Measurement 59 (2015) 145–166

Contents lists available at ScienceDirect

Measurement

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

Bearing faults diagnostics based on hybrid LS-SVM and EMD method

Xiaofeng Liu *, Lin Bo, Honglin Luo

The State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400044, PR China

ARTICLE INFO

Article history:
Received 26 March 2014
Received in revised form 3 September 2014
Accepted 16 September 2014
Available online 28 September 2014

Keywords:
LS-SVM
EMD
Bearing fault diagnosis
Noise removal
Data extrapolation
Smoothed local mean curve

ABSTRACT

In this paper, a novel method that integrates the LS-SVM and Empirical Mode Decomposition (EMD) is proposed to improve the performance of conventional EMD. The analyzed signal is preprocessed with the weighted Least Squares Support Vector Machines (LS-SVM) to suppress the interference of high-frequency intermittent components and other non-Gaussian noises. The denoised signal is extended with LS-SVM rolling forecast modeling. Next, the linear function is used to construct upper and lower envelopes of the extrapolated data in order to determine the temporary mean envelope curve which is then smoothed with the adaptive mapped LS-SVM to obtain the local mean curve. Signal decomposition is self-adaptively performed to achieve IMFs through removal of the smoothed local mean curve. The representative IMF containing fault information is selected for demodulation analysis to identify the fault characteristics. The effectiveness of the proposed method is verified by means of simulations and applications to bearing fault diagnosis.

1. Introduction

For rolling bearing fault detection, it is expected that a desired time–frequency analysis method should have good resolution in both time domain and frequency domain. Due to the limitation of Heisenberg–Gabor inequality, the time–frequency transform based on Fourier transform cannot achieve fine resolutions in both time domain and frequency domain simultaneously. Empirical Mode Decomposition (EMD) is an effective signal processing technique that can self-adaptively decompose a non-stationary signal into a sum of Intrinsic Mode Functions (IMFs) according to the inherent characteristics of the signal [1]. EMD has been used in mechanical system diagnosis applications such as bearing damage detection and gearbox fault diagnosis [2–4]. However, in the practical applications, aside from the lack of a perfect mathematical foundation, the conventional EMD methodology also faces the problems of mode mixing and end effects in algorithm implementation and application. In order to address the problems, a number of approaches have been undertaken, such as the use of the alternate extrema for the envelopes [5], the bandwidth criterion for IMF [6], the masking signal [7,8], the heuristic search optimization approach [9], doubly-iterative sifting [10], ensemble EMD [11], centroid-based sifting [12].

Support Vector Machines (SVM) is a powerful supervised machine learning tool introduced in framework of statistical learning theory. It is used in a number of applications for both noise removing and non-linear regression. It has the ability to generalize well on unknown data without the domain knowledge. Ref. [13] applied SVRM (support vector regression machines) to the analyzed time series forecasting successfully to restrain the end effects of EMD. However, the method in Ref. [13]

* Corresponding author. Tel.: +86 02365106973; fax: +86 02365106641.
E-mail address: liuxfeng0080@126.com (X. Liu).

http://dx.doi.org/10.1016/j.measurement.2014.09.037
0263-2241/

maybe lose its effectiveness for the analysis of signal interfered by high-frequency intermittent and non-Gaussian noises which make it troublesome to select right training data sets, leading to the inaccurate prediction result of SVM. Therefore, the method in Ref. [13] lacks of the robustness to noises. The same problem exists in Ref. [14]. Other combinations of EMD and SVM [15–17] are just about the post-processing the EMD results with SVM aiming to the automatic fault classification of identification. They are not about the improvement of EMD performance.

The above discussion shows that although some approaches have shown promising results in improving EMD performance, none has been widely accepted and there is still room for improvement to reach the ultimate goal of a completely reliable EMD method. In an effort to achieve this goal we propose an improved approach in this paper for bearing diagnosis combining EMD and LS-SVM techniques. This paper is organized as follows: the next section introduces EMD and LS-SVM. Afterwards, the proposed method for the fault diagnosis of rolling element bearing is developed. Next, we describe the simulation and case studies performed to validate the method, followed by our conclusions and some ideas for future work.

2. Basic principle

2.1. EMD method principle

The most appealing nature of EMD is its dependency on the data-driven mechanism which does not require a priori known basis unlike Wavelet and Fourier transform. IMFs obtained by EMD are band-limited, which can represent the features of signal and reserve the local information. The EMD method identifies all the local maxima and minima for a given input signal $x(t)$ which are connected by spline curves to form the upper and the lower envelopes, $e_{up}(t)$ and $e_{low}(t)$, respectively. The mean of the two envelopes is calculated as $m(t) = [e_{up}(t) + e_{low}(t)]/2$ and is subtracted from the signal using $q(t) = x(t) - m(t)$. An IMF $\mathrm{IMF}_i(t)$ is obtained if $q(t)$ satisfies the two conditions of IMF, these are, the number of extrema and number of zero crossings is either equal or differs at most by one, and the envelopes defined by the local maxima and minima are symmetric with respect to zero mean. This procedure is called as the sifting process. Then $x(t)$ is replaced with the residual $r(t) = x(t) - q(t)$. If $q(t)$ is not an IMF, $x(t)$ is replaced with $q(t)$. The above process is repeated until the residual satisfies the stopping criterion. At the end of this process the signal $x(t)$ would result in N IMFs and a residue signal as in Eq. (1).

$$x(t) = \sum_{n=1}^{N} \mathrm{IMF}_n(t) + r_N(t) \tag{1}$$

where n and N represents the order and the total number of IMFs respectively, and r_N denotes the final residue. The signal $x(t)$ is decomposed such that the lower-order components represent fast oscillation modes and higher-order components represent slow oscillation modes. A detailed explanation of the method is provided in Ref. [18].

While executing the EMD, there still exist some problems: (1) It is sensitive to the abnormal or false extrema points caused by high-frequency intermittent noises, so its decomposition lacks of the robustness to noises; (2) The cubic spline fitting easily results in overshoot or undershoot which lead to generation of a big error, thus the original essential structure of IMF is easy to be destroyed in EMD decomposition process; (3) The sifting rule would render it impossible to decompose intrinsically different modes whose frequencies fall in an octave.

2.2. LS-SVM principle

SVM is a novel machine learning technique based on a statistical learning theory that aims at finding optimal hyperplanes among different classes of input data or training data in high dimensional feature space, and new test data can be classified using the separating hyperplanes. Least Squares Support Vector Machines (LS-SVM) is an extension of SVM. LS-SVM changes the traditional inequality constraints to equality constraints and regards the sum of squared errors loss function as experience loss of training set. It transforms solving quadratic programming problem into solving linear equations problem. Compared with SVM, LS-SVM has a faster solution speed and higher solution accuracy. In the following sections, there will be brief description to LS-SVM.

The training dataset is assumed to be $\{x_i, y_i\}$ ($i = 1, 2, \ldots, l$) in which x_i is the input vector and y_i is its corresponding target vector. Let $z_i = \Phi(x_i)$ denote the corresponding feature space vector with a mapping function Φ from the input space to high-dimensional feature space. The hyperplane can then be defined as

$$\mathbf{w} \cdot z + b = 0 \tag{2}$$

where $\mathbf{w}$ is the weight vector defining the orientation of hyperplane and b is the bias parameter. For efficient computation purposes, in LS-SVM, one aims at minimization of the following object function. Data samples are said to be linearly separable if there exists ($\mathbf{w}$, b), such that

$$\text{Minimize}: \; J(\mathbf{w}, e) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}C\sum_{i=1}^{N} e_i^2 \tag{3}$$

Subject to: $y_i = \mathbf{w}^T z_i + b + e_i \quad i = 1, \ldots, N.$

The first term stands for the minimization of the Vapnik Chervonenkis (VC) dimension, while the second one minimizes the training errors (e_i). C is the tradeoff parameter between the terms. Define the following equation:

$$L(\mathbf{w}, b, e, \alpha) = J(\mathbf{w}, e) - \sum_{i=1}^{N} \alpha_i \{\mathbf{w}^T z_i + b + e_i - y_i\} \tag{4}$$

With Lagrange multipliers $\alpha_i \in R$. The conditions for optimality are given by: $\partial L/\partial w = 0$, $\partial L/\partial b = 0$, $\partial L/\partial e_i = 0$, and $\partial L/\partial \alpha_i = 0$. After elimination of $\mathbf{w}$ and e, one can obtain

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + I/C \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{5}$$

where $y = [y_1, y_2, \ldots, y_N]^T$; $1_v = [1, 1, \ldots, 1]^T$; $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_N]^T$; and $\Omega_{i,j} = K(x_i, x_j)$, for $i, j = 1, 2, \ldots, N$. Here