Research on Fault Feature Extraction of Rolling Bearing Based on dFIF

Xueliang Chen¹ Baojia Chen¹, * Zhengkun Chen¹

¹Hubei Key Laboratory of Hydroelectric Machinery Design and Maintenance, College of Mechanical and Power Engineering, China Three Gorges University, Yichang, Hubei 443002, China

*Corresponding author’s e-mail: cbjia@163.com

Abstract. Aiming at the problems that the traditional adaptive mode decomposition method has low decomposition efficiency, poor ability of anti-mode mixing and decomposition accuracy, a rolling bearing fault feature recognition method based on direct fast iterative filtering (dFIF) and the crest factor of envelope spectrum is proposed. dFIF first decomposes the raw rolling bearing vibration signal into a set of intrinsic mode function (IMFs), then the optimal component is chosen based on the maximal crest factor of envelope spectrum, and finally the bearing fault characteristic frequency and its frequency multiplication are extracted using envelope demodulation analysis. It is demonstrated by relevant experimental data that the proposed method has efficient decomposition speed, accurate decomposition accuracy, and the accurate selection of effective components, and can effectively realize rolling bearing fault detection.

1. Introduction

The rolling bearings are widely used in rotating machinery [1]. It is critical to ensure the safe and reliable operation of mechanical equipment if the incipient fault can be identified and the fault type and location can be accurately determined. Because the impact feature of a rolling bearing's incipient fault is weak and easily affected by the transmission path, environmental background noise, and other factors [2], it is difficult to extract the fault feature and the diagnosis effect is poor. Efficient signal processing technology can effectively remove redundant noise, improve features, and improve the accuracy and precision of early fault detection and prediction.

The first locally adaptive data-driven method, empirical mode decomposition (EMD), proposed by Huang [3], is very suitable for dealing with non-stationary and nonlinear signals, but there are some issues such as over-envelope, mode mixing, and end effect under strong background noise [4, 5]. Smith and Cicone respectively proposed local mean decomposition (LMD) [6] and adaptive local iterative filtering (ALIF) [7] recursive mode decomposition methods similar to EMD, which mitigated the influence of mode mixing and end effect. Despite the fact that the aforementioned methods and their improved versions in analysing a signal with very short data points can be decomposed well, and they have been successfully applied in the field of fault diagnosis [8-11], but they may fail to deal with long time series data. In the face of long time series, the decomposition efficiency and accuracy are greatly reduced, and the decomposition may even fall into a local cycle and fail. Attempting to overcome this problem, Cicone proposed a more novel and efficient decomposition method-direct fast iterative filtering (dFIF) [12], which realizes the instantaneous decomposition of non-stationary signals-on the basis of fast iterative filtering (FIF) [13].
According to the findings of the preceding research, the accuracy of fault diagnosis results is closely related to the effective mode decomposition method. In order to improve the diagnosis effect of rolling bearing faults, a fault diagnosis method based on dFIF decomposition and the crest factor of envelope spectrum is proposed in this paper, and the diagnosis results are compared with those of other commonly used adaptive decomposition methods.

2. Theoretical background of dFIF algorithm

The dFIF method is the online version of the FIF algorithm. FIF is obtained by introducing fast Fourier transform (FFT) into discrete iterative filtering (DIF). The detailed description and derivation process of DIF algorithm can be found in [13]. In the DIF algorithm, given a discrete signal \( f(x_j) \), sampled at \( N \) points \( x(j) = j / N - 1 \), with \( j = 0, \ldots, N - 1 \). The goal of the FIF algorithm is to decompose the signal \( [f(x_j)]_{j=0}^{N-1} \) into IMFs. The \( m \)th iteration is given by the following formula:

\[
f_{m+1}(x_i) = f_m(x_i) - \sum_{x_j=x_i-l}^{x_i+l} f_m(x_j) \omega_m(x_j - x_i) \frac{1}{N}, \quad i = 0, \ldots, N - 1
\]

The form of the above matrix can be expressed as:

\[
f_{m+1} = (I - W_m) f_m
\]

where

\[
W_m = \left[ \omega_m(x_i - x_j) \frac{1}{N} \right]_{j=0}^{N-1}
\]

\( \omega_m \) is the Fokker-Planck filter proposed in [7], whose area equal to one and support in \([-l_m, l_m]\). \( l_m \) is the half support length. The subscript \( m \) represents the number of loop iterations at this time. \( l_m \) is given by the equation (4).

\[
l_m = 2 \left\lfloor \frac{N}{\mu} \right\rfloor
\]

where the values of the tuning parameter \( \alpha \) fall in the range of \([1,3]\), \( \mu \) is the number of its extreme points, \( N \) is the length of the signal and \( \lfloor \cdot \rfloor \) rounds a positive number to the nearest integer closer to zero.

Assume that \( N_0 \) is the number of iterations required for the first IMF, for details about \( N_0 \), please refer to [13]. The first IMF can be further calculated as

\[
IMF_1 = U(I - D)^{N_0} U^T f = IDFT \left( I - \text{diag} \left( \text{DFT}(\omega) \right) \right)^{N_0} \text{DFT}(f)
\]

Let \( M \) be the calculated number of IMFs, and generate subsequent IMFs by iteratively applying the above process to the margin \( r = f - \sum_{i=1}^{M} \text{IMF}_i \). When \( r \) becomes a non-oscillating trend signal, the algorithm stops. The introduction of FFT can greatly shorten the calculation time of the IF method. Note that we briefly introduce the modified parts of the FIF in this subsection. For details of and stop criteria and the FIF method, please refer to the literature [13].

It can be concluded that the equation (5) of the iteration is converging all \( I - D \) the none-ones eigenvalues to zero and preserving the eigenvalues who are equal to 1. Therefore, a new idea was proposed in [12], that is, by ignoring non-zero eigenvalues, the iterative method can be made a direct method. In order to do this, it is necessary to accurately estimate an appropriate \( N_0 \), a threshold \( \tau \) and a value \( \kappa \) are set, as shown in equation (6).
To look for $\max(1 - \lambda_i)^{N_i} \approx \kappa$ as the goal, where $\tau$ is close to 1 and $\kappa$ small enough [13]. For details of $\tau$ and $\kappa$, please refer to [12].

According to equation (6), $N_0$ allows to compute each IMF in one step. Whereas, the number of steps to compute each IMF when using FIF is strictly more than 1 in general [13]). Thus, the dFIF is faster than FIF. In dFIF, it directly calculates the number of iterations $N_0$ required for the first IMF through the predetermined thresholds $\tau$ and $\kappa$. Then, repeat the process in pervious section to generate all IMFs and achieve the purpose of decomposition.

3. Bearing Fault Diagnosis Process

As an important part of rotating machinery, rolling bearing will have an important impact on production if it breaks down during operation. Therefore, in order to solve the problem that it is difficult to accurately extract the characteristic frequency of bearing fault under complex background noise, a rolling bearing fault diagnosis method based on the combination of dFIF and transient feature extraction transformation is proposed.

- Step1 Acquiring vibration signals of different bearing fault types;
- Step2 Applying dFIF to decompose the vibration signal;
- Step3 The maximal crest factor of envelope spectrum is applied to select the optimal mode;
- Step4 Perform the Hilbert transform on the optimal mode, and compare the dominant frequency with the calculated theoretical fault characteristic frequency in its power spectrum to determine the fault type.

The details of the crest factor of envelope spectrum are as follows:

Zhang et al. [14] proposed an index that can simultaneously reflect the presence and strength of the shock mode of the signal and can highlight the periodic characteristics of the shock occurrence—the crest factor of envelope spectrum, which is defined as the envelope spectrum The ratio of the maximum peak value to the effective value. In this paper, by obtaining the $C_E$ value of each mode, the mode with the largest $C_E$ value is extracted as the optimal mode. The formula is as follows:

$$C_E = \max(E_m(i)) \sqrt{\frac{1}{N}} \sum_{i=1}^{N} (E_m(i))^2$$

Where $E_m$ is the envelope spectrum and $N$ is the length of the envelope signal.

4. An example of experimental diagnosis of rolling bearing fault in CTGU

The presented method is finally applied to diagnosing the real-world bearing fault, so we process the experimental data collected from the rolling bearing fault testing platform at China Three Gorges University (CTGU) [2].

The structure of the rolling bearing fault testing platform and its picture are presented in figure1. The experiment system is composed of a DC drive motor, pulley, bearing pedestal, support housing, coupling, a DC load motor and other accessories. The speed signal during the working process of the testing platform was measured by a non-contact velocimeter arranged at the coupling, and the bearing vibration signal was measured by two vibration acceleration sensors mounted on the radial plane of the bearing seat in the horizontal and vertical directions.
Four rolling bearings of type 6308 were selected for the experiment, of which one was normal and three were artificially added faulty bearings. The fault types include outer ring (point damage, located in the middle of the race, about 7 mm², 0.2 mm deep; inner ring (point damage, located in the middle of the race, about 3.6 mm², 0.1 mm deep), and rolling element (point damage, about 1 mm²). In this article, only the rolling element failure is analyzed. Table 1 list the structure parameters of the rolling bearing. In the experiment, the rotational speed of the shaft is 1620 rpm, the rotation frequency is \( f_c = 27 \) Hz, the sampling frequency is 20 kHz, and the number of sampling points is 8192.

| Bearing model | Outer race diameter | Inner race diameter | Pitch diameter | Width | Number of balls | Ball diameter | Tapered contact angle |
|---------------|---------------------|---------------------|----------------|-------|-----------------|---------------|----------------------|
| 6308          | 90 mm               | 40 mm               | 65 mm          | 23 mm | 8               | 15 mm         | 0°                   |

According to table 1, the theoretical characteristic frequency of rolling element failure is calculated as \( f_c = 55 \) Hz. The time waveform, envelope spectrum of the CTGU raw rolling element fault signal is shown in figure 2(a) and (b). The fault impact components with uneven interval can be clearly observed in subplot (a), but most of the impact components are submerged in noise due to the relatively weak fault. This represents a difficulty to identify the fault from raw signal. The envelope spectrum of the vibration signal is shown in subplot (b). There are some obvious spectrum peaks located on the frequency of 181 Hz, 195 Hz, 327 Hz and 535 Hz. However, the rolling element fault characteristic frequency \( f_c \) and its harmonics can’t be observed. The reason is the time interval and position of contact at the damage site of the rolling body are not fixed, the signal transmission path is more complex, and it is easy to be affected by environmental noise and signal energy attenuation. In order to extract the
fault characteristic information better, the original vibration signals need to be further decomposed into single components for analysis. For the purpose of comparison, the analytical procedures of other prevalent adaptive mode decomposition methods are also presented, such as ALIF, LMD and EMD. The components decomposed by different decomposition methods of rolling body fault signals are shown in figure 3.

Figure 3. The rolling element fault signal decomposition results of each method (a) dFIF. (b) ALIF. (c) LMD. (d) EMD.

Table 2. Decomposition time of each method and the size of each mode index

| Methods | Mode | Time(s) |
|---------|------|---------|
|         | 1    | 2       | 3       | 4       | 5       | 6       |
| dFIF    | 0.009| 20.751  | 21.843  | 27.882  | 25.068  | 17.024  |
| ALIF    | 5.206| 23.331  | 19.645  | 17.12   | 24.601  | 28.865  |
| LMD     | 14.923| 23.544  | 19.181  | 16.276  | 24.129  | 27.344  |
| EMD     | 0.146| 21.945  | 27.569  | 15.501  | 21.05   | 17.508  | 24.457  |

The fault-related information cannot be obtained directly in component waveforms in figure 3. Hence, it is necessary to use the method proposed in this paper for further analysis. In table 2 the components in bold represent the optimal components decomposed by each method (i.e., with the maximum $E_C$).

Figure 4 shows the comparison of the envelope power spectrum of the optimal components obtained
from the proposed technique with the optimal components of ALIF, LMD and EMD. It is clear from the
time waveform that the transient shocks are obvious in the optimal components in figure 4(a) obtained
from dFIF. Besides, the peak spectrum lines basically coincide with the 2, 4, and 6 harmonics of
theoretical characteristic frequencies of rolling element faults in envelope power spectrum, which shows
that the bearing has occurred rolling element failure. In contrast, no fault related frequency can be
extracted from the optimal mode power spectrum in the analysis results of the other three decomposition
methods shown in figure 4(b)-(d). That is to say, ALIF, LMD and EMD failed to diagnose this signal.
In addition, according to table 2, when decomposing the measured bearing signals, the dFIF method has
advantages that are hard to be matched by other methods in terms of decomposition speed, and the
decomposition of the signal is completed in an instant.

Figure 4. The optimal component time domain waveform and the envelope power spectrum results
obtained by decomposing the bearing fault vibration signal by each method. (a) dFIF. (b) ALIF. (c)
LMD. (d) EMD.

5. Conclusions
A new method of bearing fault diagnosis based on dFIF decomposition is proposed in this paper. The
fault signal is first decomposed by dFIF, and the optimal mode component containing rich fault
information is chosen based on the maximal crest factor of envelope spectrum, and the fault features are
extracted using envelope demodulation analysis of the optimal component. When compared to ALIF,
LMD and EMD, this method can achieve signal decomposition and accurate feature extraction more
quickly through the verification of bearing experimental data. The method proposed can effectively extract fault features of rolling bearings that have been immersed in noise and has a specific application value.

Acknowledgments
This program is jointly supported by the projects of National Nature Science Foundation of China (51975324, 52075292), Open Foundation of State Key Laboratory of Mechanical Transmission, Chongqing University, China (SKLMT-MSKFKT-202020), Open Foundation of Hubei Key Laboratory of Hydroelectric Machinery Design & Maintenance (No.2021KJX02 & No.2020KJX02).

References
[1] Yang J, Huang D, Zhou, D and Liu, H 2020 Optimal IMF selection and unknown fault feature extraction for rolling bearings with different defect modes Measurement 157 p 107660.
[2] Chen B, Shen B, Chen F, Tian H, Xiao W, Zhang F and Zhao C 2019 Fault diagnosis method based on integration of RSSD and wavelet transform to rolling bearing Measurement 131 pp 400-11
[3] Ciccone A 2020 Iterative filtering as a direct method for the decomposition of nonstationary signals Numerical Algorithms 85 pp 811-27
[4] Ciccone A, Liu J and Zhou H 2016 Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis Applied and Computational Harmonic Analysis 41 pp 384-411
[5] Ciccone A and Zhou H 2021 Numerical analysis for iterative filtering with new efficient implementations based on FFT Numerische Mathematik 147 pp 1-28
[6] Li Y, Si S, Liu Z and Liang X 2019 Review of local mean decomposition and its application in fault diagnosis of rotating machinery Journal of Systems Engineering and Electronics 30(4) pp 799-814.
[7] HuangN E, Shen Z, Long S R, Wu M C, Shih H H, Zheng Q, Yen N C, Tung C C and Liu H H 1998 The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences 454 pp 903-95
[8] Li X, Ma J, Wang X, Wu J and Li Z 2020 An improved local mean decomposition method based on improved composite interpolation envelope and its application in bearing fault feature extraction ISA Transactions 97 pp 365-83
[9] Liu Q, Wang Y and Xu Y 2021 Synchronsqueezing extracting transform and its application in bearing fault diagnosis under non-stationary conditions Measurement 173 p 108569
[10] Pan H, Yang Y, Li X, Zheng J and Cheng J 2019 Symplectic geometry mode decomposition and its application to rotating machinery compound fault diagnosis Mechanical Systems and Signal Processing 114 pp 189-211
[11] Smith J S 2005 The local mean decomposition and its application to EEG perception data J R Soc Interface 2 pp 443-54
[12] Yang J, Huang D, Zhou D and Liu H 2020 Optimal IMF selection and unknown fault feature extraction for rolling bearings with different defect modes Measurement 157 p 107660
[13] Zhang L, Xiong G and Huang W 2015 New procedure and index for the parameter optimization of complex wavelet based resonance demodulation J. Mech. Eng 51 pp 129-38
[14] Zheng J, Cheng J and Yang Y 2013 Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis Mechanical Systems and Signal Processing 40 pp 136-53