Gear Damage Severity Evaluation Based on Cross Recurrence Quantification Analysis

Pu Wang
School of Instrument Science and Engineering
Southeast University
Nanjing, China
Email: 15150696306@163.com

Ruqiang Yan*
School of Instrument Science and Engineering
Southeast University
Nanjing, China
*Corresponding author: ruqiang@seu.edu.cn

Abstract—This paper utilizes cross recurrence quantification analysis (CRQA) as a quantitative tool for gear damage severity assessment, and proposes a new approach to simplify the procedure of embedding parameters selection. In principle, the CRQA derives quantitative measures from the cross recurrence plot (CRP) between two different systems to quantitatively describe their dependencies. In the present study, five measures are investigated as quantitative indicators for crack severity assessment of wheel gears. By analyzing experimental data acquired from the gear test system with different levels crack sizes, simplification method for parameters selection narrows down the time cost and provides the possibility of online monitoring. Based on results, several CRQA variables are found to be suited for evaluating the crack severity level.

Keywords—cross recurrence quantification analysis; embedding parameters selection; damage severity assessment

I. INTRODUCTION

Mechanical equipment like gear system, one of the core transmissions, gradually superseded the human labor with the development of science and technology. Since defects occurred in mechanical equipment always lead to disastrous accident, real-time fault diagnosis of machine is necessary [1]. As a result, researches on condition monitoring and detection of manufacturing systems advance rapidly across a wide range of application. As nonlinear dynamics reflected by difficult mechanical process can’t be described accurately based on linear method like Fourier transform, non-linear time series analysis techniques such as approximate entropy, box dimension and Lyapunov exponent growth rapidly [2]. Cross recurrence quantification analysis (CRQA), one of the non-linear methods, proposed recently performed well in short data analysis with good anti-noise ability has been widely used in many fields [3].

Shalbaf at [4] used CRQA to quantify frontal-temporal synchronization of EEG signals. Serra at [5] applied cross recurrence to identify cover song. Following these successful applications, this paper uses the cross recurrence quantification analysis to perform gear crack severity assessment, trying to build the relationship between results of CRQA and different crack severity. By analyzing the experimental data, the ability of the CRQA for damage severity assessment of gear crack is verified. In this paper, phase space reconstruction is introduced firstly, followed by the description of recurrence algorithm based on phase space reconstruction in section II. In section III, an approach to simplify reconstruction parameters selection is verified and multivariate cross recurrence quantification are applied to assessing crack severity of the gears. After that, CRQA based on single signal acquired at different locations are compared to test cross recurrence analysis’s ability in short-weak signal detection. Finally, some conclusions are drawn in section IV.

II. CROSS RECURRENCE PLOT AND CROSS RECURRENCE QUANTIFICATION ANALYSIS

A. Phase Space Reconstruction

Systems in engineering typically possess nonlinear characteristics so that traditional linear methods cannot meet the demand for describing the dynamics of complex mechanical system. Therefore, non-linear time series analysis techniques have been developed, where most of them, including the cross recurrence plot (CRP), are based upon topological analysis of the phase space of the underlying dynamics [6].

Chaos theory reveals that evolution of the system’s components couplings with each other, which indicates that a single component may contain dynamic information of the whole system. Correspondingly, an equivalent phase space trajectory that characterizes the topological structures of the original one can be reconstructed using only one observation [7]. A time delay method frequently used for reconstructing such a trajectory is based on the Takens embedding theorem [8]. Specifically, for a one-dimensional time series \( \{x_1, x_2, \ldots, x_n\} \), the reconstructed phase space is generated as:

\[
\begin{align*}
X_1 &= (x_1, x_1+\tau, x_1+2\tau, \ldots, x_1+(m-1)\tau) \\
X_2 &= (x_2, x_2+\tau, x_2+2\tau, \ldots, x_2+(m-1)\tau) \\
&\quad \vdots \\
X_i &= (x_i, x_i+\tau, x_i+2\tau, \ldots, x_i+(m-1)\tau) \\
&\quad \vdots
\end{align*}
\]

where \( i = 1, 2, \ldots, N \), \( N = n - (m - 1)\tau \) is the total number of points in the reconstructed phase space. \( \tau \) is the time delay and \( m \) is the embedding dimension. They can be calculated using the mutual information and False Nearest Neighbors (FNN) approaches, respectively [9][10].
B. Cross Recurrence Plot

Eckmann proposed the recurrence plot (RP) to study the m-dimensional phase space trajectory from a two-dimensional representation of a time series in 1987 [11], where recurrence of a state at two different time instances i and j is characterized by a two-dimensional squared matrix $R$ with dots:

$$R_{ij} = \Theta(\varepsilon - \| x_i - x_j \|), \quad i, j = 1, 2, \ldots, N$$

(2)

where $N$ represents the number of considered states, $\varepsilon$ is a threshold distance, $\| \cdot \|$ represents a norm operation, and $\Theta(\cdot)$ is the Heaviside function. Equation (2) defines states $x_i$ that is within an m-dimensional neighborhood of size $\varepsilon$ centered at $x_i$ is recurrent, and a dots $(i,j)$ is drawn in the RP [12]. It reveals the time-correlated information of system and represents system’s dynamical properties by line structure and recurrence point density.

The cross recurrence plot is a bivariate extension of the RP, which was proposed to investigate dependencies between two different systems. Those dynamical systems are represented by the trajectories $x_i$ and $y_j$ in the same d-dimensional phase space [13]. Then the cross recurrence matrix can be formed by calculating the pairwise mutual distances between the phase vectors of the two systems:

$$CR_{ij} = \Theta(\varepsilon - \| x_i - y_j \|)$$

(3)

The line structures based on CRP can reveal nonlinear interrelation between two systems.

C. Cross Recurrence Quantification Analysis

Both RP and CRP are visual approaches, the interpretation of them requires some professional knowledge. To avoid such subjectivity, recurrence quantification analysis (or CRQA) was developed to quantify RP structures [14]. Five recurrence variables, recurrence rate (RR), percent determinism (DET), maximal line length in the diagonal direction ($D_{\text{max}}$), average diagonal line length ($\langle D \rangle$), Shannon entropy of the frequency distribution of the diagonal line lengths (ENTR), calculated in CRQA procedure were defined based on diagonal line structure characteristics, and these variables are also used in the present study.

Mathematically, these five parameters can be described as follows:

$$RR = \frac{1}{N^2} \sum_{i=1}^{N} R_{ii}$$

(4)

which represents the density of recurrence points in the RP.

$$DET = \frac{\sum_{i=1}^{N} \sum_{j=1}^{i} P(i) R_{ij}}{\sum_{i=1}^{N} \sum_{j=1}^{N} P(i) R_{ij}}$$

(5)

where $P(i)$ is the histogram of the lengths of the diagonal structures in the RP, and $i_{\text{min}}$ is the lower bounds on the definition of lines. Typically, $i_{\text{min}}$ is chosen 2 for mechanical system. DET is defined as the fraction of recurrence points that form diagonal lines, which represents the similarity of two systems.

$$D_{\text{max}} = \arg \max_i P(i)$$

(6)

which is the length of the single longest diagonal within the entire RP.

$$\langle D \rangle = \frac{\sum_{i=1}^{N} \sum_{j=i}^{N} i P(i) R_{ij}}{\sum_{i=1}^{N} \sum_{j=1}^{N} P(i) R_{ij}}$$

(7)

which is the average time where two segments of the trajectory are close to each other.

$$\text{ENTR} = - \sum_{i=1}^{N} P(i) \ln P(i)$$

(8)

which reflects the complexity of the deterministic structure in the system.

In the case of CRP, RR quantifies the probability that two different systems occupy the same region in phase space. High values of RR represent high probabilities of it. DET reflects similarity the two systems are evolving. High values of DET and $\langle D \rangle$ represent a long time span of the occurrence of similar dynamics in both processes [15][16].

III. EXPERIMENTAL STUDY

A. Experimental Setup and Data Acquisition

To verify the applicability of CRQA in defect severity evaluation of wheel gears, vibration signals collected from a QianPeng gear test system shown in Fig.1 are analyzed. In the experiment, vibration signals are collected by four accelerometers with a sampling frequency of 8192Hz. Fig.2 illustrates vibration signals of four gears suffered from different degree of damage, where the crack size of the gear increased gradually.

B. Recurrence Parameters Selection

In the present study, CRPs are computed between the undamaged and different levels of damaged response data. The selection of time delay $\tau$, embedding dimension $m$, and threshold $\varepsilon$ determine the quality of reconstructed phase space and CRP directly. However, getting unified parameters from a broad variety of experimental data is difficult. Considering CRP represents interrelation between two systems, for the purpose of diagnosing the same type of fault, parameter selection can be simplified in the case of only comparing each CRP structure approximately.
Several tests are taken as follows: at first step, setting $m = 10, \varepsilon = 2.5$. Fig. 3 shows CRPs formed with different values of $\tau$. Secondly, setting $\tau = 5, \varepsilon = 1$, Fig. 4 shows CRPs formed with different values of $m$.

It is clear that, time delay and embedding dimension only affect the significant degree of CRPs structures, while do less in changing it. Therefore, the two parameters calculated within small data segment can be applied to the whole system, and the time lag $\tau$ in the present study is chosen as 1 because the two correlated signals often have different characteristic delay. Typically, in a recurrence based synchronization analysis, the recurrence point density is suggested to be the same in the individual RPs, therefore, $\varepsilon$ is chosen with RR being equal to 1% [17].

C. Multivariate Recurrence Analysis

Since vibration signals from four accelerometers are obtained at the same time, and sensors in different locations
contain different information, multivariate recurrence analysis are applied in present study.

J.M. Nichols at [16] has proposed a method to extend recurrence analysis to multivariate observations. Assume \( N \) time series \( x_i(t), n = 1, \ldots, N \) are collected from \( N \) sensors, the attractors of system can be considered as one time series with form: \( x(t) = [x_1(t), x_2(t), \ldots, x_N(t)] \), then the reconstructive phase space can be obtained based on (1). And CRPs for multivariate observations can be built according to (3).

**D. Data Analysis**

The CRQA parameters are calculated to evaluated the severity of gear fault with multivariate observations firstly. Table 1 and Fig.5 illustrates RR, DET, \( \langle D \rangle \), \( D_{\text{max}} \) and ENTR. It shows that with increase of crack size, DET, \( \langle D \rangle \) and ENTR show an decreasing trend. This can be explained that with increased crack size in the gear, dynamics of system get more and more different from the health one, and parameters revealed nonlinear interrelation between two systems will get smaller and smaller. Therefore, these three parameters DET, \( \langle D \rangle \) and ENTR can be chosen as indicators to assess the damage severity.

In fact, for each single sensor, parameters do have the same trend. Parameters calculated in 10 revolutions from four sensors are given in Fig. 6. It is clear that, ENTR and \( \langle D \rangle \) have the familiar trend since ENTR is calculated based on distribution of diagonal lines while \( \langle D \rangle \) is an average present of diagonal lines.

**TABLE I.** CRQA PARAMETERS OF VIBRATION SIGNALS FROM CRACK FAULT OF FOUR TEST GEARS (SPEED: 30Hz)

<table>
<thead>
<tr>
<th>Crack fault level</th>
<th>RR mean</th>
<th>DET mean</th>
<th>( \langle D \rangle ) mean</th>
<th>( D_{\text{max}} ) mean</th>
<th>ENTR mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (health)</td>
<td>0.0100</td>
<td>0.8052</td>
<td>3.3999</td>
<td>12.9099</td>
<td>1.5829</td>
</tr>
<tr>
<td>1</td>
<td>0.0100</td>
<td>0.7747</td>
<td>3.2059</td>
<td>9.7587</td>
<td>1.4800</td>
</tr>
<tr>
<td>2</td>
<td>0.0100</td>
<td>0.7636</td>
<td>3.1311</td>
<td>11.1192</td>
<td>1.4369</td>
</tr>
<tr>
<td>3</td>
<td>0.0100</td>
<td>0.7107</td>
<td>2.8998</td>
<td>8.8983</td>
<td>1.2853</td>
</tr>
<tr>
<td>4</td>
<td>0.0100</td>
<td>0.6814</td>
<td>2.8004</td>
<td>8.1890</td>
<td>1.2102</td>
</tr>
</tbody>
</table>

Although four sensors are located in different places, CRQA parameters perform well in each location based on signal acquired from single sensor while only having subtle differences in the degree of parameter index.
Figure 6. CRQA results from different sensors in the gear test system

IV. CONCLUSION

Through analysis of the experimental data, it can be seen that the change of DET, \( D \) and ENTR in cross recurrence quantification analysis with multivariate observations can be used to diagnose the gear damage severity. Cross recurrence quantification analysis from different sensors distributed in multiple positions has the familiar result of quantitative indicators’ trend which verified CRQA method’s ability in weak signal detection in the field of fault diagnosis. Inspired by the application in the gear damage severity diagnosis, the method can be extended to the prediction of remaining life in the further study.

ACKNOWLEDGMENT

This work has been supported in part by the National Natural Science Foundation of China (51575102).

REFERENCES


