Model-Based Intelligent Non-linear Signal Recognition for Gearbox Condition Monitoring

Hanxin Chen[1,2(✉)], Lang Huang[1,2], Yuzhuo Miao[1,2], Qi Wang[1,2], Liu Yang[1,2], and Yao Ke[1,2]

[1] School of Mechanical and Electrical Engineering, Wuhan Institute of Technology, Wuhan 430073, China
pg01074075@163.com
[2] Hubei Provincial Key Laboratory of Chemical Equipment, Intensification and Intrinsic Safety, Wuhan, China

Abstract. In this paper, a method for equipment fault diagnosis of gearbox using principal component analysis (PCA) and sequential probability ratio test (SPRT) is proposed. The method is to study and monitor the working state of the gearbox by studying the original vibration signal of the gearbox, and establish a corresponding experimental model by using the normal gear and the fault gear, respectively. Firstly, the vibration signal of the gearbox is preprocessed by wavelet packet transform (WPT). Then the time domain signal analysis method is used to extract the characteristic parameters of the vibration signal and the data is reduced by PCA. After the data are reduced in dimension, the principal element with the highest contribution rate are selected as the input parameter of SPRT. Test parameters to verify the proposed SPRT algorithm and Root Mean Square Error (RMSE). The results show that the proposed method is effective and practical.

Keywords: PCA · SPRT · RMSE · Condition monitoring

1 Introduction

Gearbox is commonly used in industrial production [1]. Gearbox is an important part of the power transmission of many mechanical devices. It is used to change the transmission ratio, direction of change, and power transmission. In the actual production life, the operating state of the gearbox is closely related to the operating state of the equipment, while the gears, shafts and other components in the gearbox work under high load for a long time, and the probability of mechanical failure is higher [2–4]. According to relevant statistics, the gear failure is about 60% of the failure of the gearbox components. Therefore, research on gear fault diagnosis in gearboxes has become very valuable and necessary [5, 6].

As early as 1901, PCA was first proposed by Pearson. The basic idea of the method is to compress multiple linear correlation variables into a multivariate statistical method of a few unrelated variables [7, 8]. Ding et al. reduced the amount of data dimension

X. Yuan et al. (Eds.): ICUIA 2020, CCIS 1319, pp. 99–106, 2020.
https://doi.org/10.1007/978-981-33-4601-7_10

by PCA [9]. PCA reduced the multidimensional correlated variable to be low dimensional independent eigenvector. PCA has advantages in the simple concept, convenient calculation. It has been widely used in numerous areas such as data compression, feature extraction, image processing, signal analysis etc. al [10].

In 1947, Wald introduced SPRT in his book. Because of its simplicity and high efficiency, this method has been widely used in the field of fault diagnosis in recent years. Ray et al. applied SPRT to fault detection and identification in nuclear power plants and aircraft [11]. Chen et al. used SPRT for gearbox fault diagnosis [12, 13]. Compared with the traditional method, Compared with the traditional method, sequential probability ratio test differs from the fixed sample test in that this algorithm requires a smaller average sample size, but it has a higher test efficiency [14].

This paper proposes a new gear multi-fault monitoring method based on PCA and SPRT. The gear vibration signal is denoised by wavelet packet transform, and then PCA is taken to remove the characteristic data of the signal and select the appropriate principal element as the substitute data of SPRT. Finally, SPRT algorithm is used to identify the fault state of the gearbox. With the aim for corroborating the potency of SPRT algorithm for fault diagnosis of the gearbox, the equipment fault diagnosis of the gearbox is diagnosed by jointing the root mean square error (RMSE) and the binary SPRT. The final consequences show that the strategy is effective and dependable.

2 SPRT Algorithm of Gearbox

In the likelihood ratio test, the change in mean and standard deviation has a greater impact on the results. The test sequence selected after pre-processing and PCA basically conforms to the Gaussian distribution, and the mean value is $c\mu$, and the standard deviation is σ. Suppose that the probability distribution of one of the groups to be tested satisfies the initial hypothesis: $H_j : \mu = \mu_j$; the probability distribution of another group of sequences to be tested satisfies the alternative hypothesis: $H_j : \mu = \mu_j$.The standard deviation σ does not change. When the both hypotheses are reality, the joint probability density function (PDF) of the two sets of sequences is as follows:

$$P_{ik}(y_k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(y_k - \mu_i)^2\right) \tag{1}$$

$$P_{jk}(y_k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(y_k - \mu_j)^2\right) \tag{2}$$

Where $P_{jk}(y_k)$ represents PDF under selective hypothesis; $P_{ik}(y_k)$ represents PDF under null hypothesis.

We could calculated the likelihood ratio of SPRT as follows:

$$\lambda_{i,j}(Y_{Sm}) = \frac{\prod_{k=1}^{n} p_{jk}}{\prod_{k=1}^{n} p_{ik}} = \frac{p_{j1}(y_1)p_{j2}(y_2)\cdots p_{jk}(y_k)}{p_{i1}(y_1)p_{i2}(y_2)\cdots p_{ik}(y_k)} \times \frac{p_{j0}}{p_{i0}} \tag{3}$$

Where $S_m = \{S_1, S_2, S_3, S_4\}$, P_{j0} and P_{i0} are prior probability functions under selective and null hypothesis. SPRT probability ratio $\lambda_{i,j}(Y_{Si})$ and $\lambda_{i,j}\left(Y_{Sj}\right)$ are computed by inputting the testing data $\left(Y_{Si}, Y_{Sj}\right)$ of the signal waveform for gearbox *Si* and *Sj* conditions to Eq. (3). SPRT probability ratios $\lambda_{i,j}(Y_{Si})$ and $\lambda_{i,j}\left(Y_{Sj}\right)$ are expressed in the following section. The Eq. (3) can be simplified:

$$\Delta_{i,j}(Y_{Sm}) = \ln \lambda_{i,j}(Y_{Sm}) = \ln \frac{\prod_{k=1}^{n} p_{jk}}{\prod_{k=1}^{n} p_{ik}} = \sum_{k=1}^{n} \ln \frac{p_{jk}}{p_{ik}} \tag{4}$$

$$E_{i,j}(Y_{Sm}) = \left(\frac{1}{k} \sum_{t=1}^{k} \left(\lambda_{i,j}(Y_{Sm}) - \lambda_{i,j}(Y_S) \right)^2 \right)^{\frac{1}{2}} \quad m = i, j \tag{5}$$

3 Experiment System

The normal gear F1 and the cracked fault gears F2, F3, and F4 were selected for the experiment. The full depth of the gear crack is $a = 2.4$ mm, the full width of the crack is $b = 25$ mm, the thickness of the gear crack is 0.4 mm, and the angle of the gear crack is 45°. In the experiment, the gearbox was in no-load state and the speed was 800 r/min. The parameters of the three sets of gears are shown in Table 1.

Table 1. Four gear failure modes

Gear	Crack parameter			
	Depth/mm	Width/mm	Thickness/mm	Crack angle/°
F1	0	0	0	45°
F2	(1/4)a	(1/4)b	0.4	45°
F3	(1/2)a	(1/2)b	0.4	45°
F4	(3/4)a	(3/4)b	0.4	45°

Figure 1 is a structural diagram of a gearbox. Vibration is generated between the running gears 3 and 4, so one of them is selected to simulate the fault. In this experiment, the gear 3 was selected for the simulation experiment.

The experimental signal acquisition is obtained from the gearbox SpectraQuest dynamic simulator. A PCB352C67 accelerator is fixed in the vertical and horizontal directions of the gearbox. The collected initial vibration signals are input into the PC and stored by the DSP20-42 signal analyzer. This paper only studies the analysis of signals in the horizontal direction. S1, S2, S3, and S4 are used to represent the signals under normal conditions and three fault conditions.

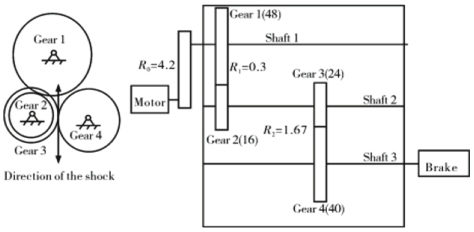

Fig. 1. Gearox structure

4 Multi-fault State Recognition of Gearbox Based on Three-Layer SPRT

4.1 Diagnosis of Gearbox in Fault State F1

The Eqs. (1–3) are used to calculate the variable $\lambda_{1,2}(Y_{S1})$. The parameters μ_0 and μ_1 are the means of the waveform under $(F1, F2)$ condition. The variables $\lambda_{1,2}(Y_{S1}(N))$, $\lambda_{1,2}(Y_{S2}(N))$, $\lambda_{1,2}(Y_{S3}(N))$, $\lambda_{1,2}(Y_{S4}(N))$ are calculated by Eqs. (4–5) by using ten waveforms (Y_{S1}) under $F1$ and one testing waveform of the $F2$, $F3$, $F4$.

Figure 2(a) shows the establishment of the unequal relationship $\lambda_{1,2}(Y_{S1}(N)) < b$. Figure 2(b) shows the variables $\left(E_{1,2}(Y_{S1})\right)$ for $\lambda_{1,2}(Y_{S1}(N))$, $N = 1, \cdots, 10$ are compared with $\left(E_{1,2}\left(Y_{Sj}\right)\right), j = 2, 3, 4$ for $\lambda_{1,2}(Y_{S1}(N))$ with $\lambda_{1,2}\left(Y_{Sj}\right), j = 2, 3, 4$. The parameters $\left(E_{1,2}\left(Y_{Sj}\right)\right), j = 2, 3, 4$ are bigger than $\left(E_{1,2}(Y_{S1})\right)$.

Figure 3(a) shows the establishment of the unequal relationship $\lambda_{1,2}(Y_{S1}(N)) < b$. Figure 3(b) shows the unequal relationship are established, which is $E_{1,3}(Y_{S1}(N)) < E_{1,3}\left(Y_{Sj}\right)$, $j = 2, 3, 4$.

Figure 4(a), the unequal relationship $\lambda_{1,4}(Y_{S1}(N)) < b$ is established. In Fig. 4(b), the parameters $E_{1,4}(Y_{S1})$, $E_{1,4}(Y_{S3})$ and $E_{1,4}(Y_{S4})$ are less than the threshold E_{C1} and RMSEs $E_{1,4}(Y_{S2})$ are more than E_C.

As shown in Figs. (2, 3 and 4), if $E_{1,2}(Y_{S1}) < E_C$ and $\lambda_{1,2}(Y_{S1}) < b$ or $E_{1,3}(Y_{S1}) < E_C$ and $\lambda_{1,3}(Y_{S1}) < b$ are satisfied, then $F1$ could be diagnosed.

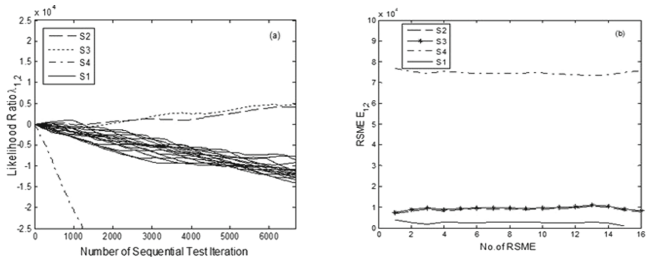

Fig. 2. Three-layer SPRT results: (a) likelihood ratio $\lambda_{1,2}(Y_{S1}(N))$, $\lambda_{1,2}(Y_{S2})$, $\lambda_{1,2}(Y_{S3})$, $\lambda_{1,2}(Y_{S4})$ versus sequential test iteration number. (b) RMSEs $E_{1,2}(Y_{S1}(N))$, $E_{1,2}(Y_{S2})$, $E_{1,2}(Y_{S3})$, $E_{1,2}(Y_{S4})$.

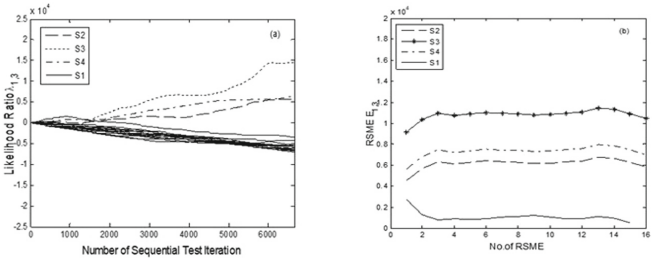

Fig. 3. Three-layer SPRT results: (a) likelihood ratio $\lambda_{1,3}(Y_{S1}(N))$, $\lambda_{1,3}(Y_{S2})$, $\lambda_{1,3}(Y_{S3})$, $\lambda_{1,3}(Y_{S4})$ versus sequential test iteration number. (b) RMSEs $E_{1,3}(Y_{S1}(N))$, $E_{1,3}(Y_{S2})$, $E_{1,3}(Y_{S3})$, $E_{1,3}(Y_{S4})$.

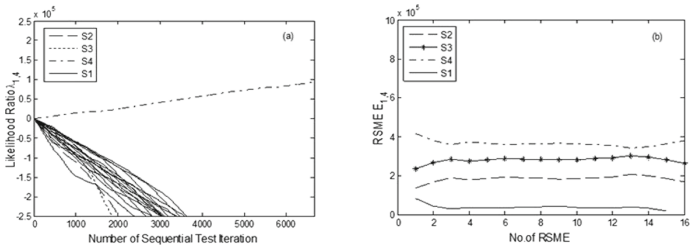

Fig. 4. Three-layer SPRT results: (a) likelihood ratio $\lambda_{1,4}(Y_{S1}(N))$, $\lambda_{1,3}(Y_{S2})$, $\lambda_{1,3}(Y_{S3})$, $\lambda_{1,3}(Y_{S4})$ versus sequential test iteration number. (b) RMSEs $E_{1,3}(Y_{S1}(N))$, $E_{1,3}(Y_{S2})$, $E_{1,3}(Y_{S3})$, $E_{1,3}(Y_{S4})$.

4.2 Diagnosis of Gearbox in Fault State F2

If the waveform is not $S1$ and falls into one of the other three waveforms that is $(S2, S3, S4)$, the variables $\lambda_{2,3}$ and $\lambda_{2,4}$ are used to identify $F2$ by classifying $(S2, S3, S4)$.

As is shown in Fig. 5(a), the inequality that $\lambda_{2,3}(Y_{S2}) < b$ is satisfied. Figure 5(b) shows that $E_{2,3}(Y_{S2}(M)) < E_C$ and $E_{2,3}(Y_{S4}(M)) > E_C$.The inequality $E_{2,3}(Y_S) < E_C$ is satisfied means the condition is $F2$ or $F3$. The variable $\lambda_{2,4}$ is used to classify the waveform $S2$ and $S4$. Figure 6(a) shows $\lambda_{2,4}(Y_{S2}(M)) < b$ is satisfied. Figure 6(b) shows RMSEs $E_{2,4}(Y_{S2}(M)) < E_C$. The condition $F2$ is classified among the three conditions by the variable $\lambda_{2,4}$. The inequalities $E_{2,4}(Y_{S2}) < E_C$ and $\lambda_{2,4}(Y_{S2}) < b$ are the index to diagnose $S2$ from $(S2, S3, S4)$, which is effective to classify $F2$ among $(F2, F3, F4)$.

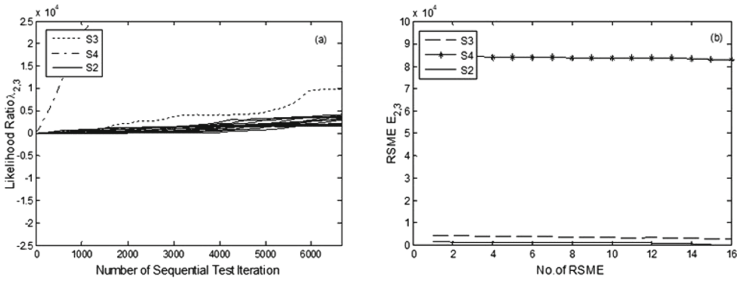

Fig. 5. Three-layer SPRT results: (a) likelihood ratio $\lambda_{2,3}(Y_{S2}(N))$, $\lambda_{2,3}(Y_{S3})$, $\lambda_{2,3}(Y_{S4})$ versus sequential test iteration number. (b) RMSEs $E_{2,3}(Y_{S2}(N))$, $E_{2,3}(Y_{S3})$ and $E_{2,3}(Y_{S4})$.

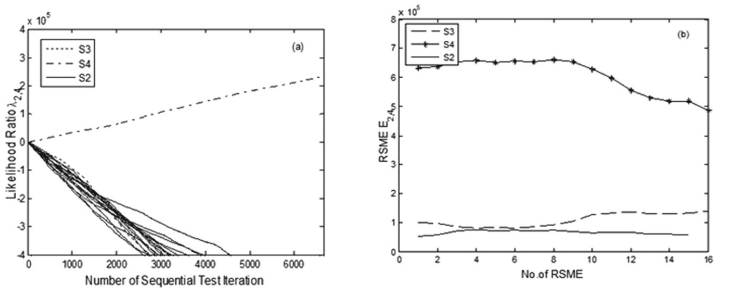

Fig. 6. Three-layer SPRT results: (a) likelihood ratio $\lambda_{2,4}(Y_{S2}(N))$, $\lambda_{2,4}(Y_{S3})$, $\lambda_{2,4}(Y_{S4})$ versus sequential test iteration number. (b) RMSEs $E_{2,4}(Y_{S2}(N))$, $E_{2,4}(Y_{S3})$ and $E_{2,4}(Y_{S4})$.

4.3 Diagnosis of Gearbox in Fault State F3

Calculate the likelihood ratios of the sequences to be tested according to the formulas (1) to (5), labeled $\lambda_{3,4}(Y_{S1}(1)) \cdots \lambda_{3,4}(Y_{S3}(10))$ and $\lambda_{3,4}(Y_{S4})$, and the test results are shown in Fig. 7(a). It can be seen from the figure that the likelihood ratio $\lambda_{3,4}(Y_{S3}(M)) < b$ of the vibration signal $S3$ in the fault state $F3$ and the likelihood ratio $\lambda_{3,4}(Y_{S4}) < b$ of the vibration signal $S4$. Figure 7(b) shows that if $E_{3,4}(Y_{S3}) < E_{3,4}(Y_{S4})$ and $E_{3,4}(Y_{S3}) < E_C$ are satisfied, the inequalities $\lambda_{3,4}(Y_{S3}(M)) < b$ and $E_{3,4}(Y_{S3}) < E_C$ are used

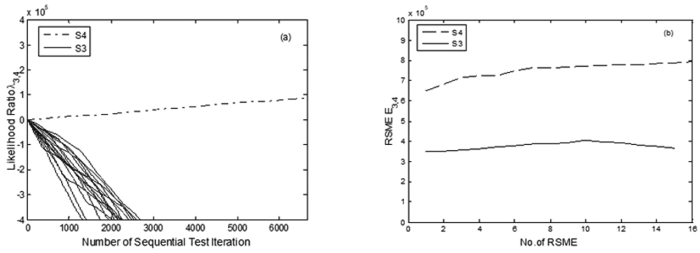

Fig. 7. Three-layer SPRT results: (a) likelihood ratio $\lambda_{3,4}(Y_{S3}(N))$, $\lambda_{3,4}(Y_{S4})$ versus sequential test iteration number; (b) RMSEs $E_{3,4}(Y_{S3}(N))$ and $E_{3,4}(Y_{S4})$.

to categorize the waveform $S3$ from the waveform $S4$. The gearbox condition $F3$ is diagnosed from the gearbox condition $F4$.

5 Conclusions

In this paper, a gearbox fault diagnosis method based on PCA and SPRT is proposed. Firstly, WPT is used to denoise the initial vibration signal, and the time domain analysis method is utilized to extract the characteristic parameters of the signal. Then PCA is used to reduce the extracted feature parameters. The principal component with the biggest contribution rate is chosen as the monitoring sequence. Finally, SPRT algorithm is used to analyze the running state of the gearbox and combine the root mean square algorithm to classify the fault. Subsequently, detailed fault diagnosis experiments and analysis were carried out on the gearbox. SPRT was used to identify different faults and combined with the root mean square error to further give the identification criteria. The experimental results show that the fault diagnosis of the gearbox is accurate and effective.

Acknowledgement. The experimental data is provided by the Reliability Research Lab in the Department of Mechanical Engineering at the University of Alberta in Canada. This work was supported by the National Natural Science Foundation of China (Grant 51775390).

References

1. Zeng, L., Chen, G., Chen, H.: Comparative study on flow-accelerated corrosion and erosion-corrosion at a 900 carbon steel bend. Materials **12**, 1780 (2020). https://doi.org/10.3390/ma13071780
2. Yang, L., Chen, H.: Fault diagnosis of gearbox based on RBF-PF and particle swarm optimization wavelet neural network. Neural Comput. Appl. **31**(9), 4463–4478 (2018). https://doi.org/10.1007/s00521-018-3525-y
3. Chen, H.: Model-based method with nonlinear ultrasonic system identification for mechanical structural health assessment. Trans. Emerg. Telecommun. Technol. e3955 (2020). http://doi.org/10.1002/ett.3955
4. Chen, H.: Nonlinear Lamb wave analysis for microdefect identification in mechanical structural health assessment. Measurement **164**, 108026 (2020). https://doi.org/10.1016/j.measurement.2020.108026
5. Zeng, L., Guo, X.P., Zhang, G.A., Chen, H.X.: Semiconductivities of passive films formed on stainless steel bend under erosion-corrosion conditions. Corros. Sci. **144**, 258–265 (2018)
6. Zeng, L., Shi, J., Luo, J., Chen, H.: Silver sulfide anchored on reduced graphene oxide as a high -performance catalyst for CO2 electroreduction. J. Power Sources **398**, 83–90 (2018)
7. Chen, H., Chen, Y., Yang, L.: Intelligent early structural health prognosis with nonlinear system identification for RFID signal analysis. Comput. Commun. **157**, 150–161 (2020). https://doi.org/10.1016/j.comcom.2020.04.026
8. Yang, L., Hanxin, C.: A novel time-frequency-space method with parallel factor theory for big data analysis in condition monitoring of complex system. Int. J. Adv. Robot. Syst. **17**(2) (2020). https://doi.org/10.1177/1729881420916948
9. Ding, J., Zhao, L., Huang, D.R.: Incipient fault feature extraction method of gearbox based on wavelet package and PCA. In: IEEE International Conference on Data Driven Control and Learning Systems, pp. 656–660 (2017)

10. Chen, H., Fan, D., Huang, J., Huang, W.: Finite element analysis model on ultrasonic phased array technique for material defect time of flight diffraction detection. Sci. Adv. Mater. **12**(5), 665–675 (2020)
11. Ray, A.: Sequential testing for fault detection in multiply-redundant systems. J. Dyn. Syst. Meas. Control Trans. ASME **111**(2), 329–332 (1989)
12. Hanxin, C., Dong, L.F., Lu, F.: Particle swarm optimization algorithm with mutation operator for particle filter noise reduction in mechanical fault diagnosis. Int. J. Pattern Recognit. Artif. Intell. https://doi.org/10.1142/S0218001420580124
13. Hanxin, C., Wenjian, H., Jinmin, H.: Multi-fault condition monitoring of slurry pump with principle component analysis and sequential hypothesis test. Int. J. Pattern Recognit. Artif. Intell. https://doi.org/10.1142/S0218001420590193
14. He, Y., Xiong, W., Chen, H.: Image quality enhanced recognition of laser cavity based on improved random hough transform. J. Vis. Commun. Image Representation (2020). https://doi.org/10.1016/j.jvcir.2019.102679