Journal of Sound and Vibration 424 (2018) 192–207

Contents lists available at ScienceDirect

Journal of Sound and Vibration

journal homepage: www.elsevier.com/locate/jsvi

Modified complementary ensemble empirical mode decomposition and intrinsic mode functions evaluation index for high-speed train gearbox fault diagnosis

Dongyue Chen*, Jianhui Lin, Yanping Li

State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China

ARTICLE INFO

Article history:
Received 2 June 2017
Received in revised form 8 March 2018
Accepted 15 March 2018

Keywords:
Complementary ensemble empirical mode decomposition
Computation cost
Intrinsic mode function
Fault diagnosis
High-speed train gearbox

ABSTRACT

Complementary ensemble empirical mode decomposition (CEEMD) has been developed for the mode-mixing problem in Empirical Mode Decomposition (EMD) method. Compared to the ensemble empirical mode decomposition (EEMD), the CEEMD method reduces residue noise in the signal reconstruction. Both CEEMD and EEMD need enough ensemble number to reduce the residue noise, and hence it would be too much computation cost. Moreover, the selection of intrinsic mode functions (IMFs) for further analysis usually depends on experience. A modified CEEMD method and IMFs evaluation index are proposed with the aim of reducing the computational cost and select IMFs automatically. A simulated signal and in-service high-speed train gearbox vibration signals are employed to validate the proposed method in this paper. The results demonstrate that the modified CEEMD can decompose the signal efficiently with less computation cost, and the IMFs evaluation index can select the meaningful IMFs automatically.

1. Introduction

The gearbox is an essential part of the high-speed train transmission system. The high-speed gearbox usually works in severe conditions, such as sudden load, dynamic unbalances state of the shaft, and overload conditions. Consequently, it is hard for gear to avoid crack, pitting, wearing, spalling and tooth surface damage. Gear faults are the main source of vibration and noise in high-speed train transmission system, and it is a threat to train operation safety. The vibration-based mechanical system fault monitoring has been developed and widely used for past decades [1]. The vibration signal of the gearbox is non-stationary and modulation, which are distinctive features for fault detection.

The EMD method, presented by Huang et al. [6], is a self-adaptive method for non-linear and non-stationary signals. In recent years, the EMD-based method has been successfully employed in gear fault diagnosis. H. Liu used EMD and multi-fractal detrended cross-correlation analysis for gearbox fault diagnosis [2]. Y. Li proposed an improved EMD method using a new envelope interpolation method instead of the cubic spline interpolation and applied to gear fault diagnosis [3]. L. Lu used EMD method to decompose a vibration signal into IMFs and then used the modified genetic algorithm for rotating machinery fault diagnosis [4]. The major drawback of EMD is the mode mixing phenomenon, caused by signal intermittency, where one IMF could contain wide disparate scale [7]. A noise-assisted data analysis (NADA) method named Ensemble EMD

* Corresponding author.
E-mail addresses: dyue_chen@163.com (D. Chen), lin13008104673@126.com (J. Lin), lyp_swj@sina.com (Y. Li).

https://doi.org/10.1016/j.jsv.2018.03.018

(EEMD) was proposed by Wu and Huang to alleviate this problem [7]. The EEMD method resolves the mode-mixing issue with the help of added white noise, and then the added noise is reduced to a small level via ensemble procedure. However, the added noise would contaminate the result of EEMD when the number of ensemble is low. The only way to reduce the added noise is to increase the ensemble number, which will lead to a heavy computation burden [18]. Another noise enhanced method CEEMD was proposed by Yeh and Huang [8], and this technique could indeed ameliorate the residue noise more effectively than EEMD under the same number of ensemble trails. However, both EEMD and CEEMD need a large ensemble number to reduce the residue noise in IMFs, and the computation of EEMD and CEEMD remains at a high level, which is a significant limitation for on-line fault diagnosis. Another barrier for practical application is the analysis of IMFs. In the recent study, fault features are derived from IMFs for rotary mechanical fault classification algorithm [4,5]. However, not all IMFs contain valid information, the selection of IMFs is a vital work. Even though the EMD-based algorithm is an adaptive method, the analysis of IMF is usually relying on the experience of users. Ricci et al. proposed a merit index, and the linear combination of two indexes was used, which allows the automatic selection of IMFs for gear fault diagnosis [9]. C. Yi presented an IMF confidence index algorithm for EEMD, which is the combination of four indexes and it has a gratifying performance on high-speed train axle bearing fault detection [20]. A default threshold was used to detect the maximum value of the signal in these two methods, and then calculated the periodicity index to measure the periodicity degree of IMFs. It is hard to set the threshold value, because of the fluctuating amplitude of the signal with fault development.

In this paper, a modified CEEMD algorithm and IMFs evaluation index are proposed to depress the computation cost and select the IMFs automatically. The noise, transient components, and abnormal components, the reason of mode-mixing phenomenon, could be extracted from the signal as the first few IMFs by using EEMD or CEEMD algorithm. After these components extracted from the signal, the rest signal is more stationary, and its extrema are more uniform than the original signal. Therefore, there is no need to do the whole EEMD or CEEMD procedure. The proposed modified CEEMD method using CEEMD extract the intermittence components from the original signal, and the EMD algorithm decomposes the residual signal. Kurtosis and autocorrelative function are used to estimate the result of CEEMD whether it is noise, transient components of signal or abnormal components. According to the study of Huang, although CEEMD has lower reconstruction error than EEMD, these two methods have similar performance in the decomposition of single IMF component when the number of added noise is the same [8]. Therefore, the kurtosis and autocorrelative function estimation can also be used to speed up EEMD with similar threshold. The IMFs evaluation index is introduced to select useful IMF automatically using Kurtosis, autocorrelative function, and proportion of IMFs energy. The IMFs evaluation index is based on the characteristic of fault signal, and hence it also applies to other EMD-based methods. This is also because the IMFs are related to the original signal, and the characteristics of each IMF obtained from EMD-based method will not change.

The remainder of this paper is organized as follows. Section 2 introduces the basic theory of CEEMD algorithm. The modified CEEMD algorithm and IMFs evaluation index are presented in detail in Section 3. Section 4 gives a simulation study to verify the effectiveness of proposed method. Section 5 illustrates the application of in-service high-speed train gearbox vibration signals. The conclusion and discussion are summarized in Section 6.

2. Basic theory of CEEMD

2.1. EMD

EMD is a data-based self-adaptive signal decomposition technique for non-linear and non-stationary signal, which decomposes a signal into a series of IMFs and a residue signal. IMFs represent the oscillation mode embedded in the signal indicating the local characteristics. IMFs are defined as following properties:

- Throughout the whole set of one single IMF, the number of extrema and the number of zeros crossing must be equal or differ by one at most.
- At any point, the mean value of the upper envelope and the lower envelope is zero.

The EMD algorithm can be defined in the following steps:

(1) Identify all the local extrema and connect all these local maxima or minima with a cubic spline as an upper envelope or lower envelope. Then obtain the mean of upper envelope and lower envelope, which is noted as $m_1(t)$.
(2) The extraction of the mean function $m_1(t)$ from the signal and obtain the first component $h_1(t)$.
(3) Repeat the step 1 and step 2 until the $h_1(t)$ respects the IMF definition, and the final $h_1(t)$ designated as $c_1(t)$, the first IMF component, which represents the highest frequency component of the signal. Then it is subtracted from the original signal:

$$r_1(t) = x(t) - c_1(t) \tag{1}$$