Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 320 (2009) 452–468
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Fault severity assessment for rolling element bearings using the Lempel–Ziv complexity and continuous wavelet transform

Hoonbin Hong, Ming Liang*

Department of Mechanical Engineering, University of Ottawa, 770 King Edward Avenue, Ottawa, Ontario, Canada K1N 6N5

Received 5 November 2007; received in revised form 15 July 2008; accepted 16 July 2008
Handling Editor: S. Bolton
Available online 29 August 2008

Abstract

This paper proposes a new version of the Lempel–Ziv complexity as a bearing fault (single point) severity measure based on the continuous wavelet transform (CWT) results, and attempts to address the issues present in the current version of the Lempel–Ziv complexity measure. To establish the relationship between the Lempel–Ziv complexity and bearing fault severity, an analytical model for a single-point defective bearing is adopted and the factors contributing to the complexity value are explained. To avoid the ambiguity between fault and noise, the Lempel–Ziv complexity is jointly applied with the CWT. The CWT is used to identify the best scale where the fault resides and eliminate the interferences of noise and irrelevant signal components as much as possible. Then, the Lempel–Ziv complexity values are calculated for both the envelope and high-frequency carrier signal obtained from wavelet coefficients at the best scale level. As the noise and other un-related signal components have been largely removed, the Lempel–Ziv complexity value will be mostly contributed by the bearing system and hence can be reliably used as a bearing fault measure. The applications to the bearing inner- and outer-race fault signals have demonstrated that the revised Lempel–Ziv complexity can effectively measure the severity of both inner- and outer-race faults. Since the complexity values are not dependent on the magnitude of the measured signal, the proposed method is less sensitive to the data sets measured under different data acquisition conditions. In addition, as the normalized complexity values are bounded between zero and one, it is convenient to observe the fault growing trend by examining the Lempel–Ziv complexity.

1. Introduction

A mechanical fault signal contains information not only about the machine health condition but also the severity of fault. Proper maintenance decision can be made only when fault severity is accurately assessed. The most commonly used measure of fault severity is the energy value in frequency or time–frequency domain. Covacece and Intronini illustrated fault developments of ball bearing of helicopter gearbox using auto- and cross-power spectrum [1]. Two indices, auto- and cross-power spectrum, increase their magnitudes as the fault develops. In the work by Loutridis [2], energy was successfully applied to the fault feature (intrinsic mode function, IMF) obtained from the empirical mode decomposition (EMD) method. Dalpiaz et al. [3] compared

*Corresponding author. Tel.: +1 613 562 5800 ext 6269; fax: +1 613 562 5177.
E-mail address: liang@eng.uottawa.ca (M. Liang).

0022-460X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2008.07.011

the cepstrum and spectral correlation density (SCD) methods to reveal the sensitivity of two indices to fault development. According to the work by Dalpiaz et al. [3], the cepstrum is insensitive to the crack evolution of a gear. On the other hand, SCD shows an increasing trend as the gear crack grows.

Yu et al. [4] adopted the EMD energy entropy for extracting bearing fault feature from a noisy signal mixture and the artificial neural network (ANN) for identifying the operating state of a roller bearing. They demonstrated the superior performance of the proposed method in both extracting fault feature and identifying bearing state by comparison with wavelet packet analysis.

Baydar and Ball [5] applied a time–frequency analysis method, smoothed pseudo-Wigner-Ville distribution (SPWVD), to evaluate gear failures. Two symptoms were reported for faulty state. First, the energy concentration appears around angular position corresponding to the fault. Second, frequency interference increases between the fundamental and higher harmonics. As a fault develops, these two phenomena appear more clearly. Boulahbal et al. [6] applied continuous wavelet transform (CWT) to the synchronous time average (STA) signal of gear crack and chip to obtain the amplitude and phase plots. The result indicated that energy concentration happened at the gear meshing frequency in the amplitude plot and phase changes drastically around the fault location.

Baydar and Ball [7] adopted both amplitude and phase plots of CWT to analyze the faulty gear states. The finding is similar to that of their previous work using SPWVD, i.e., the same two symptoms appear as fault develops. In the work by Loutridis [8], a local energy density was calculated in the time-scale domain to evaluate the relationship between energy level and gear fault development. The energy is obtained from the wavelet coefficients at a specific time instant corresponding to the largest energy value over scales. A polar wavelet amplitude map was applied to the STA signal of gear faults to display the fault severity and the circumferential location around the rotating axis [9]. In their cases, energy concentration appeared and became stronger as fault develops.

Time-domain analysis techniques were also applied for bearing fault diagnosis [10–12], which included root mean squared (rms) value, kurtosis value and crest factor. Shao and Nezu [10] investigated the relationships among fault size (or severity), the kurtosis value and the learning ratio of adaptive noise cancellation for faulty bearing signals. Williams et al. [11] used multiple sensors to monitor bearing condition in the run-to-failure test. They represented fault growth in terms of time-domain indices such as rms value, kurtosis and crest factor calculated from the data collected by multiple sensors. The time-domain result was also compared with that obtained using a frequency-domain technique, an envelope analysis. Recently, Al-Ghamd and Mba [12] applied acoustic emission (AE) technique to identify defect presence and size on a radially loaded bearing. They compared the AE and vibration results in terms of rms value, kurtosis and maximum amplitude of measured signal. The authors also revealed the relationship between the AE burst duration and fault severity (defect length).

Rubini and Meneghetti [13] compared the envelope spectrum method and the averaged wavelet amplitude spectrum approach in evaluating ball bearing's inner- and outer-race faults. Both methods were able to reveal the difference in spectral magnitude as the fault develops. According to their work, the inner race and outer race showed opposite trends in spectral magnitude as fault size increases.

Many fault evaluation methods reviewed above rely on the comparison of two different data sets related to normal and faulty conditions or less severe and more severe conditions. There are three major difficulties in the implementation of such methods:

(a) A fault of the same severity may display quite different signal strengths or magnitudes under different operating conditions or for bearings of different sizes. In particular, when non-relative measures or un-bounded (including single-end bounded) relative measures (such as kurtosis) are used, it will be difficult to specify an upper bound that would indicate a conclusive need for bearing replacement. Therefore, the signal may not be directly used for assessing fault severity. It should be noted that here the signal strengths do not necessarily mean the amplitudes directly obtained from the FFT.
(b) Different gain settings used during data acquisition lead to very different signal magnitudes and hence the use of signal "strength" in fault assessment without knowing the gain values may yield misleading results. We wish to point out that if the difference in gain setting is known, this may not cause a major problem. However, if the data sets were acquired by different people during different time periods and the gain

differences were not be well documented, it would be difficult for other users of these data sets to come up with reliable fault detection decisions.

(c) The data sets may not be available especially for new machines and new working conditions. Therefore, the comparison based methods cannot be applied under such circumstances. Even if the data can be collected, it would be very time-consuming to collect, store, manage, and retrieve them for on-line applications.

It is therefore highly desirable to develop a severity measure that is independent of working conditions/gain settings and does not require comparison between data sets. For this reason, Yan and Gao [14] proposed to evaluate fault deterioration using the Lempel–Ziv complexity. A normalized Lempel–Ziv complexity yields a non-dimensional value between zero and one [0, 1], close to zero for a pure sinusoidal signal and one for a white Gaussian noise. Yan and Gao [14] demonstrated the possibility of using the complexity as a fault severity measure and further indicated that the fault size as a severity measure is proportional to the nonlinearity of the signal: the Lempel–Ziv complexity value increases as the size of the outer race bearing fault grows. Their results also indicate that the increase of bearing speed tends to increase the complexity value and the bearing load has no appreciable effect on the complexity. With the Lempel–Ziv complexity-based fault evaluation measure, it would be possible to assess fault severity without comparing with reference data. However, several important issues are yet to be addressed, i.e.,

(a) There is no clear distinction between a mechanical fault and noise. A relatively high complexity could signify either the existence of a bearing fault or the presence of heavy noise. A signal produced by a mechanical fault has non-dimensional complexity close to 1.0 as the fault develops whereas the pure noise also has the same level of complexity.

(b) The previous complexity-based method yields quite different complexity levels for different data measurements of the same bearing. Therefore, many data sets are required to calculate the average of the complexity values of the data sets. The average complexity value is then considered a reliable indication of the fault severity. This could cause delay in maintenance decision and could become an obstacle for on-line applications.

(c) A theoretical explanation based on the cause and mechanism of bearing faults is required as for why the complexity value can be used for measuring fault severity. A connection between the complexity measure and fault severity can be established only when this is done.

In view of the above, a new version of the Lempel–Ziv complexity is developed for bearing fault severity assessment based on the continuous wavelet transform (CWT). As the Lempel–Ziv complexity is bounded between 0 and 1, it will provide a specific limit such that a complexity value close to such a limit would clearly signal the need for replacement. Nevertheless, it should be noted that the initial complexity value of a new bearing is needed in order to interpret the complexity value and its trend. The rest of the paper is organized as follows. Section 2 presents an overview of the analytical signal model for a single-point defective bearing. Section 3 describes the Lempel–Ziv complexity as a severity measure of bearing outer- and inner-race faults. The proposed algorithm based on the Lempel–Ziv complexity and CWT is detailed in Section 4. The application results of the proposed method to bearing outer- and inner-race fault signals are illustrated in Section 5. The conclusions of this work are given in Section 6.

2. A model for a single-point defective bearing

For a single-point defect of a rolling element bearing, a signal model consists of several simplified terms [15,16] as given by

$$x(t) = x_f(t)x_l(t)x_b(t) + n(t), \tag{1}$$

where $x(t)$ is a measured vibration signal, $x_f(t)$ the impulse series produced by the fault, $x_l(t)$ the modulation effect caused by the non-uniform load distribution, $x_b(t)$ the bearing-induced vibration including decay, and

$n(t)$ the mixture of unwanted signal components yielded by other parts of the machine and by noise. Each vibration term can be modeled as a sinusoidal function as described in the following subsections.

2.1. Impulses generated by fault

For a constant rotating speed, the impulses generated by a single-point defect is simulated in time and frequency domains with frequency f_f and period $T_f = 1/f_f$ as follows [15]:

$$x_f(t) = d_f \sum_{k=-\infty}^{\infty} \delta(t - kT_f), \tag{2a}$$

where d_f is the amplitude, $\delta(\cdot)$ the unit impulse function. These repetitive impulses may be approximated by a Fourier series, a sinusoidal function with frequency f_f and its harmonics, as follows [17,18]

$$x_f(t) \approx A_f \sum_{i=1}^{N_1} \cos(2\pi i f_f t + \phi_f^i), \tag{2b}$$

where ϕ_f^i is the phase of ith harmonic, A_f the amplitude, and N_1 the number of harmonics induced by impulse series. This approximation is also adopted in Ref. [16].

2.2. Non-uniform load distribution

The following Stribeck equation explains the vibration caused by non-uniform load distribution around the rolling element bearing [19]

$$x_l(\theta) = P_{\max}\left[1 - \frac{1}{2\varepsilon}(1 - \cos\theta)\right]^m, \tag{2c}$$

where $P_{\max}$ is the maximum contact pressure at a circumferential angle θ around shaft as shown in Fig. 1, ε is the load distribution factor, and $m = 1.5$ for ball bearings and $m = 10/9$ for other rolling element bearings.

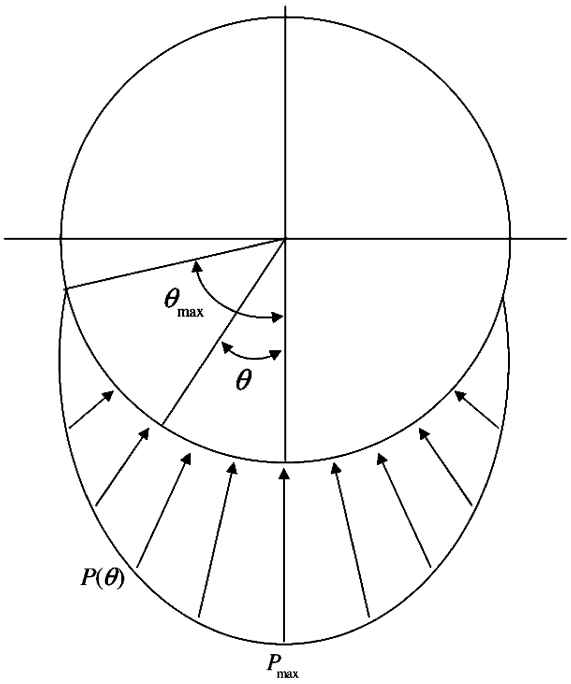

Fig. 1. Load distribution in a bearing under radial load.

The terms $P_{\max}$, $\theta_{\max}$ and ε are dependent on the bearing clearance and the applied load. For bearing with positive clearance, $\varepsilon<0.5$ and $\theta_{\max}<\pi/2$ [15]. In general, the inner race of the bearing is attached to a shaft rotating at a constant frequency f_s. Therefore, Eq. (2c) can be written in terms of time and shaft rotational frequency by using $\theta(t)=2\pi f_s t$, i.e.,

$$x_l(t)=\begin{cases} P_{\max}\left[1-\dfrac{1}{2\varepsilon}(1-\cos\,\theta(t))\right]^m, & \text{for } |\theta(t)|=|2\pi f_s t|<\theta_{\max} \\ 0, & \text{elsewhere.} \end{cases} \tag{3}$$

2.3. Bearing-induced vibration

The high-frequency resonance vibration with decay can be modeled as a superposition of all vibration modes [16]:

$$x_b(t)=\sum_{j=1}^{N_2} A_b^j \mathrm{e}^{-\alpha_b^j t}\,\cos(2\pi f_b^j t+\phi_b^j), \tag{4}$$

where $A_b{}^j$, $f_b{}^j$, $\phi_b{}^j$, and $\alpha_b{}^j$ are the amplitude, resonance frequency, phase, and damping factor of exponential decay of jth vibration mode of the bearing, respectively, and N_2 is modal order of bearing vibration.

3. Complexity as a measure of fault severity—the rationale

The complexity of a signal increases when more frequency components exist. For mechanical parts such as gears and bearings, contact pressure between two mating parts changes when a fault exists. The contact pressure variation results in amplitude and frequency modulations and hence more frequency peaks will appear in the frequency domain with their harmonics [20]. Therefore, the fault progress leads to complexity (randomness) changes because a larger fault intensifies variation of the contact pressure between two mating parts.

The signal models of different vibration sources described in the previous section are adopted to evaluate the complexities of inner- and outer-race fault signals. This study focuses on a bearing with a fixed outer race. In this case, the rotating speed of the inner race is the same as that of the shaft. It is further assumed: (a) the initial phase angles are zero, and (b) the modulation frequencies caused by a non-uniform load and the fault frequencies of impulses are lower than the resonance frequencies of bearing, i.e.,

$$f_b^j \gg f_f > f_s \quad (j=1,2,\ldots,N_2), \tag{5}$$

where $f_b{}^j$, f_f, and f_s are the resonance frequency of the bearing, the frequency of the impulses generated by the fault, and the shaft rotating frequency, respectively.

3.1. Outer-race fault

For the outer-race fault, the non-uniform load distribution does not have a strong effect on the outer-race fault signal since the relative location of fault is not affected by the non-uniform load distribution. Therefore, the non-uniform load term in Eq. (1) is negligible, i.e., $x_l(t)\approx 1$. Then, the signal model of outer-race fault is

$$x_{\mathrm{OR}}(t)=x_f(t)\,x_b(t)+n(t)=A_f\sum_{i=1}^{N_1}\cos(2\pi i f_f t+\phi_f^i)\sum_{j=1}^{N_2} A_b^j \mathrm{e}^{-\alpha_b^j t}\,\cos(2\pi f_b^j t+\phi_b^j)+n(t). \tag{6}$$

Based on the assumptions mentioned above, Eq. (6) can be written as

$$x_{\mathrm{OR}}(t)=\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} A_f A_b^j \mathrm{e}^{-\alpha_b^j t}\,\cos(2\pi i f_f t)\,\cos(2\pi f_b^j t)+n(t)=\sum_{i=1}^{N_1}\sum_{j=1}^{N_2} A_f A_b^j \mathrm{e}^{-\alpha_b^j t}\,\frac{1}{2}(X_1+X_2)+n(t), \tag{7}$$

where

$$X_1 = \cos[2\pi(f_b^j + if_f)t] = \cos\left[2\pi f_b^j\left(1 + \frac{if_f}{f_b^j}\right)t\right]$$

$$X_2 = \cos\left[2\pi(f_b^j - if_f)t\right] = \cos\left[2\pi f_b^j\left(1 - \frac{if_f}{f_b^j}\right)t\right].$$

It should be noted that the two terms, X_1 and X_2, are expressed with the frequency modulation effect.

According to Howard [21], there are three dominant factors in impact action, i.e., bearing geometry, rotational speed, and fault size. As the effects of the geometry and the rotational speed are constant for the same bearing under stationary operating conditions, the fault size is the most important factor for the impact action. Consider each term in Eq. (8). As fault size increases, the impulse associated with $x_f(t)$ gets dull since the edges of the fault are flattened by repeated impacts [13]. The dull edges can no longer produce sharp impulses and the strength of sharp impulses decreases as the fault grows. Such a phenomenon has been confirmed by the shrinking peak magnitudes in the frequency spectrum as the inner race fault expands [13]. Therefore, the contribution of term $x_f(t)$ in Eq. (8) diminishes as the fault size increases.

The resonance frequency is constant for the same bearing since it is the system's unique characteristic. Therefore, the bearing resonance vibration term, $x_b(t)$, is considered to have less contribution in Eq. (8) as fault size increases.

Based on the above, complexity of the outer-race fault is explained as follows. Consider one of rolling elements and the bearing outer race. For a healthy bearing, the contact pressure between the ball and outer race does not change much with respect to the angular position of the ball. However, for a faulty bearing, the contact pressure increases when the ball hits the fault on the surface of the outer race as the actual contact area (the nominal contact area minus the fault size) becomes smaller because of the fault. Hence, the contact pressure becomes variable when a fault exists on the outer race. The fluctuating contact pressure results in frequency variation that leads to the frequency modulation effect [20] as shown in Eq. (7). As the size of fault on the outer race increases, fluctuation of the contact pressure intensifies, which reinforces the frequency modulation effect. The stronger frequency modulation effect is evidenced by the increased number of frequency peaks in the frequency domain such as harmonics. The larger number of frequency peaks renders higher complexity. Therefore, one can expect that, for the outer-race single-point defect, the complexity will increase as the defect develops.

3.2. Inner-race fault

The radial load applied to the rotating shaft causes non-uniform load distribution around the inner race of the bearing as illustrated in Fig. 1. The non-uniform load distribution causes low-frequency amplitude modulation with a frequency of f_s as shown in Eq. (3). For simplicity, consider a bearing with $\varepsilon = 0.5$ and $\theta_{\max} = \pi/2$. The non-uniform load distribution term is accordingly

$$x_l(t) = \lim_{\varepsilon \to 0.5} P_{\max}\left[1 - \frac{1}{2\varepsilon}(1 - \cos(2\pi f_s t))\right]^m = P_{\max}\left[\cos(2\pi f_s t)\right]^m, \tag{8}$$

where m (= 1.5, if a ball bearing is considered) is negligible since it does not have any effect on the frequency component of $x_l(t)$. Eq. (8) can be rewritten as [16,18]

$$x_l(t) \approx \sum_{k=1}^{N_3} A_l^k \cos(2\pi k f_s t + \phi_l^k), \tag{9}$$

where A_l^k, ϕ_l^k, and N_3 stand for the amplitude, phase of kth harmonic and the number of harmonics, respectively. Then, the signal model for an inner-race fault is

$$\begin{aligned} x_{\mathrm{IR}}(t) &= x_f(t)\, x_b(t)\, x_l(t) + n(t) \\ &= A_f \sum_{i=1}^{N_1} \cos(2\pi i f_f t + \phi_f^i) \sum_{j=1}^{N_2} A_b^j \mathrm{e}^{-\alpha_b^j t} \cos(2\pi f_b^j t + \phi_b^j) \sum_{k=1}^{N_3} A_l^k \cos(2\pi k f_s t + \phi_l^k) + n(t). \end{aligned} \tag{10}$$

The above equation can be rewritten as

$$x_{\mathrm{IR}}(t) = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\sum_{k=1}^{N_3} A_f A_l^k A_b^j \mathrm{e}^{-\alpha_b^j t}\cos(2\pi i f_f t)\cos(2\pi f_b^j t)\cos(2\pi k f_s t) + n(t)$$
$$= \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\sum_{k=1}^{N_3} A_f A_l^k A_b^j \mathrm{e}^{-\alpha_b^j t}\frac{1}{4}(X_3 + X_4 + X_5 + X_6) + n(t), \tag{11}$$

where

$$X_3 = \cos[2\pi(f_b^j + if_f + kf_s)t]$$
$$X_4 = \cos[2\pi(f_b^j - if_f - kf_s)t]$$
$$X_5 = \cos[2\pi(f_b^j - if_f + kf_s)t]$$
$$X_6 = \cos[2\pi(f_b^j + if_f - kf_s)t].$$

The inner-race fault signal creates three major frequency components as shown in Eq. (11).

For an inner-race fault signal, a non-uniform load distribution plays an important role in the vibration signal. The vibration term corresponding to the non-uniform load distribution, $x_l(t)$, creates low-frequency amplitude modulation. The amplitude of this term is dependent on the maximum contact pressure $P_{\max}$. For a ball bearing, the surfaces of a ball and the inner race contact at a single point, which is called *point contact*. The contact point expands to an ellipse when the load is applied [19,22]. The maximum contact pressure is located at the center of the ellipse and its magnitude is [22]

$$P_{\max} = 1.5\frac{F}{A}, \tag{12}$$

where F is the force by which the components (ball and inner race) are pressed against each other and A is the actual contact area between the two components (the nominal contact area minus the fault size). If we approximate the defect area (the cavity caused by a fault) using a circle with radius r, the area corresponding to the fault would swell in proportion to n^2r^2, as the fault size (diameter) increases to $nr(n>1)$. Therefore, the actual area of contact between the ball and the inner race decreases as the fault size grows. This leads to a rapid increase of the maximum contact pressure as suggested by Eq. (12) and a stronger low-frequency amplitude modulation effect. Hence, the non-uniform load distribution dominates among three vibration sources for the inner-race fault signal as expresses in Eq. (12). This dominance indicates a low-frequency pattern formation. As a strong pattern starts to form, the randomness decreases and the orderliness increases, leading to a low complexity value [23]. Hence, the complexity values of the inner-race fault signals are expected to decrease as the fault grows, which is opposite to the outer-race case.

3.3. Lempel–Ziv complexity

Lempel and Ziv introduced an easily calculable measure of complexity of finite sequences, which adopted two basic processes: copy and insert [24]. Fig. 2 explains the procedure of calculating the Lempel–Ziv complexity for a finite sequence of length N. Consider a string $S_N = \{s_1 s_2 \ldots s_N\}$. Assume that a string up to $s_r(1<r<N)$ of complexity $c_N(r)$ can be reconstructed by simply copying and inserting some of the existing vocabulary of $S_{v,r-1} = \{s_1 s_2 \ldots s_v\}(v<r)$. To check the rest string $S_{N-r} = \{s_{r+1}\ldots s_N\}$ can be re-produced by the same approach, Lempel and Ziv introduced the following steps:

Step 1: Take $Q_{r+1} = \{s_{r+1}\}$ and ask if this term belongs to the vocabulary of $S_{v,r} = \{S_{v,r-1}s_r\}$. If so, string $Q_{r+1} = \{s_{r+1}\}$ is a simple repetition of an existing substring of $S_{v,r}$ (i.e., a simple "*copy*" of existing vocabulary can restore it), and hence the complexity remains unchanged or $c_N(r+1) = c_N(r)$.

Step 2: Read the next string and take $Q_{r+2} = \{s_{r+1}s_{r+2}\}$. Check if $Q_{r+2} = \{s_{r+1}s_{r+2}\}$ belongs to $S_{v,r+1} = \{S_{v,r-1}s_r s_{r+1}\}$ (obtained by augmenting S_v with s_{r+1}).

Step 3: If the term Q_{r+2} does not belong to $S_{v,r+1}$, increase the complexity by one, i.e., $c_N(r+2) = c_N(r+1)+1$, nullify $Q_{r+2} = \{\}$, read the next string and take $Q_{r+3} = \{s_{r+3}\}$.

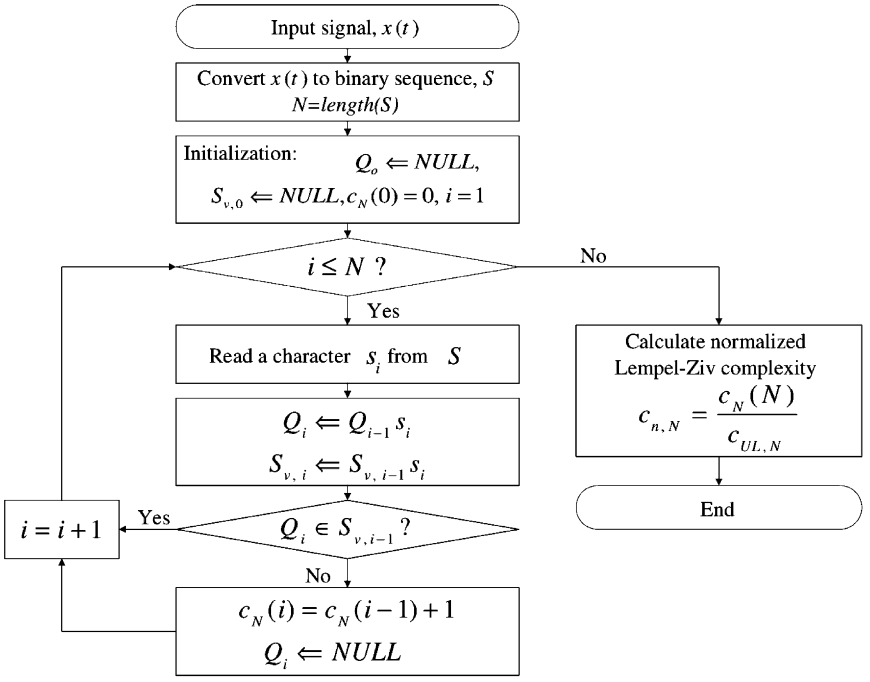

Fig. 2. Block diagram for calculating the Lempel–Ziv complexity.

Repeats the above procedure until $S_N = \{s_1 s_2 \ldots s_N\}$ is completely covered. Then, the resulting $c_N(N)$ is the complexity of a given string.

Considering an example of a finite string $S = \{0001101001000101\}$ of length $N = 16$, the above procedure can be illustrated as follows:

$$\text{Initialization}: \quad i=0, \quad S_{v,0}=\{\}, \quad Q_0=\{\}, \quad c_N(0)=0.$$

$$i=1, \quad S_{v,1}=\{0\}, \quad Q1=\{0\}, \quad Q_1 \notin S_{v,0} \rightarrow c_N(1)=1, \quad Q_1=\{\}.$$

$$i=2, \quad S_{v,2}=\{00\}, \quad Q_2=\{0\}, \quad Q_2 \in S_{v,1} \rightarrow c_N(2)=1.$$

$$i=3, \quad S_{v,3}=\{000\}, \quad Q_3=\{00\}, \quad Q_3 \in S_{v,2} \rightarrow c_N(3)=1.$$

$$i=4, \quad S_{v,4}=\{0001\}, \quad Q_4=\{001\}, \quad Q_4 \notin S_{v,3} \rightarrow c_N(4)=2, \quad Q_4=\{\}.$$

$$i=5, \quad S_{v,5}=\{00011\}, \quad Q_5=\{1\}, \quad Q_5 \in S_{v,4} \rightarrow c_N(5)=2.$$

$$i=6, \quad S_{v,6}=\{000110\}, \quad Q_6=\{10\}, \quad Q_6 \notin S_{v,5} \rightarrow c_N(6)=3, \quad Q_6=\{\}.$$

$$i=7, \quad S_{v,7}=\{0001101\}, \quad Q_7=\{1\}, \quad Q_7 \in S_{v,6} \rightarrow c_N(7)=3.$$

$$i=8, \quad S_{v,8}=\{00011010\}, \quad Q_8=\{10\}, \quad Q_8 \in S_{v,7} \rightarrow c_N(8)=3.$$

$$i=9, \quad S_{v,9}=\{000110100\}, \quad Q_9=\{100\}, \quad Q_9 \notin S_{v,8} \rightarrow c_N(9)=4, \quad Q_9=\{\}.$$

$$i=10, \quad S_{v,10}=\{0001101001\}, \quad Q_{10}=\{1\}, \quad Q_{10} \in S_{v,9} \rightarrow c_N(10)=4.$$

$$i=11, \quad S_{v,11}=\{00011010010\}, \quad Q_{11}=\{10\}, \quad Q_{11} \in S_{v,10} \rightarrow c_N(11)=4.$$

$$i=12, \quad S_{v,12}=\{000110100100\}, \quad Q_{12}=\{100\}, \quad Q_{12} \in S_{v,11} \rightarrow c_N(12)=4.$$

$$i=13, \quad S_{v,13}=\{0001101001000\}, \quad Q_{13}=\{1000\}, \quad Q_{13} \notin S_{v,12} \rightarrow c_N(13)=5, \quad Q_{13}=\{\}.$$

$$i=14,\quad S_{v,14}=\{00011010010001\},\quad Q_{14}=\{1\},\quad Q_{14}\in S_{v,13}\rightarrow c_N(14)=5.$$

$$i=15,\quad S_{v,15}=\{000110100100010\},\quad Q_{15}=\{10\},\quad Q_{15}\in S_{v,14}\rightarrow c_N(15)=5.$$

$$i=16,\quad S_{v,16}=\{0001101001000101\},\quad Q_{16}=\{101\},\quad Q_{16}\in S_{v,15}\rightarrow c_N(16)=5.$$

Now, the string shown in this example has a complexity of $c_N(16)=5$.

The complexity obtained above is equal to the number of nullifications of Q. This is, to a certain extent, affected by the length of the string, or the number of data samples N. To be robust, a complexity measure should be independent of the number of data samples N. To find a robust complexity measure, Lempel and Ziv [24] suggested a normalized Lempel–Ziv complexity measure after their names, c_{nN}, defined by

$$0\leqslant c_{nN}=\frac{c_N(N)}{c_{UL,N}}\leqslant 1, \tag{13}$$

where

$$c_{UL,N}=\lim_{N\rightarrow\infty}c_N(N)=\lim_{N\rightarrow\infty}\frac{N}{(1-\beta)\log_\kappa N}\approx\frac{N}{\log_\kappa N}, \tag{14}$$

where κ is the size of the alphabets (or distinct elements) used in string S (for binary string, $\kappa=2$), $\beta\rightarrow 0$ if $N\rightarrow\infty$ [24,25].

However, in spite the effort of Lempel and Ziv, the "normalized" c_{nN} is still affected by N when N is not sufficiently large. This can be illustrated as follows. To examine the effect of the number of data samples on the normalized complexity, a series of binary strings of different length is randomly generated based on the Gaussian distribution. Their normalized complexities are calculated following the procedure shown in Fig. 2 and are plotted in Fig. 3. Since the data are random, the normalized complexity approaches 1.0 as the string

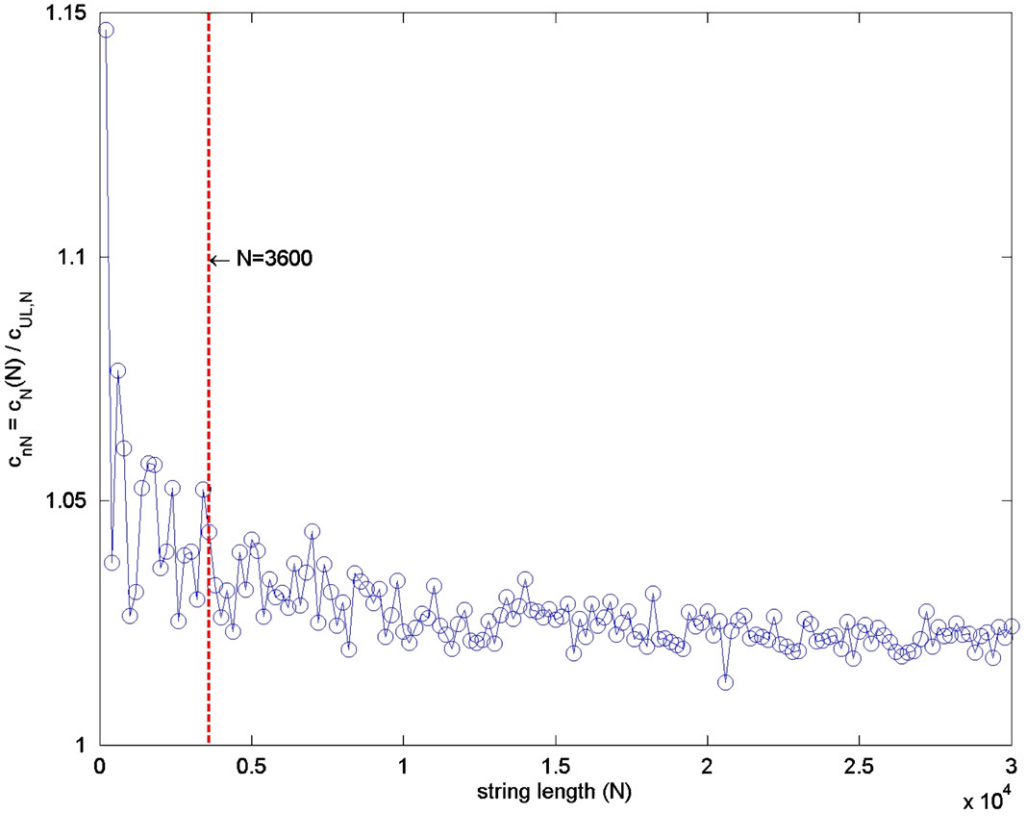

Fig. 3. The effect of data length on the normalized complexity for randomly generated data.

length increases. The normalized complexity drops below 1.05 for $N \geqslant 3600$, which is very close to the upper bound, 1, of the "normalized" complexity. This procedure therefore could be used for selecting the number of data samples to obtain a complexity index ranging roughly between [0, 1].

4. Calculating complexity of bearing fault signal

Yan and Gao [14] applied the Lempel–Ziv complexity to the cylindrical roller bearing and obtained different complexity values for different fault sizes. According to their study, the normalized Lempel–Ziv complexity approaches the upper limit, around 1.0, as a fault develops. They calculated the Lempel–Ziv complexity from the measured acceleration signal. However, the measured signal is contaminated by noise. This may increase the complexity value and it could lead to inaccurate diagnosis. As shown in Fig. 3, the pure random noise has a normalized complexity value around 1.0. As illustrated later, the Lempel–Ziv complexity of the signal of severe bearing faults is close to 1.0. Therefore, it is difficult to distinguish the sources (noise or fault) of the complexity value. To alleviate this difficulty, the following procedure is proposed, which is further detailed in Fig. 4:

Step 1: Apply continuous wavelet transform (CWT) to the measured signal.

Step 2: Obtain the energy and kurtosis in each scale level and find the best scale level by maximizing energy and kurtosis distributions over scales used in step 1.

Step 3: Calculate normalized complexity of the wavelet coefficients at the best scale level.

The rationale of finding the best scale in step 2 is to extract the fault feature and exclude the interferences from other frequency ranges. Then, the normalized complexity is calculated only based on the extracted fault feature in step 3. In this way, much of the noise components have been eliminated and hence the ambiguity between noise and severe faults can be cleared.

4.1. Continuous wavelet transform (CWT)

As shown in Fig. 4, the procedure starts with mapping the measured vibration signal into the scale-time domain by CWT. The CWT is selected since it provides finer resolution in the scale-time plane than discrete wavelet transform (DWT) does. The CWT is expressed as follows:

$$W(a,b) = \int x(t) w^*_{a,b}(t)\,\mathrm{d}t \quad (a = a_{\min}, \ldots, a_{\max}), \tag{15}$$

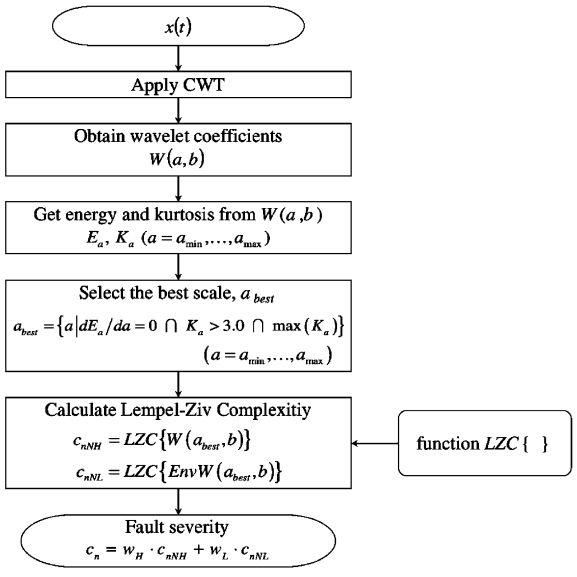

Fig. 4. Block diagram of Lempel–Ziv complexity as a measure of fault severity.

where $a_{\min}$ and $a_{\max}$ are the smallest and the largest scale levels of interest, asterisk (*) indicates complex conjugate and $w_{a,b}(t)$ is the daughter wavelet defined as

$$w_{a,b}(t) = \frac{1}{|a|^{1/2}} w\left(\frac{t-b}{a}\right), \tag{16}$$

where a is dilation (scale) factor and b is translation factor in time. The factor $|a|^{-1/2}$ is used to ensure energy preservation [26]. In this study, the Morlet wavelet as shown below is used as a mother wavelet function because of its similarity to the impulsive mechanical fault feature in shape [27,28]:

$$w(t) = \exp\left(-\frac{\gamma^2 t^2}{2}\right)\cos(\omega_0 t), \tag{17}$$

with $\gamma = 1$ and $\omega_0 = 5$. A measured signal normally contains many unnecessary signal components with Gaussian behavior such as noise and interfering vibration generated by other machine elements. As mentioned earlier, the normalized complexity of Gaussian noise approaches the upper limit of 1.0. Therefore, complexity of the signal contaminated by noise cannot truly represent the severity of fault features. By applying wavelet transform, one can select a scale level that reveals fault features with less interference of unnecessary signal components. This ensures a relevant and reliable complexity value for the fault features under a noisy environment.

4.2. Selection of the best scale

Since each scale level shows different frequency contents of the measured signal $x(t)$, only the scale level which best reflects the signal nature of the repetitive impulses should be selected. In this study, the energy and kurtosis criteria are jointly used for scale selection. The reasons are: (a) the energy level is a measure of the shape similarity of the measured signal and the daughter wavelet function, (b) unlike in the application for basis or node selection in wavelet packet transform [29], the energy criterion cannot be used alone in this study because a higher energy level is also associated with higher scale levels of CWT, and (c) kurtosis is a good indicator of impulsiveness. Hence, the joint use of energy and kurtosis criteria would take into account of both the signal-wavelet similarity and the fault signature impulsiveness. The joint use of the two would also compensate for the incapability of the energy criterion in distinguishing whether the high energy level is due to high signal-wavelet similarity or high scale level. The energy of the wavelet coefficients in each scale level can be calculated by

$$E_a = \sum_b \{W(a,b)\}^2 \quad a = a_{\min}, \ldots, a_{\max}, \tag{18}$$

where E_a represents the energy of the wavelet coefficients in scale a. Several peaks may appear in the plot of E_a. Each scale associated with a peak energy value is considered as a candidate for the best scale. However, as indicated earlier, energy alone cannot guarantee the best scale since the magnitude of the wavelet coefficients increases as scale level increases for energy conservation [26]. The energy level becomes higher as scale level increases, which may mislead the selection process of the best scale. This is further explained below using energy distribution of daughter wavelets.

The Fourier transform of the Morlet daughter wavelets can be expressed as [30]:

$$w(t) = \mathrm{e}^{-t^2/2}\mathrm{e}^{-\mathrm{j}\omega_0 t}, \quad \hat{w}(\omega) = \mathrm{e}^{-(\omega-\omega_0)^2/2}, \quad \hat{w}_{a,b}(\omega) = \sqrt{a}\,\mathrm{e}^{-(\omega/a-\omega_0)^2/2}\mathrm{e}^{\mathrm{j}\omega b}, \tag{19}$$

where $\hat{w}(\omega)$ and $\hat{w}_{a,b}(\omega)$ are the Fourier transforms of the Morlet wavelet $w(t)$, and its daughter wavelets $w_{a,b}(t)$, respectively. Fig. 5 illustrates the normalized amplitude of the daughter wavelet functions at their center frequencies ($b = 1$ is used for illustration purpose and $f_{c,a}$ is center frequency at scale a). As shown in the figure, the amplitude drastically increases with scale. Since the CWT is a series of band-pass filters with increasing amplification rate, wavelet coefficients will have larger amplitude as scale increases, which leads to higher energy levels.

To offset the above drawback of the energy criterion, the kurtosis distribution of wavelet coefficients is examined. The measured faulty signal contains impulses produced by mechanical fault, and hence, the best

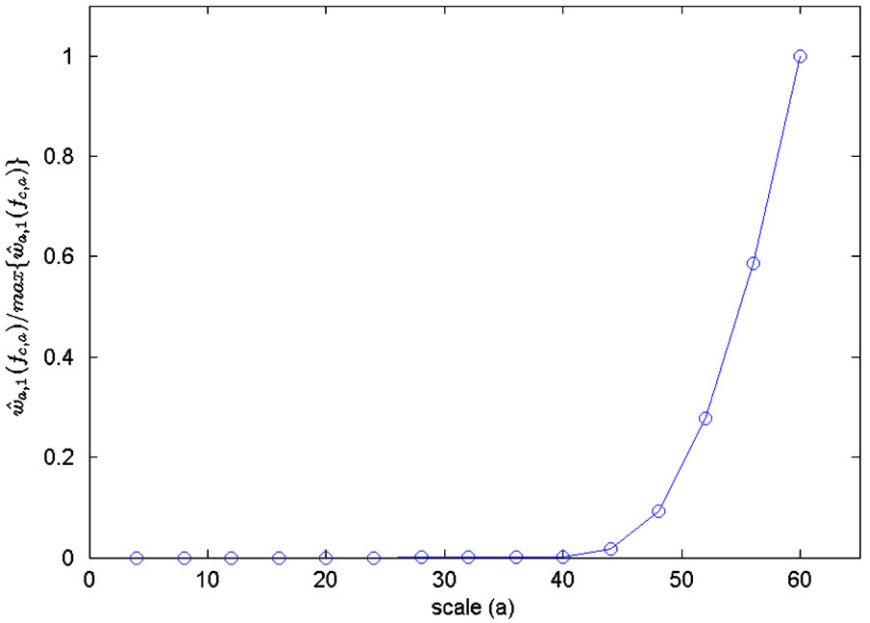

Fig. 5. Normalized amplitude distribution of daughter wavelet functions.

scale level should reveal the nature of these impulses. The energy level does not always reflect this nature since it is calculated based on the amplitude of the wavelet coefficients instead of the impulsiveness of a signal. Therefore, using energy value alone is not adequate to secure the best scale level that manifests the impulsive character of the mechanical fault signature. On the other hand, the kurtosis value is sensitive to the outliers or impulses generated by mechanical fault because it is proportional to the fourth-order higher moment defined as [31]

$$\text{kurtosis} = \frac{1}{N}\frac{\sum_{i=1}^{N}(x(i)-\bar{x})^4}{\sigma^4}, \tag{20}$$

where $\bar{x}$ and σ are, respectively, the mean value and the standard deviation of the measured signal x, and N is the number of data samples of the measured signal. A signal with a Gaussian distribution yields a kurtosis value of 3.0 whereas the kurtosis of a signal with a sub-Gaussian distribution is less than 3.0. An impulsive signal with super-Gaussian distribution has a kurtosis value greater than 3.0, a typical feature of a mechanical fault. Therefore, considering both energy and kurtosis would provide more reliable scale selection process. The best scale can then be found by

$$a_{\text{best}} = \left\{ a \left| \frac{\mathrm{d}E_a}{\mathrm{d}a} = 0 \cap \max(K_a) \cap K_a > 3.0 \right. \right\} \quad a \in [a_{\min}, a_{\max}], \tag{21}$$

where K_a is kurtosis at scale a and a_{best} is the selected best scale. The scale selected by Eq. (21) is considered to contain the signal nature closest to the impulses produced by the mechanical fault. This selection method can thus reduce the side effects of noise and un-related signals.

4.3. Calculation of Lempel–Ziv complexity

The Lempel–Ziv complexity values are calculated for the wavelet transformed signal and its envelope at the best scale a_{best}. As shown with the signal models in the previous section, the bearing fault signal consists of higher-carrier and lower-modulating frequencies. Therefore, using the Lempel–Ziv complexity for both high- and low-frequency signals is suitable for measuring bearing fault severity. The overall procedure is as follow:

Step 1: Obtain the high- and low-frequency signals from the wavelet coefficients at the best scale.

Step 2: Convert the high- and low-frequency signals into binary sequences.

Step 3: Calculate the Lempel–Ziv complexity of the binary sequences using Eq. (13).

More details are given as follows. In step 1, the wavelet coefficients at the best scale a_{best} are demodulated to obtain the high-frequency carrier signal and low-frequency envelope signal. The Hilbert transform is applied to the wavelet coefficients to find envelope signal of the wavelet coefficients, i.e., the low-frequency signal.

In the second step, both high- and low-frequency signals are converted into binary sequences. Any values greater than or equal to the mean value of the given signal are replaced with one and others with zero.

In step 3, the normalized Lempel–Ziv complexity values for the high- and low-frequency signals are calculated by

$$c_{n\text{NH}} = \left\{\frac{c_N(N)}{c_{UL,N}}\right\}_{\text{high freqeuncy}} = \text{LZC}\{W(a_{\text{best}}, b)\},$$
$$c_{n\text{NL}} = \left\{\frac{c_N(N)}{c_{UL,N}}\right\}_{\text{low frequency}} = \text{LZC}\{\text{Env}\,W(a_{\text{best}}, b)\}, \tag{22}$$

where $\text{Env}\,W(a_{\text{best}},b)$ and $\text{LZC}\{\cdot\}$ indicate the envelope signal of wavelet coefficients at the best scale a_{best} and the function for calculating Lempel–Ziv complexity, respectively. An indicator of fault severity is then obtained based on the two normalized complexity values as follows:

$$c_n = w_H c_{nNH} + w_L c_{nNL}, \tag{23}$$

where w_H and w_L are the weights of high-frequency carrier signal and low-frequency envelope signal. Equal weight is assigned to each dominant signal components. The outer-race fault signal contains two dominant signal components, i.e., bearing resonance and impulses generated by fault as shown in Eq. (6) or (7). The two signal components are corresponding to the high-frequency carrier signal and low-frequency envelope signal respectively and each receives 50% of the weight. However, as suggested by Eq. (9) or (10), there exist three dominant signal components in the inner-race signal contributed by bearing resonance, fault impulse and non-uniform load, respectively. Since the resonance frequency is usually much higher than the frequencies of the non-uniform load and fault impulses, it acts as the carrier frequency and the other two are classified as low frequencies. In summary, the weights are allocated as follows:

$$\begin{cases} w_H = 1/2 \\ w_L = 1/2 \end{cases} \text{for outer race}$$
$$\begin{cases} w_H = 1/3 \\ w_L = 2/3 \end{cases} \text{for inner race}.$$

5. Applications to bearing fault data

The proposed method is applied to the bearing fault signal. Two different sources of bearing fault signals are used to validate the proposed method: one from the Case Western Reserve University [32], and the other from Rubini and Meneghetti [13].

5.1. Application 1

The proposed method is applied to bearing fault signals obtained from the Case Western Reserve University [32]. Single-point defects of different sizes are created on the surfaces of inner and outer race of a deep groove ball bearing (SKF6205-2RS JEM). The fault sizes are 0.178, 0.356, 0.533 mm in diameter and 0.279 mm in depth. The sampling frequency is 12,000 Hz and 12,000 data samples are used. The calculated Lempel–Ziv complexity values are summarized in Fig. 6. The complexity of inner-race fault decreases as fault size increases because of the modulation effect of the non-uniform load distribution. As a fault grows, the edge of the fault cavity is smoothened out, leading to the decrease in the strength of sharp impulses. As such, the contribution of the associated frequency component to the low frequency envelope signal diminishes. On the other hand,

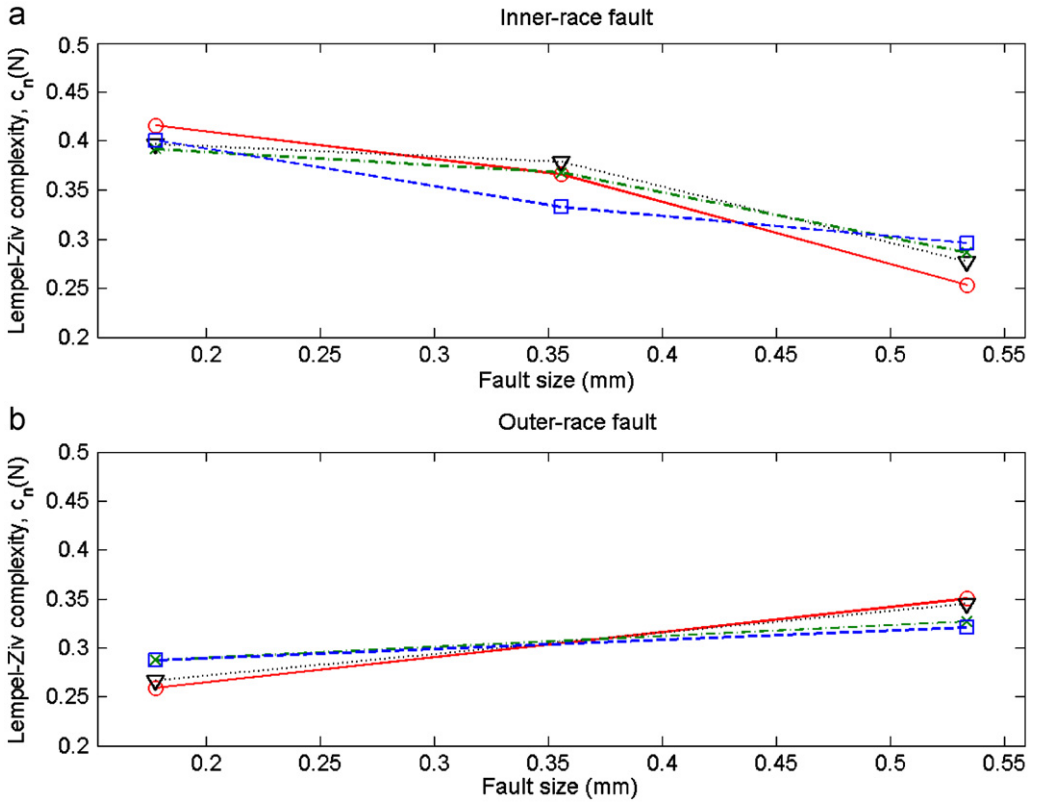

Fig. 6. Application 1. Lempel–Ziv complexity vs. fault size: (a) inner-race fault and (b) outer-race fault (note: 1797 rev/min, 1772 rev/min, 1750 rev/min, 1730 rev/min).

the actual contact area is reduced due to the growing fault area, which in turn strengthens the non-uniform load distribution effect. Because of the above, the vibration caused by the non-uniform load distribution is becoming a dominant contributor to the low-frequency envelope signal. The diminished contribution of the fault impulse and the increased effect of the non-uniform load distribution lead to the lower complexity which means more order or less randomness [23,33,34]. The complexity of outer-race fault, on the other hand, shows an increasing trend. This opposite trend is also validated by the work of Rubini and Meneghetti [13]. The outer-race fault of 0.356 mm diameter is not considered, as the data is not available. Fig. 6 clearly indicates that the Lempel–Ziv complexity is proportional (for outer-race fault) or inversely proportional (for inner-race fault) to fault size (severity) for all rotational speeds. The variances between different operating speeds are relatively small.

A closer look at the plots in Fig. 6 indicates that the complexity value is somewhat affected by shaft rotating speed. This is in agreement with the observation of Yan and Gao [14]. To further enhance the proposed method, the speed factor may be incorporated in future.

5.2. *Application 2*

To further evaluate the performance, the proposed method is applied to bearing fault signals from Rubini and Meneghetti [13]. The experimental data of three different fault sizes (small, medium, and large) are acquired for both inner and outer races of a double-row self-aligning ball bearing with 12 rolling elements per row (FAG 1204). The radial load on the bearing is set at 500 N and the shaft rotation speed is 1602 rev/min. The data are obtained at a sampling frequency of 20,000 Hz and 20,000 samples are used. The calculated Lempel–Ziv complexity values are summarized in Table 1. As shown in the table, the complexity of inner-race fault decreases and the complexity of outer-race fault increases as the faults develop. For the large outer-race

Table 1
Lempel–Ziv complexity vs. fault size (Application 2)

	Small	Medium	Large
Inner race fault	0.219	0.169	0.147
Outer race fault	0.190	0.243	N/A

fault, the size of the fault is roughly 4 mm × 4 mm in square shape, which is too large to be considered as a single-point defect compared to the ball size of 6.35 mm in diameter. In this case, the complexity value is not calculated since the best scale from CWT cannot be selected due to the low kurtosis value (below 3). The reason is that as a single-point defect becomes a distributed fault, the strength of impulsiveness weakens, leading to low kurtosis. Therefore, the selection criterion in Eq. (22) can no longer find a best scale that satisfies all conditions simultaneously. This finding could be used to predict the transition from single-point fault to distributed fault.

The above two applications based on data from different sources have shown that the proposed method is quite robust in assessing fault severity of single-point defects on the inner and outer races of a bearing. Though this method may not be used to assess the severity of distributed faults, it could reveal the transition between single-point and distributed faults. In addition, unlike the non-relative or un-bounded (including single-end bounded) measures whose values may not reveal a conclusive need for bearing replacement, the Lempel–Ziv complexity bounded between zero and one provides a specific limit such that a complexity value close to this limit would clearly signal the need for replacement. It is worth to point out that, for a healthy bearing, the "best" wavelet scale does not exist and hence the complexity value cannot be calculated. The healthy bearings are simply screened out at this stage without going to the next stage. This is an advantage of the proposed method due to the reduced computational effort.

6. Conclusion

Assessing fault severity is one of the main challenges in fault diagnosis. So far, the severity has often been evaluated based on comparison methods based on no-relative or un-bounded (including single-end bounded) measures. Though such measures may be useful, it is not an easy task to specify a safeguard limit. In addition, the magnitudes of non-relative measures of the signals are dependent on many factors such as bearing size, load, working conditions, as well as the gain used in data acquisition. Hence, the same signal magnitude does not necessarily indicate the same fault severity and such methods may not provide a reliable indication of fault severity. To reliably evaluate fault severity, many different reference data would have to be collected for different conditions if the comparison based methods are used. However, such data are often unavailable particularly for new working conditions. Even if the data can be collected, it would be very tedious and time-consuming to collect, store, maintain, and retrieve them especially for variable and less predictable working conditions. To avoid this difficulty, the Lempel–Ziv complexity, a non-dimensional index ranging roughly between 0 and 1, is adopted as a fault severity measure. As shown in this work, the value of such a normalized non-dimensional measure is mainly affected by fault size (severity) and should be less susceptible to other factors. As a result, the proposed severity measure can be used for ball bearings of different sizes under different operation environments. The performance of the method has been evaluated using experimental data from different sources. The results indicate that the fault severity can be assessed reasonably well by the proposed method. In addition, it is important to point out that the Lempel–Ziv complexity is calculated based on the best CWT scale level and hence the noise and many un-related signal components have been screened out. This makes the diagnosis result more reliable.

Acknowledgment

The authors are grateful to Professor K.A. Loparo of Case Western Reserve University, Professor R. Rubini of University of Modena and Reggio Emilia, and Professor U. Meneghetti of University of Bologna

for their kind permissions to use their bearing data. The authors would also like to thank the two reviewers for their valuable comments and suggestions. This study was supported by Natural Science and Engineering Research Council of Canada. Their support is greatly appreciated.

References

[1] M. Covacece, A. Introini, Analysis of damage of ball bearings of aeronautical transmissions by auto-power spectrum and cross-power spectrum, *Transactions of the ASME Journal of Vibration and Acoustics* 124 (2) (2002) 180–185.
[2] S. Loutridis, Damage detection in gear systems using empirical mode decomposition, *Engineering Structures* 26 (12) (2004) 1833–1841.
[3] G. Dalpiaz, A. Rivola, R. Rubini, Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears, *Mechanical Systems and Signal Processing* 14 (3) (2000) 387–412.
[4] Y. Yu, Y. Dejie, C. Junsheng, A roller bearing fault diagnosis method based on EMD energy entropy and ANN, *Journal of Sound and Vibration* 294 (2006) 269–277.
[5] N. Baydar, A. Ball, A comparative study of acoustic and vibration signals in detection of gear failures using Wigner-Ville distribution, *Mechanical Systems and Signal Processing* 15 (6) (2001) 1091–1107.
[6] D. Boulahbal, M.F. Golnaraghi, F. Ismail, Amplitude and phase wavelet maps for the detection of cracks in geared systems, *Mechanical Systems and Signal Processing* 13 (3) (1999) 423–436.
[7] N. Baydar, A. Ball, Detection of gear failures via vibration and acoustic signals using wavelet transform, *Mechanical Systems and Signal Processing* 17 (4) (2003) 787–804.
[8] S. Loutridis, A local energy density methodology for monitoring the evolution of gear faults, *NDT&E International* 37 (6) (2004) 447–453.
[9] G. Meltzer, N.P. Dien, Fault diagnosis in gears operating under non-stationary rotational speed using polar wavelet amplitude maps, *Mechanical Systems and Signal Processing* 18 (5) (2004) 985–992.
[10] Y. Shao, K. Nezu, Design of mixture de-noising for detecting faulty bearing signals, *Journal of Sound and Vibration* 282 (2005) 899–917.
[11] T. Williams, X. Ribadeneira, S. Billington, T. Kurfess, Rolling element bearing diagnostics in run-to-failure lifetime testing, *Mechanical Systems and Signal Processing* 15 (5) (2001) 979–993.
[12] A.M. Al-Ghamd, D. Mba, A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size, *Mechanical Systems and Signal Processing* 20 (2006) 1537–1571.
[13] R. Rubini, U. Meneghetti, Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearing, *Mechanical Systems and Signal Processing* 15 (2) (2001) 287–302.
[14] R. Yan, R.X. Gao, Complexity as a measure for machine health evaluation, *IEEE Transactions on Instrumentation and Measurement* 53 (4) (2004) 1327–1334.
[15] P.D. McFadden, J.D. Smith, Model for the vibration produced by a single point defect in a rolling element bearing, *Journal of Sound and Vibration* 96 (1) (1984) 69–82.
[16] Y.F. Wang, P.J. Kootsookos, Modeling of low shaft speed bearing faults for condition monitoring, *Mechanical Systems and Signal Processing* 12 (3) (1998) 415–426.
[17] P.D. McFadden, Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the meshing vibration, *Mechanical Systems and Signal Processing* 1 (2) (1987) 173–183.
[18] N. Tandon, A. Choudhury, An analytical model for the prediction of the vibration response of rolling element bearings due to a localized defect, *Journal of Sound and Vibration* 205 (3) (1997) 275–292.
[19] T.A. Harris, *Rolling Bearing Analysis*, second ed., Wiley, New York, 1984, pp. 151–152, 161–165.
[20] R.B. Randall, A new method of modeling gear faults, *Transactions of the ASME Journal of Mechanical Design* 104 (1982) 259–267.
[21] I. Howard, A review of rolling element bearing vibration—detection, diagnosis and prognosis, DSTO Aeronautical and Maritime Research Laboratory, DSTO-RR-0013, 1994.
[22] J. Brandlein, P. Eschmann, L. Hasbargen, K. Weigand, *Ball and Roller Bearings—Theory, Design and Application*, third ed., Wiley, New York, 1999, pp. 78–80, 117–118.
[23] F. Kaspar, H.G. Schuster, Easily calculable measure for the complexity of spatiotemporal patterns, *The American Physical Society Physical Review A* 36 (2) (1987) 842–848.
[24] A. Lempel, J. Ziv, On the complexity of finite sequences, *IEEE Transactions on Information Theory, IT* 22 (1) (1976) 75–81.
[25] T.M. Cover, J.A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991, pp. 319–326.
[26] P. Goupillaud, A. Grossmann, J. Morlet, Cycle-octave and related transforms in seismic signal analysis, *Geoexploration* 23 (1984) 85–102.
[27] J. Lin, L. Qu, Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis, *Journal of Sound and Vibration* 234 (1) (2000) 135–148.
[28] J. Lin, M. Zuo, K. Fyfe, Mechanical fault detection based on the wavelet de-noising technique, *Transactions of the ASME Journal of Vibration and Acoustics* 126 (1) (2004) 9–16.
[29] N.G. Nikolaou, I.A. Antoniadis, Rolling element bearing fault diagnosis using wavelet packets, *NDT&E International* 35 (3) (2002) 197–205.

[30] A. Abbate, C.M. DeCusatis, P.K. Das, *Wavelets and Subbands—Fundamentals and Applications*, Birkhäuser, Boston, 2002, pp. 107–118.
[31] A. Hyvärinen, J. Karhunen, E. Oja, *Independent Component Analysis*, Wiley, New York, 2001.
[32] Case Western Reserve University, Bearing Data Center (Seeded Fault Test Data), 〈http://www.eecs.cwru.edu/laboratory/bearing/welcome_overview.htm〉 (Last visited October 2007).
[33] S. Wolfram, Cellular automata as models of complexity, *Nature* 311 (4) (1984) 419–424.
[34] S. Wolfram, Origins of randomness in physical systems, *The American Physical Society Physical Review Letters* 55 (5) (1985) 449–452.