Signal Processing 109 (2015) 95–109

Contents lists available at ScienceDirect

Signal Processing

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

EMD interval thresholding denoising based on similarity measure to select relevant modes

Gongliu Yang [a,b], Yuanyuan Liu [a,*], Yanyong Wang [a], Zhanlong Zhu [b]

[a] School of Instrumentation Science and Opt-electronics Engineering, Beihang University, Beijing, China
[b] School of Instrumentation Science and Engineering, Southeast University, Nanjing, China

ARTICLE INFO

Article history:
Received 15 April 2014
Received in revised form
23 August 2014
Accepted 31 October 2014
Available online 13 November 2014

Keywords:
Empirical mode decomposition (EMD)
Probability density function (pdf)
Intrinsic mode functions (IMFs)
Interval thresholding
Signal denoising

ABSTRACT

This paper introduces a novel EMD interval thresholding (EMD-IT) denoising, where relevant modes are selected using a l_2-norm measure between the probability density function (pdf) of the input and that of each mode, thresholds are estimated by the characteristics of fractional Gaussian noise (fGn) through EMD. To solve the problem of more relevant modes included when the signal is corrupted by fGn with the H increase, a modified l_2-norm method was given. The computational complexity of EMD-IT denoising is also analyzed. And the time complexity of it is equal to that of EMD. Numerical simulation and real data test were carried out to evaluate the effectiveness of the proposed method. Other traditional denoisings, such as correlation-based EMD partial reconstruction (EMD-PR), EMD direct thresholding (EMD-DT) and NeighCoeff-db4 wavelet denoising are investigated to provide a comparison with the proposed one. Simulation and test results show its superior performance over other traditional denoisings in whole.

1. Introduction

Empirical mode decomposition (EMD), first introduced by Huang et al. in [1], has been widely used to analyze the non-stationary and non-linear signal processes by adaptively decomposing any signal into oscillatory components called intrinsic mode functions (IMFs), where wavelet thresholding has been the dominant techniques for many years. The fundamental reasoning of wavelet thresholding [2–4] is that all coefficients lower than a threshold are set to zero, according to the fact that the energy of a signal and noise spread among wavelet coefficients in wavelet domain. A main drawback of this approach is that the basis functions are predefined, leading to mismatch varying nature of signals [5]. In contrast to wavelet thresholding, EMD expresses the signal as an expansion of basis functions that are derived directly from the signal itself [6,7]. The decomposition is based on the sequential extraction of energy associated with various intrinsic time scales of the signal starting from finer temporal scales to coarser ones. As a powerful adaptive decomposition tool, EMD is well suited to estimate the noise or frequency in measurement domains, apart from the specific applications such as biomedical, watermarking, and audio processing.

Recently, the statistical characteristics of white Gaussian noise and fractional Gaussian noise (fGn) through EMD have been revealed in [8–12]. According to these characteristics, each mode can be classified based on its energy density spread function or power spectral density (PSD) criteria. Consequently, many EMD-based denoisings are provided to remove noises from observed data. In [13,14], Boudraa et al. proposed a signal denoising scheme with each pre-filtered IMFs to estimate the signal. However, this study is limited to signals corrupted by white Gaussian random noise. Boudraa et al. have later proved that EMD filtering based on partial reconstruction of relevant modes performs in an adaptive

* Corresponding author.
E-mail address: liuyuanyuan1773@163.com (Y. Liu).

http://dx.doi.org/10.1016/j.sigpro.2014.10.038
0165-1684/

way in [15]. Some extension work could be also found in [16,17]. This EMD denoising makes use of partial reconstruction, the relevant modes kept and irrelevant modes discarded, the evident disadvantage of which is missing some useful information in discarded modes. And it is disastrous for noise removal when the selection of relevant modes is incorrect. In [18–20], a series of novel EMD-based denoisings inspired by standard wavelet thresholding are developed and tested in various signals by Kopsinis and Mclanglin, where EMD interval thresholding, termed EMD-IT, is provided. Different from EMD direct thresholding(EMD-DT), EMD-IT considers the zero-crossing interval as a whole to perform thresholding, which can effectively avoid the discontinuity of the reconstructed signal. Qu et al. provided a novel EMD-based mode cell filtering(MCF) method in [21], where the threshold is obtained using the statistical characteristics of the amplitudes of the mode cell through EMD. However, the selection of relevant modes is unresolved like the previous work. Up to now, whichever EMD-based denoisings you select, it is necessary to determine, which IMFs are pure noise, pure signal, or contain both. So, the problem is raised to resolve urgently.

Boudraa and Cexus [15] put forward a consecutive mean squared error (CMSE) criterion to select relevant modes, but in some cases CMSE criterion can be trapped in a local minima. Papers [16,17] use a correlation-based method to discriminate whether the IMF is relevant or not. But for the noisy signal with different signal-to-noise power ratio (SNR), the method is very unstable because of too strong or weak correlation between the noisy signal and the first mode. In [18–21], the relevant modes were selected based on experience, actually. To avoid these shortcomings, Komaty et al. put forward a probabilistic similarity measure between the probability density function (pdf) of the input signal and that of each mode to determine the relevant modes in [22,23]. And the best results were obtained by the geometric similarity measures [23], especially the l_2-norm. The key idea of pdf-based filtering strategy is to find the first local maximum. However, the position of the first local maximum proves to move forward by simulation when the Hurst parameter of fGn is closer to 1. Given the problems above, a modified EMD-IT method combined with the similarity measure is developed in this paper. Numerical simulation and real data test were carried out to evaluate the effectiveness of this method.

The outline of this paper is as follows. Section 1 is the introduction; the EMD algorithm is reviewed briefly in Section 2; Section 3 gives a criterion of selecting relevant modes; EMD-based denoisings are discussed in Section 4, firstly IMF thresholding-based denoising is given, then the principle of selecting threshold is introduced, and finally the computational complexity of EMD-based algorithm is analyzed; results and discussions of the proposed method applied in the simulation and real data are presented in Section 5; the last Section is the conclusion.

2. Brief review of EMD

EMD can adaptively break down any signal $x(t)$ into a number L of IMFs, termed $h^{(i)}(t)$ $(1 \le i \le L)$. Those basic IMFs

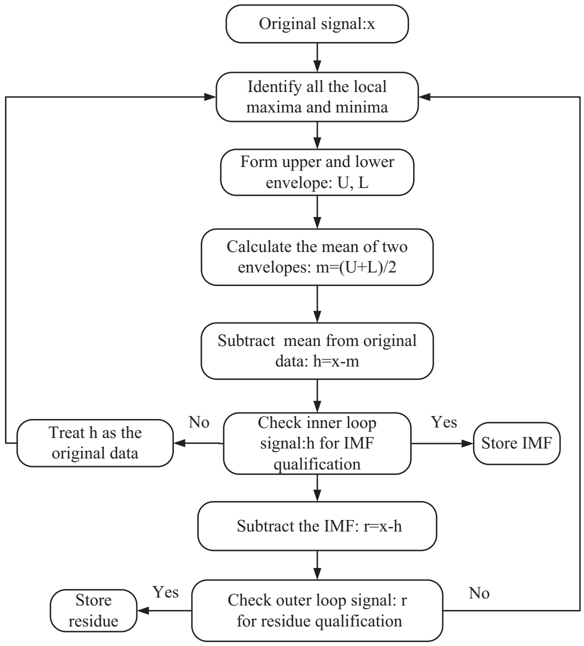

Fig. 1. Pictorial representation of empirical mode decomposition.

are obtained through a sifting process according to the following steps [1,5,24] shown in Fig. 1.

The extracted modes are nearly orthogonal to each other, which form a complete set because accumulating all modes with the residual can restore the decomposed signal. The signal can be expressed as follows:

$$x(t) = \sum_{i=1}^{L} h^{(i)}(t) + r_L(t) \tag{1}$$

Fig. 2(a) depicts as an example the EMD of a Bumps signal with length 2048. It is contaminated by white Gaussian noise, where the SNR is fixed to 5 dB. EMD results in nine IMFs and the last residual shown in Fig. 2(b) and (c).

3. Criterion of selecting relevant modes

Consider a noiseless signal $y(t)$ contaminated by an additive noise $n(t)$

$$x(t) = y(t) + n(t) \tag{2}$$

The denoising is to find an estimate $\tilde{x}(t)$ of the observed signal $y(t)$. For EMD-based denoising, one of the important steps is to discriminate between relevant and irrelevant modes. EMD denoising based on partial reconstruction, called EMD-PR, is given by

$$\tilde{x}(t) = \sum_{i=k_{th}}^{L} h^{(i)}(t) + r_L(t) \tag{3}$$

The k_{th} can be determined by an estimation of correlation coefficient between the original data and decomposition modes. The estimated $\tilde{x}(t)$ can be rewritten as

$$\tilde{x}_m(t) = x(t) - \sum_{i=1}^{m} h^{(i)}(t) \tag{4}$$