ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Procedia Technology 23 (2016) 187 – 192

Procedia Technology

3rd International Conference on Innovations in Automation and Mechatronics Engineering, ICIAME 2016

Fault Classification of Ball Bearing by Rotation Forest Technique

S. Kavathekar[a,]*, N. Upadhyay[a], P.K. Kankar[a]

[a]*Machine Dynamics and Vibration Lab,Mechanical Engineering Discipline, PDPM Indian Institute of Information Technology Design and Manufacturing, Jabalpur, 482005, India*

Abstract

Bearing failure is one of the most common causes of breakdown in rotating machines. The machine learning techniques such as Support vector machines (SVM), Artificial neural network (ANN) are widely used for fault classification. These methods are slow and sometime give inaccurate results. Therefore, the search for new classifier techniques is a necessity to increase the classification efficiency with less computation time. In this study, a classifier ensemble is used for fault classification called Rotation forest. Data obtained from Case Western Reserve University have been used to extract time-based statistical features. In all κ subsets are formed by randomly bifurcating the feature set. Principal Component Analysis (PCA) is used on each subset. All principal components are saved to preserve the transformation in the data. The novel features are calculated using κ axis rotations. This results in improved efficiency of fault classification.

Peer-review under responsibility of the organizing committee of ICIAME 2016

*Keywords:*Bearing fault classification; Statistical features; Bootstraping; PCA; Rotation Forest;

1. Introduction

Bearing is mostly used in all kinds of rotating machinery. Due to continuous long run and fatigue loading, the fault will occur on the bearing surfaces and these faults are the major cause of failure of rotating system. It is important to identify those faults in initial state, before serious damages can take place.

* Corresponding author. Tel.:+91-9823239455;
*E-mail address:*sohamkavathekar@gmail.com

Peer-review under responsibility of the organizing committee of ICIAME 2016
doi:10.1016/j.protcy.2016.03.016

Nomenclature

A	Training data *[N×n]*
B	class lable *[N×1]*
β	set of class label *[1×d]*
P_i	classifier
L	Number of classifiers
S	Features vector *[1×N]*
κ	Number of subset of each classifiers
Q	Number of feature in each disjoint feature subset
F_i	Rotation Matrix

Bearings faults are of two types, localized bearing fault like pits, spalls, dents etc. and other is distributed fault like surface waviness of outer and inner race, off sized rolling element etc. [1]. A large number of methodologies have been developed in previously published literature [3-7].

Each Fault has certain vibration signature, and has different frequency of characterization. In the initial stage of fault, the vibration signatures are less in amplitude as compared to the maximum amplitude. So, only spectral analysis is not enough for analyzing vibration signals. Further, the statistical features of signals are acquired such as Skewness, Kurtosis and Root Mean Square (RMS) etc. [5]. These statistical features are analyzed using Machine Learning Techniques to diagnose the fault.

Fault classification of mechanical component assist maintenance engineer to take decision about maintenance, repair and replacement activities. The time domain features of vibration signals had been used as inputs to the machine learning techniques. For classification, ensemble techniques are competitive and producing good results.

In this study, a new technique known as rotational forest is introduced for fault classification of the bearings and then compare the results with already well-established machine learning methods such as ANN and SVM. This study is divided into five sections. In section 2 brief description of the proposed Rotation Forest algorithm, section 3 the general description about experimental setup and selected features, section 4 contains results of this study and compared results of Rotation Forest with ANN, SVM, Random Forest and Decision Tree, section 5 drawn some brief conclusion based on this study.

2. Rotation forest algorithm

Rotation forest is an ensemble which is derived from decision tree classifier. Trees are grown using J-48 algorithm with pruning [2]. In this section Rotation Forest is briefly explained. The algorithm is represented in Fig. 1.

2.1. Algorithm

Let, $A_1 = [a_1, \ldots, a_n]^T$ be a data point having n features. Let, A be the set of such data containing many data points as A_1. Dimension of matrix A is $N \times n$. Consider a second vector, B having class labels for data, $B = [b_1, \ldots, b_n]^T$, where b_j taken value from vector $\beta = \{\beta_1, \ldots, \beta_c\}$. Taking $P_1, \ldots, P_L$,the classifiers in the ensemble, with L number of classifiers which we need to pick in advance. Let, S be the feature set classifiers. Classifiers $P_1, \ldots, P_L$ are trained in parallel, this process is similar in the case of Bagging and Random Forest.

For construction of single classifier P_i, we need to follow following procedure,

1. Split the feature vector S randomly in κ subsets. The subsets may be intersecting or disjoint. To increase diversity of classifier we choose disjoint subsets. In simple words, $Q = n/\kappa$, where Q is number of features in each of the κ subsets and n is total number of features.
2. $S_{i,j}$ represents the j^{th} feature subset for training classifier P_i for each subset selecting at random a non-empty subset of classes and later draw a bootstrap sample of size 75% of data count.

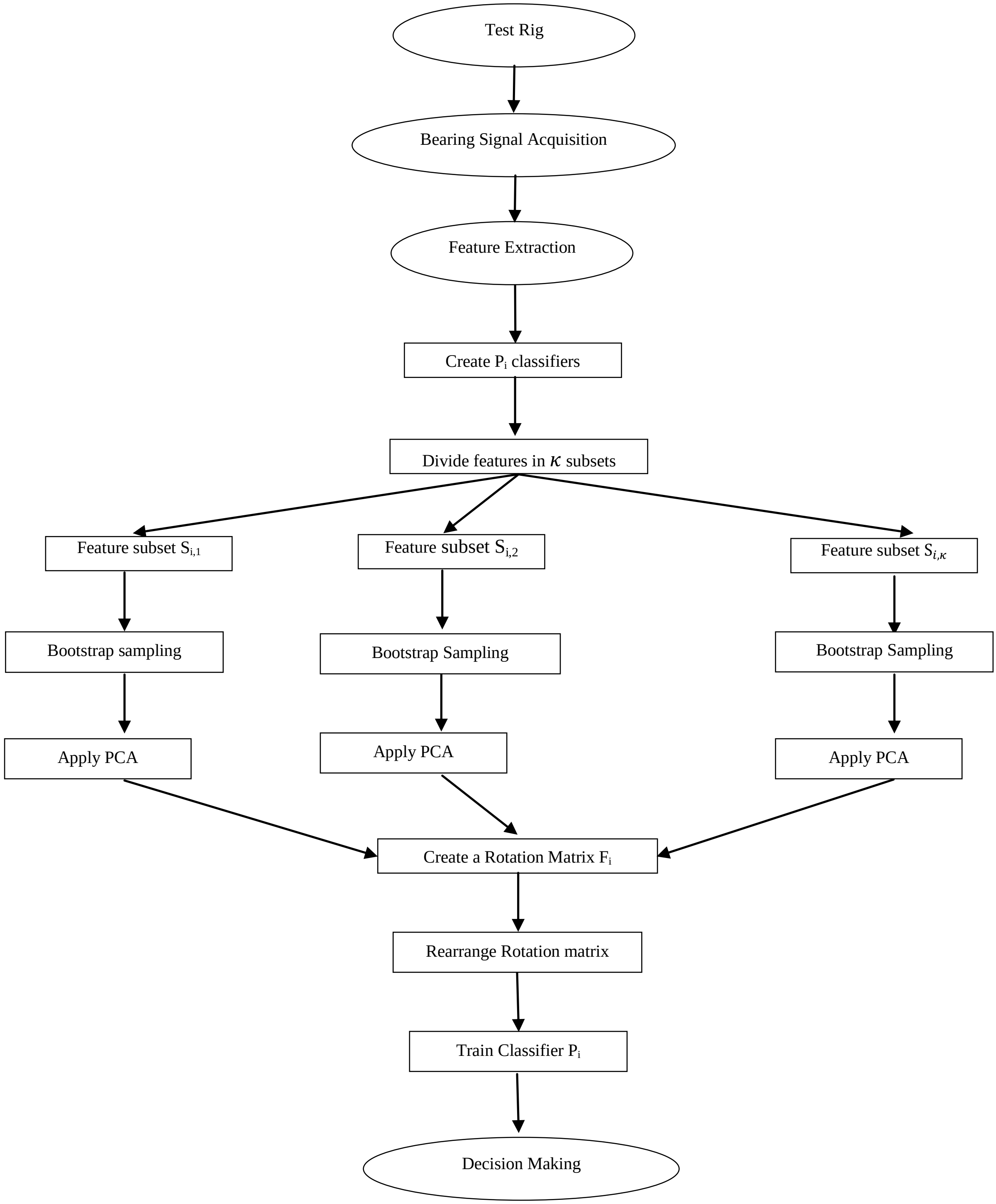

Fig. 1 Construction of one classifier in Rotation forest.

3. Run Principal Component Analysis (PCA) on Q feature in $S_{i,j}$ and choose subset of A. The coefficients of PCA, $m_{i,j}^{(1)}, \ldots, m_{i,j}^{Q_j}$, each of size $Q \times 1$ are saved. There is a possibility that, some of the Eigen values are zero. So that, the number of vectors getting after PCA may have less than Q vectors i.e. $Q_j < Q$. The benefit of application of PCA on each subset over applying on the complete set is to get variation of PCA coefficients in the subset of same feature selected for different classifiers.
4. Coefficients obtained in the form of vectors are arranged in distributed 'Rotation' matrix F_i

$$F_i = \begin{bmatrix} m_{i,j}{}^{(1)}, \ldots, m_{i,j}{}^{(Q_1)} & \cdots & [0] \\ \vdots & \ddots & \vdots \\ [0] & \cdots & m_{i,j}{}^{(1)}, \ldots, m_{i,j}{}^{(Q_K)} \end{bmatrix}_{(n \times \Sigma_j Q_j)} \tag{1}$$

The training set for classifier P_i is computed by rearranging columns of rotation matrix so that they correspond to original features. This result in training set for classifier P_i which is given as XR_i^m.

3. Experimental Setup and Features Selected

Experimental data is taken from Case Western Reserve University (CWRU) [8]. The setup has a 2HP motor, a dynamometer, bearing support. Data is acquired using an accelerometer at 12000 and 48000 Hz sampling frequency. In this study, 48000 Hz sampling frequency data is used to calculate time series statistical data. The Fig. 2 below shows the setup used. The load was varied from 0 HP to 3 HP with fault dimensions varying from 0.1761 to 0.7044 mm.

3.1. Features Selected

The total 10 features have been extracted from signal obtained from drive end. These features include mean, standard deviation, variance, root mean square value, skewness, kurtosis, minimum value, peak value, crest factor and form factor. Details of these features are explained by Vakharia et al. [5].

3.2. Efficiency of Classifier: The ratio of Correctly classified instances to Total number of instances is called efficiency of classifier.

$$Efficiency = \frac{Correctly\ classified\ Instances}{Total\ number\ of\ Instances}$$

4. Results

The comparative study was done using time series features of Case Western Reserve University [8]. Total 55 cases are considered which covers 4 cases of Healthy bearing, 11 cases of Inner race (IR) defects, 12 cases of Ball defect, and 28 cases of Outer race (OR) defects. The results are calculated using different machine learning techniques like ANN, SVM, Decision Tree, Random Forest, ANN and SVM using Cluster Membership filter, and Rotation Forest. Results are displayed in the form of confusion matrix.

Table1. Confusion Matrix using ANN classifier

Classified as	Healthy	IR	Ball	OR
Healthy	3	0	1	0
IR	0	2	3	6
Ball	0	2	8	2
OR	0	5	4	19

Table2. Confusion Matrix using SVM classifier

Classified as	Healthy	IR	Ball	OR
Healthy	0	0	0	4
IR	0	0	0	11
Ball	0	0	0	12
OR	0	1	1	26

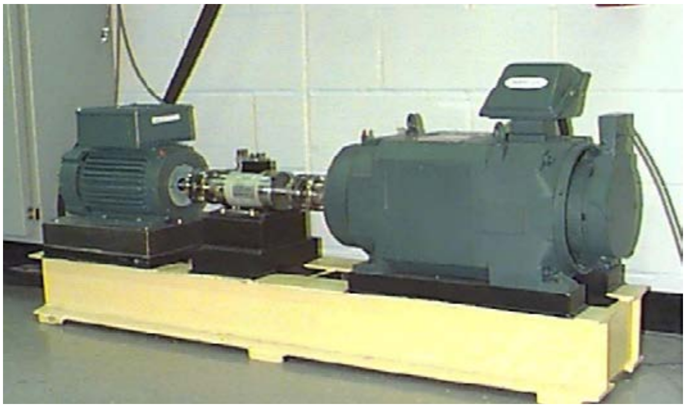

Fig. 2 Experimental setup of CWRU Test rig [8]

Table3. Confusion Matrix using ANN with Cluster Membership filter classifier

Classified as	Healthy	IR	Ball	OR
Healthy	4	0	0	0
IR	1	5	0	5
Ball	0	0	7	5
OR	0	1	2	25

Table4. Confusion Matrix using SVM with Cluster Membership filter classifier

Classified as	Healthy	IR	Ball	OR
Healthy	4	0	0	0
IR	0	0	1	10
Ball	1	0	7	4
OR	0	1	2	25

Table5. Confusion Matrix using Decision Tree classifier

Classified as	Healthy	IR	Ball	OR
Healthy	3	1	0	0
IR	0	4	1	6
Ball	0	2	6	4
OR	0	4	4	20

Table6. Confusion Matrix using Random forest classifier

Classified as	Healthy	IR	Ball	OR
Healthy	4	0	0	0
IR	0	4	1	6
Ball	1	0	7	4
OR	0	2	4	22

Table7. Confusion Matrix using Rotation Forest classifier

Classified as	Healthy	IR	Ball	OR
Healthy	4	0	0	0
IR	0	6	1	4
Ball	0	0	10	2

OR	0	2	4	22

The efficiency of classification ANN and SVM is 57.14% and 46.42% respectively. After applying Cluster membership filter and again repeating ANN and SVM, classification efficiency achieved is 73.21% and 64.28 % respectively. The diversity achieved in classifiers like ANN and SVM is less. As a result ends up in lower classification efficiency, solely and with use of Cluster Membership filter. As base classifier used in Rotation Forest algorithm is Decision Tree, it was used for classification which gave classification efficiency of 58.92 %. Also, Random Forest Algorithm is applied having efficiency 66.07%. As in used Rotation Forest, Random Forest uses Bootstrapping for data sampling which increase Diversity. Since Rotation of axes is done using PCA, Rotation Forest gives more efficiency than Random Forest i.e. 75%.

5. Conclusions

Feature extraction and classification of bearing defects using time series data and machine learning technique is recently reported in literature. It is important to select the efficient machine learning technique for any classification problems. In this study, Rotation Forest method is used for bearing fault classification. The proposed Rotation Forest algorithm is a novel ensemble classifier that builds large number of decision trees with applying PCA, which will solve the feature selection problem as compared to the other machine learning techniques. The diversity provided by Rotational Forest algorithm is highest as compared to other classifiers. This process is fast enough to be used in real time fault classification. Comparing the classification result with ANN, SVM, Decision tree and Random Forest, proposed technique gives better efficiency. In future we can apply this method with combination of different signal decomposition method to increase the efficiency of classification, and further we can apply for fault diagnosis.

6. Acknowledgement

Authors are thankful to Case Western Reserve University Bearing Data Centre for their variety of data in respect of fault sizes and speed along with loads.

References

[1]. Kankar P K, Sharma S C, Harsha S P. Fault diagnosis of ball bearings using machine learning methods. Expert Systems with Applications, 2011; 38(3): pp.1876–1886.
[2]. Rodríguez J J, Kuncheva L I, Alonso C J. Rotation forest: A New classifier ensemble method. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006; 28(10): pp.1619–1630.
[3]. Sharma A, Amarnath M, Kankar P K. Feature extraction and fault severity classification in ball bearings. Journal of Vibration and Control, 2014; (February): pp.1–17.
[4]. Tiwari R, Gupta V K, Kankar P K. Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier. Journal of Vibration and Control, 2013; Vol. 21(3) :461–467.
[5]. Vakharia V, Gupta V K, Kankar P K. A comparison of feature ranking techniques for fault diagnosis of ball bearing. Soft Computing. 2015;DOI 10.1007/s00500-015-1608-6.
[6]. Yang B S, Di X, Han T. Random forests classifier for machine fault diagnosis. Journal of Mechanical Science and Technology, 2008; 22(9): pp.1716–1725.
[7]. Yang B S, Lim D S, Tan A C C. VIBEX: An expert system for vibration fault diagnosis of rotating machinery using decision tree and decision table. Expert Systems with Applications, 2005; 28: pp.735–742.
[8]. Loparo K A, Bearing vibration data set. Case Western Reserve University. Available at: www.eecs.cwru.edu/laboratory/bearing (accessed September 2013)