http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (16): 3870-3874, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.3870.3874

Comprehensive Evaluation Study on the Methods for Restraining the End Effects in EMD

[1]XiaojunZhu, [1]Jingxian Hu and [2]Shiqin Lv
[1]College of Computer Science and Technology, Taiyuan University of Technology, Taiyuan, 030024, China
[2]College of Mathematics, Taiyuan University of Technology, Taiyuan, 030024, China

Abstract: For the endpoint effect of the Empirical Mode Decomposition (EMD), researchers have proposed some suitable suppression algorithms so far. In order to compare the decomposition results of those algorithms quantitatively and objectively, after in-depth analysis of the causes of endpoint effects, this study proposes a synthetic evaluation indicator system, which gives attention to both decomposition efficiency and decomposition results. Through allocating reasonable and feasible weights to decomposition rate, correlation coefficient of IMF components and the original signal and the energy difference before-and-after decomposition, this proposed system constructs an evaluation function. Finally, by utilizing this proposed system to evaluate currently used endpoint effect solutions with simulated signals, the results have shown that this proposed system is an effective way of evaluating those suppressing methods.

Key words: Empirical Mode Decomposition (EMD), endpoint effect, synthetic evaluation, correlation coefficient, energy difference

INTRODUCTION

After extensive research and analysis, in 1998, Norden E. Huang and others put forward a new time-frequency analysis method for nonlinear and non-stationary signals, which named as Hilbert-Huang Transform. This method is consists of the EMD (Qiao and Chen, 2012) and the Hilbert Transform. Since, its introduction, the HHT method has been successfully applied to many fields, such as mechanical fault diagnoses, seismic data analysis, biomedical signals processing, etc. However, HHT also suffers from some defects. Indeed, the endpoints of the signal are not necessarily extreme points, so it is uncertainty in fitting the envelopes of the signal by using the cubic spline interpolation function method. This might lead to the endpoint effect to affect the results of signal analysis (He *et al.*, 2012).

In order to restrain the endpoint effect produced in HHT, international and national researchers have been introduced some efficient methods, including mirror extension method (Zhao and Huang, 2001), envelope extension method (Qiao and Chen, 2012), period continuation method (Wen *et al.*, 2011), polynomial fitting method (Liu *et al.*, 2004), AR predict method (Cheng *et al.*, 2005) and so on. There is no doubt that those methods can reduce the impact of the endpoint effect to some extent. But only from IMF components' endpoints to evaluate the restraining effect of a certain method is unreasonable. We need some more related indicators to do further quantitative evaluation.

The rest of the study is organized as follows. Section 2 is the description of common methods for the endpoint effect. In section 3 we present the synthetic evaluation indicator system. Section 4 focuses on experiments and analysis of results. Finally, we describe the conclusions in Section 5.

SYNTHETIC EVALUATIONINDICATOR SYSTEM

After understanding the above introduction, then how to evaluate the effectiveness of each method? Or, how to measure one method is more effective than another? For these problems, we can from the waveforms at the endpoints and the number of the IMFs to do the qualitative analysis.

Then, construct a nonlinear signal as follows:

$$S(t) = \sin(2\pi 90t) + (2 + 0.3\sin(2\pi 5t))\cos(2\pi 13t + 2) \quad (1)$$

Figure 2 shows the EMD results of the simulated signal generated by (a) Original EMD method, (b) Mirror

Corresponding Author: XiaojunZhu, College of Computer Science and Technology, Taiyuan University of Technology, Taiyuan, 030024, China

extension method and (c) Waveform matching method. From the results, we can see there are five IMF components with each method and the last component is the residue. From the results shown in Fig. 2a, there has been appearedserious endpoint effect since the second component. With the iterations, this effect will pass down and ultimately lead to more serious endpoint wing phenomenon in the last two components. It can be seen that, both of the mirror extension method and the waveform matching method have represented their effectiveness in restraining the endpoint effect of EMD. But, the result with the latter method is more efficient and the tip waveforms of the IMFs are more stable. That is, the waveform matching method will achieve a more ideal result than the mirror extension method.

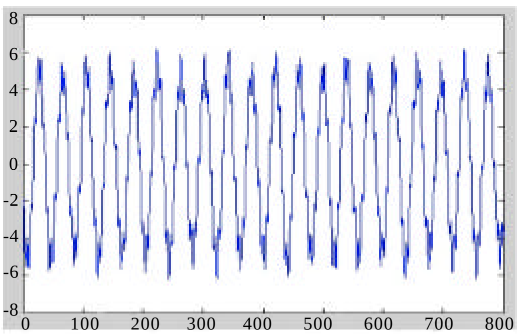

Fig. 1: Waveform of the simulated signal S (t)

Evaluation indicators of restraining results: The effectiveness of one method in restraining the endpoint effect is not only about the tip waveforms, but also about the number of the pseudo-components and the monotonicity of the residue. So, a detailed discussion of quantitative decomposition results is necessary. In fact, Xu Bao-jie (Xu *et al*., 2006) used decomposing time of the EMD to evaluate the decomposition results of those various extending methods. Wang Qiu-sheng (Wang and Duan, 2006) proposed using the correlation coefficient of each IMF and original signal to measure the accuracy of the IMF component. Ren Da-qian (Ren *et al.*, 2007) introduced using energy difference before-and-after the decomposition to evaluate the results of those methods. In reality, each of the three methods above can be used to measure the results of those restraining methods to some extent. However, there also exist some flaws. Therefore, this study will consider the three aspects mentioned above to construct a comprehensive evaluation indicator system.

Computing time: Indeed, high speed is needed in some fields, such as seismic signal analysis, clinical disease diagnosis and so on, so computational speed should be concerned while focusing on restraining results.

Correlation coefficient: From the study of the principle of HHT and signal analysis, we know that each IMF has its own specific physical significance and it is related to the

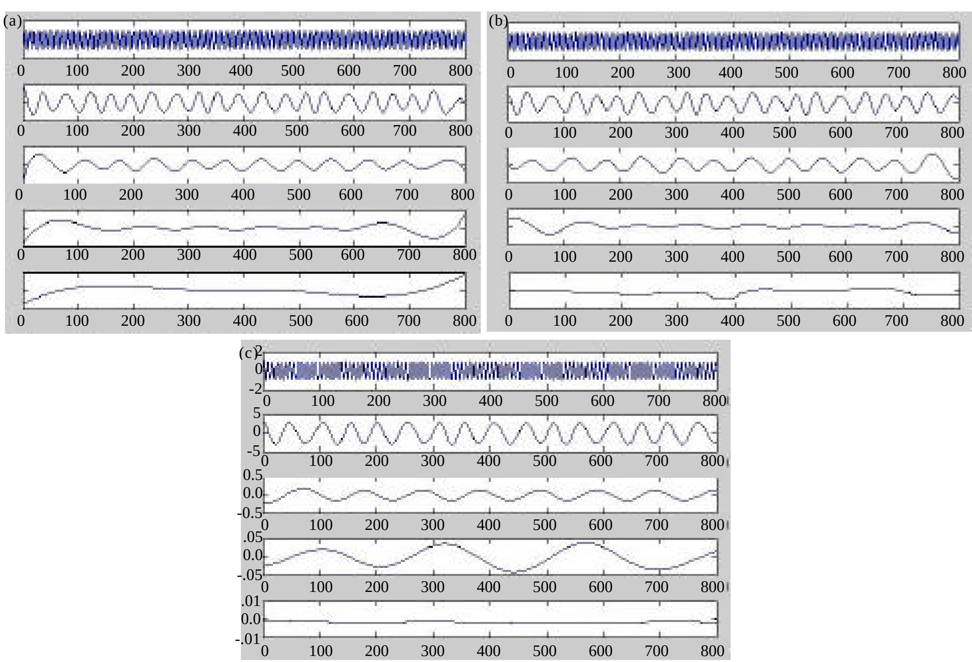

Fig. 2 (a-c): EMD results of S (t) with different method: (a) Original method, (b) Mirror extension method and (c) Wave matching method original signal. So, we can judge the effects of the restraining methods by calculating the correlation coefficients of the IMFs and the original signal.

After the processing of an original signal X (t) using the EMD method, we obtain:

$$X(t) = \sum_{i=1}^{n} c_i(t) + r_n \tag{2}$$

where, $c_i(t)$ is the i-th IMF component and r_n is a residue. And the correlation coefficient of the I-th IMF and the original signal is defined as follows:

$$\rho_i(X(t), c_i(t)) = \frac{\mathrm{cov}(X_i(t), c_i(t))}{\sqrt{E(X(t))}\sqrt{E(c_i(t))}} \tag{3}$$

where, cov () is a covariance function and E () is a variance function. The higher the value of ρ_i, the stronger the relevance.

Pseudo-components produced by the EMD also can be eliminated with this method. The threshold can be set to one tenth of the maximum correlation coefficient. And the pseudo-components below the threshold will be eliminated (Shen *et al.*, 2009). The original signal X (t) can be decomposed into k IMF components which include h pseudo-components. As a result, the number of the effective IMFs can be expressed as t = k-h and the total correlation coefficient can be defined as:

$$\rho' = \frac{\sum_{i=1}^{t} \rho_i}{t} \tag{4}$$

And the higher the value of ρ', the greater the restraining influence.

Energy difference: The pseudo-components produced by EMD method will increase the amount of IMF components, so the energy difference before-and-after the decomposition also can be an indicator to evaluate the restraining result. However, the energy of the original signal itself should be taken into account while using the energy difference indicator. Therefore, we can evaluate the decomposition results according to the ratio of the energy difference to the energy before the decomposition. From the study of signal analysis, the average energy of a signal could be expressed by the root mean square value as follows:

$$RMS_i = \sqrt{\frac{\sum_{i=1}^{k+1} c_i^2(t)}{k+1}} \tag{5}$$

where, k represents the order of the IMF component:

$$RMS_i = \sqrt{\frac{\sum_{i=1}^{k+1} X^2(n)}{N}} \tag{6}$$

where, N is the number of sampling points of the X (t). According to Eq. 5 and 6, θ can be expressed in the following form:

$$\theta = \frac{|RMS_p - RMS_i|}{RMS_p} \tag{7}$$

The value of θ can be used to describe the restraining effects. Obviously, the smaller the value of θ, the smaller the impact of the endpoint effect and vice versa. And if its value is close to zero, it means that the energy of the signal has no significant change before-and-after the EMD processing.

Synthetic evaluation indicator model: After the above analysis of several factors affecting the endpoint effect of EMD, how to weigh the importance of each factor is a question which needs careful consideration.

Normalization of evaluation indicators: Generally, evaluation indicator may include a "minimum type" indicator, a "maximum type" indicator, a "center type" indicator and "interval type" indicator. When the evaluation indicators are inconsistent, we should make them consistent before the comprehensive evaluation.

Therefore, it can be handled as follows.

Make t' = 1/t and θ' = 1/θ to maximize the t and θ separately. Here, t represents computing time and θ represents energy difference. In this way, all of the evaluation indicators are consistent into a "maximum type" indicator.

Non-dimension of evaluation indicators: The dimensions of the computing time t, the correlation coefficient ρ' and the energy difference θ are different and their units and orders of magnitude are also different. So, it has incommensurability. If making direct weighted evaluation, it will cause a phenomenon that larger numbers eat smaller numbers and it ultimately influences the evaluation effects. To eliminate the influence, those evaluation indicators should become dimensionless.

Non-dimensional technology is to make evaluation indicator standardization. Methods such as standard deviation method, extremum method, efficacy coefficient method, normalization method and so on are commonly used. According to the actual situation, the extremum method is used in this study.

The principle of extremum method is as follows.

Suppose there are n evaluation indicators $x_1, x_2, \ldots, x_n$. These data are carried on uniformization to be "maximum type" indicators and there are k groups of records as x_{ij} (i = 1, 2, …k; j = 1, 2, …n).

After that, those data are dealt with the dimensionless method.

Set:

$$x_{ij}' = \frac{x_{ij} - m_j}{M_j - m_j} (i=1,2,...,k;\ j=1,2,...,n) \tag{8}$$

Where:

$$M_j = \max_{1\le i\le n}\{x_{ij}\}, m_j = \min_{1\le i\le n}\{x_{ij}\} (j = 1, 2, ..., n)$$

So, see $x_{ij}' \in [0, 1]$ as the non-dimensional quantitative indicator.

Because the three indicators t, ρ' and θ are independent each other, the linear weighted summation method cloud be used to set their weights. To decompose a signal using EMD method, at the same time of achieving higher decomposition quality, we should pay attention to the decomposition efficiency. Therefore, we can set their weights at the rate of 7:3. The decomposition efficiency is determined by computing time. And the decomposition quality is determined together by energy difference and correlation coefficient, so we can set their weights with the proportion of 4:3. Ultimately, the weights of the evaluation indicators are as follows: {t':ρ':θ} = {0.4:0.3:0.4}. From the foregoing, the general decomposition quality degree can be defined as:

$$T = \sum_{i=1}^{3} a_i b_i \tag{9}$$

where, α_i is the evaluation indicator parameter and b_i is the weight of the relevant indicator.

EXPERIMENTS

Several commonly used methods were introduced in section 1 and we knew that they had their own characteristics. In section 2, the indicators which can measure the restraining methods for the endpoint effect were discussed. Next, some of the methods will be analyzed through simulation experiments to discuss their advantages and disadvantages in restraining the endpoint effects. At first, structure a non-stationary signal and then deal it with the mirror extension method, the envelope extension method, the wave matching method, the polynomial fitting method and the AR prediction method, respectively.

Table 1: Effect comparison of extension method

Restraining methods	Computing time	Correlation coefficient	Energy difference
Mirror extension method	0.2743	0.6086	0.0197
Envelope extension method	0.2221	0.5349	0.0311
Wave matching method	0.8633	0.8931	0.0021
Polynomial fitting method	0.7742	0.7895	0.0034
AR prediction method	1.9285	0.7802	0.0029

Table 2: Comparison results of comprehensive evaluation indicators

Restraining methods	0.3t'	0.3ρ'	0.4θ'	T
Mirror extension method	0.2355	0.0617	0.0168	0.3140
Envelope extension method	0.3000	0	0	0.3000
Wave matching method	0.0482	0.3000	0.4000	0.7482
Polynomial fitting method	0.0582	0.2132	0.2344	0.5058
AR prediction method	0	0.2054	0.2819	0.4873

Through analyzing those data in Table 1, we can summarize as follows. Firstly, from the computing time, the envelope extension method is the best and the AR prediction method is the least in efficiency. Secondly, the correlation coefficient of the wave matching method is the biggest and that of the envelope extension method is the smallest. Finally, the energy difference of the wave matching method is the smallest, but that of the envelope extension method is the biggest.

First, make those data in Table 1 into consistent and dimensionless (Le and Cotoni, 2010) with the comprehensive evaluation indicator in this study. Then weighing method is used to get results as shown in Table 2 from which we can see that the wave matching methodachieves the best results. And the general decomposition quality indicator T of it is equal to 0.7482 which explains this point nicely.

CONCLUSION

Some solutions have been proposed for the endpoint effect of EMD. A restraining method is superior or not, that can be reflected from the decomposition time, the correlation coefficient of each IMF and the original signal and the energy difference before-and-after the decomposition. In this study, several commonly used restraining methods were presented and a comprehensive evaluation indicator system is constructed by assigning a reasonable weight to each factor. The simulation results have shown that the evaluation indicator system in this study can identify the strengths and weaknesses of those methods.

ACKNOWLEDGMENT

This study is Supported by the R and D Infrastructure and Facility Development Program of

Shanxi Province (grant number 2011091003-0103) and the Natural Science Foundation for Young Scientists of Shanxi Province (grant number 2013021016-3)

REFERENCES

Cheng, J.S., D.J. Yu and Y. Yang, 2005. Research on the problem of the endpoint effect of empirical mode decomsition. J. Vibration Shock, 24: 40-42.

He, Z., Y. Shen and Q. Wang, 2012. Boundary extension for hilbert-huang transform inspired by gray prediction model. Signal Process., 92: 685-697.

Le, B. and V. Cotoni, 2010. Validity diagrams of statistical energy analysis. J. Sound Vibration, 329: 221-235.

Liu, H., Y. Zhang and J. Cheng, 2004. Dealing with the end issue of EMD based on polynomial fitting algorithm. Comput. Eng. Appl., 16: 83-86.

Qiao, L.H. and S.S. Chen, 2012. Bidimensional empirical mode decomposition and developed 2D hilbert spectrum characters for the segmentation of texture. Adv. Inform. Sci. Ser. Sci. Adv. Inst. Convergence Inform. Technol., 4: 319-327.

Qiao, L.H. and S.S. Chen, 2012. The construction of two dimensional hilbert huang transform and its application in image analysis. Int. J. Digital Content Technol. Appl. Adv. Inst. Convergence Inform. Technol., 6: 236-244.

Ren, D.Q., Z.T. Wu and G.B. Yan, 2007. Evaluation of the EMD endpoint effect and its window based method. Manuf. Automation, 29: 21-24.

Shen, L., J.S. Li and J.H. Wang, 2009. Application of improved hilbert-huang transform method in gear fault diagnosis. J. Aerospace Power, 24: 1899-1903.

Wen, S.G., F. Wang, C.D. Wu and Y.Z. Zhang, 2011. Using a gaussian process to improve and utilize empirical mode decomposition. J. Northeastern Univ. (Nat. Sci.), 32: 468-471.

Wang, Q.S. and D.H. Duan, 2006. Endpoint effects processing technique of empirical mode decomposition. Comput. Measurement Control, 16: 1673-1675.

Xu, B.J., J.M. Zhang and X.L. Xu, 2006. A study on the method of restraining the ending effect of Empirical Mode Decomposition (EMD). Transa. Beijing Inst. Technol., 26: 195-200.

Zhao, J.P. and D.J. Huang, 2001. Mirror extending and circular spline function for empirical mode decomposition method. J. Zhejiang Univ., 2: 247-252.