Measurement 102 (2017) 230–243

Contents lists available at ScienceDirect

Measurement

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

Parameter free and reliable signal denoising based on constants obtained from IMFs of white Gaussian noise

Weichao Kuang, Bingo Wing-Kuen Ling *, Zhijing Yang

Faculty of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China

ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form 11 January 2017
Accepted 8 February 2017
Available online 12 February 2017

Keywords:
Denoising
Additive white Gaussian noise
Empirical mode decomposition
Intrinsic mode function
Parameter free
Reliable
Constants

ABSTRACT

This paper proposes a denoising method based on the constants obtained from the intrinsic mode functions (IMFs) of the noise. In particular, the model based constant is calculated using the analytical form of both the energy and the mean period of the first IMF of the noise. For practical situations, the corresponding constants are obtained using the sums of logarithms of the real energies and real mean periods of the IMFs of noise. Since the corresponding constants obtained using the practical data of the high order IMFs suffer from large fluctuations, these obtained values are unreliable. To address this issue, the relative percentage errors between the corresponding constants obtained using the practical data and the calculated model based constant are computed only using the second IMF to the sixth IMF. In this case, the corresponding constants obtained using the practical data of the selected IMFs are more reliable and less fluctuated. Next, the upper bound and the lower bound of the corresponding constants obtained using the practical data of the selected IMFs are computed. As no parameter is required to be predefined in the proposed algorithm, the proposed algorithm is more reliable than the existing algorithms. Computer numerical simulation results also show that the proposed algorithm outperforms the existing algorithms.

Contents

1. Introduction

Many signals obtained from practical systems are contaminated by additive white Gaussian noises. These noises could destruct the structures of the signals which cause the degradations of the system performances. Therefore, suppressing the noises while preserving the basic structures of the signals is important for many applications such as for power system applications [1], industrial turbine applications [2] and biomedical science applications [3].

Currently, linear time invariant methods such as the lowpass filtering technique are widely used for the denoising applications because they are easy to be implemented. However, these methods assume that the signals are stationary while most of practical signals are nonstationary. Also, the averaging effect introduced by the

* Corresponding author.
E-mail addresses: kkuangweichao@163.com (W. Kuang), yongquanling@gdut.edu.cn (B.W.-K. Ling), yzhj@gdut.edu.cn (Z. Yang).

http://dx.doi.org/10.1016/j.measurement.2017.02.011
0263-2241/

weighted sum in the convolution operator results to the failure of tracking the sudden jumps in the signal. On the other hand, the wavelet thresholding methods [4] are found to be effective for denoising some nonstationary signals. However, these wavelet approaches require to predefine the basis functions and the predefined basis function does not perform well for all practical signals. Thus, the methods may not achieve satisfied performances for some signals.

Recently, a nonlinear and adaptive signal decomposition method referred to the empirical mode decomposition (EMD) approach is proposed [5]. Since the EMD approach decomposes the signal to a finite set of IMFs based on the local time characteristics of the signal, it is potentially useful in signal denoising. There are several existing EMD based signal denoising methods. These methods are based on deciding the IMFs whether they are dominated by the signal or dominated by the noise. Then, the denoised signal is reconstructed using the signal dominated IMFs. In [6], five different denoising strategies for selecting the signal dominated IMFs are proposed. The strategies are based on different metrics such as the Pearson's correlations, the number of iterations and the entropy. It is proposed in [7] that a threshold region is defined based on the detrended fluctuation analysis (DFA). By comparing the DFA slope of each IMF to the threshold region, the signal dominated IMFs and noise dominated IMFs are decided. In [8], the signal dominated IMFs and the noise dominated IMFs are clustered based on the mutual information between the autocorrelation function of the noisy signal and that of each IMF. Moreover, some other related methods based on the correlation dependent threshold [9,10], relative entropy [11] and optimization [12] are proposed. However, most of these methods require to predefine the thresholds for making the decision. Inappropriate thresholds will cause poor results. In order to overcome this problem, some threshold free denoising methods are proposed. A consecutive mean squared error (CMSE) criterion is proposed to select the IMFs [14]. This criterion is based on the fact that the energies of the IMFs of the pure noise decrease as the indices of the IMFs increase [15–17]. Therefore, the first local minimum of the plot of the energies against the indices of the IMFs was used for discriminating the signals and the noises. This method is a noise dependent method because it employs the empirical characteristics of the IMFs of the pure noise for performing the selection of the IMFs. However, for some cases, the IMF with its index equal to the first local minimum in the plot of the energies against the indices of the IMFs may not be dominated by signal. In this case, employing the CMSE criterion for performing the denoising operation may not be effective [13]. Besides, a statistical based EMD approach was proposed. A probabilistic similarity measure between the probability density function (PDF) of the input signal and that of each IMF is employed for deciding whether the IMFs are dominated by the signal or dominated by the noise. Here, the best probabilistic similarity measure is defined using the l_2 norm criterion [13]. The objective is to select the IMFs whose PDFs catch the dominated features of the noise free signal. Therefore, this method can be considered as a signal dependent method because it is not based on the empirical characteristics of the IMFs of the pure noise. However, this approach is not suitable for various kinds of signals. This is because some IMFs dominated by the signal may also have large l_2 norm distances.

Furthermore, it is proposed in [17] that the uniformly distributed white noise is employed to decide whether the IMFs are dominated by the signal or dominated by the noise. The constants of the pure white noise are estimated based on the real energies and the numbers of peaks of the IMFs. Here, the analytical form of the constants has not been exploited. Instead, the spread lines for the constants obtained from the IMFs of the noisy samples are modeled by an exponential function. The components with constants inside the spread lines are regarded as noises and vice versa as signals. However, the spread lines diverge quickly for high order IMFs. Hence, the constants corresponding to the high order signal components may locate between the lines easily. Therefore, this method is unreliable. Moreover, it is very difficult to select an appropriate confident limit which controls the separation between the spread lines. The spread lines at high confident limit suffer from a large separation. As a result, the constants obtained from some high order IMFs dominated by signal can locate between the lines. On the other hand, the constants obtained from some low order IMFs dominated by noise may locate out of the spread lines at low confident limit. This is because the separation between the spread lines is small. Since different confident limit will result to different discrimination results, an inappropriate confident limit will cause the wrong discrimination results.

To address the above issues, this paper proposes a noise dependent method so that it is suitable for various kinds of signals. The model based constant is calculated using the analytical form of both the energy and the mean period of the first IMF of the additive Gaussian distributed white noise. The relative percentage errors between the corresponding constants obtained using the practical data and the calculated model based constant are computed only using the second IMF to the sixth IMF. This is because the corresponding constants obtained using the practical data based on these low order IMFs of noisy samples are relatively concentrated and reliable. The details will be discussed in detail in Section 3. Also, our selection criterion is not based on the first local minimum of the plot of the energies against the indices of the IMFs. Therefore, the corresponding disadvantages are avoided. Moreover, our proposed method is not required to predefine an appropriate confident limit. This is because the upper and the lower bounds of the corresponding constants are determined automatically. Therefore, our proposed method is a parameter free approach.

The outline of this paper is as follows. Section 2 briefly reviews the procedures for performing the EMD. Section 3 presents our proposed method for performing the parameter free and reliable signal denoising algorithm based on the constants obtained from the IMFs of the white Gaussian noise. Some computer numerical simulation results are presented in Section 4. Finally, a conclusion is drawn in Section 5.

2. Brief review on the EMD

Performing the EMD of signals is to iteratively detect the envelopes of the signal as well as to perform the sifting of the signals. Since this decomposition is based on the local time scale characteristics of the signals, it is applicable to nonlinear and nonstationary signal representations. The definition of the IMF is as follows [5]:

Definition (IMF): A function is considered as an IMF if it satisfies the following two conditions: (1) the total number of extrema and the total number of zero crossing points are equal, or their difference is no more than 1; and (2) its local mean is zero.

Given a signal $x(n)$, where $n = 1, 2, 3, \ldots, N$, the procedures for performing the EMD of $x(n)$ are as follows [5]:

(1) Initialization: Denote $r_0(n) = x(n)$ and $i = 1$.
(2) Compute the ith IMF denoted as $c_i(n)$ using the following iterative procedures:
   (a) Denote $d_0(n) = r_{i-1}(n)$ and $j = 1$.
   (b) Identify all the local maxima and minima of $d_{j-1}(n)$.
   (c) Generate the upper and the lower envelopes of $d_{j-1}(n)$ denoted as $e_{up}(t)$ and $e_{low}(n)$, respectively, using the cubic spline interpolation (typically).
   (d) Calculate the local mean denoted as $m(n) = \frac{e_{up}(n)+e_{low}(n)}{2}$.
   (e) Performing the sifting operation by $d_j(n) = d_{j-1}(n) - m(n)$.