The International Journal of Advanced Manufacturing Technology
https://doi.org/10.1007/s00170-018-2902-0

ORIGINAL ARTICLE

Adaptive online dictionary learning for bearing fault diagnosis

Yanfei Lu [1] · Rui Xie [2] · Steven Y. Liang [1,3]

Received: 7 August 2018 / Accepted: 18 October 2018

Abstract

One of the most common parts to maintain system balance and support the various load in rotating machinery is the rolling element bearing. The breakdown of the element in bearings leads to inefficiency and sometimes catastrophic events across various industries. The main challenge over the last few years for fault diagnosis of bearings is the early detection of fault signature. In this paper, an adaptive online dictionary learning algorithm is developed for early fault detection of bearing elements. The dictionary is trained using a set of vibration signal from a heavily damaged bearing. The enveloped signal of the bearing is obtained using the Kurtogram and split into several sections. The K-SVD algorithm is implemented to the dictionaries corresponding to the enveloped signal. OMP is implemented with the calculated dictionaries to obtain the sparse representation of the testing signal. Then the envelope analysis is implemented to obtain the fault signal from the recovered signal by the dictionaries. The adaptive algorithm is added to the dictionary learning to update the dictionary based on newly acquired data with the weighted least square method. Without retraining the dictionaries using the K-SVD algorithm, the computation speed is significantly improved. The proposed algorithm is compared with a traditional dictionary learning algorithm to show the improvement in detection of new fault frequency and improved signal to noise ratio.

Keywords Ball bearing · Fault diagnosis · Dictionary learning · Adaptive algorithm

1 Introduction

Rolling element bearing, the most common component in rotating machineries, serves in various industrial applications. Because of the commonality of usage and the importance of normal operation, the study of diagnosis and prognosis of bearings has drawn great attention for researchers and industries. Once the crack in the bearing subsurface grows to a certain size, the degradation process is accelerated because the crack has become unstable. Therefore, the early detection of bearing degradation is critical in maintaining continuous operation of the machinery. In the last decade, the concept of condition-based maintenance (CBM) was proposed to better accommodate the industry requirements of machineries. A critical step to perform CBM is to accurately diagnose the severity and location of a fault [1–3].

A general guideline for bearing diagnosis proposed by Randall and Antoni is shown in [4]. In the proposed tutorial, a phenomenon of strong masking signals from machine components is reported. Because of a strong background noise generated by the machine component and acquisition system, various signal processing techniques are implemented to reduce the system noise and extract the important information. The most widely used methods are categorized into three different types: time-domain, frequency-domain, and time-frequency analysis. The early research mainly focuses on the time-domain analysis. The most general method includes calculating the root mean square (RMS), skewness, kurtosis, and peak values of the acquired vibration signal [5–7]. The abovementioned methods still serve as the foundations of bearing diagnosis nowadays. Over the last few decades, the method of discrete/random separation is developed to remove the effect of small fluctuations in rotating speed for various machines. Bonnarot et al. implemented the unsupervised noise cancelation in the angular domain to resample the data in the time domain for the removal of random noise [8]. The method significantly improves the signal to noise ratio of the

✉ Yanfei Lu
ylu318@gatech.edu

1 George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

2 Department of Statistics at the University of Georgia, Athens, GA 30602, USA

3 College of Mechanical Engineering, Donghua University, Shanghai 201620, China

Published online: 31 October 2018

vibration signal of bearings. McFadden implemented the time averaging method to remove the undesirable noise, but the proposed model is only limited to certain conditions [9]. Sawalhi and Randall implemented the autoregressive (AR) model to remove the regular gear meshing noise in [10]. The residual signal after the AR model reveals the periodical information of the rotating component within the gearbox. In general, the time-domain method has several disadvantages. The parameters of the time-series model various over the degradation process of rotating machineries as indicated in [11]. Therefore, an adaptive algorithm is required in general [12]. Although various time synchronous average methods are proposed, the method still relies on auxiliary devices to acquire information other than the vibrational signal [13]. Although several tacholess order tracking method can be implemented in limited conditions [14–16]. The true tacholess order tracking method is still under development for early fault detection.

The frequency domain approach usually involves the Fourier transform, Hilbert transform, and various demodulation techniques. The Hilbert transform is the relationship between the real and imaginary parts of the Fourier transform [4]. Rai and Mohant demonstrate the effectiveness of using Hilbert transform incorporated with the empirical mode decomposition in bearing diagnosis [17]. The proposed method successfully extracted the fault signature while the traditional fast Fourier transform cannot achieve. Modulation occurs when a carrier signal's amplitude or frequency varies with respect to time [18]. Stack et al. implemented an amplitude modulation detector for the detection of bearing fault [19]. The method reduces its difficulty by only requiring the knowledge of bearings characteristic fault frequencies. Gong and Qiao implemented the current demodulation technique to identify the bearing fault [20]. The proposed method has the advantage of ease of implementation because only the current need to be monitored.

The most widely used signal processing technique nowadays is the time-frequency domain analysis. The Short-Time Fourier transform, which is equivalent to performing Fourier transform in a moving data window, is the simplest to implement [21]. Another popular technique is the wavelet analysis. Qiu et al. has successfully implemented the wavelet filter to detect weak signature in rolling element bearings [22]. Lei et al. combined the wavelet packet transform with Kurtogram to improve the performance of fault extraction [23]. Segreto et al. implemented the wavelet packet transform in combination with neural network to monitor vibration during the machining process [24]. A review of the applications of the wavelet transform can be found in [25]. A widely used technique for signal denoising nowadays is the Kurtogram. Kurtogram provides critical information for the spectral kurtosis analysis. Antoni first proposed using the spectral kurtosis (SK) to detect the transients of vibration signal in [26]. Later, Antoni implemented the fast Kurtogram in combination with envelope analysis to reduce the computation complexity while maintaining accurate result [27].

Online dictionary learning [28] is an efficient tool for signal denoising and feature extraction. The technique is first implemented to analyze multi-dimensional data [29]. It has the capability of adaptation to dynamic train sets and fast convergence. Over the years, various technologies are developed to train the dictionary and find the sparse representation of the signal to be analyzed. Aharon et al. presents the K-SVD method to calculate the dictionaries based on the training data and parameter [30]. Various versions and modifications to improve the computation speed of the K-SVD algorithms are presented in [31–33]. Once the dictionaries are obtained, a sparse representation of the signal needed to be processed is calculated by the orthogonal matching pursuit (OMP) algorithm. A wide variety of the OMP algorithms are listed in [34–36]. The dictionary learning methods have been used for bearing fault diagnosis [37, 38], which restricted diagnosis signal to shift-invariant in time and assumed hidden Markov model structure.

In this paper, an adaptive online dictionary learning algorithm for bearing fault diagnosis is proposed. Comparing to the previous work using dictionary learning methods, our proposed work is free from assumptions on the shift-invariant property and allow online update for newly collected signals. The model uses the Kurtogram and envelope analysis to remove undesirable system noise and locate the fault frequency band. The dictionaries are first trained using a set of vibrational data with distinctive fault frequencies. After the training of the dictionaries, a new set of training data is input to the diagnostic model. A weighted least square algorithm is implemented to update the weights of the dictionaries rather than recalculating the dictionaries using the K-SVD algorithm. The diagnostic model is compared with a wavelet diagnostic model to demonstrate its capability of early fault detection.

2 Diagnostic model

The proposed diagnostic model is shown in Fig. 1. The vibration signal of the bearing is first denoised using the Kurtogram and envelope analysis [27]. The filtered signal is input into the online dictionary learning algorithm. The online dictionary learning algorithm calculates the desired dictionaries based on selected parameters. The dictionaries are tested using a series of data belonging to the same training set to validate successful training of the dictionary. The testing bearing vibration signal is used as the input for the OMP algorithm, and the OMP outputs the sparse representation of the testing signal. Once the dictionary training yield satisfactory result, the training for the dictionary is terminated. The dictionary update starts with a new set of training data. The sparse representation of the training data is obtained through the OMP algorithm. A

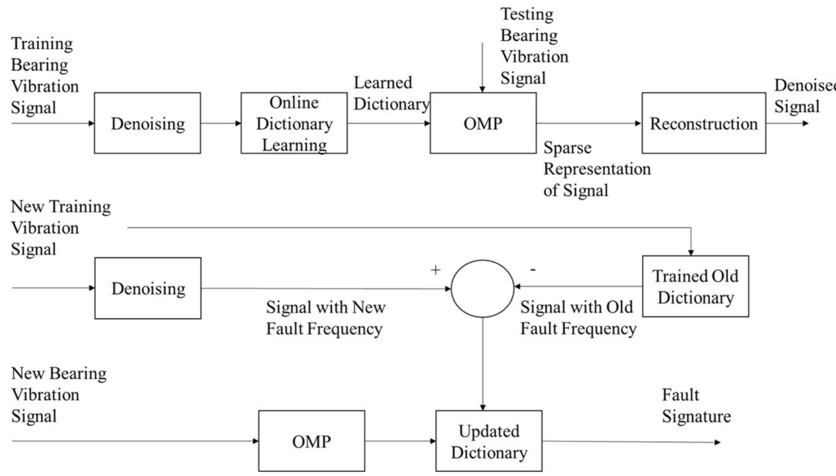

Fig. 1 Adaptive bearing diagnostic model

weighted least square algorithm is implemented in the dictionary update process. The updated dictionary is able to distinguish additional fault frequency in the bearing vibration signal.

2.1 Kurtogram

Spectral Kurtosis is an ideal tool for bearing diagnosis because of its capability to detect weak transient signal regardless of the additive noise. The implementation of the SK assumes that the vibration signal of bearings can be separated as shown in (1):

$$y(t) = x(t) + n(t) \tag{1}$$

where $y(t)$ is the measured vibration signal, $x(t)$ is the fault signal with transients, and $n(t)$ is the stationary system noise. The SK is generally defined as (2) [39]:

$$SK_X(f) = \frac{\left\langle |H(n,f)|^4 \right\rangle}{\left\langle |H(n,f)|^2 \right\rangle^2} - 2 \tag{2}$$

where $H(n,f)$ is the complex envelope of signal at frequency f, and $\langle|H(n,f)|^4\rangle$ is the temporary average of the envelope signal. The SK of a non-stationary signal is presented by (3) [39]:

$$SK_Y(f) = \frac{SK_X(f)}{\left[1 + \frac{1}{SNR(f)}\right]^2} \tag{3}$$

where $SK_Y(f)$ is the SK of the signal $y(t)$, $SK_X(f)$ is the SK of the signal $x(t)$, and $SNR(f)$ is the signal to noise ratio. When the signal to noise ratio is high, $SK_Y(f)$ is approximately equivalent to $SK_X(f)$. When the signal to noise ratio is low, $SK_Y(f)$ approaches zero. Therefore, by searching the whole frequency-domain, the SK is capable of distinguishing the fault signal from system noise.

The fast Kurtogram examines a dyadic grid in the $(f, \Delta f)$ plane instead of searching the whole plane to increase the efficiency of computation. The mean of the vibration signal is first removed, and the signal is filtered by an auto-regressive filter of length of 100. A bandpass prototype filter $h(n)$ with cutoff frequency of $f_c = 1/8 + \varepsilon$, where $\varepsilon \geq 0$, is created to establish the frequency bands to be filtered. The low-pass and high-pass filters from $h(n)$ are represented by,

$$h_l(n) = h(n)e^{j\pi n/4}, h_h(n) = h(n)e^{j3\pi n/4} \tag{4}$$

The two filters are used to decompose the signal iteratively with each level consisting of 2^k bands. The signal from the ith filter is denoted by $c_k^i(n)$. The low-pass and high-pass filtered signals are represented as $c_{k+1}^{2i}(n)$ and $c_{k+1}^{2i+1}(n)$, respectively, at a decomposition level of K-1. The number of filtered signal is increased by a factor of 2 at each level; meanwhile, the respective length of signal is decreased by a factor of 2. The central frequency of the complex envelope of signal $x(n)$ is represented by,

$$f_i = \left(i + 2^{-1}\right)2^{-k-1} \tag{5}$$

and the bandwidth is:

$$(\Delta f)_k = 2^{-k-1} \tag{6}$$

The kurtosis is computed for all $c_k^i(n)$ for $i = 0, ..., 2^k - 1$, $k = 0, ..., K - 1$. Based on (2), the Kurtogram equals to:

$$K_k^i = \frac{\left\langle \left|c_k^i(n)\right|^4 \right\rangle}{\left\langle \left|c_k^i(n)\right|^2 \right\rangle^2} - 2 \tag{7}$$

The envelope analysis has gained popularity in bearing diagnosis because of its effectiveness and simplicity of implementation [40]. Because of the well-known knowledge of this method, the detailed methodology presented in [41, 42] is ignored here. The enveloped analysis in this paper is executed by filtering the signal using a bandpass filter around the center frequency and squaring the filtered signal after the Kurtogram. The filtered signal is then low-pass filtered to obtain the envelope signal.

2.2 Online dictionary learning

The sparse representation of signals is a popular research topic in multi-dimensional signal processing and data compression in the recent years. By implementing an over-complete dictionary $D \in \mathbb{R}^{n\times k}$ which contains n as the number of rows and k the number of columns of data, a signal $y \in \mathbb{R}^n$can be represented as a linear combination of the dictionaries denoted as $y = Dx$, where $x \in \mathbb{R}^k$is the sparse representation of the signal y. The object is to find either:

$$\min_x \|x\|_0 \ subject \quad to\, y = Dx \tag{8}$$

or

$$\min_x \|x\|_0 \ subject \quad to\, \|y-Dx\|_2 \le \varepsilon \tag{9}$$

where ε is the error tolerance, which is selected to be 0.01 to ensure relatively low computation effort while maintaining reconstruction accuracy. The online dictionary learning algorithm aims to find the sparser representations of the signal by solving the optimization problem defined in:

$$\min_{D,x} \|y-Dx\|_F^2 \ subject \quad to \forall i\, \|x_i\|_0 \le K \tag{10}$$

where K is the targeted sparsity [36]. The algorithm consists of two alternating steps: (1) given current dictionary estimate, calculating the sparse representation matrix x which is normally done by using the OMP introduced in section 2. Updating dictionary one atom at a time to optimize the target function defined in (10) for each atom individually. The resulting problem is a rank-1 approximation task shown in (11):

$$\{a, b\} \coloneqq \underset{a,b}{Argmin} \|E-ab^T\|_F^2 \ subject \quad to\, \|a\|_2 = 1 \tag{11}$$

where E is the error matrix, a is the updated atom and b^T is the new coefficients row in x. The problem can be directly solved by singular value decomposition (SVD) [43] or iterative updated method such as gradient descent or Newton methods [44].

2.3 OMP

The OMP algorithm implemented takes an $m \times n$ data matrix y, an $m \times l$ dictionary D and the desired sparsity level of k and outputs a $l \times n$ sparse representation of the data matrix y. The estimate for the ideal signal is obtained by Dy. The general procedure usually involves the following:

1. Initialize the residual defined by $e = y - Dx$ as $e = y$, initialize a set Λ_p containing p elements from $\{1, ..., l\}$ and set $\Lambda_0 = \varnothing$ and counter $t = 1$.
2. Find the index λ_t that solves the optimization problem defined in (12):

$$\lambda_t = Argmax_{i=1,\dots,l} |e_{t-1}, \varphi_j| \tag{12}$$

where φ is the columns of D.

3. Augment Λ_p and the matrix of chosen atoms: $\Lambda_p = \Lambda_{p-1} \cup \{\lambda_t\}$ and $D_t = \left[D_{t-1}\, \varphi_{\lambda_t}\right]$.
4. Solve a least square problem to get the new signal estimation:

$$\hat{y}_t = Argmin_y \|y-D_t\hat{y}\|_2 \tag{13}$$

5. Calculate the new approximation of the data $a_t$and the new residual as

$$a_t = D_t\hat{y}_t \tag{14}$$

$$e_t = y-a_t \tag{15}$$

6. Increase t, and return to Step 2 until $t < p$.
7. The estimate Dy has nonzero indices at the components listed in Λ_p and the value of the estimation in component λ_i equals the ith component of $\hat{y}_t$.

2.4 Dictionary update

The dictionary weight update is performed using an iteratively reweighted least square algorithm [45]. The error E between the new signal $Y$and the reconstructed signal $D_{old}X$ using the old dictionary is shown in (16):

$$E = Y-D_{old}X \tag{16}$$

The weight is selected as shown in (17):

$$W = diag\left[\frac{1}{\|e_1\|_2^2}, \dots, \frac{1}{\|e_n\|_2^2}\right] \tag{17}$$

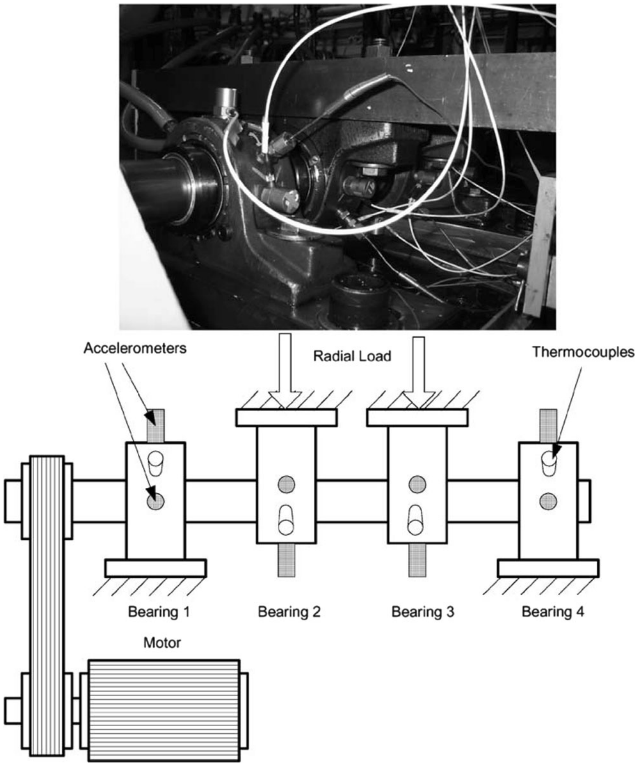

Fig. 2 Experimental setup for monitoring of bearing vibration [22]

where e_n is the individual term within E. The updated dictionary is:

$$D_{new} = YWX^T\left(XWX^T\right)^{-1} \tag{18}$$

The updating scheme is executed iteratively until the average value of the error E converges to the desirable value. The desirable value should be set with consideration of the actual applications. In this paper, E is selected to be 0.50. By continuous updating the dictionary during the degradation process of bearings, more accurate diagnostic information can be obtained.

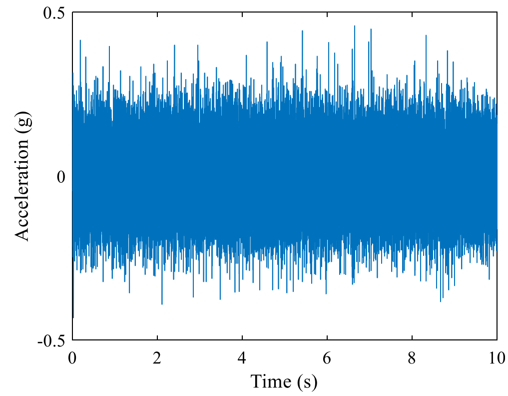

Fig. 4 Dictionary training data

3 Experiment

To validate the proposed diagnostic method, the experimental data from Qiu et al. [22] is used. The bearing run-to-failure test setup is shown in Fig. 2 with four Rexnord ZA-2115 double row roller bearings mounted. The bearings have a total of 32 rollers. The pitch diameter is 2.815 in., and the roller diameter is 0.331 in. The tapered contact angle is 15.2°. The shaft is driven by an AC motor at 2000 RPM. A 6000 lbs. radial load is applied to the shaft and bearing through a spring mechanism as shown in the experimental setup. A magnetic plug is installed to collect the debris in the lubricant. Once the accumulated debris exceed the preset value, the test will stop. The data is taken every 10 min and the sampling frequency is 20,480 Hz. The tested bearing had experienced outer race failure by the end of the test. The outer race failure frequency is calculated to be 236.4 Hz.

The overall vibration signal is shown in Fig. 3. To simulate an online monitoring scenario, the bearing vibration data from 3900 min to 3990 min are used to train the dictionary. The training data is shown in Fig. 4. To test the updating algorithm for the dictionary learning, the vibration data from 9750 min is used.

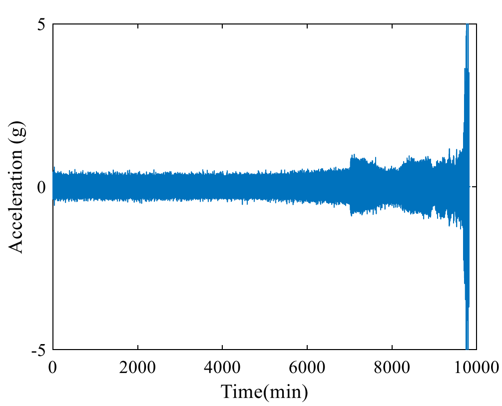

Fig. 3 Vibration signal of bearing run-to-failure test

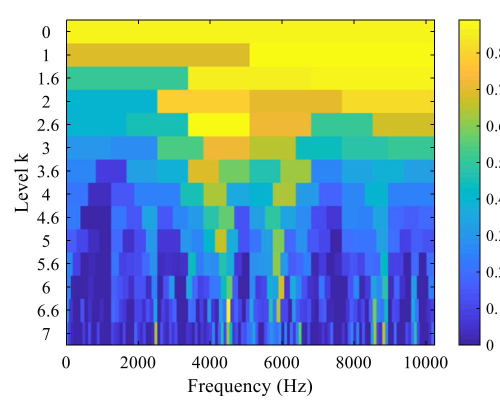

Fig. 5 Kurtogram of the training signal

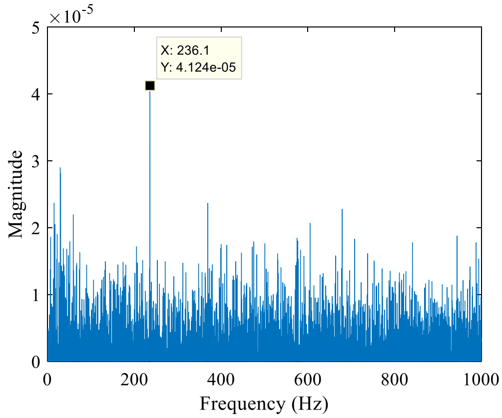

Fig. 6 Amplitude spectrum of filtered data

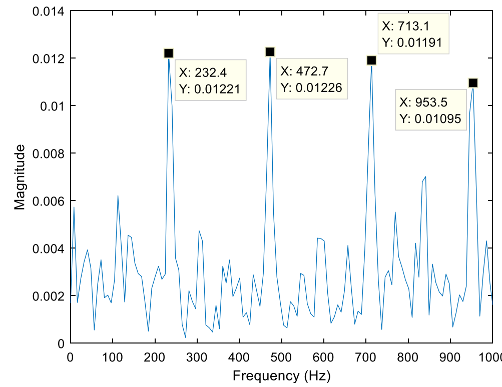

Fig. 8 Frequency spectrum of filtered data using new dictionaries

4 Results

The training data are first analyzed by the Kurtogram. The mean is first subtracted from the vibration signal and a 16th order Hanning window low-pass filter with the cutoff frequency equivalent to 40% of the Nyquist frequency based on the result from [26] is implemented to remove the random noise of the signal. The center frequency of the system resonance is determined to be 7680 Hz by the sampling rate and the decomposition level, and the bandwidth is calculated to be 5120 Hz [27]. The bandpass filter is set to keep signals within 2560 Hz and 12,800 Hz. The Kurtogram for the training data is shown in Fig. 5. Based on the magnitude of the Kurtogram, the filtered signal in level 1 is selected, and the amplitude spectrum of the filtered data is shown in Fig. 6.

It can be observed that by using the Kurtogram at the early degradation stage, the fault signature can be extracted. The filtered data is input into the K-SVD algorithm to train the initial dictionary parameter. The number of elements in each linear combination is set from 1 to 20. The number of dictionary element is set to be 50 with 15 iterations of executing the K-SVD algorithm. The whole training is repeated for 20 times and for each time, the amplitude at the fault frequency and its second harmonic is recorded. A criterion is established as:

$$Factor = \frac{A_{33\pm5} + A_{236\pm5} + A_{472\pm5}}{A_{all}} \tag{19}$$

where $A_{33\pm5}$ denotes the sum of amplitude of shaft rotating frequency interval, $A_{236\pm5}$ represents the sum of amplitude of the fault frequency interval, $A_{472\pm5}$ is the sum of amplitude of the harmonic frequency interval, and A_{all} is the sum of amplitude of all the frequencies together. The dictionary which yields the largest factor is selected for future signal denoising.

To demonstrate the effectiveness of the adaptive online dictionary algorithm, a new set of data with distinctive fault frequency, second, third and fourth harmonics are used in the dictionary learning algorithm. The filtered signal without updating the dictionary parameter is shown in Fig. 7 with an average error of 0.59.

It can be observed that the dictionaries are still able to recover the significant information from the training set used before. However, it is not able to detect the added third and fourth harmonics in the signal. After the dictionaries are updated using the proposed iteratively reweight least square method. The added harmonics can be detected using the new dictionaries as shown in Fig. 8 with an average error of 0.4203.

By comparing Figs. 7 and 8, it can be concluded that the update of the dictionary improves the detection of fault signature in frequency domain and increase the amplitude of the detected fault by a significant amount. In addition, more fault signatures are revealed by using this proposed dictionary updating method rather than recomputing the dictionary, which is time-consuming and cost-inefficient.

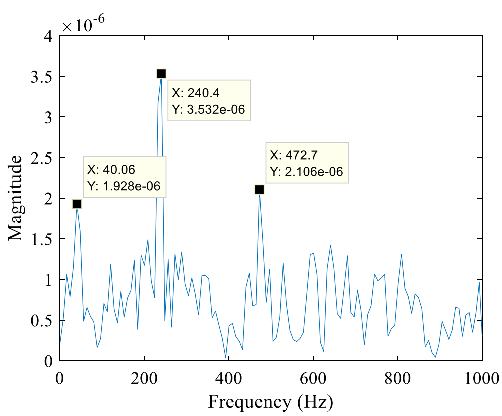

Fig. 7 Frequency spectrum of filtered data using old dictionaries

5 Conclusions

In this paper, an adaptive online dictionary learning bearing diagnostic model is proposed. The Kurtogram is initially implemented to find the weak fault signatures. The filtered signal is used to train the initial dictionaries using the K-SVD algorithm with parameters selected by trial and error. The trained dictionary is able to capture the features obtained by the Kurtogram without reperforming the analysis for a second time. As the damage of the bearing becomes more severe, new fault or harmonics are present in the system. The dictionary updating algorithm has the capability of distinguishing the newly added fault feature while the old dictionary is not able to detect the newly developed damage. The adaptive online dictionary learning also has the benefits of short computation time because the dictionary training algorithm such as the K-SVD does not need to be performed with new training data. The proposed model creates a framework for the online monitoring of bearing degradation process. It is more robust and flexible than a fixed model such as the wavelet denoising, and it can combine the benefits of multiple denoising techniques. Future research will be performed on the detection of multiple fault signatures and vary the structures of the dictionaries.

Author contributions Y.L. and R.X. created the model and analyzed the data; S.Y.L. provided feedback of the concept; Y.L. and R.X. wrote the paper.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Lu Y, Li Q, Pan Z, Liang SY (2018) Prognosis of bearing degradation using gradient variable forgetting factor RLS combined with time series model. IEEE Access 6:10986–10995
2. Liang SY, Li Y, Billington SA, Zhang C, Shiroishi J, Kurfess TR, Danyluk S (2014) Adaptive prognostics for rotary machineries. Procedia Eng 86:852–857
3. Kurfess TR, Billington S, and Liang SY (2006) *Advanced diagnostic and prognostic techniques for rolling element bearings*, in *Condition monitoring and control for intelligent manufacturing*. Springer. p. 137–165
4. Randall RB, Antoni J (2011) Rolling element bearing diagnostics—a tutorial. Mech Syst Signal Process 25(2):485–520
5. Igba J, Alemzadeh K, Durugbo C, Eiriksson ET (2016) Analysing RMS and peak values of vibration signals for condition monitoring of wind turbine gearboxes. Renew Energy 91:90–106
6. Li Y (1999) *Dynamic prognostics of rolling element bearing condition*. Georgia Institute of Technology
7. Samanta B, Al-Balushi K (2003) Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. Mech Syst Signal Process 17(2):317–328
8. Bonnardot F, Randall R, and Antoni J Enhanced unsupervised noise cancellation using angular resampling for planetary bearing fault diagnosis
9. McFadden P, Smith J (1984) Vibration monitoring of rolling element bearings by the high-frequency resonance technique—a review. Tribol Int 17(1):3–10
10. Sawalhi N, Randall R (2008) Helicopter gearbox bearing blind fault identification using a range of analysis techniques. Aust J Mech Eng 5(2):157–168
11. Lu Y, Li Q, and Liang SY (2017) *Adaptive prognosis of bearing degradation based on wavelet decomposition assisted ARMA model*. in *Technology, Networking, Electronic and Automation Control Conference (ITNEC),* 2017 *IEEE 2nd Information*. IEEE
12. Bai M, Huang J, Hong M, Su F (2005) Fault diagnosis of rotating machinery using an intelligent order tracking system. J Sound Vib 280(3–5):699–718
13. Wang J, Peng Y, Qiao W (2016) Current-aided order tracking of vibration signals for bearing fault diagnosis of direct-drive wind turbines. IEEE Trans Ind Electron 63(10):6336–6346
14. Wang Y, Xie Y, Xu G, Zhang S, Hou C (2017) Tacholess order-tracking approach for wind turbine gearbox fault detection. Front Mech Eng 12(3):427–439
15. Zhao M, Lin J, Wang X, Lei Y, Cao J (2013) A tacho-less order tracking technique for large speed variations. Mech Syst Signal Process 40(1):76–90
16. Zhao M, Lin J, Xu X, Lei Y (2013) Tacholess envelope order analysis and its application to fault detection of rolling element bearings with varying speeds. Sensors 13(8):10856–10875
17. Rai V, Mohanty A (2007) Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform. Mech Syst Signal Process 21(6):2607–2615
18. Randall RB (2011) Vibration-based condition monitoring*: industrial,* aerospace *and automotive applications*. John Wiley & Sons
19. Stack JR, Harley RG, Habetler TG (2004) An amplitude modulation detector for fault diagnosis in rolling element bearings. IEEE Trans Ind Electron 51(5):1097–1102
20. Gong X, Qiao W (2013) Bearing fault diagnosis for direct-drive wind turbines via current-demodulated signals. IEEE Trans Ind Electron 60(8):3419–3428
21. Griffin D, Lim J (1984) Signal estimation from modified short-time Fourier transform. IEEE Trans Acoust Speech Signal Process 32(2): 236–243
22. Qiu H, Lee J, Lin J, Yu G (2006) Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. J Sound Vib 289(4–5):1066–1090
23. Lei Y, Lin J, He Z, Zi Y (2011) Application of an improved kurtogram method for fault diagnosis of rolling element bearings. Mech Syst Signal Process 25(5):1738–1749
24. Segreto T, Caggiano A, Karam S, Teti R (2017) Vibration sensor monitoring of nickel-titanium alloy turning for machinability evaluation. Sensors 17(12):2885
25. Yan R, Gao RX, Chen X (2014) Wavelets for fault diagnosis of rotary machines: a review with applications. Signal Process 96:1–15
26. Antoni J (2006) The spectral kurtosis: a useful tool for characterising non-stationary signals. Mech Syst Signal Process 20(2):282–307
27. Antoni J (2007) Fast computation of the kurtogram for the detection of transient faults. Mech Syst Signal Process 21(1):108–124
28. Mairal J, Bach F, Ponce J, and Sapiro G (2009) *Online dictionary learning for sparse coding*. In *Proceedings of the 26th annual international conference on machine learning*. ACM
29. Mairal J, Bach F, Ponce J, Sapiro G (2010) Online learning for matrix factorization and sparse coding. J Mach Learn Res 11(Jan): 19–60
30. Aharon M, Elad M, Bruckstein A (2006) $ rm k $-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Signal Process 54(11):4311–4322
31. Rubinstein R, Peleg T, Elad M (2013) Analysis K-SVD: a dictionary-learning algorithm for the analysis sparse model. IEEE Trans Signal Process 61(3):661–677
32. Wang L, Lu K, Liu P, Ranjan R, Chen L (2014) IK-SVD: dictionary learning for spatial big data via incremental atom update. Comput Sci Eng 16(4):41–52
33. Jiang, Z., Z. Lin, and L.S. Davis. *Learning a discriminative dictionary for sparse coding via label consistent K-SVD*. In *Computer*

Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. 2011. IEEE
34. Tropp JA, Gilbert AC (2007) Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theory 53(12):4655–4666
35. Donoho DL, Tsaig Y, Drori I, Starck J-L (2012) Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. IEEE Trans Inf Theory 58(2):1094–1121
36. Rubinstein R, Zibulevsky M, Elad M (2008) Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit. Cs Technion 40(8):1–15
37. Zhou H, Chen J, Dong G, Wang R (2016) Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model. Mech Syst Signal Process 72:65–79
38. Yang B, Liu R, Chen X (2017) Fault diagnosis for a wind turbine generator bearing via sparse representation and shift-invariant K-SVD. IEEE Trans Ind Inform 13(3):1321–1331
39. Antoni J, Randall R (2006) The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. Mech Syst Signal Process 20(2):308–331
40. Borghesani P, Ricci R, Chatterton S, Pennacchi P (2013) A new procedure for using envelope analysis for rolling element bearing diagnostics in variable operating conditions. Mech Syst Signal Process 38(1):23–35
41. Antoni J (2007) Cyclic spectral analysis of rolling-element bearing signals: facts and fictions. J Sound Vib 304(3–5):497–529
42. Randall RB, Antoni J, Chobsaard S (2001) The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. Mech Syst Signal Process 15(5):945–962
43. Golub GH, Reinsch C (1970) Singular value decomposition and least squares solutions. Numer Math 14(5):403–420
44. Nocedal J, Wright SJ (2006) Conjugate gradient methods. Numerical optimization:101–134
45. Daubechies I, DeVore R, Fornasier M, Güntürk CS (2010) Iteratively reweighted least squares minimization for sparse recovery. Commun Pure App Mathematics 63(1):1–38