Study on Ways to Restrain End Effect of Hilbert-Huang Transform

Guang Yang[1] Xiang-Bin Sun[1] Ming-Xi Zhang[1]
Xiao-Li Li[2] Xin-Rong Liu[3]

[1] Electrical &Electionic Experment Training Center, Yantai University,
Yantai, 264005, China
ygyantai@163.com

[2] Training Center, State Grid Shanghai Municipal Electric Power Company,
Shanghai, 200433, China
zhongshanlinggu@163.com

[3] Department of computer, Shandong Institute of Business and Technology,
Yantai, 264005, China
dbdlxy@tom.com

Received 21 January 2014; Revised 11 April 2014 ; Accepted 21 July 2014

Abstract. The Hilbert Huang Transform (HHT) is an advanced algorithm, but its end effect problem limits its applications severely. A new method to deal with the end effect is proposed in this paper. A new computation parameter is applied to estimating the similarity between two waveforms. Both the shape factor and amplitude factor are considered in the new parameter, and the more suitable end extending waveform can be acquired. The window function is designed and added to the extended data wave, and that can reduce the influence of end effect further. In the simulation experiments, time-frequency spectrums and marginal spectrums are given, and by contrasting the spectrums we can draw the conclusion that this new method to deal with the end effect is feasible and it can achieve better results than the traditional algorithms.

Keywords: end effect, similarity, end extending, window function

1 Introduction

Now days, Hilbert Huang Transform (HHT) has been applying to many fields such as geophysics, biomedicine and engineering. Due to the advantage in analyzing non-linear and non-stationary signals, HHT is more and more being paid attention to [1]. But the end effect of it has an impact to the computational accuracy, and that has limited its application widely.

Domestic and foreign scholars have already done a great deal of theoretical and empirical research on the end effect of HHT. The research method can be roughly divided into three categories: the waveform extending method, the extremum extending method and the data forecasting method. [2][3] give a extending method based on waveform matching, but the shape similarity is not considered when the matching degree is calculated. A mirror extending method is discussed in [4][5], but the extended waveform is not accurate. An extremum extending method is studied in [6], and the computing method is too complex. In [7], the model of time series is built, and the extending data is forecasted. But the method is not good at deal with the non-linear and non-stationary signals. A process method for end effects of hilbert-huang transform based on support vector regression machines is studied by [8]. This method is super to the neural net algorithm method, but the complicated calculations cause a slow matching speed.

The methods above show the representative methods for dealing with the end effect, but that still have a common defect and the end point after being improved is uncertain. Usually, the end point is not the extreme point and the end of envelope is still divergent. That still can lead to decomposition error, and with the process of decomposition, the error will diffuse to all the decomposition processes. So, the more effectively end effect methods still need to be studied.

Research shows that the waveform extending method has better effect on restraining the end effect. In this paper, a new end extending method is studied, and the realization process of the algorithm is introduced in detail.

2 End Effect

HHT include two parts. One is the empirical mode decomposition (EMD); the other is spectral analysis of Hilbert. In EMD, the envelope of maximum and the envelope of minimum are obtained by cubic spline interpolation [9][10][11]. In most cases, the endpoint of the signal is not the extreme point, and this lead to the extreme envelope become divergent at the end part, and this causes significant errors to the decomposition results. The errors can cause the waveform distortion at the end of intrinsic mode function (IMF)[13][14]. Because the next IMF is associated with the previous ones, the errors will be gradually spread with the decomposition process. Finally, the all data sequence is influenced by the errors. That is the end effect of EMD [15]. The formula 1 is used to illustrate this phenomenon.

$$X(t) = 0.8 \cdot \cos(10 \cdot \pi \cdot t) + \cos(4 \cdot \pi \cdot t) + \sin(2 \cdot \pi \cdot t) \tag{1}$$

Fig.1 gives the curve of $X(t)$ in formula 1 ($0.4 \leq t \leq .3.6$), and the extreme envelope are given. The two red dotted lines are the envelope of maximum and the envelope of minimum. From Fig.1 we can see that the end of the minimum envelope is diverging, and the minimum envelope can not include all data. That will influence the results of EMD decomposition. Fig. 2 gives the curve of $x(t)$ and three IMFs by the decomposition process. The three IMFs should show the three periodic functions in formula 1. We can see that the ends of the three curves are all distorted, and the distorted parts are all mark up by the red circles. The distorted degree of IMF1 is the smallest, and that of IMF3 is the biggest. That is due to the errors are conducted from IMF1 to IMF3, and the errors increase with the conductive process.

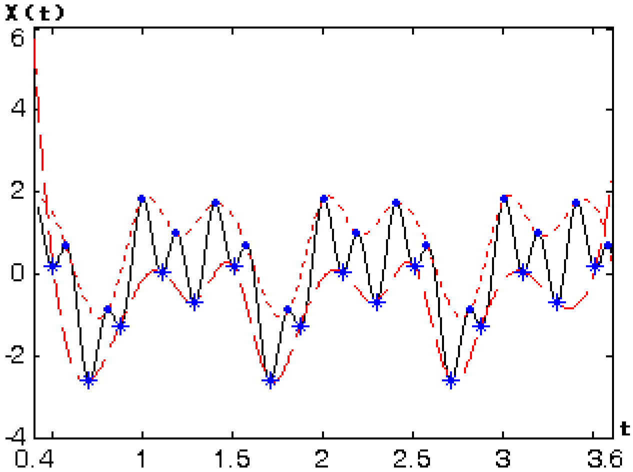

Fig. 1. Extreme Envelope of *X(t)*

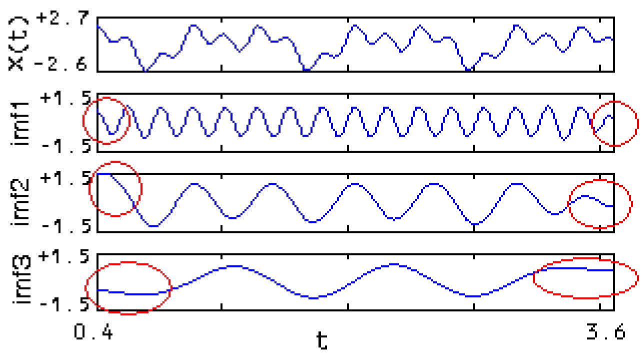

Fig. 2. Result of EMD Decomposition

The problem of end effect also exists in the spectral analysis of Hilbert Transform [16][17]. When we get the time-frequency spectrum and the marginal spectrum of each IMF, The sampling signal cycle is usually not integral, and that will result to a phenomenon of frequency spectrum leakage when the Fourier transform is processed. That causes the end of time-frequency spectrum become divergent, and the end effects come into being [18]. The result of data analyses is affected. To avoid the influence of the end effect, the sampling signal cycle must be insured to be integral, but that is impossible in a practical application because of the randomness and uncertainty of sampling signal [19][20].

In order to deal with the problem of end effect, two factors should be considered. One is how to avoid the problem of the envelope end-point singularity; the other is how to obtain the trendency of signal endpoint. The ultimate aim is to eliminate error of the envelope that is obtained by cubic spline interpolation.

At present, the most effective way to reduce the impact of end effect is the end extending method. In this paper, a new end extending method is proposed, and the viability and effectiveness are verified by a series of experiments.

3 The End Extending

3.1 Fundamental Concepts

To determine endpoint curve's tendency is the crux of end extending. The change tendency of original signal not only is related to the shape of terminal curve, but also the change of inside curve [21]. If we find the trend of a waveform is similar to the terminal curve, the waveform can be selected as the end extending waveform. The similarity of two segments of the curve is related to the shape and the amplitude of the waveform. But there are seldom end extending approaches consider two factors at the same time. In this paper, both shape and amplitude are considered, and the most suitable end extending waveform can be obtained.

Some fundamental concepts are explained as below:

1) The matching distance of two curves is defined as follows:

$$md=\sum_{i=1}^{n}|x(t_1+i\cdot\Delta t)-y(t_2+i\cdot\Delta t)| \tag{2}$$

Where *md* is matching distance of *x(t)* and *y(t)*; *i* is the number of sampling point and $\triangle t$ is sampling interval.

It can be seen from this formula that the smaller the value of *md* is, the more similar the two curves are.

2) The similarity coefficient of two curves is defined as below:

$$sc(x(t),y(t))=\frac{\mathrm{cov}(x(t),y(t))}{\sqrt{\sigma(x)}\cdot\sqrt{\sigma(x)}} \tag{3}$$

Where *sc(x(t),y(t))* is the similarity coefficient of *x(t)* and *y(t)*; *cov(·)* is covariance, and $\sigma(\cdot)$ is variance. $-1\leq sc\leq 1$.

From above formula, we can see that the bigger the value of *sc* is, the more similar the two curves are. When *sc*=1, the waveform is identical, and on the contrary, when *sc*=-1, the waveform is completely opposite.

3) The matching distance only gives the difference degree of two waveforms' amplitude. The similarity coefficient only gives the similarity of two waveforms, but the difference of amplitude between two curves is not considered. So, the new parameter *sd* is established in this paper, and in parameter *sd*, both the matching distance and the similarity coefficient are considered. The parameter *sd* is defined as below:

$$sd=\frac{md_{uni}}{sc+\varepsilon} \tag{4}$$

Where ε is a small positive number, and it can avoid the denominator is zero when *sc*=0. The md_{uni} is normalized matching distance, and it is given as below:

$$md_{uni}=\frac{md}{\max(md)} \tag{5}$$

Through analysis, we know that the smaller the value of *sd* is, the more similar the two curves are.
4) In the process of end extending, a new window function is designed, and it can avoid further influenced by the end effect. The window function is given as below:

$$wf(t) = \begin{cases} 1-(1-t/A)^4 & -A \le t < 0 \\ 1 & 0 \le t < L \\ 1-(t/A)^4 & L \le t < t+A \end{cases} \quad \textbf{(6)}$$

Where *wf(t)* is the window function, *A* is the *t* axis length of the extending waveform. *L* is the t axis length of original signal. In this paper, the *t* axis length of both left and right extending waveform is same. Fig. 3 gives the waveform shape of the window function.

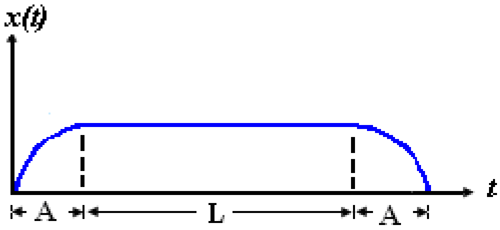

Fig. 3. Window Function

3.2 New End Extending Algorithm

In this paper, we will focus on the end extending of left end point, and the end extending of right end point works in the same way. Suppose the original signal is *x(t)*, and Fig.4 gives a section of *x(t)*. The collection of all maximum t-coordinate is: $\{M_{max}|\ M_0, M_1, \cdots M_k\}$, and the collection of all minimum t-coordinate is: $\{N_{max}|\ N_0, N_1, \cdots N_k\}$. S_0 is the t-coordinate of left end point. The curve segment includes S_0, M_0 and N_0 is called Ψ_0, and Ψ_0 is used for matching.

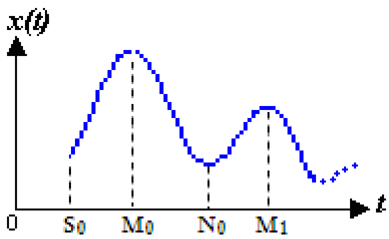

Fig. 4. Waveform Diagram

The end extending process is illustrated as follows:
1) Moving Ψ_0 from the left of the curve to the right, until M_0 overlap with $M_i(i=1,2,\cdots k)$. The curve segment that overlaps with Ψ_0 is called $\Psi_i(i=1,2,\cdots k)$. The *sd* of Ψ_0 and Ψ_i is calculated respectively according to the formula 4. Finding out a Ψ_p that the minimum value of *sd* is corresponds to, and Ψ_p is the matching waveform we want to find. M_p is the maximum of Ψ_p.
2) A curve segment that on the left of M_p is picked up as the extending waveform and the extending waveform is moved to the head of S_0 to achieve the left end extending. The length of the extending waveform should be determined in advance.
By the same way, the right end extending can be achieved also.
3) After the end extending process, the window function should be added to the extended curve according to the formula 7:

$$y(t)=\begin{cases} x(t)(1-(1-t/A)^4) & -A\le t<0 \\ x(t) & 0\le t<L \\ x(t)(1-((t-L)/A)^4) & L\le t<t+A \end{cases} \tag{7}$$

. Where *x(t)* is the data signal after being extended, and *A* is the t axis length of the extending waveform. *L* is the *t* axis length of original signal before extending. *y(t)* is the waveform after the window function is added.

Fig.5 gives a data signal after being extended, and the curve segments that are included in two red boxes are the extending waveforms. Fig.6 gives the curve after the window function is added to. S_0 is the left end point of original signal, and S_k is the right end point.

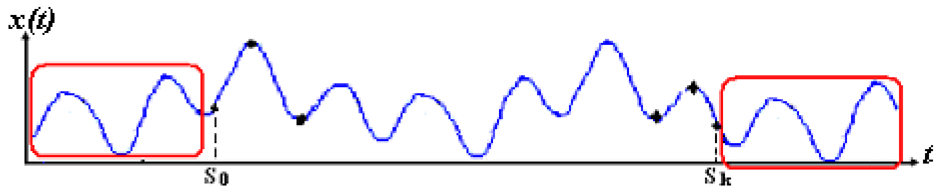

Fig. 5. A Data Signal after Being Extended

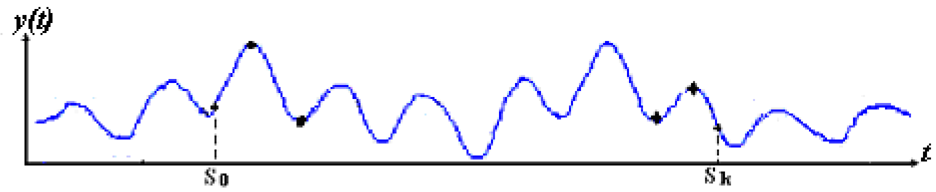

Fig. 6. The Signal after Window Function Is Added

4) To the extended signal, the EMD and Hilbert transform are run. By cutting out the extended segment part from the extended signal result, the final result for original data is obtained.

In the above algorithm, the waveform matching extension is achieved at first, and the window function is added followed. By the process, the junction of the original signal and its extending waveforms can have a smooth transition, and the abrupt frequency hopping at the boundary is avoided. In order to further reduce the extension error, the window function is added. The window function can weaken the error at both extending curve ends, and it can solve the problem of end-point divergency at envelope's end point. This can prevent the error diffuse to the internal signal, and the better decomposition process can be acquired.

4 The Simulation Experiments

In order to verify the effectiveness of the above algorithm, simulation experiment is implemented. Formula 8 is designed to test the algorithm.

$$\begin{aligned} \mathrm{x(t)} = (1+0.2\cdot\sin(15\cdot\pi\cdot \mathrm{t}))\cdot\cos(60\cdot\pi\cdot\mathrm{t}+ \\ 0.5\cdot\sin(30\cdot\pi\cdot\mathrm{t}))+\sin(240\cdot\pi\cdot\mathrm{t}) \end{aligned} \tag{8}$$

We use the interval of *x(t)* ($t\in[0.02\ \ 0.21]$) to process the EMD decomposition and HHT transform. Sample frequency is 2560, A=0.0037.

Fig.7 shows the envelopes for the curve of formula 8 ($t \in [0.02\ \ 0.21]$). From Fig.7 we can see that the enve-

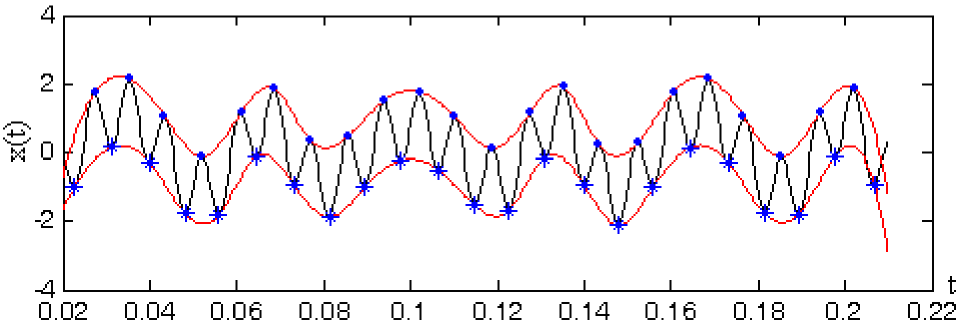

Fig. 7. Curve of *x(t)*

lope of right endpoint is divergent, and the maximum envelope can not include all data near the right endpoint. The problem of end-point singularity is considered to be the main role in the process of end effect. Because the envelopes will be used to evaluate the IMFs, and the error of the envelopes will influence the all process of EMD and Hilbert transform. So, the end effect will generate inevitably.

In Fig.8 , the curve that has been extended using traditional algorithm(introduced by [2]) is given($t \in [-0.03\ \ 0.26]$), and the envelopes are shown too. From Fig.8 we can see that the envelope of left endpoint is divergent and the right endpoint envelope can not include all data. Though, the end effect caused by this situation can be weakened by the end extending algorithm, but that still influence the final result to some extent.

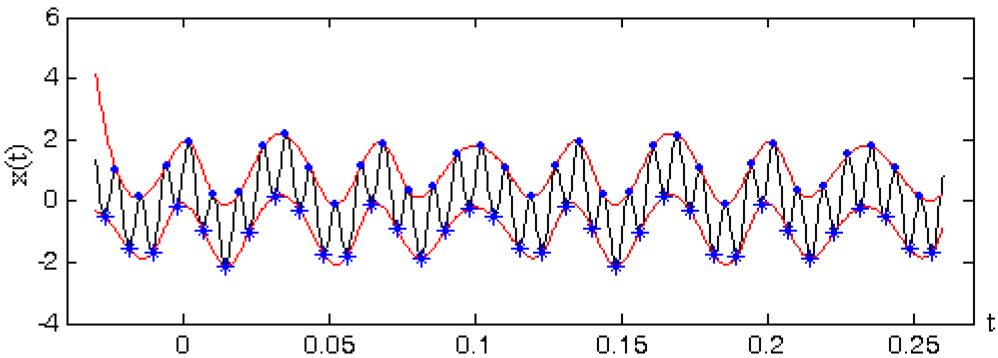

Fig. 8. Extended Curve by Traditional Algorithm

Fig.9 shows the extended curve by algorithm of this article($t \in [-0.03\ \ 0.26]$), and the corresponding envelopes are also given. As can be seen from Fig.9, all the data of the curve can be included by envelopes and the envelope shape at two ends of the curve is convergence. In Fig.9, the shape factor and amplitude factor are embodied in the initial stages of the curve, and the divergent phenomena of envelopes can be corrected effectively at the ends of the curve. That can play an important role in restraining the end effect.

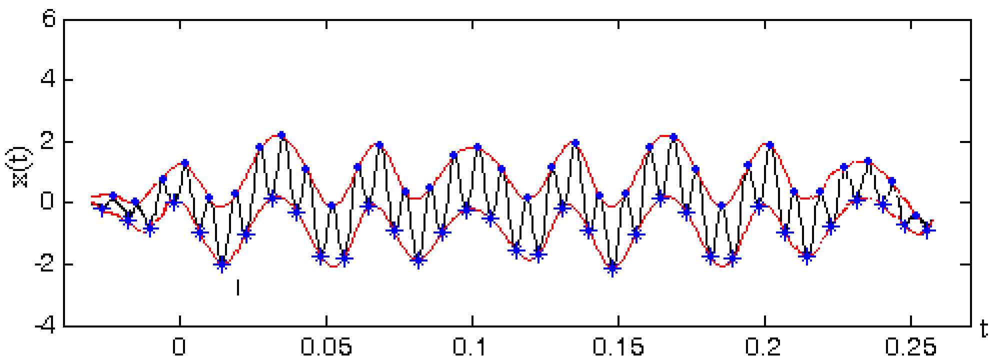

Fig. 9. Extended Curve by New Algorithm

Fig.10 is the decomposition of EMD before end extending. From Fig.10 we can see that four IMFs are decomposed, and the frequency of IMF1 is the highest. Because of the impact of end effect, the end waveform of IMF1 is distortion. Fig. 11 is the time-frequency spectrum of original signal before end extending, and Fig.12 is the IMF1 marginal spectrum of it. Form Fig.11 we can see that the end of IMF1 frequency spectrum has several distortion, and the frequency has anomalous change. In this paper, only the marginal spectrum of IMF1 is given

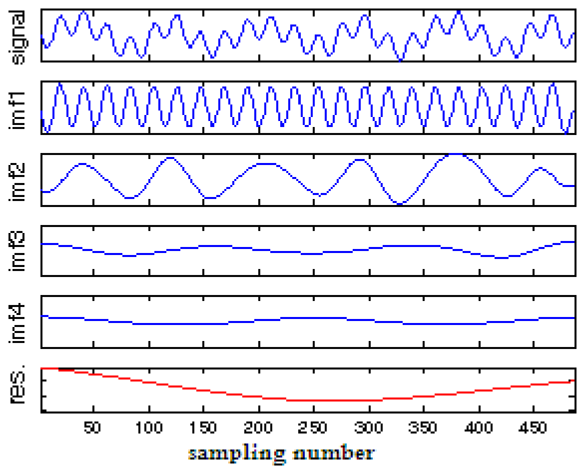

Fig. 10. EMD Decomposition before End Extending

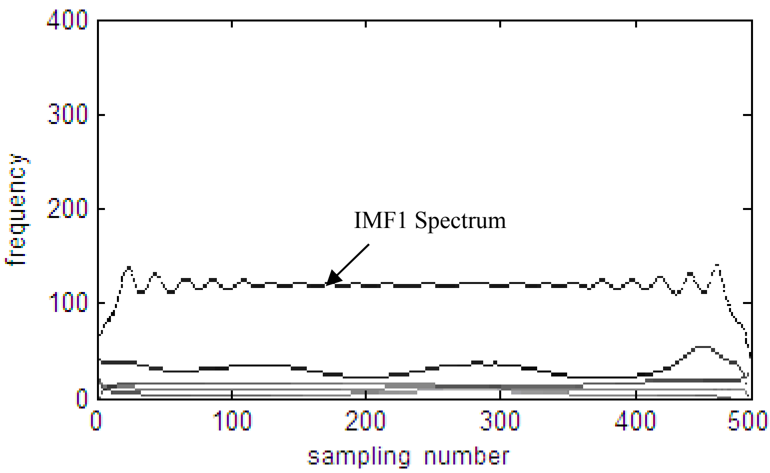

Fig. 11. Original Time-frequency Spectrum

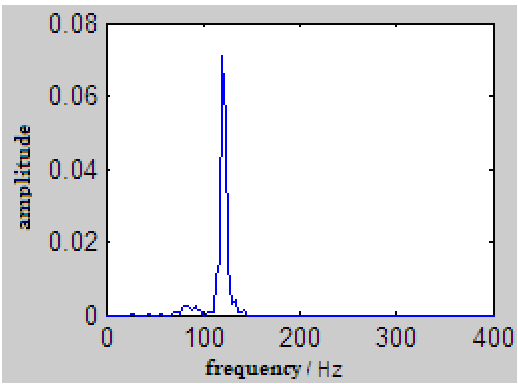

Fig. 12. Original Marginal Spectrum of IMF1

so as to show the result clearly. From Fig.12 we can see that the frequency amplitude is the maximum near 120Hz, and it is the main frequency bands of IMF1. There is also some low frequency part between 50Hz to 100Hz, but the frequency amplitude is very small. This low frequency part is caused by end effect.

The process of end extending is fulfilled according to the algorithm mentioned above. After end extending and adding the window function, the EMD decomposition and Hilbert transform are implemented. Finally, the results are picked out according the actual data length, and the additional sections are deleted. Fig.13 is the Time-frequency spectrum of the final result, and Fig.14 is the IMF1 Marginal spectrum of it. From Fig.13 we can see that the distortion at the end of IMF1 frequency spectrum has been rectified. The low frequency part between 50Hz to 100Hz has almost disappeared in Fig.14. That can demonstrate that the end extending algorithm has a good effect.

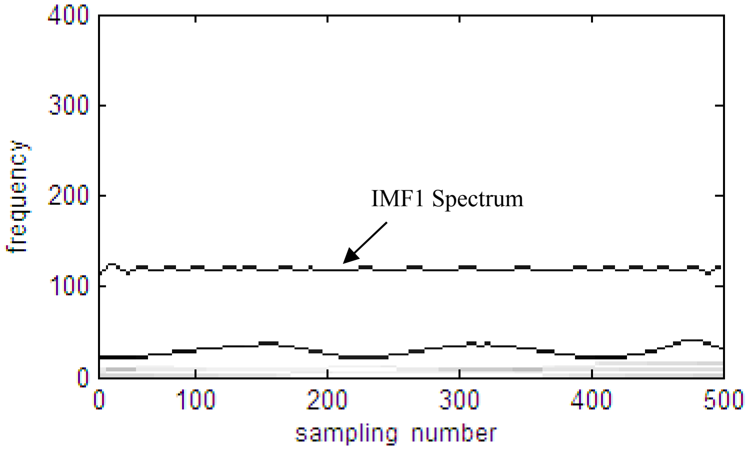

Fig. 13. Time-frequency Spectrum by Improved Extending Algorithm

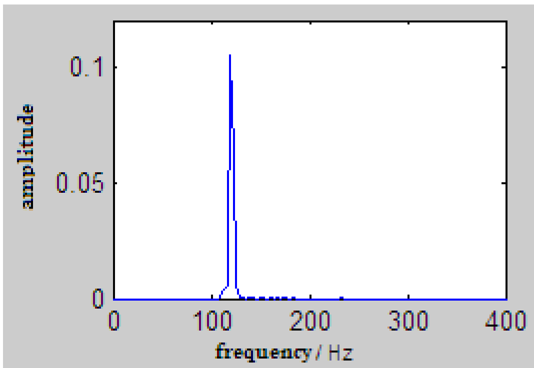

Fig. 14. Marginal Spectrum of IMF1 by Improved Algorithm

The traditional end extending algorithm is introduced by [2], but only the amplitude similarity factor is considered when the waveform similarity is calculated and the window function is not considered. Fig.15 is the time-frequency spectrum by the algorithm, and Fig.16 is the IMF1 Marginal spectrum of it. From the Fig.15, we can see although the algorithm having gained some effect, there is still a little divergent phenomenon in the end of IMF1 frequency spectrum. Comparing with Fig.13 and Fig.14, there is also more the low frequency part before 100Hz in Fig.15 and Fig.16. The reason is that both shape and amplitude factor are considered in the new

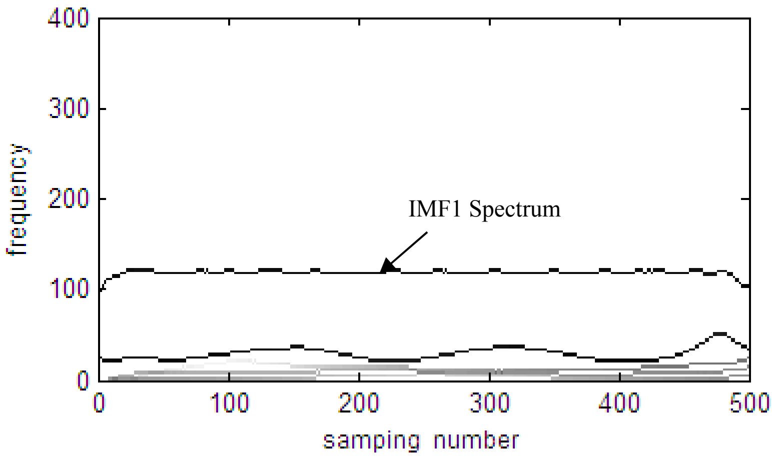

Fig. 15. Time-frequency Spectrum by Traditional Extending Algorithm

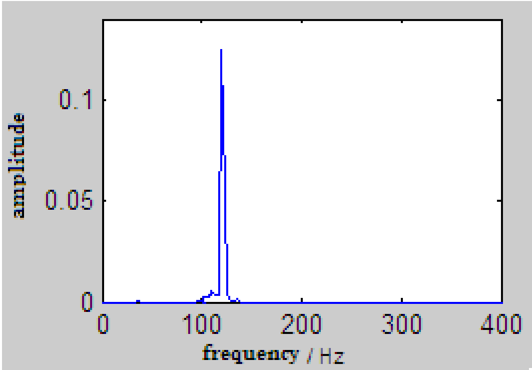

Fig. 16. Marginal Spectrum of IMF1 by Traditional Algorithm

algorithm, and the window function is added to the extended waveform. All of that can play a role in avoiding end effect. So, we can see that the algorithm in this paper has a better performance.

5 Conclusions

The end effect is a serious problem in practical application of HHT. For example, when we analyze a vibration signal, more than 10 IMFs usually can be decomposed, and the more IMFs are decomposed, the more serious the end effect plays impact to.

After discussing the end effect HHT in detail, a new method for the end effect is discussed in this article. The traditional method is improved. Both shape and amplitude factors are considered, and the variation tendency of the waveform can be more precise analyzed. This can make the extended waveform has a more exact embodiment to the change trend of original curve. In order to reduce the envelopes endpoint divergence after being extended, the window function is added to the algorithm. This can reduce the influence of end effect further. The contrast for the extended curve envelopes between the new algorithm and traditional algorithm is given. By analyzing the relevant time-frequency spectrum and marginal spectrum in detail, the experiments have verified that the proposed algorithm is superior to traditional algorithm, and better results can be obtained.

Acknowledgement

This work has been supported by the Shangdong Science and Technology Development Project of China under Grant 2014GGB01358 and Yantai Science and Technology Development Project of China under Grant No. 2014YT061069.

References

[1] Riling G, Flandrin P, "One or Two Frequencies? The Empirical Mode Decomposition Answers," *IEEE Trans. on Signal Processing*, Vol. 56, No. 1, pp. 85-95, 2008.

[2] Shao Chenxi, Wang Jian, Fan Jinfeng, Yang Ming, Wang Zi Cai, "A Self-Adaptive Method Dealing with the End Issue of EMD," *Acta Electronic Sinica*, Vol. 35, No. 10, pp. 1944-1948, 2007.

[3] Hu Aijun, An Liansuo, Tang Guiji, "New Process Method for End Effects of Hilbert-Huang Transform," *Chinese Journal of Mechanical Engineering*, Vol. 44, No. 4, pp. 154-158, 2008.

[4] Hao J P, HuangDJ, "Mirror Extending and Circular Spline Function for Empirical Mode Decomposition Method," *Journal of Zhejiang University*, Vol. 2, No. 3, pp. 247-252, 2001.

[5] Liu Chao, Li Zhengjie, Tong Ningning, Zheng Zhiren, "Improved EMD and Its Simulation Analysis," *Aerospace Electronic Warfare*, Vol. 27, No. 2, pp. 58-61, 2011.

[6] Guo Mingwei, Ni Shihong, Zhu Jiahai, Zhang Zhipeng, "HHT/EMD End Extension Method in Vibration Signal Analysis," *Journal of Vibragtion and Shock*, Vol. 31, No. 8, pp. 62-66, 2012.

[7] Yang Jianwen, Jia Minping, "Study on Processing Method and Analysis of End Problem of Hilbert-Huang Spectrum," *Journal of Vibration Engineering*, Vol. 19, No. 2, pp. 283-288, 2006.

[8] Cheng Junsheng, Yu Dejie, Yang Yu, "Process Method for End Effects of Hilbert-huang Transform Based on Support Vector Regression Machines," *Chines Journal of Mechanical Engineering*, Vol. 42, No. 4, pp. 23-31, 2006.

[9] Guang Yang, Zziqiang Hu, Xinrong Liu, "A Novel Strategy for Wind Speed Prediction in Wind Farm," *TElKOMNIKA Indonesia Journal of Electrical Engineering*, Vol. 11, No. 12, pp. 1001-1008, 2013.

[10] Ding Kang, "Development in Vibration Signal Analysisand Processing Methods," *Journal of Vibration Engineering*, Vol. 16, No. 1, pp. 1-10, 2003.

[11] Huang Daji, Zhao Jinping, Su Jilan, "Practical Implementation of the Hilbert-Huang Transform Algorithm," *Acta Oceanologica Sinica*, Vol. 25, No. 1, pp. 1-11, 2003.

[12] Shi Peiming, Ding Xuejuan, Li Geng, Han Dongying, "An Improved Method of EMD and Its Applications in Rotating Machinery Fault Diagnosis," *Journal of Vibration and Shcok*, Vol. 32, No. 4, pp. 186-190, 2013.

[13] Yann NY, Ying Lei, Pan SW, "System Identification of Linear Structures Based on Hilbert-Huang Spectral Analysis," *Earthquack Engineering & Structure Dynamics*, Vol. 32, No. 3, pp. 1533-1554, 2003.

[14] Qi K, He Z G, Zi Y Y, "Cosine Window-based Boundary Processing Method for EMD and Its Application in Rubbing Fault Diagnosis," *Mechanical Systems and Signal Processing*, Vol. 21, No. 7, pp. 2750-2760, 2007.

[15] Cruz-Rolddn F, Amo-Lpez P,Maldonado-Bascbn S, "An Efficient and Simple Method for Designing Prototype Filters for Cosine-Modulated Pseudo-QMF Banks," *IEEE Signal Processing Letters*, Vol. 9, No. 1, pp. 29-31, 2002.

[16] Zhang Haiyong, "A New Method for Non-stationary Signal Analysis-Local Wave Analysis," *Journal of Electronics and Information Technology*, Vol. 25, No. 10, pp. 1327-1334, 2003.

[17] Huang N E, Shen Z, Long S R, "The Empirical Mode Decomposition and the Hilben Spectrum for Nonlinear and Non-stationary Time Aeries Analysis," *Proceedings of the Royal Society of London Series A*, Vol. 454, No. 10, pp. 903-995, 1998.

[18] Gai Qiang, Zhang Haiyong, Xu Xiaogang, "Study of adaptive frequency multiresolution analysis of the Hilbert-Huang transform," *Journal of Electronics*, Vol. 33, No. 3, pp. 563-566, 2005.

[19] Wen Tzeng Huang, Sun Yen Tan, Che Wun Chiou, Chiu Ching Tuan, Chih Hsiang Chang, "On-line Error Detection in a Polynomial Basis Multiplier over GF(2^m) Using Self-Checking Alternating Logic," *Journal of Computers*, Vol. 24, No. 2, pp. 46-58, 2013.

[20] Rybczyflski J, "The Possibility of Evaluating Turbo·set Beating Misalignment Defects On the Basis of Bearing Trajectory Features," *Mechanical Systems and Signal Processing*, Vol. 25, No. 2, pp. 521-536, 2011.

[21] Y.Y. Hua, J.-M. Lin, C.W. Chiou, C.-Y. Lee, Y. H. Liu, "A Novel Digit-serial Dual Basis Systolic Karatsuba Multiplier over GF(2m)," *Journal of Computers*, Vol. 23, No. 2, pp. 80-94, 2012.