Measurement 46 (2013) 3247–3260

Contents lists available at SciVerse ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Comparative study of decision tree classifier and best first tree classifier for fault diagnosis of automobile hydraulic brake system using statistical features

R. Jegadeeshwaran *, V. Sugumaran

School of Mechanical and Building Sciences, Department of Mechanical Engineering, VIT University Chennai Campus, Chennai, Tamil Nadu, India

ARTICLE INFO

Article history:
Received 26 February 2013
Received in revised form 8 April 2013
Accepted 23 April 2013
Available online 27 May 2013

Keywords:
Decision tree
Statistical features
Feature extraction
Feature selection
Best first tree
Pruning

ABSTRACT

Hydraulic brakes are the most important components in automobile. It requires advanced supervision and fault diagnosis to improve the safety of passenger, reliability and economy. Condition monitoring is one of the major division through which the reliability of such components could be monitored. The condition of the brake components can be monitored by using the vibration characteristics which will reveal the condition of the brake systems. In this paper machine learning algorithm using vibration monitoring is proposed as a possible solution to this problem. From the hydraulic brake test set up, the vibration signals were acquired by using a piezoelectric transducer and data acquisition system. C4.5 decision tree algorithm was used to extract statistical features from vibration signals. Feature selection was also carried out. Since no much of methodologies are available to find the effective number of features for a given problem, a detailed study is needed to find the best possible number of features. Hence the effect of number of features was studied by using decision tree. The selected features were classified using C4.5 decision tree algorithm and Best first decision tree algorithm with pre pruning and post pruning techniques. The results are discussed and conclusions of the study are presented.

1. Introduction

Brakes are the most important control components responsible for the safety and stability of the vehicle. Every automobile should be equipped with an efficient brake system to bring the vehicle to rest within a reasonable distance even under the most adverse conditions. In addition to the main requirement that a vehicle shall stop within a reasonable distance, it is also desirable that the retardation should be smooth and free from shudder and the rate of retardation shall be proportional to the pedal effort. This means that whilst the effort required by the driver to operate the brakes shall not be excessive. The brake system should be very reliable to promote the highest degree of safety on the road. It is not that easy to maintain a brake system. There are many things that must be taken into account. The very important idea of maintenance is safety, not alone for the person driving but also for the others moving on the road. The motivation behind this study is to stop accidents due to faulty brakes. Since there are moving components involved, they are bound to get faulty due to various reasons, viz. wearing, air leak, fade, etc. When such things occur, the effectiveness of the brake reduces resulting in accidents. Hence it is necessary that they should be monitored all the time and diagnosed when faults occur. Monitoring of brakes is a separate area of concern in the contemporary automotive world.

Condition based monitoring is the process of monitoring a parameter of condition in machinery. A failure will indicate some significant change in its physical structure. The use of condition monitoring allows maintenance or other actions to be scheduled, to avoid the consequences of failure. Condition monitoring systems can only measure the deterioration of the condition when failure occurs.

* Corresponding author. Tel.: +91 44 39931335; fax: +91 44 39932555.
E-mail addresses: krjegadeeshwaran@gmail.com (R. Jegadeeshwaran), v_sugu@yahoo.com (V. Sugumaran).

0263-2241/$ - see front matter
http://dx.doi.org/10.1016/j.measurement.2013.04.068

Interference in the early stages of deterioration is usually much more cost effective and life saving than allowing the brakes to fail.

Machine fault diagnosis is a branch of study concerned with finding faults arising in machine components. To identify the most probable fault, many methods were used including vibration analysis, oil particle analysis, thermal imaging, etc. Amongst them vibration analysis is most commonly used one. Comparison of the vibration spectra of faulty signal conditions Vs good signal conditions will provide the information required to make a decision when intervention is required for maintenance. The vibration signals are processed and analyzed by using wavelet analysis, spectral analysis and waveform analysis. The results of such analysis are used to determine the original cause of the fault through root cause failure analysis.

Fault diagnosis involves three main steps namely, feature extraction, feature selection, and feature classification. Features can be mainly statistical features [1], histogram features [2] and wavelet features [3,4]. In this study statistical feature were used.

Many techniques were used for feature selection including principal component analysis (PCA) [5], genetic algorithm (GA) [6], decision tree (DT) [7], fuzzy and artificial neural network [8–10]. Principal component analysis (PCA) is one of the pattern identification techniques in data of high dimension. It reduces dimensionality of data by performing a covariance analysis betw1een factors. Hence PCA models have trouble with high dimensional data or large numbers of data points, and it is not obvious how to deal properly with incomplete data set, in which some of the data points are missing. In a study, Sakthivel et al. used decision tree (DT) to identify the best features from a given set of samples for classification [10]. Based on this technique, the most important features were identified. Decision trees are simple to understand and it can be represented more compactly as an influence diagram. Hence, in the present study decision tree was used for feature selection.

A number of classifiers are available for feature classification namely, Support vector machine (SVM), Proximal support vector machine (PSVM), Naïve Bayes (NB) and Bayes Net (BN), Artificial neural network (ANN), Fuzzy, Decision tree, etc. Sakthivel et al. developed a fault classification model for mono block centrifugal pump using support vector machine and proximal support vector machines [11]. Sugumaran et al. used support vector machine (SVM) and proximal support vector machine (PSVM) to classify the faults in roller bearing using statistical features [12]. Yuan and Chu developed a SVM based fault diagnosis model [13,14]. However, the size of the patterns increases, the training time increases and also the computational complexity increases in SVM based model.

In recent studies, machine learning approaches such as Bayes net and Naïve Bayes were reported for fault diagnosis of critical components [15]. Addin and Sapuan studied about damage detection technique in engineering materials [16]. Tool condition monitoring using Naïve Bayes and Bayes net algorithms were also studied [17]. Bayes net and Naïve Bayes were successfully applied for finding faults in mono block centrifugal pump, and was proved that the classification accuracy was more than the classification done by decision tree [18].

The robustness and effectiveness of fuzzy classifier depends on the fuzzy rules. Rajakarunakaran et al. developed a fault classification model for a centrifugal pump using artificial neural network approach [19]. Wang et al. also developed a neural network model for a centrifugal pump with frequency domain signal to detect faults at early stages [8]. Eventhough it gave very good result, training of an artificial neural network classifier was complex and time consuming process.

To overcome the above difficulties, researchers need to identify a classifier model which will give a better classification accuracy with simple training operation. It should do both feature selection and feature classification simultaneously. In many fault diagnosis applications, C4.5 decision tree algorithm was successfully used for both feature selection and feature classification. A novel hybrid system based on C4.5 algorithm was proposed by Polat and Gunes, to classify the multi-class problems [20]. Condition monitoring of roller bearing using decision tree was reported by Sugumaran et al. Statistical features were extracted from vibration signals [9].

Jin Yi et al. apply the virtual instrumentation technology (Lab VIEW and MATLAB) to examine the hydraulic brake system and put forward the fault diagnosis system to vehicle hydraulic system. The diagnosis system thus effectively collects the characteristic signals of hydraulic system to analyze and compare the fault signals [21].

Many classification algorithms have been used for fault classification in various elements. In order to find a better algorithm, a detailed study is needed. This study particularly focuses the performances of best first tree classifier algorithms in the fault classification of automobile hydraulic brake system. The best first tree is highly sensitive classifier algorithm. Hence it can be used for feature classification effectively. However, classification of faults in automobile hydraulic brake system using best first tree and decision tree algorithm has not been attempted. Hence an effort was made in the present study to classify the faults in hydraulic brake using decision tree and best first tree. The flow chart of the fault diagnostic system is shown in Fig. 1.

Contributions in the present work are the following:

(1) The procedure of fault diagnosis of hydraulic brake system was illustrated. From the brake fault simulator experimental setup the following fault conditions were simulated and the vibration signals were recorded. Air in the brake fluid, brake oil spill on disc brake, drum brake pad wear, disc brake pad wear (even) – inner, disc brake pad wear (even) – inner and outer, disc brake pad wear (uneven) – inner, disc brake pad wear (uneven) – inner and outer, reservoir leak, drum brake mechanical fade.
(2) From the time domain signals, a set of statistical features were extracted and the order of importance and effective number of features were found using decision tree.