COMPRESSIVE SENSING STRATEGY FOR CLASSIFICATION OF BEARING FAULTS

*H.O.A. Ahmed[1], M. L. Dennis Wong[2], A. K. Nandi[1,3] **

[1]Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, UB8 3PH, United Kingdom

[2]Heriot-Watt University Malaysia, Precinct 5, 62200 Putrajaya.

[3] The Key Laboratory of Embedded Systems and Service Computing, College of Electronic and Information Engineering, Tongji University, Shanghai, P.R. China

ABSTRACT

Owing to the importance of rolling element bearings in rotating machines, condition monitoring of rolling element bearings has been studied extensively over the past decades. However, most of the existing techniques require large storage and time for signal processing. This paper presents a new strategy based on compressive sensing for bearing faults classification that uses fewer measurements. Under this strategy, to match the compressed sensing mechanism, the compressed vibration signals are first obtained by resampling the acquired bearing vibration signals in the time domain with a random Gaussian matrix using different compressed sensing sampling rates. Then three approaches have been chosen to process these compressed data for the purpose of bearing fault classification these includes using the data directly as the input of classifier, and extract features from the data using linear feature extraction methods, namely, unsupervised Principal Component Analysis (PCA) and supervised Linear Discriminant Analysis (LDA). Classification performance using Logistic Regression Classifier (LRC) achieved high classification accuracy with significantly reduced bandwidth consumption compared with the existing techniques**.**

Index Terms — Bearing Fault Classification, Compressive Sensing, Linear Discriminant Analysis, Principal Component Analysis, Machine Condition Monitoring.

1. INTRODUCTION

Rolling element bearings are most critical components in rotating machinery and their failures account for more major failures in the machine. In fact, bearing faults are often a warning sign of other defects in the machine. Thus, rolling element bearings condition monitoring and faults diagnosis has involved considerable attention of researchers over the past decades. Many characteristics features can be measured from vibration signals that make it the best choice for machine condition monitoring. Vibration analysis can be performed in three main categories of waveform data analysis including, time domain [1 – 5], frequency domain and time- frequency domain [6], [7]. The time-frequency domain has been developed for non-stationary waveform signals which are very common when machinery fault occurs. Several time-frequency analysis techniques have been proposed and applied to machinery fault diagnosis, for example, wavelet transform, adaptive parametric time-frequency analysis based on atomic decomposition, and non-parametric time-frequency analysis including Hilbert-Huang transform, local mean decomposition, energy separation and empirical mode decomposition (EMD) [8 - 10]. All these techniques are generally based on Nyquist sampling theorem in which the sampling rate must be at least twice the maximum frequency present in the signal. Nyquist theorem based methods provide a means of measuring a large amount of data to ensure the accurateness of machine condition monitoring, which means large storage and time for signal processing are needed.

To overcome these hindrances, several dimensionality reduction methods have also been widely studied in machine fault diagnosis in order to identify a lower – dimensional space features that efficiently represents the high-dimensional data while retaining the essential information of the machine defects. Among various dimensionality reduction methods, Principal Component Analysis (PCA) [11], Independent Discriminant Analysis (ICA) [12] and Linear Discriminant Analysis (LDA) [13] are among the most popular methods that have been successfully utilized in machine fault diagnosis. For instance, Jiying et al. [14] proposed a condition monitoring method for rolling element bearing based on PCA that effectively identify different conditions of rolling element bearings. Miao et al. [15] proposed a joint bearing fault detection and feature extraction method using EMD and ICA. Giabattoni et al. [16] introduced a novel LDA-based algorithm for bearing data dimensionality reduction and fault detection that shows improvement in the classification accuracy when the classes are overlapped.

These techniques may reduce the time and storage space required. However, since our acquisition systems based on Nyquist sampling rate, we still need to measure a large amount of vibration data. Consequently, the number of machines that can be monitored in remote areas may be limited due to bandwidth limitations.

*Corresponding author: Asoke.Nandi@brunel.ac.uk

Professor Nandi is a Distinguished Visiting Professor at Tongji University, Shanghai, China. This work was partly supported by the National Science Foundation of China grant number 61520106006

Compressive sensing (CS) also called compressed sensing or compressed sampling [17] is a new technique that supports sampling below Nyquist rate. By applying compressive sensing, a signal can be reconstructed using extremely fewer measurements than needed by the Nyquist sampling Theorem. CS has been applied to a large variety of applications including Medical Imaging (MI), Seismic Imaging (SI) and Radio Detection and Ranging [18]. The basic idea is that many images or signals have sparse representations in some domain, e.g., Wavelet Transform (WT), can be reconstructed from fewer measurements under certain conditions. The literature on CS shows a variety of approaches to reconstruct signals from these fewer measurements [19 - 22]. These approaches may not be practical in all situations; for instance, bearing vibration signal is always acquired for faults detection and estimation. Thus, as long as it is possible to detect faulty signals in the compressed domain, then we do not necessarily reconstruct the signal for fault diagnosis.

Effects of compressive sensing on the classification of bearing faults after reconstructing the original signal are discussed in [23]. There are relatively few recent studies on compressed sensing based methods for bearing fault diagnosis. Tang et al. [24] proposed sparse representation classification strategy combined with CS. Zhang et al. [25] suggested several over-complete dictionaries that can be effective in signal sparse decomposition for each vibration signal state. Another learning dictionary basis for extracting impulse components is described by Chen et al. in [26]. Tang et al. [27] proposed an interesting approach to detect the characteristic harmonics from sparse measurements through a compressive sensing pursuit strategy during the process of incomplete reconstruction.

A key limitation of these investigations is that they tended to focus on sparse representations and signal reconstruction rather than learning from the compressed measurements. Moreover, they achieved poor classification results with very small compressive sampling rates. The purpose of this paper is to present a new approach for bearing faults diagnosis from fewer measurements based on CS without reconstructing the original signal. Thus, the proposed method is expected to reduce the bandwidth consumption for remote machine condition monitoring while achieve high classification accuracy.

2. PROBLEM DESCRIPTION

In this paper, we explore the possibility of learning from fewer measurements in the compressive domain based on CS to investigate the current state of rolling element bearings. There are six conditions: a brand new condition (NO) and a worn condition (NW), inner race fault (IR), an outer race (OR) fault, rolling element (RE) fault, and cage (CA) fault. The vibration data have been acquired from experiments on a small test rig that consists of a DC motor driving the shaft through a flexible coupling and supported by two Plummer bearing blocks where several damaged bearings were inserted in one of the plummer. The resulting vibrations in the horizontal and vertical planes were measured using two accelerometers. The output of the accelerometers was fed back through a charge amplifier to Loughbrough Sound Images DSP32 ADC card (using a low-pass filter with a cut-off 18 kHz), and sampled at 48 kHz, giving a slight oversampling. The machine was run at a series of 16 different speeds ranging from 25 and 75 rev/s, and ten-time series were taken at each speed. This results in a total of 160 examples of each condition and a total of 960 raw data files to work with. Each raw data file is a time – series of 6000 samples representing amplitude of vibrations. Fig. 1 describes some typical time series plots for the six conditions defined above.

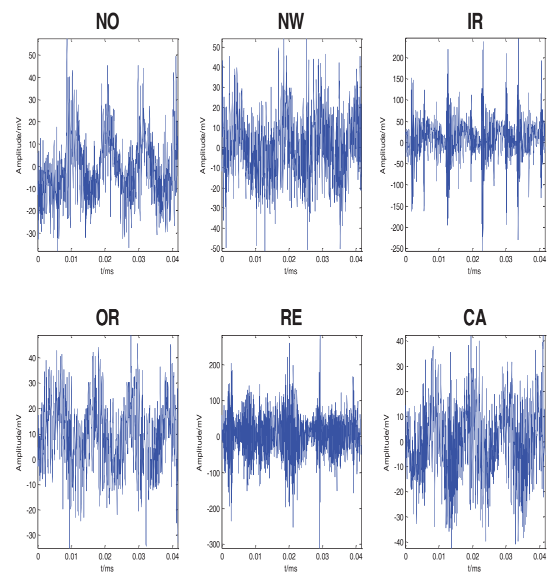

Fig. 1. Characteristic vibration signals for the six different conditions

Based on the fault condition, the damage modulates the vibration signals with their own patterns. Both inner and outer race fault conditions have a fairly periodic signal; the faulty signal of rolling element may or may not be periodic, depend on some factors including the loading of the bearing, the degree of defects to the rolling element, and the track that the ball describes within the raceway. The cage fault generates a random distortion that also depend on the degree of damage and the bearing loading.

3. COMPRESSED SENSING

CS model [12] for a signal $x \epsilon R^{Nx1}$ enables the reconstruction of the signal from fewer measurements that contain all the information of the signal, i.e., transmitting the signal sampling into information sampling. In fact, CS is an extension of sparse representation and special case of it, when only minor measurements are available than the whole number signal samples to be reconstructed. To obtain a set of sparse representations of the signal x, first, we need to use sparsifying transform matrix $\psi \in R^{NxN}$ and the equation that describes this is as follows:

$$x = \psi s \quad (1)$$

where $s \in R^{Nx1}$ is a column vector with k *nonzero* coefficients and represent the sparse elements. Based on compressive sampling theory when the measurement matrix ϕ is incoherent with the sparsifying transform ψ the signal x can be recovered from its compressed measurements y that can be computed by the following equation:

$$y = \phi\psi s \quad (2)$$

where $\phi \in R^{MxN}$ is the measurement matrix and $y \in R^{Mx1}$ is the compressed measurements. To estimate the vector s, we need to solve the optimization problem using L_1-norm; and the estimation $\hat{s}$ of s can be given by the equation:

$$\hat{s} = \underset{s \in R^N}{argmin} \|s\|_{L_1} \quad (3)$$

4. THE PROPOSED APPROACH

We started by obtaining the compressed vibration signals from the high dimensional original bearing vibration signal through CS framework. The vibration signals are not sparse in time domain, but it can be represented efficiently in some domain, e.g., Fourier, Wavelet, etc. Hence, wavelet transform is used as a sparsifying transform ψ and Gaussian matrix that satisfy Restricted Isometry Property (RIP) used as a measurement matrix ϕ. The size of the measurement matrix (MxN) depends on the compressed sampling rate (α) that is significantly lower than Nyquist sampling rate, i.e., (M<<N). The compressed measurements y from equation (2) has enough information to recover the original signal. In our case, this compressed measurement will be the received data signal and can be used for fault classification without reconstructing the original signal.

Three different approaches to use these compressed measurements are proposed here. These are: (1) CS: using the compressed measurement (y) directly for fault classification, (2) CS-PCA: extract features from y using the linear unsupervised learning algorithm PCA, and (3) CS-LDA: selecting features from y using the linear supervised learning algorithm LDA.

Approach I: Bearing condition diagnosis can be conducted directly from the received compressed data (y) obtained from equation (2). Then we apply LRC on y to classify bearing status, as shown in the following

$$y \xrightarrow{\text{LRC}} h_\theta^i(y) = p(class = i|y;\, \theta) \quad (4)$$

where $h_\theta^i(y)$ is a Logistic Regression Classifier (LRC) that can be trained for each class i = 1, 2, 3, ..., n to predict the probability (p) that class is equal to i, and n is the number of classes. This approach will be referred to as CS.

Approach II: We apply PCA [6] on y to extract features for classification. PCA technique transform y with a d – dimensional space to $\hat{y}$ with a $\hat{d}$- dimensional space and this can be computed by the following equation

$$\hat{y} = \sum_{k=1}^{\hat{d}} a_k \; e_k \quad (5)$$

where e_k are the eigenvectors and a_k are the estimates of the compressed data y. Then we apply LRC on $\hat{y}$. The process is represented in the following

$$y \xrightarrow{\text{PCA}} \hat{y} \xrightarrow{\text{LRC}} h_\theta^i(y) = p(class = i|y;\, \theta) \quad (6)$$

This method will be referred to as CS-PCA.

Approach III: We apply LDA [8] on y to select features for classification. LDA method finds a subspace $\hat{y}$ from y by collecting the samples from the same class and expanding the margin of samples from different classes. This can be done by maximizing the Fisher criterion (FC), i.e., the ratio of the between the class scatter (S_B) to the within class scatter (S_w) and this can be given by the following equation

$$\hat{y} = \underset{\hat{y}}{arg\,max}\, FC(\hat{y}) \quad (7)$$

where

$$FC = \frac{|W^T S_B W|}{|W^T S_w W|} \quad (8)$$

and

$$S_B = \frac{1}{N}\sum_{i=1}^{n} N_i\,(\mu^i - \mu)(\mu^i - \mu)^T \quad (9)$$

$$S_w = \frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{N_i}(y_j^i - \mu^i)(y_j^i - \mu^i)^T \quad (10)$$

N is the total number of samples, n is the number of classes and μ^i is the mean vector of class i. The LDA algorithm finds the vector W that maximizes the FC in equation (8). Then we apply LRC on $\hat{y}$ from equation (7) such that

$$y \xrightarrow{\text{LDA}} \hat{y} \xrightarrow{\text{LRC}} h_\theta^i(y) = p(class = i|y;\, \theta) \quad (11)$$

We call this approach CS-LDA.

5. EXPERIMENTAL SETUP

The condition of each bearing is known before we recorded the vibration signals. First we obtain the bearing compressed vibration signals using different compressed sampling rate (α) ranging from 0.05 up to 0.4 with 256, 512,

1024 and 2048 compressed measurements of the bearings original vibration signal. These different sets of compressed measurements used for the fault diagnosis purpose. Half of this data is used for training and the other half is used to assess the generalization of classification performance.
Classification performance is evaluated using the linear Logistic Regression Classifier (LRC) with regularization parameter value (λ = 0.1) and by averaging the results of classification accuracy from ten experiments for each compressed data set. The computing platform for the experiment in this paper is a laptop, with an Intel® Core ™ i3-5010U CPU @ 2.1 GHZ processor, 8GB RAM, running on 64-bit Windows 10. MATLAB R 2013a is used as the main testing platform, with necessary classifier toolboxes.

6. EXPERIMENTAL RESULTS

In order to verify the validity of our proposed approach, various experiments for the three approaches described in section 4 were conducted using different compressed data sets that were obtained by utilizing different compressed sampling rates. In these experiments, both PCA and LDA were used to extract a low dimensional space with 64 features, i.e., 64 principal components for PCA and 64 most discriminant features for LDA. Table 1 summarizes the classification results for all experiments. It is apparent from Table 1 that choices of small values α (0.05 and 0.1) can lead to high classification accuracies in both CS and CS-PCA, unlike CS-LDA that achieved low classification results although it obtained 100% classification accuracy for α = 0.2 and 0.4.
For further verification of the performance of the proposed method, we conducted several experiments using two well-known dimensionality reduction techniques, namely PCA and LDA to extract features from the raw vibration data. For the classification purpose, we used the same classifier LRC that used to obtain the results in Table 1. The classification results of these methods are shown in Table 2 and can be compared to the classification results of our proposed approach in Table 1 which shows the possibility to achieve high classification performance with only a few compressed measurements comparable to the classification performance obtained using the high dimensional vibration signal.

Table 1. Classification accuracies (%) and related standard deviations (in brackets) for compressed sensed datasets.

	Compressed Sensed data sets Sampling rates (α) / Number of measurements (M)			
	0.05 M=256	0.1 M=512	0.2 M=1024	0.4 M=2048
CS	98.2 (0.3)	98.6 (0.3)	98.9 (0.2)	99.0 (0.3)
CS-PCA	98.8 (0.7)	98.5 (0.4)	98.7 (0.6)	98.8 (0.7)
CS-LDA	72.5 (1.5)	89.8 (3.5)	100 (0.0)	100 (0.0)

Table 2. Summary of classification accuracies (%) and their related standard deviations for the raw vibration using PCA and LDA feature extraction methods.

	Feature Extraction method	
	PCA	**LDA**
Raw Vibration	99.6 ± 0.4	100 ± 0.0

For additional comparison of our approach performance to some of compressed sensing based methods for bearing fault classification Table 3 presents summary of classification accuracies obtained using our proposed method and two published methods [25, 27] where low-dimensional compressed vibration signal used for bearing fault diagnosis. As illustrated in Table 3, our dataset contains six classes which is more than the other two methods considered. Yet our proposed method outperforms the other two methods.

Table 3. Classification performance on different compressed measurements, comparing our work to a number of compressed sensing based methods. N = Vibration signal length. n = Number of classes. α = Sampling rate. M = Compressed measurement dimension. ACA = average classification accuracy.

	N	**n**	**α**	**M**	**ACA (%)**
Our method (CS) Our method (CS-PCA)	5120	6	0.02	100	98.1 98.3
Zhang et al. [25] (Gaussian random matrix)	512	4	0.2	100	92.0 (Normal) 79.0 (IR) 90.0 (OR) 84.0 (RE)
Tang et al. [27]	2000	4	0.05 0.1 0.8	100 200 1600	72.0 80.0 96.0

Taken together, these results show the possibility to reduce the bandwidth consumption by up to 95% for remote machine condition monitoring while achieve good fault diagnosis performance comparable to fault classification performance from high dimensional vibration signal.

7. CONCLUSION

From the outcomes of our investigations, it is possible to conclude that bearing faults can be classified well from vibration data sampled below Nyquist rate without reconstructing the signal. For clarity, we have chosen three approaches to use these data for the purpose of classification these includes: using the data directly as the input of classifier and extracting features from the data using PCA and LDA. We have been able to achieve high levels of classification accuracy while reducing bandwidth requirements compared with other existing techniques. On the basis of the promising findings presented in this paper, the next stage of our research will be to explore the possibility to design CS based hardware that allows us to collect the bearing vibration data in a compressed form, i.e., perform the compression before Analog to Digital Conversion (ADC) stage in data acquisition system.

8. REFERENCES

[1] L Zhang and A K Nandi, “Fault classification using genetic programming,” *Mechanical Systems and Signal Processing*, vol. 21, no. 3, pp. 1273-1284, 2007.

[2] L Zhang, L B Jack and A K Nandi, “Fault detection using genetic programming,” *Mechanical Systems and Signal Processing*, vol. 19, no. 2, pp. 271-289, 2005.

[3] H Guo, L B Jack and A K Nandi, “Feature generation using genetic programming with application to fault classification,” *IEEE Transactions on Systems,* Man, and Cybernetics, Part B, vol. 35, no. 1, pp. 89-99, 2005.

[4] L B Jack and A K Nandi, “Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms,” *Mechanical Systems and Signal Processing*, vol. 16, no. 2-3, pp. 373-390, 2002.

[5] L B Jack and A K Nandi, “Genetic algorithms for feature selection in machine condition monitoring with vibration signals,” *IEEE Proceeding,* Part VIS, vol. 147, no. 3, pp. 205-212, 2000.

[6] A. Jardine, D. Lin, D. Banjevic, “A review on machinery diagnostics and prognostics implementing condition-based maintenance,” *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483-1510, 2006.

[7] A. K. Nandi, C. Liu, M. Wong, "Intelligent Vibration Signal Processing for Condition Monitoring." *Proceedings of the International Conference Surveillance 7, 2013*.

[8] R. Yan, R. Gao, and X. Chen. "Wavelets for fault diagnosis of rotary machines: a review with applications." *Signal Processing.* Vol. 96, Part A, pp. 1-15, 2014.

[9] F. Zhipeng, M. Liang, and F. Chu. "Recent advances in time–frequency analysis methods for machinery fault diagnosis: a review with application examples." *Mechanical Systems and Signal Processing,* vol.38.1, pp. 165-205, 2013.

[10] W. Cong, M. Gan, and C. Zhu. "Non-negative EMD manifold for feature extraction in machinery fault diagnosis." *Measurement.* Vol. 70, pp. 188-202, 2015.

[11] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” *Chemom. Intell. Lab. Syst.*, vol. 2, no. 1–3, pp. 37–52, 1987.

[12] P. Comon, “Independent component analysis, A new concept,” *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994.

[13] S. Balakrishnama and a Ganapathiraju, “Linear Discriminant Analysis - a Brief Tutorial,” *Compute*, vol. 11, pp. 1–9, 1998.

[14] L. Jiang, Z. Liu, J. Cui, and Z. Li, “Condition monitoring of rolling element bearing based on Phase-PCA,” *2013 25th Chinese Control Decis. Conf. CCDC 2013*, pp. 1082–1086, 2013.

[15] Q. Miao, D. Wang, and M. Pecht, “Rolling element bearing fault feature extraction using EMD-based independent component analysis,” *2011 IEEE Conf. Progn. Heal. Manag.*, pp. 1–6, 2011.

[16] L. Ciabattoni, G. Cimini, F. Ferracuti, A. Freddi, G. Ippoliti, and A. Monteriù, “A novel LDA-based approach for motor bearing fault detection,” *Proceeding - 2015 IEEE Int. Conf. Ind. Informatics, INDIN 2015*, pp. 771–776, 2015.

[17] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[18] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, “Compressive sensing: From theory to applications, a survey,” *J. Commun. Networks*, vol. 15, no. 5, pp. 443–456, 2013.

[19] F. Renna, R. Calderbank, L. Carin, and M. R. D. Rodrigues, “Reconstruction of signals drawn from a gaussian mixture via noisy compressive measurements,” *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2265–2277, 2014.

[20] D. Needell and J. A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, 2009.

[21] M. Mishali and Y. Eldar, “Blind multiband signal reconstruction: Compressed sensing for analog signals,” *Signal Process. IEEE Trans.*, Vol. 57, no. 3, pp. 993-1009, 2009.

[22] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” *Inf. Theory, IEEE Trans.*, vol. 55, no. 5, pp. 2230-2249, 2009.

[23] M. L. D. Wong, M. Zhang, and A. K. Nandi, “Effects of compressive sensing on classification of bearing faults with Entropic features” *EUSIPCO, IEEE*, pp. 2296–2300, 2015.

[24] G. Tang, Q. Yang, H. Q. Wang, G. gang Luo, and J. wei Ma, “Sparse classification of rotating machinery faults based on compressive sensing strategy,” *Mechatronics*, vol. 31, pp. 60–67, 2014.

[25] X. Zhang, N. Hu, L. Hu, L. Chen, and Z. Cheng, “A bearing fault diagnosis method based on the low-dimensional compressed vibration signal,” *Adv. Mech. Eng.*, vol. 7, no. 7, pp. 1–12, 2015.

[26] X. Chen, Z. Du, J. Li, X. Li, and H. Zhang, “Compressed sensing based on dictionary learning for extracting impulse components,” *Signal Processing*, vol. 96, no. PART A, pp. 94–109, 2014.

[27] G. Tang, W. Hou, H. Wang, G. Luo, and J. Ma, “Compressive sensing of roller bearing faults via harmonic detection from under-sampled vibration signals,” *Sensors (Switzerland)*, vol. 15, no. 10, pp. 25648–25662, 2015.