, as they are more effective for vibrations that exhibit stationary and periodic behaviour (Wang & Chen 2008).

8.2 MACHINE LEARNING METHODS IN MACHINE FAULT DIAGNOSIS

The pattern recognition technique was first implemented in the year 1989 to automate the fault diagnosis (Li & Wu 1989). The pattern recognition or machine learning technique classifies the group of objects on the basis of subjective similarity measures. Wang et al., (1989) proposed an automatic fault diagnosis system for ball bearing based on pattern recognition technique and time domain signal processing.

The fault identification using machine learning techniques has three phases. They are feature extraction, feature selection/reduction and feature classification. In feature extraction, statistical features, histogram features, wavelet features, etc., are extracted from the collected vibration data. From the extracted features, the prominent features are selected using feature selection techniques such as independent component analysis, principal component analysis, decision tree etc., in the second phase. The machine learning process has two stages in the third phase. In the first stage, the classification algorithms are trained with the help of various fault signal data. In the second stage, the trained algorithm is tested with the help of test data. The cross validation is one of the methods of training and testing process. The model is constructed with the help of part of the training data and remaining data is used for prediction by the trained model (Wold 1978).
It is understood from the literature Saimurugan et al., (2011) that the machine learning technique is highly helpful to automate the machinery fault diagnosis. The procedure for machine learning based fault diagnosis of a reciprocating machine in the present work is shown in Figure 8.1

![Flowchart](image)

**Figure 8.1 Flowchart for fault diagnostic procedure**

In this study, statistical feature have been used for extracting the fault information from the vibration signals. The number of features were reduced with the help of decision tree analysis and it will remove the redundant information data and reduce the computation time. Then the same decision tree and support vector machine classification algorithm were used for detection of machine faults. The support vector machine techniques are discussed in the following section.
8.3 EXPERIMENTAL SET UP

In order to capture the vibration signal, the piezoelectric accelerometer was mounted on top of the bearing housing. The bearing housing was machined in such a way that, the transducer has to be mounted closer to the bearing housing. The area selected for mounting the accelerometer is made flat and smooth to ensure effective coupling between the sensors and bearing housing. The speed sensor was used to capture the rotating speed of the compressor. The locations of the accelerometer and speed sensor are shown in Figure 8.2.

Figure 8.2 Experimental setup and accelerometer
8.4 DATA ACQUISITION SYSTEM

A piezoelectric accelerometer was mounted on the flat surface using the direct adhesive mounting technique for collecting the vibration data. The acceleration is one of the parameters to characterize vibration. The accelerometer gives the acceleration in the form of the voltage signal. The voltage signal is directly proportional to the acceleration of the machine. The accelerometer was connected to the signal-conditioning unit (LMS SCADAS Mobile is shown in Figure 8.3), where the acquired signals go through a charge amplifier and an analogue-to-digital converter. The vibration signal in digital form was fed to the computer through an Ethernet interface. The software LMS Test. Lab Rev 8A was used for recording the signals directly in the computer’s secondary memory. The signals are then read from the memory, replayed and processed to extract statistical features.

![Figure 8.3 LMS SCADAS Mobile –Signal conditioner unit](image)

The vibration signals are stored in the computer as *.LMS file format and the same has to be exported as MATLAB file format.

8.4.1 Specifications of Accelerometer and Speed sensor

The specification of the accelerometer is used in this experiment as follows:

Make : Bruel&KajaerInc
Model : 4513(-B)-001
Weight : 8.6 grams
Voltage Sensitivity (@ 160 Hz): 100 ±10% mV/g
Measuring Range (±pk) : 490 m/s²
Mounted Resonance Frequency: 32 kHz
Amplitude Response ±10% : 1 to 10000 Hz
Transverse Sensitivity : < 5 %
Temperature range : -51 to 100°C
Max. Operational Shock (peak): 5000 g pk

The specification of speed sensor used in this experiment is given below:

Make : Monarch Instrument, USA
Operating Envelope : Up to 3 feet and 45° from reflective target
Type : ROS-P Remote optical sensor
Speed range : 1 to 250,000 RPM
Illumination Source : Visible red LED, 5Vdc @ 30 mA
Operating Temp : -10° C to 70° C
Output Signal : 5 to 0 Vdc TTL compatible pulse
On-Target Indicator : Green LED on end cap
Material : 303 Stainless steel
Lens : Acrylic plastic
8.5 EXPERIMENTAL PROCEDURE AND FAULT SIMULATION

Initially, the vibration was measured with good bearing from the accelerometer mounted nearer to the non-drive end side of the bearing housing. During testing, the rotational speed of 982 RPM and working pressure of 10 bar was maintained. In order to simulate the bearing fault, a small cut was made to produce a defect of 4X0.9 mm in the outer raceway as shown in Figure 8.4. As part of the experimental measurement, the vibration signal was captured with good and faulty bearings with the same operating condition.

![Defect created in the outer raceway](image)

**Figure 8.4 Faulty bearing**

In the experimental study, the sampling frequency of 25000 Hz and sample length of 8192 was selected. The statistical analysis of the data is meaningful if the number of samples is sufficiently large. But if the number of the sample increases, computation time will get increased. The 10000 sample length can be selected for statistical measures. In the wavelet feature extraction, it is preferable to choose a sample length in the order of $2^n$. The sample length of 8192 was chosen which is near to 10000. The vibration signals from the piezoelectric pickup were taken after allowing some initial
run and then stored in LMS file format in a computer through the Ethernet interface.

8.6 VIBRATION SIGNAL OF GOOD AND FAULTY BEARING

The vibration signal measured in the normal condition is as shown in Figure 8.5 and the signal measured with a faulty bearing at the same location with the same operating condition is shown in Figure 8.6. The acceleration magnitude is high with the faulty bearing when compared with good one. The size of the defect created artificially in the bearing outer race is slightly higher than the spalling or flaking defects of the bearing. Therefore, the magnitude of the defective signal is higher than the good one. The huge vibration data acquired from the rotating machine cannot be directly used for machine learning. The required fault information is extracted from these huge data in the form of features such as statistical features, histogram features, wavelet features etc.

![Figure 8.5 Vibration signal measured with good bearing](image)
8.7 STATISTICAL FEATURES

The sound and vibration signals from the rotating machine carry lot of information about the various fault conditions. The statistical analysis of the acquired vibration signals for various fault conditions gives statistical parameters. These parameters bring out the information about the various faults from the time domain signals. These parameters are called statistical features, which are used for detection of faults. The eleven statistical features have been chosen for this analysis. They are mean, median, standard error, standard deviation, sample variance, kurtosis, skewness, sum, maximum, minimum and range. The eleven statistical parameter values are extracted from good and fault signal is presented in Table A 4.1 in Appendix 4. The statistical features yield better classification accuracy in fault diagnosis of rolling element bearings (Kankor 2011b). These statistical features are explained below.
(a) Standard error: Standard error is a measure of the amount of error in the prediction of y for an individual x in the regression, where \( \bar{x} \) and \( \bar{y} \) are the sample means and ‘n’ is the sample size.

\[
\text{Standard error of the predicted } y = \sqrt{\frac{1}{(n-2)} \left[ \sum (y - \bar{y})^2 \right] - \left[ \sum \left( x - \bar{x} \right) \left( y - \bar{y} \right) \right]^2 \left( x - \bar{x} \right)^2} \quad (8.1)
\]

(b) Standard deviation: This is a measure of the effective energy or power content of vibration signals. The following formula is used for computation of standard deviation.

\[
\text{Standard deviation} = \sqrt{\frac{\sum (x - \bar{x})^2}{(n-1)}} \quad (8.2)
\]

(c) Sample variance: It is the variance of the signal points and the following formula is used for computation of sample variance.

\[
\text{Sample variance} = \frac{\sum (x - \bar{x})^2}{(n-1)} \quad (8.3)
\]

(d) Kurtosis: Kurtosis indicates the flatness or the spikiness of the signal.

\[
\text{Kurtosis} = \left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left( \frac{x - \bar{x}}{s} \right)^4 \right\} - 3\left( \frac{n-1}{n-2} \right)^2 \left( \frac{n-3}{n-2} \right) \quad (8.4)
\]

Where ‘s’ is the sample standard deviation.
(e) Skewness: Skewness characterizes the degree of asymmetry of a distribution around its mean. The following formula was used for computation of skewness.

\[
\text{Skewness} = \frac{n}{(n-1)} \sum \left(\frac{x_i - \bar{x}}{s}\right)^3
\]

(f) Range: It is the difference between maximum and minimum signal point values for a given signal.

(g) Minimum value: It refers to the minimum signal point value in a given signal.

(h) Maximum value: It refers to the maximum signal point value in a given signal.

(i) Sum: It is the sum of all signal point values in a given signal.

(j) Mean: It refers to the arithmetic mean of the signal point values for a given signal.

(k) Median: It refers to the middle value of the ordered signal point values in a given signal.

These eleven statistical features were extracted from vibration signals. Among these eleven features, the predominant features that will have the high discriminating ability are elicited with the help of dimensionality reduction techniques.

In this work, the decision tree algorithm was used as dimensionality reduction techniques. These reduced features were given as an input to the
various classifiers for the classification of different fault conditions. The classifiers employed in this study are decision tree and support vector machine. These classifiers and data reduction techniques have been described in the below sections.

8.8 DECISION TREE

Decision tree represents the information in the signal as features, in the form of a tree. The classification is done through the decision tree with its leaves representing the different fault classes. The sequential branching process ending up with the leaves here is based on conditional probabilities associated with individual features. C4.5 algorithm introduced by Quinlan (1993) is one of the widely used algorithms to generate decision tree. The eleven statistical features were given as an input to C4.5 algorithm. J48 algorithm (A WEKA implementation of C4.5 algorithm) is widely used in the decision tree process (Saimurugan et al., 2011). The decision tree starts with a root node and branches out to a number of nodes, then finally ends with the leaf node. The nodes in the decision tree specify the attributes and the branches indicate attribute values. The leaf node in the tree represents the set of features and its class, which is not partitioned further. In the decision tree, the nodes are arranged from root node to a leaf node in the descending order of importance of attributes.

8.8.1 Application of Decision tree

The extracted statistical features were given as an input to the decision tree algorithm. The Figure 8.7 shows the result of a decision tree process for the outer race fault condition of vibration signals at 982 rpm.
Figure 8.7 Decision tree

The root node attribute contains more information about the various classes and the remaining attributes in the internal nodes are positioned in the order of their information gain. The numbers within the parenthesis of the leaf node indicates the contribution of feature set for identification of the corresponding class. In this fault class, the standard error is the prominent feature for classification of faults. The classification efficiency is 99.3671% and correctly classified instances are 157 out of 158 instances. The confusion matrix is given as follows.

The classification result for the speed of 980 rpm is presented in the form of confusion matrix and the same is shown in Table 8.1.

Table 8.1 Confusion matrix from DT results

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$b$</th>
<th>Classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>79</td>
<td>0</td>
<td>Good</td>
</tr>
<tr>
<td>$b$</td>
<td>1</td>
<td>78</td>
<td>Bad</td>
</tr>
</tbody>
</table>

The interpretation of the confusion matrix is as follows:
• The diagonal elements in the confusion matrix show the number of correctly classified data points. In the first row, the first element shows the number of data points that belong to ‘a’ (good bearing) and classified by the classifier as ‘a’ class.

• In the second row, the first element shows one data point belongs to ‘b’ misclassified as ‘a’ (good bearing) class.

8.9 SUPPORT VECTOR MACHINE (SVM)

In this work, support vector machine was used as a classifier to compare results with decision tree results. SVM is based on statistical learning theory and it is originally developed by Vapnik (Vapnik 1995, Cortes 1995). SVM belongs to a class of supervised learning algorithm, which constructs an optimal hyper plane for linearly separable patterns to classify the data into two categories, and it extends to patterns that are not linearly separable by transformations of original data to map into a new space with the help of kernel functions. The kernel functions such as linear, polynomial, sigmoid and radial basis function can be used in SVM.

The decision tree results have been verified by using SVM techniques with the help of DTREG software as given in the software manual (Sherrod 2003).

DTREG builds classification and regression decision trees, neural networks, support vector machine, Group method of data handling polynomial networks, gene expression programs, discriminant analysis and logistic regression models that describe data relationships and can be used to predict values for future observations.

One of the most useful application of statistical analysis is the development of a model to represent and explain the relationship between
variables. DTREG accepts a dataset containing of number of rows with a column for each variable. One of the variable is “target variable” whose value has to be modelled and predicted as a function of the “predictor variables”.

There are two types of methods for analysing and modelling the data. One is supervised learning and other is unsupervised learning. Supervised learning requires input data that has both predictor (independent) variables and a target (dependent) variable whose value is to be estimated. By various means, the process learns how to model (predict) the value of the target variable based on the predictor variables. Decision trees, regression analysis and neural networks are examples of supervised learning.

Eleven statistical features were extracted with the sample length of 8192 for each data set. Total dataset of 158 were used for training the algorithm. The classification process has two stages: training and testing. Training is the process of learning to label from the examples. Testing is the process of checking how well the classifier has learnt to label the unseen examples. There are two main ways of classification of data. The one way is to train the algorithm by passing all the dataset with their class and test the trained algorithm by sending only the particular class dataset for identification of the class to which the dataset belongs. In the next one, input all the dataset with their class to the algorithm. The algorithm gets trained with the dataset and does the cross fold validation with the help of the same dataset. In this work, 10 times cross fold validation was used for the classification process.

The selected feature (Standard error) was given as an input to the SVM and classification efficiency is 99.367 %. This is calculated from the confusion matrix for the speed of 982 rpm as shown in Table 8.2.
The number of correctly classified data set: 157
Incorrectly classified dataset : 1
Total number of data set : 158
Classification efficiency : 157/158= 99.367%

Table 8.2 Confusion matrix from SVM results

<table>
<thead>
<tr>
<th>Category</th>
<th>Good</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>79</td>
<td>0</td>
</tr>
<tr>
<td>Bad</td>
<td>1</td>
<td>78</td>
</tr>
</tbody>
</table>

The above result (99.367%) was compared with DT result(99.367%) and found there was no deviation.

The possibilities of using SVM are explored in this work and used to detect overall defects in the compressor. In order to determine the exact fault in the components of the bearings, number of defects and its combinations have to be simulated practically to train the SVM. To avoid using complex SVM techniques to find out the exact faults in the bearing components, another technique called adaptive noise cancelling method was explored.
8.10 BEARING FAULT DIAGNOSIS BY ADAPTIVE FILTERING TECHNIQUES

8.10.1 Introduction to Adaptive filter and LMS algorithm

There are several methods used for detecting the bearing defects from the vibration measurement like time domain analysis, frequency domain analysis, enveloping, cepstrum techniques, time frequency analysis, higher order spectral analysis, Haar transform and adaptive noise cancellation. In case of severe noise applications or when the signal is masked by the noise or vibration from other sources, the advanced signal processing techniques have to be deployed to detect the faults. A parameter used to determine the contribution of original vibration signal and the noise is the signal to noise ratio. Time domain and Frequency domain are usually not effective in the study of vibration signal when the signal to noise ratio is poor. In order to improve the signal to noise ratio some signal processing techniques have to be employed. In order to improve the diagnostic ability and overall detection of conventional methods like time domain analysis, Spectrum analysis and cepstrum analysis in the presence of severe noise an Adaptive Noise Cancellation (ANC) techniques can be implemented. In this method, the noise canceller output is fed to the adaptive filter and the filter weights are adjusted through an LMS adaptive algorithm to minimize the total output power of the system (Widrow et al., 1975). ANC techniques require minimum two inputs. A primary input containing the corrupted signal and a reference input containing noise correlated in some unknown way with the primary noise and the second input reference signal greatly depends on the location of the vibration sensor. The reference input is adaptively filtered and subtracted from the primary input to obtain the signal estimate.
The LMS algorithm is simple, robust and most popular algorithm for adaptive filtering applications (Alexander1986, Samuel & David 1996). Its simplicity and low computational complexity make this algorithm an attractive solution for many practical problems (Sakai&Hinamoto2005). As the signal continues into the filter, the adaptive filter coefficients adjust themselves to achieve the desired results, such as identifying an unknown filter or cancelling noise in the input signal. Many adaptive filtering algorithms are iterative methods for solving error minimization problems. The most commonly used adaptive systems are those based on LMS adaptive algorithm (Huaqing & Peng 2008). The flow chart for fault diagnosis by adaptive filtering is shown in Figure 8.8.

![Figure 8.8 Flowchart for fault diagnostics by adaptive filtering](image-url)


8.10.2 Adaptive Noise Cancelling

A signal $S$ corrupted with noise $N_o$ is received at the primary sensor while the bearing runs under faulty conditions as shown in Figure 8.9. The signal $S$ is the fault signal from the bearing and $N_o$ is the noise of the normal bearing. A reference noise $N_1$, which is related to noise $N_o$ in some unknown way, but not correlated with signal $S$, is received by the reference sensor in normal state. The filter output ‘$y$’ is then adaptively filtered to match $N_o$ as close as possible. Then the filter output is then subtracted from the primary input $S+N_o$ to produce the system output $\varepsilon$ (Yimin Shao & KikuoNezu 2005).

\[ \varepsilon = S + N_o - y \]  
\[ \varepsilon^2 = S^2 + (N_o - y)^2 + 2S(N_o - y) \]  
\[ \text{Taking expectation of both sides of the Equation 8.7} \]

\[ E(\varepsilon^2) = E(S^2) + E((N_o - y)^2) \]  
\[ Y = W^TX, \]

\[ E(\varepsilon^2) = E(S^2) + E((N_o - W^TX)^2) \]
Where $W$ is the weight vector and $X$ is input vector. The signal power $E(S^2)$ will be unaffected as the filter weights are adjusted to minimize $E(\varepsilon^2)$:

$$\min E(\varepsilon^2) = E(S^2) + \min E((N_o - W^T X)^2) \quad (8.10)$$

For optimal set of filter weights, the output $y$ is the best least-square estimate of primary noise $N_o$. From the following equation

$$\varepsilon - S = N_o - y \quad (8.11)$$

The adjustment of the filter weights to minimize the output power causes $\varepsilon$ to be best least square estimate of the signal $S$.

The ANC filter is easily implemented by computer software. The implemented method is same as that LMS algorithm. This algorithm is shown by the following equation.

$$W_{j+1} = W_j + 2\lambda \varepsilon_j X_j \quad (8.12)$$

Where $W_j$ is the weight vector at the $j^{th}$ instant of time, $X_j$ is the reference input vector at the $j^{th}$ instant of time, $\varepsilon_j$ is the error signal at the $j^{th}$ instant of the output of the noise canceller, and $\lambda$ is the gain constant that regulates the speed and stability of adaptation. To ensure the convergence of the adaptive algorithm, the value of $\lambda$ should be chosen such that

$$0 < \lambda < 1/(L+1) \quad \text{(Signal power)}$$

Where $L$ is the index of the last filter weight.

### 8.11 NOISE CANCELLING BY ADAPTIVE FILTERING

The signal captured with a faulty bearing as the test signal to be diagnosed and the signal detected with good bearing at the same location with the same working condition as the reference signal. The signal measured with a faulty bearing as shown in Figure 8.10 is corrupted with noise. There are
impulses in the vibration signal of the normal state because of the reciprocating mechanism. The effect of noise signal in the signal measured for the diagnosis is stronger than the fault signal from the bearing flaw. The signal shown in Figure 8.11 is filtered from the source signal through adaptive noise cancelling method. To remove the noise from the measured signal and to perform the FFT analysis, a code has been developed in MATLAB software and it is given in Appendix 5.

Figure 8.10 Signal measured with faulty bearing

Figure 8.11 Signal after noise cancelling
8.11.1 Fast Fourier Transform (FFT) spectrum of filtered signal

When the bearing spins, any irregularities in the race way surfaces or in the roundness of the rolling elements excite periodic frequencies called fundamental defect frequencies. These are

1. **FTF** – Fundamental Train Frequency (Frequency of the cage)

2. **BSF** – Ball Spin Frequency (circular frequency of each rolling element as it spins)

3. **BPFO** – Ball Pass Frequency of the Outer race (frequency created when all the rolling elements roll across the defect in the outer race)

4. **BPIO** – Ball Pass Frequency of the Inner race (frequency created when all the rolling elements roll across the defect in the inner race)

Fundamental defect frequencies depend upon number of rolling elements, the shaft rotation speed, ball diameter, pitch diameter, and the contact angle. Therefore, an accurate rotation speed is necessary to perform the analysis. In order to find the faulty elements of the bearing, the corresponding pass frequencies are calculated by using the below formulae and compared with FFT spectrum of the filtered signal shown in Figure 8.12.

Bearing Frequencies Formulae:

1. \( BPFO = \frac{z}{2} (N/60) (1-B_d/P_d) \cos \beta \)

2. \( BPFI = \frac{z}{2} (N/60) (1+B_d/P_d) \cos \beta \)

3. \( BSF = \frac{z}{2} (N/60) (1-[B_d/P_d]^2) \cos \beta \)

4. \( FTF = \frac{1}{2} (N/60) (1-B_d/P_d) \cos \beta \)
Where:

\( z \) - Number of balls or rollers

\( B_d \) - Ball diameter (mm)

\( P_d \) - Bearing Pitch diameter (mm)

\( \beta \) - Contact angle, ball to race

\( N \) - Rotational speed (RPM)

In this case the outer race defect is simulated in 6211 bearing, therefore

\[
BPFO = \left( \frac{10}{2} \right) \left( \frac{982}{60} \right) \left( 1 - \frac{14.287}{77.6725} \right) (\cos 0)
\]

\( \beta = 0 \) for deep groove ball bearing

=66.78 Hz

The harmonics of the above frequency are 133.56 and 200.34 etc.,

**Figure 8.12 FFT Spectrum of the filtered signal**
FFT spectrum of the filtered single is shown in Figure 8.12 exhibits many harmonics of the BPFO. Therefore, the adaptive filtering techniques are more suitable for reciprocating compressor bearing fault detection.

8.12 CONCLUSION

In this chapter, fault diagnosis of rolling element bearing in reciprocating compressor was explored by using vibration signal. Since vibration signals in the reciprocating compressor contains stronger noise signal than fault signal, diagnosis of fault is more difficult than other rotating machines. Therefore, it was decided to extract the statistical features from the vibration signals with good and bad bearing to detect the faults. The eleven extracted statistical features were reduced with the help of dimensionality reduction techniques such as decision tree. The technique with its optimum number of features was chosen using a decision tree algorithm. SVM methods were used to validate the classification efficiency of the DT algorithm. These methods can be used to detect the faults in the compressor in general. In order to detect the exact faulty elements, an alternate method called adaptive noise cancelling has been explored. The above method is most useful for the new designers to foresee any deterioration of the bearing components during testing and field trial.