Mechanical Systems and Signal Processing 110 (2018) 1–27

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Adaptive iterative generalized demodulation for nonstationary complex signal analysis: Principle and application in rotating machinery fault diagnosis

Zhipeng Feng *, Xiaowang Chen

School of Mechanical Engineering, University of Science and Technology Beijing, Beijing 100083, China

A R T I C L E I N F O

Article history:
Received 17 August 2017
Received in revised form 23 January 2018
Accepted 5 March 2018

Keywords:
Adaptive iterative generalized demodulation
Surrogate test
Time-frequency analysis
Rotating machinery
Nonstationary

A B S T R A C T

Effective identification of signal frequency contents and their time variability is a key to success in rotating machinery fault diagnosis under nonstationary conditions. Fine time-frequency resolution and free from both inner and outer interferences are necessary to achieve this purpose. Iterative generalized demodulation (IGD) can separate a nonstationary multi-component signal into constituent mono-components, and derive quality time-frequency representation based on Hilbert spectrum of each mono-component, thus offering an effective approach to nonstationary complex signal analysis. Nevertheless, it requires prior expertise knowledge about signal structure to ascertain true constituent components and construct proper demodulation phase functions. This leads to a major difficulty in real signal analysis tasks, where expertise knowledge is usually unavailable and signal time-frequency structure identification via visual observation is susceptible to noise interferences. To address this issue, the capability of surrogate test technique in recognizing true signal components is exploited, and thereby an adaptive iterative generalized demodulation (AIGD) is proposed. Proper demodulation phase functions corresponding to constituent components are derived accordingly, and Hilbert spectra of all constituent mono-components are superposed to generate time-frequency representation (TFR). This method features good merits, such as fine time-frequency resolution and free of both inner and outer interferences. Additionally, it neither relies on prior expertise knowledge about signals, nor needs to construct any basis function. It is highly adaptive to reveal the frequency contents and track their time variability of a signal, and provides an effective approach to nonstationary complex multi-component signal analysis. It is illustrated and validated via numerical simulations, lab experiments of a planetary gearbox and rolling bearings, and on-site tests of a hydraulic turbine in a hydraulic power plant.

1. Introduction

Rotating machinery includes rotors, gears, bearings etc., and is widely used in many applications. Vibration signal analysis is commonly used for condition monitoring and fault diagnosis of rotating machinery. It relies essentially on effectively identifying constituent frequency components, detecting their presence, and tracking changes in their magnitudes. However, a rotating machinery vibration signal is usually a mixture of multiple components, including those generated from various

* Corresponding author.
E-mail address: fengzp@ustb.edu.cn (Z. Feng).

https://doi.org/10.1016/j.ymssp.2018.03.004
0888-3270/

mechanical parts (such as rotors or shafts, gears and bearings) and inevitable background noise. Additionally, a rotating machinery often works under nonstationary conditions in practice, particularly during variable speed processes. In this situation, vibration signals feature strong nonstationarity yet contain rich information about its health condition. Therefore, how to effectively unveil frequency contents and their time variability is a major topic in rotating machinery fault diagnosis [1–4].

Time-frequency analysis (TFA) represents a signal in joint time-frequency domain, thus providing a potential approach to vibration feature extraction of rotating machinery under nonstationary conditions. However, conventional TFA methods suffer from shortcomings. For instance, linear time-frequency representations (TFRs), such as short time Fourier transform (STFT) and continuous wavelet transform (CWT), are subject to Heisenberg uncertainty principle and limited time-frequency resolution, since they are an integral transform with any basis or window function. Bilinear TFRs are based on Wigner-Ville distribution (WVD), and therefore suffer from both outer and inner interferences, because they involve a double integral and quadratic term of a signal [2–6]. To improve the time-frequency readability of TFRs, reassignment methods [7] reallocate diluted energy to the center of gravity of time-frequency distribution, and synchrosqueezing transform [8] compresses diluted energy in frequency/scale domain aided by instantaneous frequency estimation. However, these post-processing methods may fail when signal components are closely neighboring to each other on time-frequency plane.

Hilbert transform based TFR has unique merits: (1) Fine time-frequency resolution. It does not involve an integral transform with any basis or window function, but calculates instantaneous frequency as a time derivative of instantaneous phase via Hilbert transform based analytic signal approach, and allocates signal energy to its instantaneous frequency, thus highlighting local signal characteristics. (2) Free of both outer and inner interferences. It is a superposition of constituent mono-components' TFR, rather than a double integral involving quadratic terms of multiple components. These merits overcome the shortcoming of conventional TFA methods, and make it a leading trend in nonstationary signal analysis [9,10].

Hilbert transform based TFR requires effective mono-component decomposition of signals. Empirical mode decomposition (EMD) and its variants can adaptively extract constituent mono-components of a signal by means of numerical approximation. However, they lack rigorous mathematical formulation. Additionally, spectral overlaps between components and existence of intermittences in signals may induce mode mixing, and the resultant pseudo mono-component will mislead further signal analysis [9].

Recently, iterative generalized demodulation (IGD) has been demonstrated to be an effective mono-component decomposition method. Inspired by the idea of generalized demodulation by Olhede and Walden [10], Feng and his colleagues [11–15] proposed IGD, and successfully extracted vibration features of a hydraulic turbine and gear fault features of planetary gearboxes in joint time-frequency domain. Shi et al. [16] further extended it to rolling bearing fault diagnosis. IGD essentially maps each signal component into a constant frequency component by multiplying with a demodulation function (a conjugate complex exponential function of each component's instantaneous phase), then separates it via filtration one by one, and finally recovers each original component by multiplying with an inverse demodulation function (a complex exponential function of each component's instantaneous phase). A successful application relies on proper design of demodulation phase functions. For each component, its instantaneous phase is essentially a time integral of its instantaneous frequency. Therefore, IGD requires the instantaneous frequency of each component known a priori, for the sake of proper demodulation phase function, and the suitability of demodulation functions depends on instantaneous frequency estimation. Consequently, effective identification of signal constituent components and accurate estimation of their instantaneous frequencies play a vital role in application of IGD. However, IGD relies on expertise knowledge about signal structure to identify constituent mono-components and estimate their instantaneous frequencies via time-frequency analysis. This leads to a major hurdle in complex signal analysis, particularly in the case when prior knowledge on signal characteristics is unknown.

Surrogate tests based on time-frequency analysis can adaptively identify the true constituent components of a nonstationary signal [17–19]. Iatsenko et al. [17] exploited surrogate tests to identify the main component and relevant harmonic and/or subharmonics, and defined a nonlinear mode as a sum of these components. They developed nonlinear mode decomposition (NMD) method, which is very useful to reveal the dynamic nature of a nonlinear system. Although a nonlinear mode is usually multi-component in nature, and thereby does not meet the mono-component requirement by instantaneous frequency calculation, the signal component identification via surrogate tests offers an effective approach to address the dependence issue of IGD on expertise knowledge.

In this paper, different from the goal of NMD to discover nonlinear nature of dynamic systems, we aim to develop a quality time-frequency analysis method, so as to clearly reveal the time-frequency details of each constituent component in nonstationary signals, for rotating machinery fault diagnosis under nonstationary conditions. This involves identification of true signal components and separation of each mono-component. This paper addresses these two issues through surrogate test and IGD algorithms respectively, and thereby derives a quality time-frequency analysis method of high time-frequency resolution and free of outer and inner interferences. We utilize surrogate test technique to identify true frequency components, thus meeting the requirement of a priori known instantaneous frequencies by IGD. This eliminates the effect of one's subjective factors on signal component identification and thus addresses a major issue inherent with IGD. Based on the identified signal components via surrogate tests, we design proper demodulation phase functions for IGD. This enables an adaptive mono-component decomposition. To accurately reveal the local frequency details of each mono-component, we calculate instantaneous frequency finely via Hilbert transform based analytic signal approach. Finally, we calculate the Hilbert spectrum of each mono-component, and construct the TFR of raw signal by superposing all mono-components' Hilbert spectra. This adaptive iterative generalized demodulation (AIGD) based time-frequency analysis method has the same merits as IGD