Measurement 127 (2018) 533–545

Contents lists available at ScienceDirect

Measurement

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

Kurtogram manifold learning and its application to rolling bearing weak signal detection

Yi Wang[a,b,*], Peter W. Tse[c], Baoping Tang[a,b], Yi Qin[a,b], Lei Deng[a,b], Tao Huang[a,b]

[a] The State Key Laboratory of Mechanical Transmission, Chongqing University, Chongqing 400044, PR China
[b] College of Mechanical Engineering, Chongqing University, Chongqing 400044, PR China
[c] Center for System Informatics and Quality Engineering, Department of Systems Engineering & Engineering Management, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, China

ARTICLE INFO

Keywords:
Kurtogram
Manifold learning
Signal detection
Dimensionality reduction
In-band noise

ABSTRACT

The kurtogram method has been proposed for weak signal enhancement and been validated very powerful. The essence of the kurtogram method is de-noising by optimal band-pass filtering, however, the in-band noise are left unprocessed. As a result, the fault induced transient impulses, which spread within a wide frequency band, would be still contaminated by noise. Aiming at the flaws encountered by the kurtogram method, this paper proposes a novel kurtogram manifold learning method for signal de-noising in whole time-frequency space. In this method, the sub-signals split by kurtogram are fused to build a high dimensional transient impulse feature space, manifold learning is conducted to mine the transient impulse feature space. On this basis, the in-band noise can be filtered out, thereby the transient impulse features will be uncovered. The experimental validation results exhibit the proposed method outperforms kurtogram method and is effective for rolling bearing weak signal detection.

1. Introduction

Rolling element bearings are widely used in rotary machineries and provide the support of rotary components [1–4]. Since bearing failure is one of the foremost causes of breakdowns in rotary machinery, therefore, detection of rolling bearing fault especially at its early stage, is essential to prevent unexpected accidents [5].

In most cases, rotary machines are very difficult for directly inspecting, due to the huge machine size, severe environment or restrictions of disassembly [6–9]. Since the vibration signals which collected from the surface of the machine contain plenty of fault characteristics. Therefore, vibration analysis can be viewed as an alternative approach to learn the condition of rotary equipment [10–14]. When a local defect formed in the rolling bearing, the rolling elements will strike the defect surface such as inner race or outer race, periodic or quasi-periodic vibration shocks are generated to stimulate resonant frequencies of the structure between sensors and bearings [15]. The response signal collected by the vibrational sensors will be composed of a chain of transient impulses which spread within a wide frequency band [16]. However, the defect at its early stage, the slight stimulated signal collected by sensors will be seriously contaminated due to the heavy noise [17]. In this circumstance, the weak transient features excited by bearing fault should be effectively extracted prior to bearing condition assessment, such as fault detection and diagnosis, condition monitoring, remaining useful life prediction, etc.

To enhance the original vibration signal, a band-pass filter is usually used to obtain transient impulse signal from the resonance frequency band. In recent years, Antoni [18] analysed the spectral kurtosis thoroughly to conduct the band-pass filtering process adaptively. Afterwards, a kurtogram method based on short-time Fourier transform (STFT) and spectral kurtosis is proposed [19]. To reduce the computation burden of the STFT-based kurtogram method for online band-pass filtering, a fast kurtogram method based on 1/3-binary tree was proposed [20]. Afterwards, the optimal resonance frequency band selection has attracted a great deal of attention, and many approaches have been proposed to improve the spectral kurtosis and kurtogram [21–24]. In order to tackle the shortcomings of the original kurtogram method, Sawalhi et al. [21] developed a method to hone impulses and increase the kurtosis values by the combination of minimum entropy deconvolution and spectral kurtosis. Considering the fast kurtogram method can only obtain an approximate estimation, Zhang [22] integrated the fast kurtogram method with genetic algorithm for optimal central frequency and bandwidth selection. Aiming at the negligence of matching the filtered signal with actual transient features in directly

* Corresponding author at: College of Mechanical Engineering, Chongqing University, Chongqing 400044, PR China.
E-mail address: wycqdx@cqu.edu.cn (Y. Wang).

https://doi.org/10.1016/j.measurement.2018.06.026
Received 9 April 2018; Received in revised form 12 June 2018; Accepted 14 June 2018
Available online 18 June 2018
0263-2241/

band-pass filtering algorithm, improved versions of kurtogram [23,24] were put forwarded and in which wavelet packet transform (WPT) was used to obtain a better filtering result.

In a word, the basic idea of the above mentioned optimal bandpass filtering methods is using metrics to find the occurrence of transient impulse components, and to confirm the central frequency and the band width of the bandpass filter. Due to the transient impulses spread within a wide frequency band, the energy of spectrum components will be dispersed with low amplitude. Therefore, when bearing fault just formed, the fault induced transient impulses with small amplitude will be heavily contaminated by background noise. As a result, by optimal band-pass filtering, only the noise outside the selected frequency-band are wiped off, while the in-band noise are left unprocessed. Besides, the transient features spread within a wide frequency band cannot be effectively uncovered by bandpass filtering.

To address this issue, a WFM method [25] was proposed, the WPT was combined with manifold learning to suppress background noise in entire time-frequency space for signal enhancement. In WFM method, the number of wavelet function is manually selected by experience before conducting the algorithm. By multi-level WPT decomposition, the weak transient impulses are represented in a high-dimensional waveform feature space. Then, the non-linear dimension reduction method, i.e. manifold learning is applied to the waveform feature space for intrinsic principle waveform feature mining. And finally, the noise is suppressed in entire time-frequency space, thereby the weak transient impulses are uncovered. However, the transient impulse features, which embedded in low-dimensional space, is easy to be affected by the dominance of the weak transient impulse features in the high-dimensional manifold. If the mother wavelet matches very well with the actual transient impulse features hidden in a signal, then the transient characteristics will be uncovered and easy to be extracted by WFM from the high-dimensional waveform feature space. In practical application cases, the mother wavelet function which used to match the transient characteristics is hard to be appropriately selected because of the lack of prior knowledge (i.e. the damping coefficient and resonance frequency) about the rotary system. If the mother wavelet function is highly deviated from the actual attenuate waveform, the high-dimensional waveform feature space will with poor transient information significance. As a result, the extracted signal mined by WFMmethod may be still contaminated by noise or even highly deviate from the actual transient characteristics.

In the kurtogram algorithm, a series of adaptive band-pass filters based on STFT or FIR filters are adopted to split the entire frequency band into different sub-spaces. Therefore, though the shortcomings the adaptive band-pass filters have in signal representation, they are free from parameter selection and more applicable when compared with WPT based filtering. Besides, the experimental results in the literatures [19–24] shows the weak impulse features can be effectively detected by kurtogram based on optimal band-pass filtering. Therefore, by applying kurtogram method, the original signal will be split into several time-frequency spaces which carrying different transient features. The high dimensional transient impulse feature space, which depict the original signal in different time-frequency spaces, can be obtained through fusing the sub-signals. On this basis, this paper proposes a kurtogram manifold learning method (KML) to overcome the drawbacks encountered in kurtogram method [20], and WFM method [25]. Its main innovative idea is investigating the effectiveness of weak transient signal extraction, by using manifold learning to mine the high-dimensional transient impulse feature space which obtained from the kurtogram. The experimental results exhibit that the proposed KML method outperforms the kurtogram method and is validated effective in weak signature extraction, besides, exhibits some improved characteristics when compared with WFM method. It should be noted that this paper mainly concentrated on vibrational signal noise suppression for weak signal enhancement, while the further processing and application of the extracted signal for signal demodulation, condition monitoring and remaining useful life prediction is an open topic, which is beyond the scope of this paper.

The remainder of our paper is organized as follows. In Section 2, the theory of kurtogram is reviewed and the algorithm for transient impulse feature space construction by kurtogram is described, then the main procedures about manifold learning is described, afterwards, the detailed steps of the proposed method is demonstrated. The effectiveness of proposed KML method is experimentally validated by two experimental cases of bearing signals in Section 3. Finally, our conclusions are summarized in Section 4.

2. Kurtogram manifold learning for weak transient impulse signal detection

2.1. Theoretical backgrounds and characteristics of kurtogram

On the basis of spectral kurtosis (SK) [26], the kurtogram was first introduced by Antoni and Randall [19]. The kurtosis is used as a metric in SK to measure the similarity between a stochastic process and a Gaussian process to detect the existence of transient impulses. The kurtogram is formed by calculating the SK value of the filtered signal component in each frequency band. Therefore, SK represents the significance of the transient features in different frequency band. On this basis, SK can be used as a tool to indicate the existence of transient signals and their accurate frequency-band.

On the basis of Wold–Cramér decomposition of a non-stationary signal, vibrational response $y(t)$ of a system with response $h(t,s)$ excited by $x(t)$ can be represented as follows:

$$y(t)=\int_{-\infty}^{+\infty}e^{j2\pi ft}H(t,f)\mathrm{d}X(f) \tag{1}$$

where $\mathrm{d}X(f)$ is the orthonormal increment of frequency and $H(t,f)$ is the transfer function which varying with time, and can be obtained by Fourier transform of the response function $h(t,s)$ of the investigated system. $H(t,f)$ can also be described as the complex envelope or complex demodulate of response signal $y(t)$ at frequency f, i.e. $e^{j2\pi ft}H(t,f)\mathrm{d}X(f)$ is the signal at time instant t of an infinite band-pass filter centred on frequency f. For vibration signal of rotating machinery, $H(t,f)$ is rather stochastic than deterministic. In this circumstance, the time varying transfer function should be represented as $H(t,f,w)$, where w is the variable of random time. For a conditionally non-stationary process (CNS), under which H is time-stationary and independent operation of process x. Then the SK value of the CNS process is calculated based on fourth-order spectral cumulate [18].

$$C_{4y}(f)=S_{4y}(f)-2S_{2y}^2(f) \tag{2}$$

where $S_{2y}(f)$ represent the time-average of $S_{2y}(t,f)$ and $S_{2y}(t,f)$, and is the matrices of energy fluctuation of the complex envelope. And then, SK is obtained by calculating the 4th-order cumulate depicted in Eq. (2). It characterize the peakiness of the probability density function of the considered process at instantaneous frequency f. On this basis, SK is defined in Eq. (3)

$$K_y(f)=\frac{C_{4y}(f)}{S_{2y}^2(f)}=\frac{S_{4y}(f)}{S_{2y}^2(f)}-2,\ f\neq 0 \tag{3}$$

In the kurtogram algorithm, the estimation of SK at specified frequency f is conducted by calculated the STFT of the observed signal $y(t)$ and is demonstrated in Eq. (4) [23].

$$F_y(\tau,f)=\int_{-\infty}^{+\infty}y(t)w^*(t-\tau)e^{-j2\pi ft}\mathrm{d}t=\langle y(t),w(t-\tau)e^{j2\pi ft}\rangle \tag{4}$$

where $w^*(t-\tau)$ denotes the window function in the STFT algorithm and the superscript symbol $*$ represents the conjugation.

In conclusion, the general idea of the kurtogram method is using SK as a metrics to find the presence of transient impulses, and filtering the background noise with optimal selected central frequency and