Mechanical Systems and Signal Processing 54-55 (2015) 259–276

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis

Yi Wang [a], Guanghua Xu [a,b,*,1], Lin Liang [a], Kuosheng Jiang [a]

[a] School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, PR China
[b] State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, PR China

ARTICLE INFO

Article history:
Received 28 November 2013
Received in revised form
1 May 2014
Accepted 2 September 2014
Available online 14 October 2014

Keywords:
Rolling element bearing
Wavelet packet transform
Manifold learning
Permutation entropy
Fault diagnosis

ABSTRACT

The kurtogram-based methods have been proved powerful and practical to detect and characterize transient components in a signal. The basic idea of the kurtogram-based methods is to use the kurtosis as a measure to discover the presence of transient impulse components and to indicate the frequency band where these occur. However, the performance of the kurtogram-based methods is poor due to the low signal-to-noise ratio. As the weak transient signal with a wide spread frequency band can be easily masked by noise. Besides, selecting signal just in one frequency band will leave out some transient features. Aiming at these shortcomings, different frequency bands signal fusion is adopted in this paper. Considering that manifold learning aims at discovering the nonlinear intrinsic structure which embedded in high dimensional data, this paper proposes a waveform feature manifold (WFM) method to extract the weak signature from waveform feature space which obtained by binary wavelet packet transform. Minimum permutation entropy is used to select the optimal parameter in a manifold learning algorithm. A simulated bearing fault signal and two real bearing fault signals are used to validate the improved performance of the proposed method through the comparison with the kurtogram-based methods. The results show that the proposed method outperforms the kurtogram-based methods and is effective in weak signature extraction.

1. Introduction

Rolling element bearings are widely used in rotating machinery, faults occurring in bearings may lead to fatal breakdowns in rotating machinery and such failure can be catastrophic, resulting in costly downtime. Therefore, it is significant to accurately diagnose the existence of faults at an early stage. Vibration signals collected from bearings contain rich information on machine health conditions [1]. Hence, it is possible to obtain vital characteristic information from vibration signals through the use of advanced signal processing techniques due to their intrinsic advantage of revealing bearing failure [2].

The theoretical background of bearing failure mechanism has been covered quite comprehensively during the past decades. The transients or transient signals are generally referred to as the signal components composed of exponentially

* Corresponding author at: School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, PR China. Tel.: +86 29 8266 4257.
E-mail address: xugh@mail.xjtu.edu.cn (G. Xu).
[1] Tel./fax: +86 29 8339 5054.

http://dx.doi.org/10.1016/j.ymssp.2014.09.002

decaying ringing that lasts a short period of time and spans within a wide frequency range [3,4]. The signature of a defected bearing consists of transients or transient signals that occur approximate periodically at a characteristic frequency and the duration between two adjacent impulses has a slight randomness [5]. Fault diagnosis is achieved by detecting the weak transient signals and alerting maintenance personnel before the fault develops into a catastrophic failure.

As for the vibration signal of rolling element bearing, signal modulation effect and noise are two major barriers in incipient defect detection for bearing fault diagnosis [6]. In order to overcome the barrier of modulation, Randall [7] proposed a Fast Fourier Transform (FFT)-based Hilbert transform. Though FFT-based Hilbert transform offers an effective technique for signal demodulation, this method failed to address how to enhance the weak transients from a noisy signal, especially in early stage defects detection for bearing failure diagnosis.

To enhance the signal-to-noise ratio of the original signal, a signal representation method is first needed for the identification of transient feature and background noise. For transient feature representation and identification, the time–frequency representation (TFR) is the most frequently used method, through which the transient feature can be represented in time–frequency space [8]. The wavelet transform (WT) is actually a kind of TFR method, as the scale of the WT corresponds to the frequency [9]. The wide applications of WT are based on its distinct features such as the freedom of mother wavelet selection and the avoidance of cross items in traditional TFRs (e.g., Wigner–Ville distribution, Pseudo-Wigner–Ville). Though WT has some distinct features compared with the traditional TFR methods, WT cannot effectively split the high frequency band containing rich fault modulation information [10]. The way of overcoming this difficulty is to extend WT to wavelet packet transform (WPT). As one of the famous TFR techniques, WPT has the well-known properties of being orthogonal, complete, and local. Based on the above-mentioned distinct features, WPT can be used as an excellent signal representation method to enhance the weak transients from a noisy signal.

In recent years, several WPT-based optimal frequency-band selection methods for transient signal extraction have been proposed. Aiming at the shortcomings of the original kurtogram method [11], Lei et al. [1] proposed an improved kurtogram method to overcome the shortcomings by replacing the filter of the original kurtogram with WPT filter. Wang et al. [12] proposed an enhanced kurtogram method, the major innovation of which is that kurtosis values are calculated based on the power spectrum of the envelope of the signals extracted from the wavelet packet nodes at different WPT levels. Some other optimal filter band selection based transient signal extracting methods have also been proposed, such as protrugram [13], sparsogram [14] and adaptive spectral kurtosis [15]. The basic idea of the optimal filter band selection based methods is to exploit the possibility of using a metrics (such as kurtosis, sparsity, etc.) as a measure to discover the presence of transient impulse components and to indicate the frequency band these occur. Considering the fact that transient signals span within a wide frequency range, only selecting the reconstructed signal in one frequency band may leave out some important information. Besides, at the early stage of bearing defect development, the weak transient signal of a defective bearing with a wide spread frequency band can be easily masked by noise in each frequency band [16]. By the optimal frequency band selection methods, only the noises outside the selected frequency band are removed from the original signal, while those inside the selected frequency band cannot be wiped off effectively. As a result, the performances of those methods are poor in the presence of low signal-to-noise ratio.

Recently, manifold learning has emerged in nonlinear feature extraction, due to its capability in effectively identifying low-dimensional nonlinear intrinsic structure embedded in high-dimensional data. The technique can be realized through several algorithms including locally linear embedding (LLE) [17], isometric feature mapping (IsoMap) [18], local tangent space alignment (LTSA) [19], and Laplacian eigenmaps (LE) [20], etc. Many studies have been conducted by applying manifold learning to the machinery fault diagnosis [21–32]. In recent years, the application of manifold learning in mechanical fault diagnosis can be divided into two categories. On one hand, manifold learning is used to extract non-linear features for fault classification [21–28], on the other hand, this technique is applied to extract the transient signals from the noise contaminated signal [29–32]. The above mentioned studies have demonstrated that manifold learning is effective to extract the intrinsic manifold features related to non-linear dynamics of the mechanical system.

In this paper, the reconstructed signal in each time–frequency subspace is considered as waveform feature (WF). The binary WPT separates the signal waveform information into different time–frequency subspaces layer by layer, forming a kind of high-dimensional WF space. In other words, WF space is a fusion of different frequency band signals. Consider the merits of WF space in signal representation and the manifold learning in nonlinear intrinsic structure extraction, a novel waveform feature manifold (WFM) technique is proposed. Aiming at the shortcomings of the optimal filter band selection methods, the WFM technique intends to mine the nonlinear WF structure by manifold learning to obtain transient signal feature for bearing fault diagnosis. Minimal permutation entropy is used as a criterion to optimize the selection of the number of neighbourhoods in the manifold learning algorithm. In order to overcome the barrier of modulation, the envelope analysis is used for further demodulation of the extracted transient signal. The frequency signatures of the envelope spectrum are used to diagnose the type of bearing fault by identifying its characteristic frequency. As typical optimal filter band selection methods, the original kurtogram, the improved kurtogram and the enhanced kurtogram are used in this paper for comparison study, and to demonstrate the improved performance of the proposed WFM method.

The outline of this paper is as follows. Section 2 first reviews the theory of WPT and presents the algorithm to construct the waveform feature space from the binary WPT tree, then the main steps of manifold learning is introduced and a minimum permutation criterion is demonstrated, and finally, the algorithm of the proposed method is presented. And the