A New Rolling Bearing Work Condition Monitoring Method Based on Back Propagation Network

Qilu Wu[1,2], Yuxi Chen[1,2], and Wenxin Hu[1,2](✉)

[1] Artificial Intelligence Research Institute, Shenzhen MSU-BIT University, Shenzhen 518172, Guangdong, China
huwenxin@smbu.edu.cn

[2] Guangdong-Hong Kong-Macao Joint Laboratory for Emotional Intelligence and Pervasive Computing, Shenzhen MSU-BIT University, Shenzhen 518172, Guangdong, China

Abstract. Rolling bearings are among the most used elements in rotating machines, thus effective bearing work condition monitoring is necessary to avoid sudden machine failures. Vibration signal analysis is a common way. In order to automatically provide accurate diagnosis results, a new method based on Back Propagation (BP) network is proposed. Wavelet ridge transform has been applied to filter part environmental noise, and then a new demodulation method combing Hilbert transform method and autocorrelation curve has been introduced to overcome the remaining noise. The autocorrelation curve of envelop signal always shows its common and relatively apparent periodicity rule when defect occurs. At last, a Back Propagation (BP) network is used and trained to identify the periodicity and give the label whether the bearing is normal (label = 0) or defective (label = 1). Thus, bearing condition can be effectively monitored. At the same time, the characteristic frequency of the defect can be directly obtained according to the autocorrelation curve. Compared to existing methods based on complicated networks, this method operates simpler. Simulated signals and experimental defect signals illustrated the effectiveness and superiority of this method.

Keywords: Rolling bearings · Work condition monitoring · Back Propagation (BP) network · Autocorrelation curve

1 Introduction

Rolling bearing has been widely used in rotating machines such as turbine, and its failure will cause costly downtime. So effective bearing condition monitoring techniques are required to prevent this kind of damage, and among them vibration signal analysis is a common way [1, 2]. Many theoretical researches [3–5] about the bearing failure mechanism have proved that the damage bearing signature consists of exponentially decaying ringing that occurs periodically at the characteristic frequency. Also, the defective bearing vibration signal is usually considered being amplitude modulated at the characteristic defect frequency. Thus, the demodulation and following defect diagnosis are the main research contents.

V. C. M. Leung et al. (Eds.): Qshine 2023, LNICST 574, pp. 26–37, 2024.
https://doi.org/10.1007/978-3-031-65123-6_3

However, the vibration signal is collected by the sensor installed on the bearing housing, so the environmental noise is inevitably introduced. The effective information is usually drowned especially on the early stage of the defect development. So how to reduce the interference of the noise before demodulation is a key problem. Many different methods have been proposed in the past. Wavelet threshold method has been widely used [6–8]. This method computes the wavelet decomposition of the signal and then adjusts the wavelet coefficients according to a given threshold. Through wavelet reconstruction, a de-noising signal is obtained. The noise and original signal can be separated effectively when selecting a suitable threshold, but how to select this threshold and perform thresholding is still an open issue. Wavelet filter-based denoising method is introduced and proved performing better when facing the impulse-like defect signal [9]. In order to select optimal parameters for the wavelet filter, a two-step optimization process is proposed, which is a little time-consuming. What's more, a patch-based denoising method has been proposed which was called nonlocal means (NLM) based on the abundant signal redundant information and the rich similar patches [10]. But the noise variance of the bearing fault signal needs to be estimation in advance according to the characteristics of Gauss noise.

Due to the amplitude-modulated effect, the spectrum of defect signals consists of a harmonic series of frequency components present at the bearing defect frequency with the highest amplitude around the resonance frequency. In order to achieve effective signal demodulation and defect diagnosis, Fast Fourier transform (FFT) method [11], Hilbert transform method [12–14] and wavelet transform method [15, 16] have been used. The wavelet transform method owns its extraordinary time-frequency representation capability compared with FFT method, but it relies on appropriate wavelet basis function which is similar to the original signal. Besides, the envelop spectrum of the defect signal can be obtained with Hilbert transform, and then its characteristic features can be extracted from the envelop spectrum. Besides, Diego et al. have proposed a one-class v-SVM method to discriminate between normal and faulty conditions [17]. However, these methods will all be influenced by the remaining noise.

In recent years, in order to automatically provide accurate diagnosis results, many methods based on artificial neural networks (ANNs) have been proposed, which employ signal processing techniques for extracting features and further input the features to ANNs for classifying faults. Some studies focus on building suitable manual features such as root mean square (RMS) [18], Rényi entropy based feature [19], features extracted by locality preserving projections and Gaussian mixture models [20] and so on [21, 22] to discriminate between normal and faulty conditions. However, it is hard to find a feature suitable to all issues. On the other hands, more and more complicated networks such as MD-CUMSUM and NARX neural network [23], convolutional and long-short-term memory network [24], hierarchical gated recurrent unit network [25] and multi-scale dense gate recurrent unit network (MDGRU) [26] have been used to mine the useful information from the raw data, which are called data-driven methods [27, 28]. High nonlinear characteristics are learned by deep layers gradually, in the last, the difference between the normal and defect condition can be learned by these networks. Thus, a large quantity of data is needed. Also, aiming to promptly process the massive data and complicated networks, there is some requirement on the computing ability of the device.

In this paper, a de-noising method based on wavelet ridge transform has been proposed, which performs like a low-pass filter. Then, a new demodulation method combing Hilbert transform method and autocorrelation curve has been introduced to overcome the remaining noise. The vibration signal is usually considered being amplitude modulated at the characteristic defect frequency when defect occurs, however, it is still hard to recognize the periodicity of the envelop signal due to the remaining noise. But the autocorrelation curve of envelop signal always shows its common and relatively apparent periodicity rule. A Back Propagation (BP) network is used and trained to identify the periodicity and give the label whether the bearing is normal (label = 0) or defective (label = 1), which input the whole autocorrelation curve to the network rather than several manual parameters. Thus, bearing condition can be effectively monitored. At the same time, the characteristic frequency of the defect can be directly obtained according to the autocorrelation curve. Compared to above methods based on complicated networks, this method operates simpler. Simulated signals and experimental defect signals illustrated the effectiveness and superiority of this method.

2 Principle

2.1 De-noising Method Based on Wavelet Ridge Transform

When choosing analytic wavelet $\varnothing(t) = g(t) * \exp(iw_0 t)$, the continuous Wavelet Transform of a vibration signal $f(t) = A(t) * \cos(\varphi(t))$ is defined as:

$$W(a, b) = \frac{\sqrt{a}}{2} A(b) \exp[i\varphi(b)] \left\{ G\left[a\left(\xi - \varphi'(b)\right)\right] \right\} \quad (1)$$

where a is the scale parameter, and b is a shift parameter related to t. A(b) is the instantaneous amplitude at point b, and $\varphi'(b)$ represents its instantaneous frequency. G represents the Fourier Transform. Also, $\xi = w_0/a$.

According to Eq. (1), the modulus of W(a, b) can be calculated as:

$$\frac{|W(a, b)|}{\sqrt{a}} = \frac{A(b)}{2} \left|G\left[a\left(\xi - \varphi'(b)\right)\right]\right| \quad (2)$$

Thus, for a fixed point b, when $\xi = \varphi'(b)$, this is $\frac{w_0}{a_b} = \varphi'(b)$, $W(a_b, b)$ get its maximum called ridge value, where a_b represents the scale at the ridge point associated with position b. The de-noising signal can be obtained by applying inverse wavelet transform to the ridge curve [29]. This method performs like a low-pass filter, but it avoids selecting suitable threshold and filter compared to traditional wavelet method.

A sinusoidal signal and a modulation signal both with additive white noise were shown in Fig. 1(a). Figure 1(b) shows that the signals after wavelet ridge transform de-noising fit well with the original signal, verifying the effectiveness of this method.

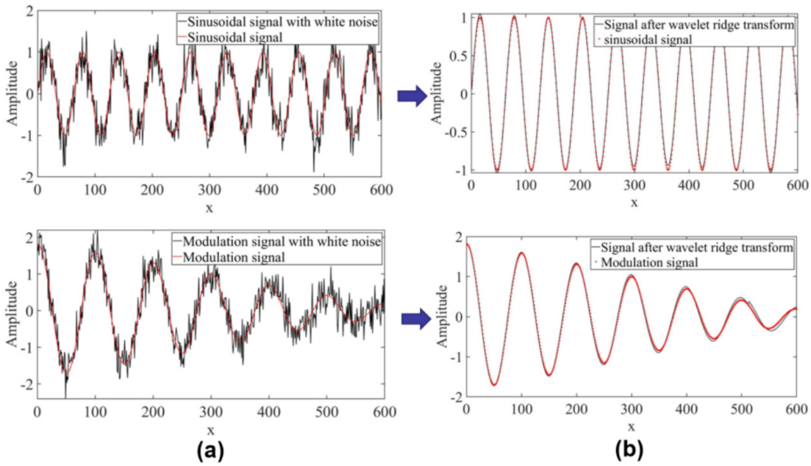

Fig. 1. (a) Original signals and (b) signals after wavelet ridge transform.

2.2 Defect Diagnosis Method

For a vibration signal $f(t) = A(t) * \cos(\varphi(t))$, its Hilbert transform is defined as:

$$\widehat{f(t)} = H[f(t)] = \frac{1}{\pi}\int_{-\infty}^{+\infty} f(\tau)\frac{1}{t-\tau}d\tau \tag{3}$$

So $\widehat{f(t)}$ is a linear function of f(t). It is obtained from x(t) applying convolution with $(\pi t)^{-1}$ as shown in the following formula:

$$\widehat{f(t)} = (\pi t)^{-1} * f(t) \tag{4}$$

Many theoretical researches have proved that the Hilbert transform of the original signal f(t) represents its harmonic conjugate. By introducing its harmonic conjugate, a new complex analytical signal can be described by the expression:

$$g(t) = f(t) + j\widehat{f(t)} \tag{5}$$

And the envelop signal A(t) can be calculated as:

$$A(t) = \sqrt{(f(t))^2 + (\widehat{f(t)})^2} \tag{6}$$

Thus, defect characteristic frequency can be obtained through its envelop signal. However, the effective information still can be drowned when the remaining noise is too strong. According to the theoretical solution of the defect bearing vibration signal, a simulation signal adding white noise is used to illustrate this issue. The simulation signal is described as following:

$$y = [\sum\nolimits_{n=0}^{N} \delta(t - nT)] * [e^{-\gamma 2\pi ft}\cos(2\pi f\sqrt{(1 - r^2)}t)] \tag{7}$$

where δ represents impulse function. T = 1/300, which means the defect frequency is 300 Hz. f represents the bearing natural frequency, and here is 3000 Hz. γ represents the damping ratio, and here is 0.1.

As the figure shows, Fig. 2(a) and Fig. 3(a) represents the simulated defect bearing vibration signal adding 10 dB white noise and 20 dB white noise, respectively. Figure 2(b) gives the envelop signal of Fig. 2(a), and its periodicity is obvious, so the defect frequency can be obtained directly. However, the signal of Fig. 3(b) is disordered, which is envelop signal of Fig. 3(a).

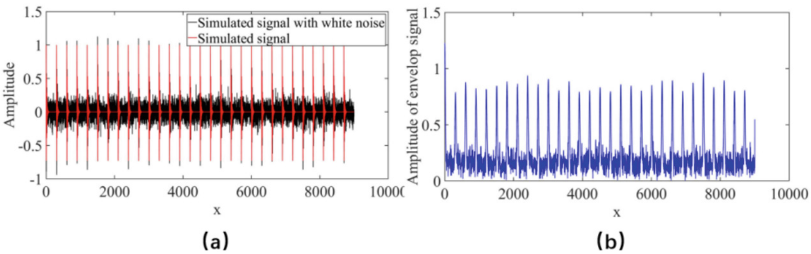

Fig. 2. (a) Simulated defect bearing vibration signal adding 10 dB white noise and (b) its envelop signal.

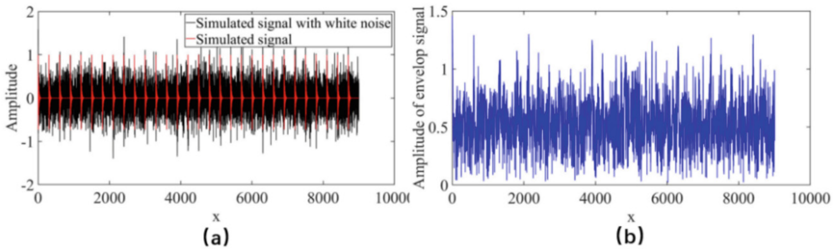

Fig. 3. (a) Simulated defect bearing vibration signal adding 20 dB white noise and (b) its envelop signal.

Facing this problem, a method based on autocorrelation curve is proposed. The autocorrelation curve of the envelop signal is calculated, which owns stronger periodicity and the same period. The autocorrelation curve of Fig. 3(b) is drawn in Fig. 4(a), and it exists relatively apparent periodicity. The detailed profile of the marked region in (a) is shown in (b), from which the defect impulse frequency 300 Hz can be easily seen.

2.3 Work Condition Monitoring Method

The accompanying periodicity of autocorrelation curve when defect happens can be used to monitor the work condition of bearing, which is the key to prevent sudden failure. The costly damage can be avoided if the defect can be distinguished once it happens. Regarding as a binary classification problem, the BP neural network is chosen to

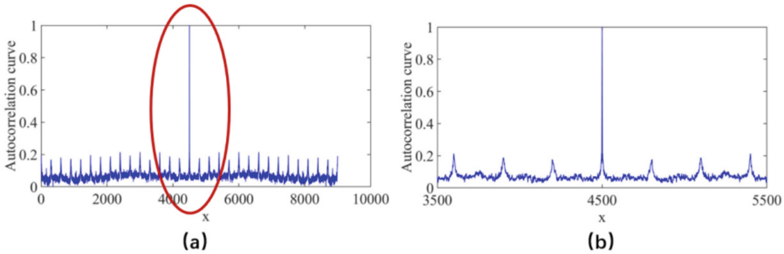

Fig. 4. (a) Autocorrelation curve and (b) detailed profile of the marked region.

distinguish the periodicity of autocorrelation curve corresponding to the work condition of bearing. The autocorrelation curve data is used as the input of the network, and the output is 0 or 1, which represents normal and defective, respectively. After training, the network can be used to give label directly according to new autocorrelation curve, this is to say, identifying the work condition. Thus, costly downtime due to bearing sudden failure can be avoided successfully.

The more layers and more neurons a BP network have, the more complex corresponding relationship can be learned. Considering there is two condition needed to be identified, a four-layer network was used. As Fig. 5 shows, the input layer consisted of 2000 neurons corresponding to the number of the autocorrelation curve data used for training. The number of neurons in the two hidden layers were set to 100 and 100, respectively. The output layer consisted of 2 neurons, and two output values represent the probability to be 0 and 1, respectively. At last, the label was given according to the one owning larger probability. The cross-entropy [27] known as multi-class log loss was used as loss function. All connecting weights were initialized with random numbers from a truncated Gaussian distribution and the values of the biases were initialized as constant.

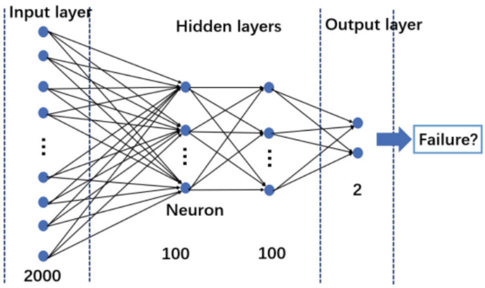

Fig. 5. BP network schematic diagram.

In the training processes, an input was first propagated through the network. The difference between the calculated and the desired outputs was then backpropagated from the output neurons to the first layer of the network, thereby adjusting network weights in the direction opposite to that of the derivative of the network error with respect to each network weight. By repeating this procedure multiple times for each data in a training set, the network can be taught to map the inputs to the correct outputs. Here the learning rate was set to 0.001.

3 Experiment

The experiment was conducted by the NSF I/UCR Center for Intelligent Maintenance Systems. As the Fig. 6 shows, four bearings were installed on a shaft. The rotation speed was kept constant at 2000 RPM by an AC motor. A radial load of 6000 lbs is applied onto the shaft and bearing by a spring mechanism. All bearings are force lubricated. Rexnord ZA-2115 double row bearings were installed on the shaft. PCB 353B33 High Sensitivity Quartz ICP accelerometers were installed on the bearing housing (One accelerometer for each bearing). Sensor placement is also shown in Fig. 7.

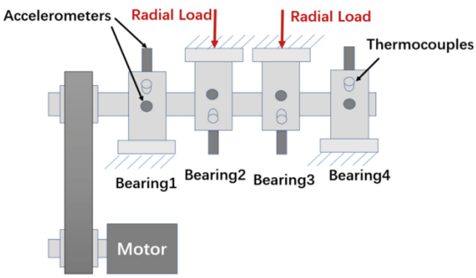

Fig. 6. Experiment schematic diagram.

The data set 1 describes a test-to-failure experiment (Bearing 1 happens defect.), and each bearing records 900 groups of data at different time. Each group consists of 20,480 points with the sampling rate set at 20 kHz. The record process of data set 2 is the same, the only difference is bearing 3 happens defect.

Based on data set 1, vibration signal of bearing 1 before failure is shown in Fig. 7(a), and vibration signal of bearing 1 after failure is shown in Fig. 7(b). Figure 7(c) and (d) give corresponding envelop signal after de-noising, respectively. The disordered envelop signals show the information is highly masked by the remaining noise. The normalized autocorrelation curve of Fig. 7(c) is drawn in Fig. 8(a), and the detailed profile of the marked region in Fig. 8(a) is shown in Fig. 8(b). The autocorrelation curve of Fig. 7(d) is drawn in Fig. 9(a), also, the detailed profile of the marked region in (a) is shown in (b). Through comparison, the Fig. 9 shows apparent periodicity, which means defect exists. Besides, the impulse frequency can be calculated.

Considering the key problem is to real-time diagnosis, this is to find the sudden change moment, normalized autocorrelation curves at different moments (every 30 groups) are shown in Fig. 10. Results show there exist two kinds of curve, and change begins in time series of No.19. The shape of autocorrelation curve changes suddenly once the defect happens. Thus, the difference can be used as failure sign.

A BP neural network was then used to learn and discriminate the difference. 90 groups of normalized autocorrelation curve data were used as the input of the network,

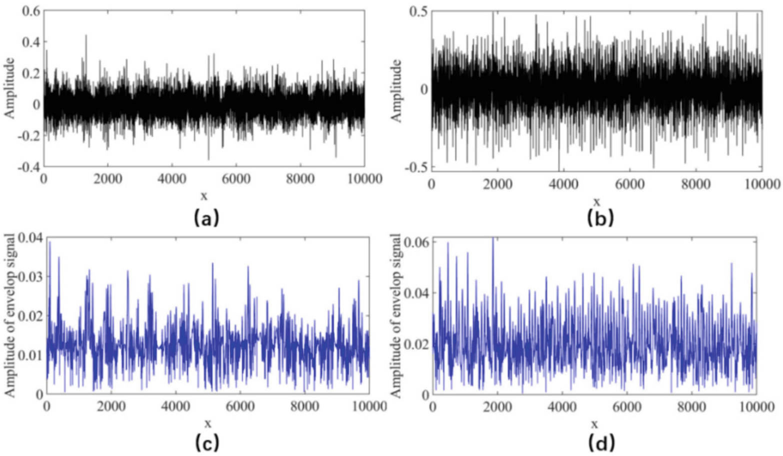

Fig. 7. (a) Vibration signal of bearing 1 before failure and (c) its envelop signal; (b) vibration signal of bearing 1 after failure and (d) its envelop signal.

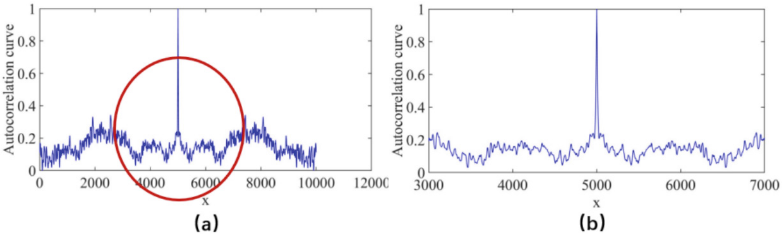

Fig. 8. (a) Normalized autocorrelation curve and (b) its detailed profile of the marked region.

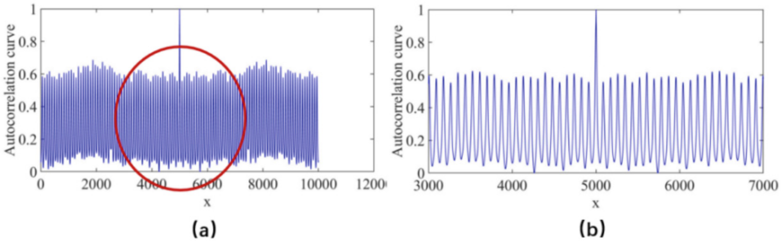

Fig. 9. (a) Normalized autocorrelation curve and (b) its detailed profile of the marked region.

and artificially given labels were set as the output of the network. The training procedure repeated 1000 times, and the learning rate was set to 0.001. After training, the network could be used to discriminate and give label directly.

For data set 2, vibration signals of bearing 3 before and after failure were also used to verify the proposed method. The normalized autocorrelation curves of envelop signal

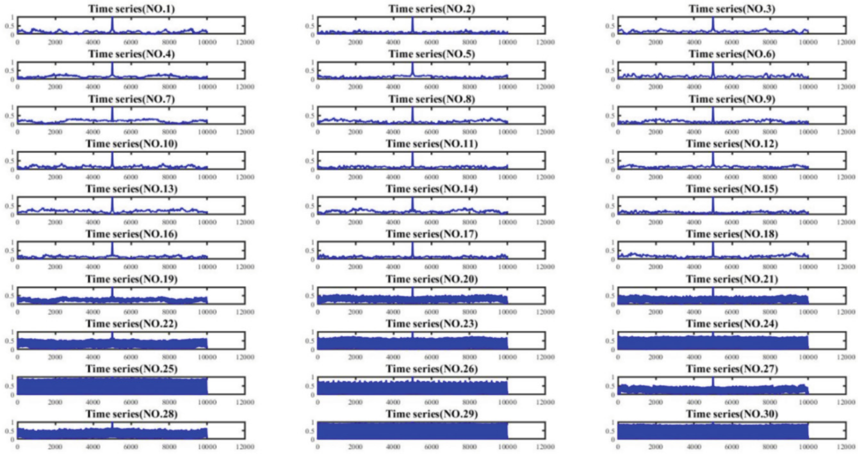

Fig. 10. Normalized autocorrelation curve.

showed same regularity, changing from Fig. 11(a) to Fig. 11(b). Also, the trained network gave correct labels 0 and 1, respectively.

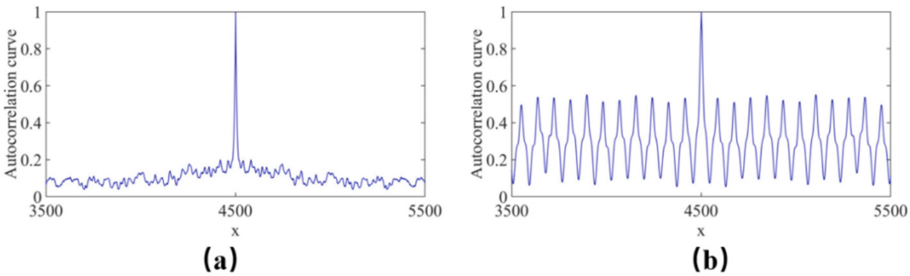

Fig. 11. Normalized autocorrelation curves (a) before failure and (b) after failure.

More data coming from data set 1 (bearing 1) and data set 2 (bearing 3) were used to test the trained network, and the test results were recorded in Table 1. Note that the validation data includes un-trained data and partly trained data.

Table 1. Accuracy results

	Data set 1	Data set 2
The number of test data	90	100
The number of correct labels	90	99
Accuracy	100%	98.89%

As the Table 1 shows, the results verified the high accuracy and generalization ability of the trained network, which also illustrated the effectiveness of the proposed monitoring method.

4 Conclusion and Discussion

A new rolling bearing work condition monitoring method based on Back Propagation network has been described in this paper, and simulation results and experimental results verified its effectiveness and high accuracy.

A de-noising method based on wavelet ridge transform has been proposed, which performs like a low-pass filter. Noise can't be filtered completely, especially the information is totally and highly masked by the noise. Considering the autocorrelation curve of envelop signal always owns relatively apparent periodicity rule when defect occurs, a new demodulation method combing Hilbert transform method and autocorrelation curve has been introduced to overcome the remaining noise. Also, experimental results verify that the stage of autocorrelation curve changes suddenly once the defect happens, thus, it can be used as failure sign. What's more, the characteristic frequency of the defect can be directly obtained according to the autocorrelation curve.

Aiming to achieve real-time monitoring, a Back Propagation (BP) network is used and trained to identify the periodicity and give the label whether the bearing is normal (label = 0) or defective (label = 1), which input the whole autocorrelation curve to the network rather than several manual parameters. Compared to above methods based on complicated networks, this method also owns generalization ability in a degree but operates simpler.

Acknowledgment. The work is supported by the National Natural Science Foundation of China (under grants 12102267).

References

1. Peng, B., Bi, Y., Xue, B., Zhang, M., Wan, S.: A survey on fault diagnosis of rolling bearings. Algorithms **15**(10), 347 (2022)
2. Wu, G., Yan, T., Yang, G., Chai, H., Cao, C.: A review on rolling bearing fault signal detection methods based on different sensors. Sensors **22**(21), 8330 (2022)
3. McFadden, P.D., Smith, J.D.: Model for the vibration produced by a signal point defect in a rolling element bearing. J. Sound Vib. **96**(1), 69–82 (1984)
4. Jiang, Z., Zhang, K., Xiang, L., Yu, G., Xu, Y.: A time-frequency spectral amplitude modulation method and its applications in rolling bearing fault diagnosis. Mech. Syst. Signal Process. **185**, 109832 (2023)
5. Zhao, S., Liang, L., Xu, G., Wang, J., Zhang, W.: Quantitative diagnosis of a spall-like fault of a rolling element bearing by empirical mode decomposition and the approximate entropy method. Mech. Syst. Signal Process. **40**(1), 154–177 (2013)
6. Liu, Z., Zhang, L., Carrasco, J.: Vibration analysis for large-scale wind turbine blade bearing fault detection with an empirical wavelet thresholding method. Renewable Energy **146**, 99–110 (2020)

7. Lin, J.: Gearbox fault diagnosis using adaptive wavelet filter. Mech. Syst. Signal Process. **17**(6), 1259–1269 (2003)
8. Yan, R., et al.: Wavelet transform for rotary machine fault diagnosis: 10 years revisited. Mech. Syst. Signal Process. **200**, 110545 (2023)
9. Qiu, H., Lee, J., Lin, J., Yu, G.: Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. J. Sound Vibr. **289**(4–5), 1066–1090 (2006)
10. Quan, Y., Ji, H., Shen, Z.: Data-driven multi-scale non-local wavelet frame construction and image recovery. J. Sci. Comput. **63**(2), 307–329 (2015)
11. Rai, V.K., Mohanty, A.: R: Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert-Huang transform. Mech. Syst. Signal Process. **21**(6), 2607–2615 (2007)
12. Ramu, S.K., Irudayaraj, G.C.R., Subramani, S., Subramaniam, U.: Broken rotor bar fault detection using Hilbert transform and neural networks applied to direct torque control of induction motor drive. IET Power Electronics **13**(15), 3328–3338 (2020)
13. Deng, L., Hou, Z., Liu, H., Sun, Z.: A fractional hilbert transform order optimization algorithm based DE for bearing health monitoring. In: 2019 Chinese Control Conference (CCC), pp. 2183–2186. IEEE (2019)
14. Peng, Z.K., Tse, P.W., Chu, F.L.: A comparison study of improved Hilbert-Huang transform and wavelet transform: application to fault diagnosis for rolling bearing. Mech. Syst. Signal Process. **19**(5), 974–988 (2005)
15. Zhang, K., Ma, C., Xu, Y., Chen, P., Du, J.: Feature extraction method based on adaptive and concise empirical wavelet transform and its applications in bearing fault diagnosis. Measurement **172**, 108976 (2021)
16. Wang, S., Huang, W., Zhu, Z.K.: Transient modeling and parameter identification based on wavelet and correlation filtering for rotating machine fault diagnosis. Mech. Syst. Signal Process. **25**(4), 1299–1320 (2011)
17. FernáNdez-Francos, D., MartíNez-Rego, D., Fontenla-Romero, O., Alonso-Betanzos, A.: Automatic bearing fault diagnosis based on one-class ν-SVM. Comput. Ind. Eng. **64**(1), 357–365 (2013)
18. Huang, R., Xi, L., Li, X., Liu, C.R., Qiu, H., Lee, J.: Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. Mech. Syst. Signal Process. **21**(1), 193–207 (2007)
19. Boškoski, P., Gašperin, M., Petelin, D., Juričić, Đ: Bearing fault prognostics using Rényi entropy based features and Gaussian process models. Mech. Syst. Signal Process. **52**, 327–337 (2015)
20. Yu, J.: Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models. Mech. Syst. Signal Process. **25**(7), 2573–2588 (2011)
21. Siegel, D., Al-Atat, H., Shauche, V., Liao, L., Snyder, J., Lee, J.: Novel method for rolling element bearing health assessment—a tachometer-less synchronously averaged envelope feature extraction technique. Mech. Syst. Signal Process. **29**, 362–376 (2012)
22. Benkedjouh, T., Medjaher, K., Zerhouni, N., Rechak, S.: Remaining useful life estimation based on nonlinear feature reduction and support vector regression. Eng. Appl. Artif. Intell. **26**(7), 1751–1760 (2013)
23. Rai, A., Upadhyay, S.H.: The use of MD-CUMSUM and NARX neural network for anticipating the remaining useful life of bearings. Measurement **111**, 397–410 (2017)
24. Hinchi, A.Z., Tkiouat, M.: Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network. Procedia Comput. Sci. **127**, 123–132 (2018)
25. Li, X., Jiang, H., Xiong, X., Shao, H.: Rolling bearing health prognosis using a modified health index based hierarchical gated recurrent unit network. Mech. Mach. Theory **133**, 229–249 (2019)

26. Ren, L., Cheng, X., Wang, X., Cui, J., Zhang, L.: Multi-scale dense gate recurrent unit networks for bearing remaining useful life prediction. Futur. Gener. Comput. Syst. **94**, 601–609 (2019)
27. Pan, Y., Er, M.J., Li, X., Yu, H., Gouriveau, R.: Machine health condition prediction via online dynamic fuzzy neural networks. Eng. Appl. Artif. Intell. **35**, 105–113 (2014)
28. Jia, F., Lei, Y., Lin, J., Zhou, X., Lu, N.: Deep neural networks: a promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mech. Syst. Signal Process. **72**, 303–315 (2016)
29. Li, S., Su, X., Chen, W.: Wavelet ridge techniques in optical fringe pattern analysis. JOSA A **27**(6), 1245–1254 (2010)