Indian Journal of Science and Technology, *Vol 9(34), DOI: 10.17485/ijst/2016/v9i34/100957, September 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

On-Road Testing of A Vehicle for Gearbox Fault Detection using Vibration Signals

M. Saimurugan*, T. Praveenkumar, B. Sabhrish, P. Sachin Menon and S. Sanjiv

Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham University, Amrita University, Coimbatore – 641112, Tamil Nadu, India; m_saimurugan@cb.amrita.edu,praveent.mtp@gmail.com,sabhrish1395@gmail.com, sachin.menon94@gmail.com, sanjivsubramanian010@gmail.com

Abstract

Gearbox is one of the most important components in an automobile, enabling power transmission from the engine to the wheels. Gears and bearings are prone to failure. The impending case of failure can be predicted by performing vibration analysis of a gear box, usually done by acquiring data in lab conditions. **Objective:** This paper proposes an idea to enable fault detection in the gearbox by acquiring data under on road conditions without having to remove the gearbox, thereby simplifying the condition monitoring of a gearbox. **Methodology:** The experimental studies were conducted on the gearbox in a test vehicle run in real time conditions and the vibration data from the gearbox was acquired using a piezoelectric accelerometer for different conditions of gearbox. The acquired time domain data was normalized and its statistical features were extracted. The classification of the fault class was done by using decision tree (J48) algorithm. **Findings:** Classification efficiencies as high as 99% were obtained by using decision tree algorithm. Further, normalization of raw data was found to increase the efficiency of the classifier. This observation can be used to make decision trees more efficient. **Improvements:** This paper has highlighted the concept of on road testing for two fault conditions. Further research work on other fault conditions can be done.

Keywords: Decision Tree, Gearbox, Normalization, On-road Testing, Statistical Features, Vibration

1. Introduction

Everyday technological advances are being made in the field of automobile engineering. The most common need of a customer in the case of an automobile is a vehicle with long life and less maintenance. Although maintenance cannot be completely minimized, the best possible solution is to predict the impending failure so that adequate maintenance, if possible, can be done. Hence predictive analysis is an important technological advancement. Condition monitoring is a form of predictive analysis which uses data acquired from the component to predict failure.

A gearbox consists of mechanical components that are essential in power transmission. The main components that can fail in a gearbox are the gears and bearings. Various bearing faults have been simulated previously and their vibrations were used to simulate fault signals which can be used for prognostic algorithms. Propagation of cracks in the outer race of the bearing and tooth chipping in gears are the common faults that occur in a gearbox. Studies showing the effect of fault in the outer race and their effects on the gear have been done and was validated using experiments[1,2]. The effect of tooth chipping due to excessive load and the resulting effect of fault on gear mesh stiffness were also studied by analysing the vibrations and an analytical model was developed to represent a time varying gear mesh stiffness. Comparison between the analytical and finite element model was done[3]. Fault diagnosis were done on both helical and spur gearboxes by using multimodal feature from the obtained vibrations and high efficiencies were achieved using multimodal deep support vector classification using fault classes such as faulty bearing and faulty gear[4].

**Author for correspondence*

Vibrational analysis has been found to give consistent and accurate results. Conditional monitoring of a gearbox used in wind turbine was performed by combining vibrational and acoustic signals and worst case standard deviation for the signals were calculated to avoid false alarms and the results can be used to develop a correlation between health of machinery and the vibrations at a given time[5]. Vibrational analysis for fault diagnosis of a gearbox by using discrete spectrum correction has been done and its effectiveness was validated using simulation and experiments[6]. Time domain based diagnostic algorithm have been developed for improving the efficiency of extracting information from the vibrations and also to determine the position of fault[7]. Vibration signals along with decision tree classifiers have been used to monitor the health of bearings and high accuracy of 95.64 % has been achieved[8] and the kurtosis values were used to detect the presence of gear tooth cracks and the statistical model developed was found to be efficient[9].This paper uses vibrational analysis for predicting the most common bearing and gear faults. The vibration signals from the gearbox are acquired and the raw data is normalised. Features were extracted from the data and statistical feature extraction was preferred. Statistical features are frequently used for the condition monitoring of hydraulic brakes and Self Aligning Carrying Idler and high accuracy of 99 % was obtained[10,11]. Statistical features of the acquired vibration signals have been used to strengthen the predictive analysis of lifetime of bearings[12].

The extracted features were used to train a classifier. Decision tree algorithm is a popular method used for classification of vibration signals. Decision tree is a widely used classifier due to its simplicity and a lot of work has been done on improving the attribute selection process by avoiding overfitting and complexity[13]. Various Decision tree algorithm have been used to detect misfire in the engine and their accuracies were also compared and 100% accuracy was achieved by using Linear model tree[14]. Fault diagnosis of gearbox was done by using decision tree for classification and high efficiency was obtained[15]. The condition monitoring of a monoblock pump using vibrational analysis and decision tree has produced good results[16]. Fault diagnosis of a single point cutting tool by using J48 algorithm resulted in 90% efficiency[17]. Decision tree was preferred owing to its simple operation and less computational time. It was observed that the condition monitoring of gearbox was conventionally performed under lab conditions. This paper deals with the condition monitoring of an automobile gearbox based on the data acquired under real time test conditions.

The main contribution of this paper is as follows: Vibration signals from the gear box of a test vehicle run on real time conditions were acquired using a piezoelectric accelerometer. The data was obtained for different conditions as shown in Table 1.The data was normalised and the prominent statistical features were extracted. Classification based on different fault conditions was done using decision tree algorithm.

In chapter 2, the experimental procedure to obtain the vibration signals and data normalisation is discussed. In chapter 3, statistical feature extraction process is explained. Chapter 4 pertains to feature classification using decision tree algorithm. Chapter 5 describes the results inferred from the experiment. Chapter 6 conveys the conclusion drawn from the experiment. Figure 1 shows the methodology followed.

2. Experimental Studies

The experiment was performed on the gearbox of the test vehicle. The test vehicle used was Hyundai Santro Figure 2. The gearbox is a 5 speed manual synchromesh

Table 1. Different conditions and fault classes

Condition	Fault class
Good gearbox	None
Faulty gearbox	Crack on the inner race of the bearing
Faulty gearbox	Tooth chipping on the gear.

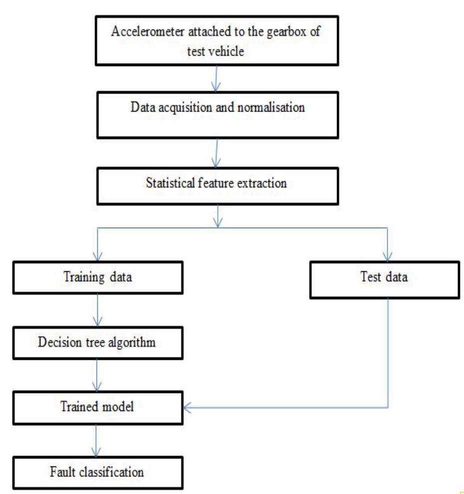

Figure 1. Methodology flow chart.

gearbox with helical gear. The gear ratios are mentioned in Table 2.

Various test conditions of the gearbox are shown in Table 3. For each combination of gear and speed, three conditions of the gearbox namely good gearbox, faulty bearing and faulty gear were chosen. Faults were manually induced in the bearing and gear.

A crack was induced in the outer race of the bearing using Electric Discharge Machining (EDM) technique as shown in Figure 3. Tooth chipping was performed on a tooth of the 3rd gear by grinding operation as shown in Figure 4.

The accelerometer was fitted to the bottom surface of the gearbox as shown in Figure 5. The test road was devoid of potholes, speed breakers and other obstructions.

For each combination of gear, speed and gearbox condition, vibration signals were acquired for 100 seconds

Table 2. Specifications of gearbox

Gears	Gear Ratio
First	3.25 : 1
Second	2.05 : 1
Third	1.37 : 1
Fourth	1.03 : 1
Fifth	0.84 : 1
Reverse	3.58 : 1

Figure 2. The test vehicle.

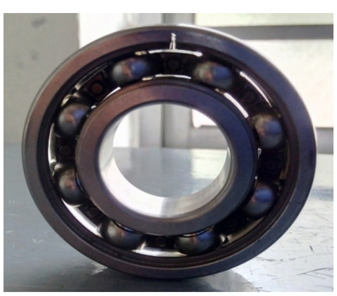

Figure 3. Crack induced in the outer race of the bearing.

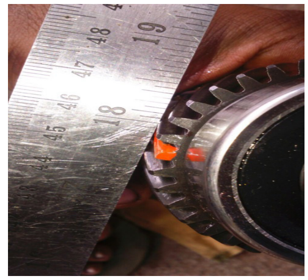

Figure 4. Tooth chipping induced in the gear.

Figure 5. Accelerometer fitted to the bottom of gearbox.

from the gearbox when the test vehicle was run on the test road. The data acquired had a sampling rate of 3200 per second. Hence, 320000 datum points were obtained for each test condition. The total number of test conditions were 30 and 9600000 datum points were obtained (Table 3).

The data acquired was then normalised. Normalization is done to standardize or limit the range of the input data. It is done to make the processing quicker, to prevent the output from skewing towards a particular data set due to higher range and also to avoid a localoptima i.e. to avoid the possibility of classifier providing the solution with respect to a neighbouring data set rather than providing a solution with respect to all data (global optima).

$$\mathbf{x}' = \frac{\mathbf{x} - \min(\mathbf{x})}{\max(\mathbf{x}) - \min(\mathbf{x})} \tag{1}$$

where, x is the time-domain datum and x' is the normalised datum. Normalized time domain data is shown in Figure 6. Similarly, raw data before normalization is shown in Figure 7.

3. Feature Extraction

It is difficult to process and interpret from the huge data acquired from the sensor. Hence feature extraction technique was used. Feature extraction reduces the number of data to be monitored thus reducing the process

Table 3. Test conditions

Gear	Speed [kmph]	Condition
1	10	Good Gearbox
		Faulty Bearing
		Faulty Gear
	20	Good Gearbox
		Faulty Bearing
		Faulty Gear
2	20	Good Gearbox
		Faulty Bearing
		Faulty Gear
	30	Good Gearbox
		Faulty Bearing
		Faulty Gear
3	30	Good Gearbox
		Faulty Bearing
		Faulty Gear
	40	Good Gearbox
		Faulty Bearing
		Faulty Gear
4	40	Good Gearbox
		Faulty Bearing
		Faulty Gear
	50	Good Gearbox
		Faulty Bearing
		Faulty Gear
5	50	Good Gearbox
		Faulty Bearing
		Faulty Gear
	60	Good Gearbox
		Faulty Bearing
		Faulty Gear

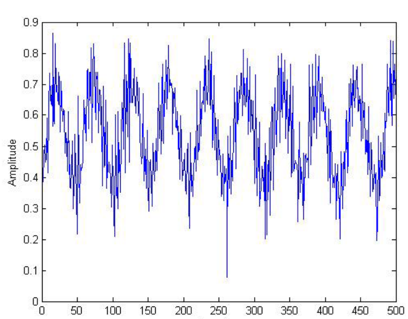

Figure 6. Normalized Data.

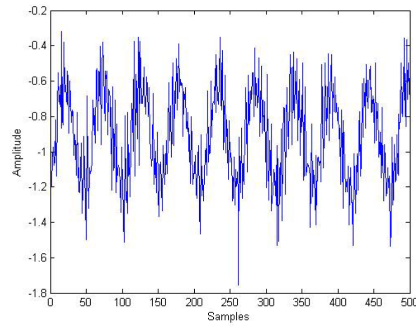

Figure 7. Data without normalization.

time. Feature extraction helps to remove the factors that are repetitive and eliminates the irrelevant data. Statistical features are used to represent a large quantity of data by meaningful factors with minimal loss of information. Eight statistical features namely variance, skewness, crest factor, root mean square (RMS), kurtosis, RMS*kurtosis, peak value and standard error were extracted. The description of the extracted features is mentioned in Table 4.

where

s = Standard deviation

$\bar{y}$ = Mean of the sample

n = Number of values in the sample.

y = time domain datum.

4. Feature Classification

Classification is the process of categorizing a new data point into its respective sub population. Decision tree is a commonly used classifier. The decision tree algorithm is simple, non-parametric, computationally faster and easy to interpret. Decision tree classification has been done in vibration based gearbox fault diagnosis and was found to give good results. Decision tree has a great potential in deriving the rules from the feature set[18]. Decision tree reduces the domain knowledge required by identifying the relevant features[19]. Decision tree consists of number of branches starting from a root and ending in number of leaves. This tree like structure is induced with nodes wherein each node acts as an input feature. Here the root represents the feature with the highest weight and the leaf represents the irrelevant features or the features with the least weightage. Figure8 shows a part of the decision tree formed by using raw data. In this decision tree kurtosis has the highest weightage and the weightage decreases as we go down the tree. Similarly Figure 9 shows a part of the decision tree formed using normalized data. Variance

Table 4. Extracted features

Feature Description	Formula
RMS is the square root of the mean of the squares of the sample.	$\sqrt{\frac{\sum_{i=1}^{n} y_i^2}{n}}$
Kurtosis provides a description about the distribution of the samples. It indicates whether the data is flat or reaches a peak with respect to a normal distribution.	$\sum_{i}^{n} \frac{(y_i - \bar{y})^4}{n.s^3}$
Crest factor is the ratio between the peak value and the RMS value. This is a dimensionless quantity which is used to represent the factor by which the peaks are greater than the RMS. It is to be noted that crest factor cannot be less than 1.	$\frac{y_{peak}}{y_{rms}}$
Skewness helps in description of the distribution of samples. Skewness is a measure of how symmetric the distribution is. For a perfectly symmetric distribution (like normal distribution) the skewness is 0.	$\sum_{i}^{n} \frac{(y_i - \bar{y})^3}{n.s^3}$
Variance depicts how the data varies with respect to the mean of the sample.	$\sum_{i=0}^{n} \frac{(y_i - \bar{y})^2}{n}$
Standard error represents the deviation of sample mean with respect to the true mean of the population	$\frac{s}{\sqrt{n}}$

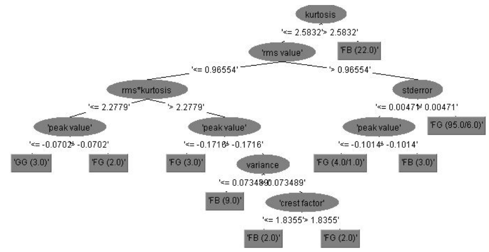

Figure 8. A part of the Decision tree using raw data.

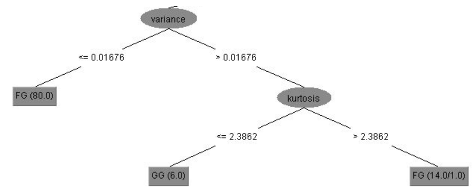

Figure 9. A part of the Decision tree using normalized data.

was found to have the maximum weightage in this case. Thus decision tree shows a descending order of weightage from the root to the leaf. The features extracted from the time-domain data were used to train a decision tree and the classification efficiency was obtained.

5. Results and Discussion

Test has been carried for the conditions as mentioned in Table 1 and the time-domain data acquired were stored. The time-domain data acquired from the accelerometer for different test conditions were then normalised. This huge data acquired for each test condition was then extracted with the help of variables such as variance, skewness, crest factor, root mean square (RMS), kurtosis, RMS*kurtosis, peak value and standard error was stored.

The extracted data corresponding to different conditions of the gearbox for 1st gear tested at 10kmph were combined to a single file. A decision tree was trained using the combined data and the classification efficiency was determined. Table 3 and Table 4 represent the test results in the form of a confusion matrix when the data was processed without and with normalisation respectively. In Table 3 and Table 4, GG represents Good Gearbox, FG represents Faulty Gear, FB represents Faulty Bearing, N represents Normalised and WN represents Without Normalisation. Data representation corresponding to the confusion matrices shown in Table 5 and Table 6 is shown in Figure 10.

The same procedure is repeated for other combinations as mentioned in Table 1.

In Table 5 and Table 6, the diagonal elements represent the number of correctly classified instances. In Table 4,the value in the first row and first column represent the number of instances that were correctly classified as 1_10_GG_N whereas the second and third column of first row represent the number of instances that were misclassified as 1_10_FG and 1_10_FB respectively. Hence, 95 instances were correctly classified whereas 5 instances were misclassified as 1_10_FG_N.

The classification efficiencies for all the test conditions with and without normalisation are mentioned in Table 5. From Table 7 it is seen that classification efficiency of all the test conditions was obtained to be above 90% when the data was normalised whereas the classification efficiency of the test conditions was significantly low when data was not normalised. Thus, normalisation was found to increase the classification efficiency and decision tree can be used for automated fault diagnosis.

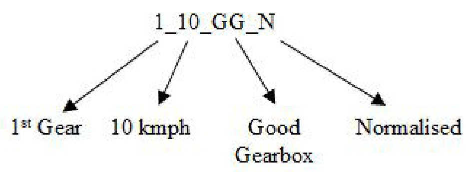

Figure 10. Data representation.

Table 5. Confusion matrix for Gear 1 tested at 10 kmph (without normalisation)

Condition	1_10_GG_WN	1_10_FG_WN	1_10_FB_WN
1_10_GG_WN	80	1	19
1_10_FG_WN	4	87	9
1_10_FB_WN	16	11	73

Table 6. Confusion matrix for Gear 1 tested at 10 kmph (with normalisation)

Condition	1_10_GG_N	1_10_FG_N	1_10_FB_N
1_10_GG_N	95	5	0
1_10_FG_N	3	97	0
1_10_FB_N	0	1	99

Table 7. Classification efficiency

Condition	Classification efficiency of decision tree (%)	
	Without normalisation	With normalisation
1_10	80	97
1_20	84.3333	97.6667
2_20	67.3333	99.3333
2_30	59	99.3333
3_30	47	95.3333
3_40	61	99.3333
4_40	64.3333	99.6667
4_50	60	98.3333
5_50	55.6667	99
5_60	79.3333	90.3333

6. Conclusion

Thus tests were conducted on the gearbox in a test vehicle run in real time conditions and data was acquired for different conditions of gearbox. The time-domain data acquired was extracted to obtain statistical features and decision tree classifier was used to classify the different conditions. Also, the time-domain data was normalised and features were extracted from the normalised data. The features obtained from the normalised data were then used to train the decision tree algorithm. The classification efficiency of the decision tree was found to increase significantly when features extracted from normalised data were used. It is to be noted that data normalization has highly improved the efficiency of the decision tree classifier. Also, the high classification efficiency shows that vibration monitoring of an automobile gearbox in a test vehicle run in real time conditions is feasible.

7. References

1. Sawalhi N, Randall RB. Simulating gear and bearing interactions in the presence of faults: Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults. Mechanical Systems and Signal Processing. 2008 Nov; 22(8):1924–51.
2. Sawalhi N, Randall RB.Simulating gear and bearing interactions in the presence of faults: Part II: Simulation of the vibrations produced by extended bearing faults. Mechanical Systems and Signal Processing. 2008 Nov; 22(8):1952–66.
3. Chaari F, Fakhfakh T, Haddar M. Analytical modeling of spur gear tooth crack and influence on gear mesh stiffness. European Journal of Mechanics. A/Solids.2009 May–Jun; 28(3):461–8.
4. Li C,Sanchez R-V, Zurita G, Cerrada M, Cabrera D, Vasquez RE. Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. Neurocomputing. 2015 Nov; 168:119–27.
5. Soua S, Van Lieshout P, Perera A, Gan T-H, Bridge B. Determination of the combined vibrational and acoustic emission signature of a wind turbine gearbox and generator shaft in service as a pre-requisite for effective condition monitoring. Renewable Energy. 2013; 51:175–81.
6. Hea G, Dinga K, Li W, Jiao X. A novel order tracking method for wind turbine planetary gearbox vibration analysis based on discrete spectrum correction technique. Renewable Energy. 2016 Mar; 87:364–75.
7. Hong L, Dhupia JS. A time domain approach to diagnose gearbox fault based on measured vibration signals. Journal of Sound and Vibration. 2014 Mar; 333(7):2164–80.

8. Satishkumar R, Sugumaran V. Vibration based health assessment of bearings using random forest classifier.Indian Journal of Science and Technology. 2016 Mar; 9(10):1–6. DOI: 10.17485/ijst/2016/v9i10/85303.
9. Yin J, Wang W, Manc Z, Khoo S. Statistical modelling of gear vibration signals and its application to detecting and diagnosing gear faults. Information Sciences. 2014 Feb; 259:295–303.
10. Jegadeeshwaran R, Sugumaran V. Fault diagnosis of automobile hydraulic brake system using statistical features and support vector machines. Information Sciences. 2015 Feb; 52:436–46.
11. Muralidharan V, Ravikumar S, Kangasabapathy H. Condition monitoring of Self Aligning carrying Idler (SAI) in belt-conveyor system using statistical features and decision tree algorithm. Measurement. 2014 Dec; 58:274–9.
12. Satishkumar R, Sugumaran V. Estimation of remaining useful life of bearings based on support vector regression. Indian Journal of Science and Technology. 2016 Mar; 9(10):1–7. DOI: 10.17485/ijst/2016/v9i10/88997.
13. Karabadji NE, Seridi H, Khelf I, Azizi N, Boulkroune R. Improved decision tree construction based on attribute selection and data sampling for fault diagnosis in rotating machines. Engineering Applications of Artificial Intelligence. 2014 Oct; 35:71–83.
14. Sharma A, Sugumaran V, Babudevasenapati S. Misfire detection in an IC engine using vibration signal and decision tree algorithms. Measurement. 2014 Apr; 50:370–80.
15. Saravanan N, Ramachandran KI. Fault diagnosis of spur bevel gear box using discrete wavelet features and Decision Tree classification. Expert Systems with Applications. 2009 Jul; 36:9564–73.
16. Sakthivel NR, Sugumaran V, Babudevasenapati S. Vibration based fault diagnosis of monoblock centrifugal pump using decision tree. Expert Systems with Applications. 2009 Oct; 37(6):4040–9.
17. Gangadhar N, Kumar H, Narendranath S, Sugumaran V. Fault diagnosis of single point cutting tool through vibration signal using Decision Tree algorithm. Procedia Materials Science. 2014 Sep; 5:1434–41.
18. Saravanan N, Cholairajan S, Ramachandran KI. Vibration-based fault diagnosis of spur bevel gear box using fuzzy technique. Expert Systems with Applications. 2009 Mar; 36(2):3119–35.
19. Saimurugan M, Ramachandran KI, Sugumaran V, Sakthivel NR. Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine. Expert Systems with Applications. 2011 Apr;38(4):3819–26.