Advances in Adaptive Data Analysis
Vol. 2, No. 4 (2010) 521–543

DOI: 10.1142/S1793536910000616

SMOOTHING EMPIRICAL MODE DECOMPOSITION: A PATCH TO IMPROVE THE DECOMPOSED ACCURACY

SUN-HUA PAO, CHIEH-NENG YOUNG
and CHIEN-LUN TSENG

Industrial Technology Research Institute Chutung,
Hsinchu, Taiwan, ROC 31040
itri960484@itri.org.tw

NORDEN E. HUANG

Research Center for Adaptive Data Analysis,
National Central University, Zhongli,
Taiwan, ROC 32001
Norden@ncu.edu.tw

Hilbert-Huang Transformation (HHT) is designed especially for analyzing data from nonlinear and nonstationary processes. It consists of the Empirical Mode Decomposition (EMD) to generate Intrinsic Mode Function (IMF) components, from which the instantaneous frequency can be computed for the time-frequency Hilbert spectral Analysis. Currently, EMD, based on the cubic spline, is the most efficient and popular algorithm to implement HHT. However, EMD as implemented now suffers from dependence on the cubic spline function chosen as the basis. Furthermore, due to the various stoppage criteria, it is difficult to establish the uniqueness of the decomposition. Consequently, the interpretation of the EMD result is subject to certain degree of ambiguity. As the IMF components from the classic EMD are all approximations from the combinations of piece-wise cubic spline functions, there could also be artificial frequency modulation in addition to amplitude modulation. A novel Smoothing Empirical Mode Decomposition (SEMD) is proposed. Although SEMD is also an approximation, extensive tests on nonlinear and nonstationary data indicate that the smoothing procedure is a robust and accurate approach to eliminate the dependence of chosen spline functional forms. Thus, we have proved the uniqueness of the decomposition under the weak limitation of spline fittings. The natural signal length-of-day 1965–1985 was tested for the performance in nonstationary and nonlinear decomposition. The resulting spectrum by SEMD is quite stable and quantitatively similar to the optimization of EMD.

Keywords: Empirical mode decomposition; smoothing EMD; nonlinear and nonstationary time series; instantaneous frequency; Hilbert spectral analysis.

1. Introduction

Empirical Mode Decomposition (EMD) was first proposed by Huang *et al.* (1998) for analyzing data from nonlinear and nonstationary processes. Unlike the convectional data analysis method, the data would be decomposed, in time domain, into a set of finite number of Intrinsic Mode Function (IMF), which are amplitude and frequency modulations functions. The procedure of EMD consists of 'sifting,' which remove the local mean determined from the envelopes constructed by spline fitting through the local extrema. The IMF admits well-behaved Hilbert Transformation, from which the instantaneous frequency could be computed and Hilbert Spectral Analysis (HSA) would result. The combination of EMD and HSA is designated by NASA as Hilbert-Huang Transform [HHT, see, for example, Huang and Attoh-Okine, 2005 and Huang and Shen, 1999]. HHT has initiated a brand new approach and methodology in data analysis and informatics.

Ever since the introduction of EMD, it has found wide area of applications in engineering, science, biomedical and even financial studies. To ensure the sifting procedure could be implemented precisely, many *ad hoc* criteria have been adopted. Furthermore, methods for finding the optimal local mean for decomposition have been discussed [Wu and Huang, 2009; Riemanschneider *et al.*, 2005; Deering *et al.*, 2005; Chen *et al.*, 2006; Olhede and Walden, 2004; Hou *et al.*, 2009; Lin *et al.*, 2009]. Additionally, the original time series analysis method has also been extended to image and multi-dimensional data analysis [Nunes *et al.* (2003a, 2003b, 2005); Liu *et al.*, 2004; Qin *et al.*, 2006; Xu *et al.*, 2007; Nunes and Delechelle, 2009; Bhuiyan *et al.*, 2009; Yuan *et al.*, 2009; Wu *et al.*, 2009]. While the applications have advanced on a broad front, the mathematical foundation is still left untreated. It is an urgent and unsolved problem. Even on a less level, the problem with regard to the optimal spline basis had attracted very little improvement and attention. To attack the fundamental mathematics is a much harder problem. In this paper, we will examine the limited problem on the implement of the popular EMD algorithm based on cubic spline interpolation. Through our efforts, we have established the uniqueness of EMD under the weak limitation of spline fitting.

The present implement EMD is firmly based on cubic spline. As discussed by Wang *et al.* (2010), there is a conflict between strict definition of IMF and the cubic spline implementation: It would force the IMFs to be constant amplitude functions. This conflict could only be circumvented by keeping the iterative sifting steps at a low number. As pointed out by Wang *et al.* (2010), for a data, $x(t)$, the final set of IMFs are all combination of piecewise cubic functions:

$$c_1(t) = x(t) - (m_{1,1} + m_{1,2} + \cdots + m_{1,k}) \tag{1}$$

and

$$c_i(t) = r_{i-1} - r_i = \sum_{j=1}^{k_1} m_{i-1,j} - \sum_{j=1}^{k_2} m_{i,j} \tag{2}$$

where $c_i(t)$ is the ith IMF, $m_{i,j}$ is the jth iteration for the ith spline functions, and r_j is the jth partial residuals. Although the cubic spline guarantees continuity in data and its first and second derivatives, it could not also guarantee the smoothness and steadiness in the derivatives of the phase function. To improve upon the smoothness, we proposed here a Smoothing EMD method (SEMD), which is designed to eliminate the need for the fixed cubic spline approach. Without the drawback of cubic spline fitting, the SEMD could use any form of spline, linear or cubic, and enable all IMFs to converge to the same IMF decomposition. Thus the smoothing approach could help to reveal more intrinsic properties and deepen our understanding of the underlying physical properties.

Furthermore, we will demonstrate that SEMD always produces a good approximation to the original EMD; therefore, it is a robust alternative to the cubic spline based sifting. The most important consequence of this smoothing approach is to establish the uniqueness of IMF obtained, for this new approach is totally independent of the spline functional form selected for the decomposition, yet the resulting IMFs all converge to the same function. As such, this new SEMD has deepened our understanding of the empirical mode decomposition method. This paper consists of the following sections in addition to the introduction. In the next section, we offer a brief review of EMD method and a detailed description of the SEMD method. Then, we have a section of demonstrations of the SEMD with various examples; and, finally, a section of discussion and conclusion.

2. Empirical Mode Decomposition

The core operation of EMD is to find a precise local mean, $m(t)$, from which one can deduce the local variation of the given data, $x(t)$, that is defined as an intrinsic mode function (IMF), $c(t)$:

$$c(t) = x(t) - m(t) \tag{3}$$

In order to estimate the true mean, the method proposed by Huang *et al.* (1998) was an iterative approach as given in Eqs. (1) and (2), the procedure known as sifting. In the sifting, the mean was derived from the mean of the envelopes constructed from extrema of the data. For a given physical input $x(t)$, the EMD procedure can be summarized as follows:

(1) identify every extremum, including minimum and maximum, of $x(t)$.
(2) calculate envelope $e_{\min}(t)$ (respectively, $e_{\max}(t)$) by interpolating between minima (respectively, maxima).
(3) find the mean of the envelopes as $m(t) = (e_{\min}(t) + e_{\max}(t))/2$.
(4) extract the IMF $c(t)$ by subtracting $m(t)$ from $x(t)$, i.e., $c(t) = x(t) - m(t)$.
(5) repeat the above procedure by treating $c(t)$ so produced as data and produce the next level of sifting.

Once the $c(t)$ satisfies the definition of IMF according to a pre-set stoppage criterion, the refined $c(t)$ is designated as an IMF. It should be pointed out that the so called 'local mean' defined from the mean of envelopes through extrema is really a median.

However, as pointed out by Hou *et al.* (2010) that the median rather than the 'mean' should be the quantity we should abstract. Either mean or median, it is difficult to find the true value from a physical input only through limited extrema even if using recursively sifting. Specifically, the median is given by

$$m(t) = \sum_{i=1}^{S} m_i(t) + e_s(t) \tag{4}$$

where s is the recursive number of sifting and $e_s(t)$ is the left error from the true mean after the sth iteration.

2.1. *Details of the SEMD algorithm*

As mentioned above, the key idea of iterative sifting process is designed to remove the riding waves and, more importantly, to force the local median serving as a reference line so that the wave forms would be symmetric with respect to it. As the local mean and the subsequent IMF components all depend on the spline functional form, which would be a decisive factor in determine the final resulting IMFs. Cubic spline was selected as an over all optimum to retain continuity up to second derivative. Here we will try to propose a modified method, designated as SEMD, to remove the restriction on cubic spline. We will show that through the smoothing procedure, we could reach the same result whether we use a linear or a cubic spline to construct the initial envelopes. The approach is to remove the low frequency part first and gradually move to higher frequency parts. The procedure to obtain the relative lower frequency parts of the envelopes mean executed by n-point moving average is named smoothing. For a physical input $x(t)$, the proposed SEMD procedure can be implemented as follows:

(1) identify every extremum, including all the minimum and maximum, of $x(t)$.
(2) calculate envelope $e_{\min}(t)$ (respectively, $e_{\max}(t)$) by interpolating between minima (respectively, maxima), shown in Fig. 1.
(3) compute the mean $m(t) = (e_{\min}(t) + e_{\max}(t))/2$.
(4) smooth $\bar{m}(t)$ with a given window, shown in Fig. 2.
(5) extract the $c(t)$ by subtracting the smoothed $\bar{m}(t)$.
(6) replace $x(t)$ by $c(t)$ and return to step 1 until the standard deviation (SD) of smoothed $\bar{m}(t)$ is less then the pre-set stoppage criterion, shown in Fig. 3.

In this smoothing EMD, there are two kinds of end effects. The first one is the spline fitting; this is treated in the same way as given in Wu and Huang (2009). Additionally, there is another end effect needs our attention in the smoothing of the mean envelopes: We need to extend the data beyond the boundaries to achieve a running mean for the smoothing. To treat the ends, we propose to extend the points near boundaries by mirror of mean envelope at ends. Although the mirror would produce a cusp point at the boundary, the smoothing procedure would patch

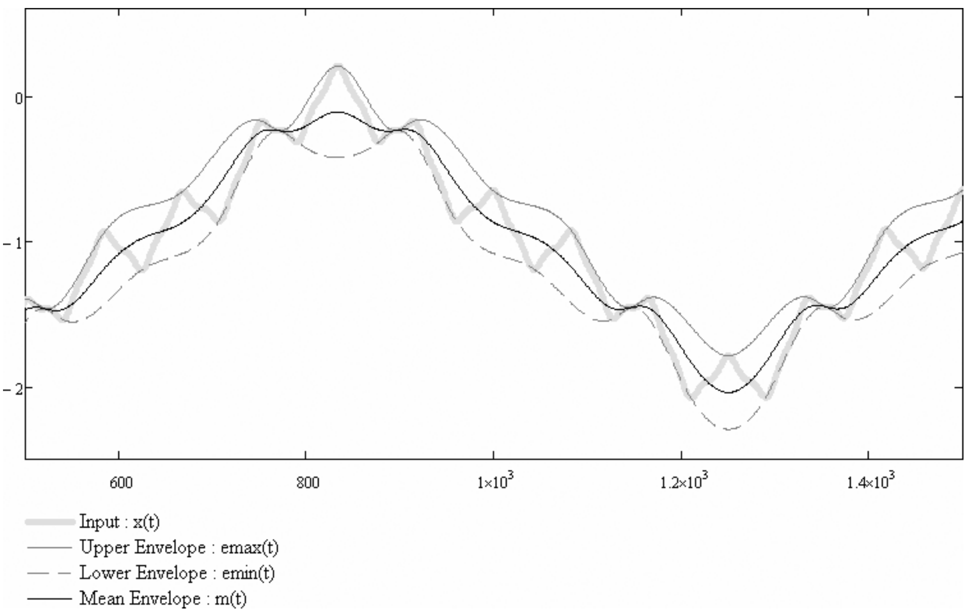


Fig. 1. Construct the envelopes by Cubic Spline passing through all extrema of Input $x(t)$. These envelopes can be replaced by other functions, for example, piecewise straight lines.

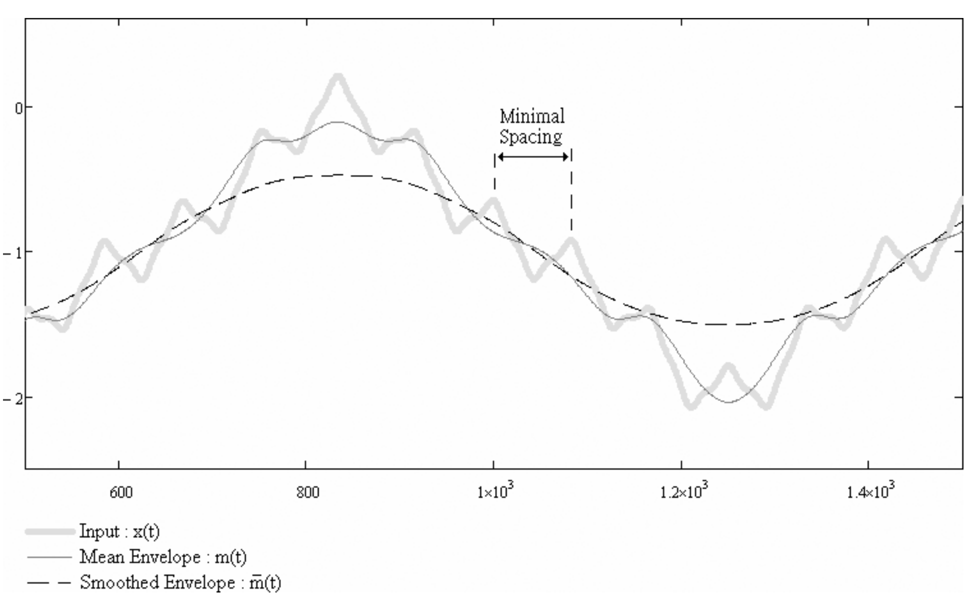


Fig. 2. Smoothed envelope is come from the moving average of mean $m(t)$ with a given window. The smoothing window is altered adaptively by minimal spacing which is the least distance between extremes of $x(t)$.

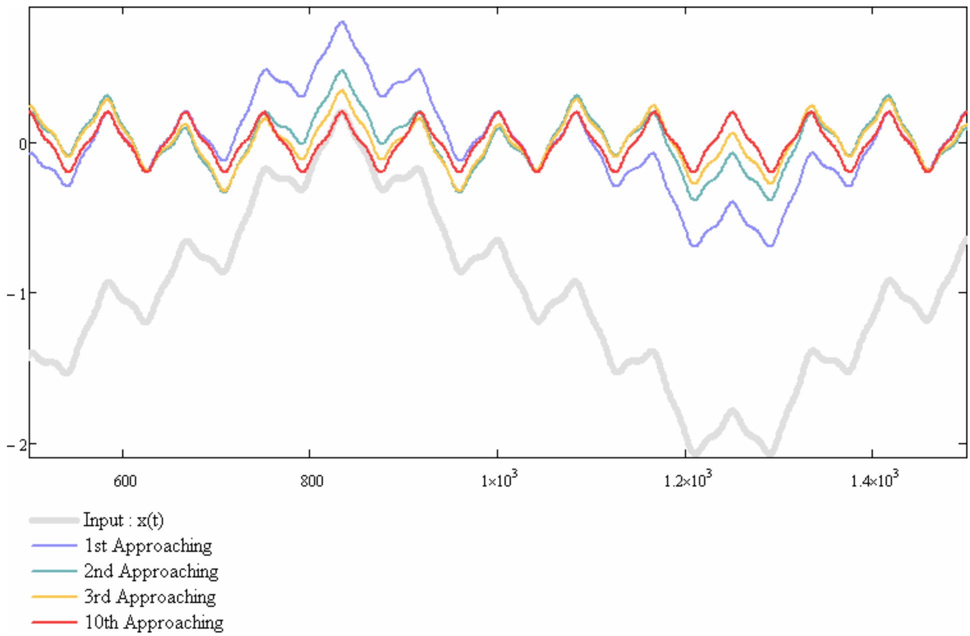


Fig. 3. The Smoothing-Sifting is the procedure to remove the smoothed mean from original data $x(t)$ recursively.

over it and the final result would not influence the low frequency smoothed envelope mean. Once the $c(t)$ is considered as symmetry according to the required stoppage criterion, the refined $c(t)$ is an IMF.

2.2. *The implement of the adaptive smoothing in SEMD*

Now, let us analyze the proposed smoothing procedure in details. Mathematically, the mean envelope, $m(t)$, given above could be written in the Fourier sense as

$$m(t) = \sum_{i=0}^{\infty} (a_i \cos \omega_i t + b_i \sin \omega_i t) \tag{5}$$

This expression is perfectly legitimate mathematically, for Fourier expansion exists for any function linear or nonlinear, stationary or nonstationary. Typically, a continuous-time sinusoid $f(t) = \cos \omega t$ sampled every ΔT seconds ($t = k\Delta T$) results in a discrete-time sinusoidal for, $f[k] = \cos \omega k \Delta T = \cos \Omega k$, where $\Omega = \omega \Delta T$. Note that the discrete-time sinusoids $\cos \Omega k$ have unique waveforms only for the value of frequencies in the range $\Omega \leq \pi$ (fundamental frequency range). Thus Eq. (5) with N sampled data can be rewritten as

$$m[k] = \sum_{i=0}^{N-1} (a_i \cos \Omega_i k + b_i \sin \Omega_i k) \tag{6}$$

The n-point box-car, with n as an odd integer, smoothing of $m[k]$ is given by

$$\bar{m}[k] = \frac{1}{n} \sum_{l=-(n-1)/2}^{(n-1)/2} m[k+l] \tag{7}$$

When we apply the smoothing procedure repeatedly on the same mean envelope signal, say R times, Eq. (6) can be generalized as

$$\begin{aligned} \bar{m}[k] &= \sum_{i=0}^{N-1} \left[\frac{1}{n} \sum_{l=-(n-1)/2}^{(n-1)/2} \cos(\Omega_i \cdot l) \right]^R \cdot (a_i \cos \Omega_i \mathbf{k} + b_i \sin \Omega_i \mathbf{k}) \\ &= \sum_{i=0}^{N-1} F(\Omega_i, n, R) \cdot (a_i \cos \Omega_i \mathbf{k} + b_i \sin \Omega_i \mathbf{k}). \end{aligned} \tag{8}$$

Compared with $m[k]$, the phase spectrum of $\bar{m}[k]$ is unchanged after smoothing, but the amplitude spectrum of $\bar{m}[k]$ has changed with its value depending on Ω, n, and R denoted as F (Ω, n, R). The F (Ω, n, R) is plotted and shown in Fig. 4. for a given window n and R times smoothing.

Obviously, the smoothed mean has "damped" amplitude with the amplitude of higher frequency suppressed. As a result, only the relative lower frequency parts

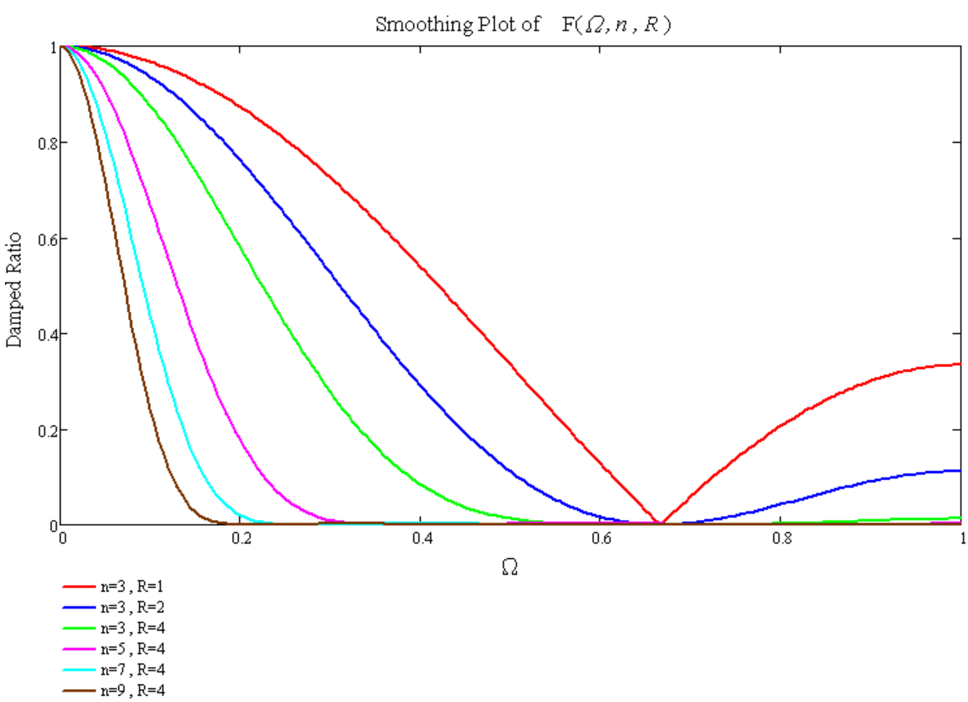


Fig. 4. The correlation between digital frequency Ω and damped ratio R at a given window n and R times smoothing.

of the envelope mean are retained and removed without distorting the phase after smoothing. Theoretically, the more smoothing you choose, the more number of times of sifting is required.

Smoothing process is designed to explore the foundation of EMD: How can an accurate mean envelope be found. Here, we have shown that adequate smoothing as proposed above could generalize the envelope definition from a cubic spline to any form and still retain the capability of extracting the identical estimation of the mean envelope. We can achieve this uniqueness property of the envelope mean because the initial envelope mean is determined directly come from the interpolation of few extrema; therefore, the lower frequency part is more reliable than high frequency part. In each sifting process, the proposed smoothing procedure here will guarantee that only the reliable part of envelope mean will be removed successively. Granted that theoretically, this smooth procedure would inevitably leave some higher harmonics out of the resulting envelope mean, if the signal is highly nonlinear with many phase locked harmonics content, the final result we obtained through this procedure however would still be as good an approximation as one would obtain through the cubic spline sifting procedure. With this smoothing procedure, however, the decomposition would be more controllable and also have definitive mathematic expressions. The most important advantage of the present approach is that adequate smoothing will make EMD independent of spline selection. In the other words, all kinds of envelope fittings will always converge to the same smoothed mean result. Furthermore, the final results would be much smoother. A minor disadvantage is that the smoothing procedure might be more time consuming than the classic EMD.

Just as in the classic EMD, there is still the analogous problem of the stoppage criterion: Here we have to determine the window size of the moving average, and the number of times we iterate the smoothing procedure. The damping function is helpful to define the adequate smoothing and written as Eq. (8). According the sifting process of EMD, the spectrum of the removed mean is limited in the quantity of found extrema. Hence, the number of extrema determines the band location Ω of the decomposed signal. When the band location Ω is given, the damped ratio (5%–50%) of the highest spectrum of the envelope mean is reliable enough to be removed. After giving the value of damped ratio, the suitable smoothing window n could be found through Eq. (9). It is very easy to search the adequate smoothing widow n by pre-designed lookup table. The smoothing window n must be an odd integer greater than 1. The minimal spacing is the smallest interval between two adjoining extrema. The example is displayed on Fig. 5.

$$\text{Damped Ratio} = \left(\frac{1}{n} \sum_{l=-(n-1)/2}^{(n-1)/2} \cos(\Omega \cdot l) \right)^R \quad \text{where } \Omega = \frac{\pi}{\text{Minimal Spacing}} \tag{9}$$

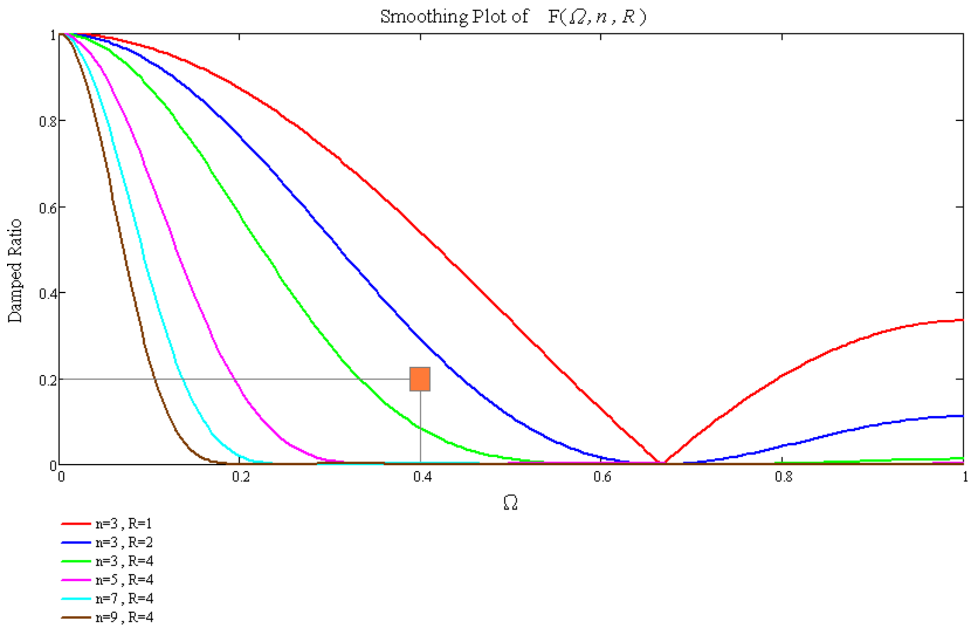


Fig. 5. For example: Trust only 20% (**Damped Ratio = 0.2**) of smoothed mean at digital frequency $\Omega = 1.257$, therefore, the curve $n = 3$ and $R = 4$ is conservatively chosen (Green line).

The Hilbert Transformation was originally based on linear system. In order to inspect the results in nonlinear and nonstationary testing, the improved Hilbert Transformation by Direct Quadrature was performed in our spectral comparison.

3. Experiments

Having explained the smoothing procedure, we will demonstrate the power of the new smoothing EMD (SEMD) with the following examples.

(1) Composite sinusoidal and nonlinear Duffing waves

First, we will test the decomposition of the composite consisted of sinusoid signals as $f(t)$ and Duffing type wave signals as $g(t)$ and to compare the performance between EMD and SEMD. Unlike the sinusoidal waves, the Duffing type wave is an intra-wave frequency modulated wave. Both test signals are a combination of three pure and stationary signals at different frequency band. The second IMF could be pure only when first IMF (high frequency) and third IMF (trend) are removed thoroughly. In order to check how precisely the estimated mean could be removed by the different procedures, the second IMF is chosen as the target for the comparison. The test signals are given as

$$f(t) = 0.5 \cdot \cos\left(\frac{600\pi}{5000}t\right) + 0.5\sin\left(\frac{120\pi}{5000}t\right) + \sin\left(\frac{\pi}{1024}t - \frac{\pi}{2}\right)$$

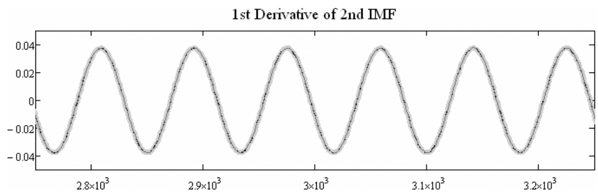
$$\begin{aligned} g(t) = 0.5 \cdot \cos&\left[\left(\frac{1200\pi}{5000}t\right) + 0.3\sin\left(\frac{2400\pi}{5000}t\right)\right] \\ &+ 0.5 \cdot \cos\left[\frac{120\pi}{5000}t + 0.3\sin\left(\frac{240\pi}{5000}t\right)\right] + \sin\left(\frac{2\pi}{5000}t + \frac{\pi}{2}\right) \\ &\qquad \text{where } 0 \leq t < 4999 \end{aligned}$$

After the decomposition by EMD and SEMD, we find both decomposed IMFs are distorted at ends by the end effects; therefore, it would be meaningless to measure the quality of the IMF through their standard error. Instead, we propose here to compare the higher order of derivatives to test the smoothness of the resulting IMF components. Granted that in the Hilbert-Huang Transform, the requirement on the smoothness is only on the derivative of the phase function to obtain the instantaneous frequency, smoothness of the IMF, nevertheless, is still a desired quality of the results.

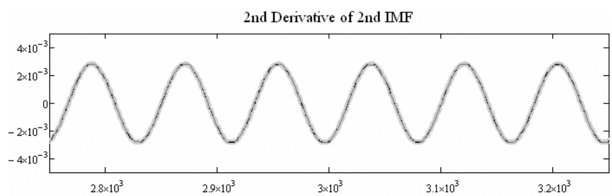
Because of the cubic spline based sifting, EMD makes all IMFs as a linear combination of cubic spline. Thus the resulting IMFs could only serve as an approximation. That is the base and the basic foundation of the EMD method: all IMFs are only constructed by several linear combinations of cubic spline no matter how many times of sifting are applied shown in Figs. 6(c) and 7(c). In contrast the IMF produced by SEMD is well converged and differentiable even at its 7th derivative as shown in Figs. 6 and 7. As only first phase function derivative of IMF is used for Hilbert spectral analysis, the decomposed error in EMD seems to be not any problem in applications. It is really hard to discriminate their differences at the first derivative by a naked eye in Figs. 6(a) and 7(a). However, their differences are so apparent in frequency domain shown in Figs. 8 and 9. As the signal $f(t)$ is consisted of sinusoidal functions, the resulting signal should be a sine wave with a constant frequency. The result given by SEMD indeed fulfilled this expectation; while the result from EMD showed a constant range of values, but the error appeared in the second significant decimal places, which could be seen in the expanded scale of the constant frequency reference. For the nonlinear wave case shown in Fig. 9, the difference is almost invisible. More detailed comparisons using ratio test perhaps could still show the differences, but for any practical purpose there is no difference at all, because the intra-wave modulations has a finite value that would mask the small differences.

(2) A saw-tooth function

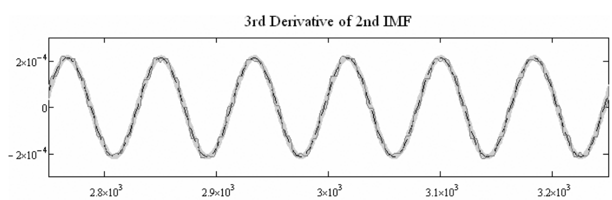
Next, we will test the signal consisted of a high and low frequency sinusoidal waves with an intermediate frequency component of saw-tooth function as shown in Figs. 10(a)–10(d).



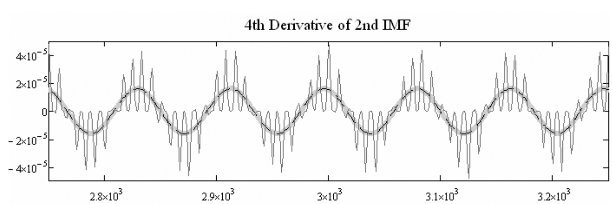
(a)



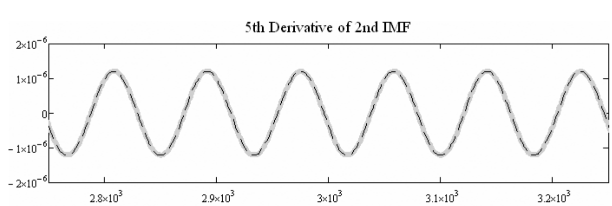
(b)



(c)

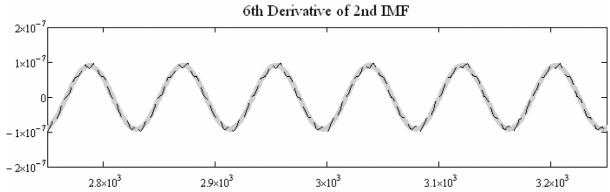


(d)

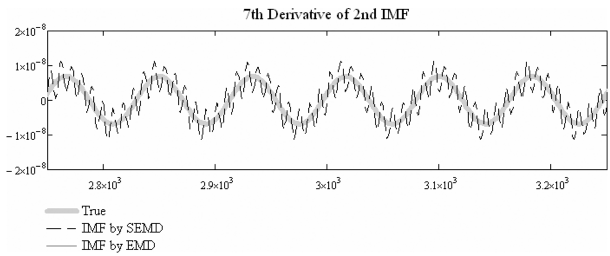


(e)

Fig. 6. Comparison of 2nd IMF in test signal $f(t)$ in different order derivative.

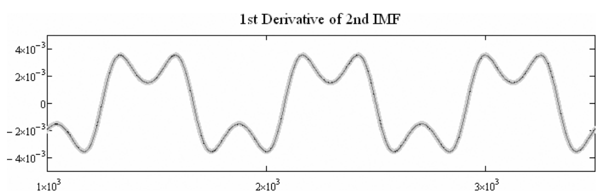


(f)

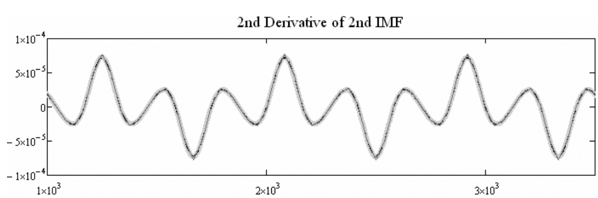


(g)

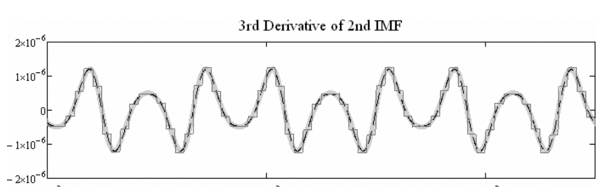
Fig. 6. (*Continued*)



(a)

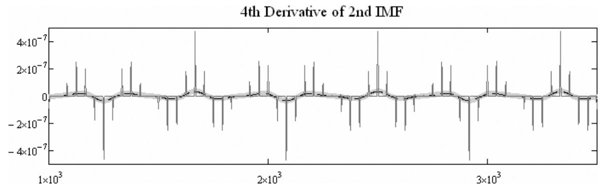


(b)

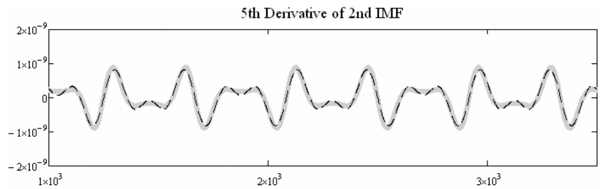


(c)

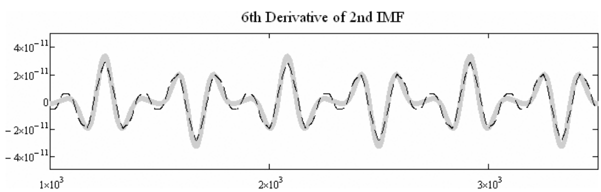
Fig. 7. Comparison of 2nd IMF in test signal $g(t)$ in different order derivative.



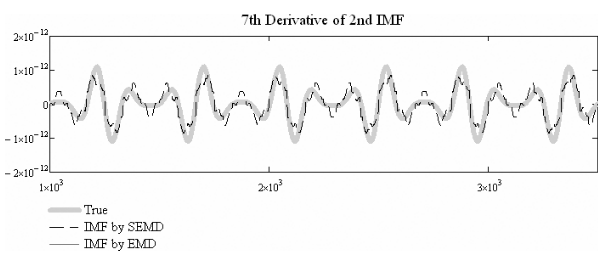
(d)



(e)



(f)



(g)

Fig. 7. (*Continued*)

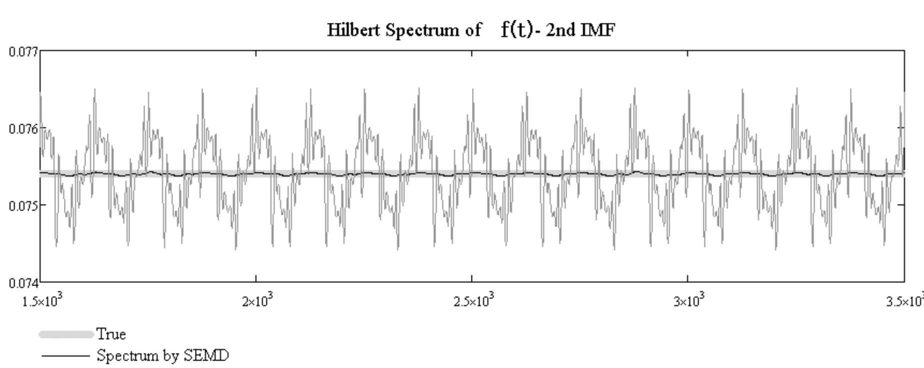


Fig. 8. The Hilbert spectrum of second IMF decomposed by SEMD and EMD in test signal $f(t)$.

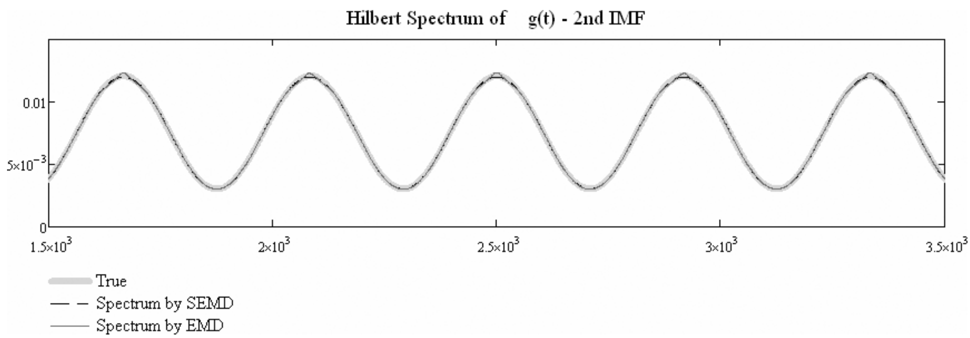


Fig. 9. The Hilbert spectrum of second IMF decomposed by SEMD and EMD in test signal $g(t)$.

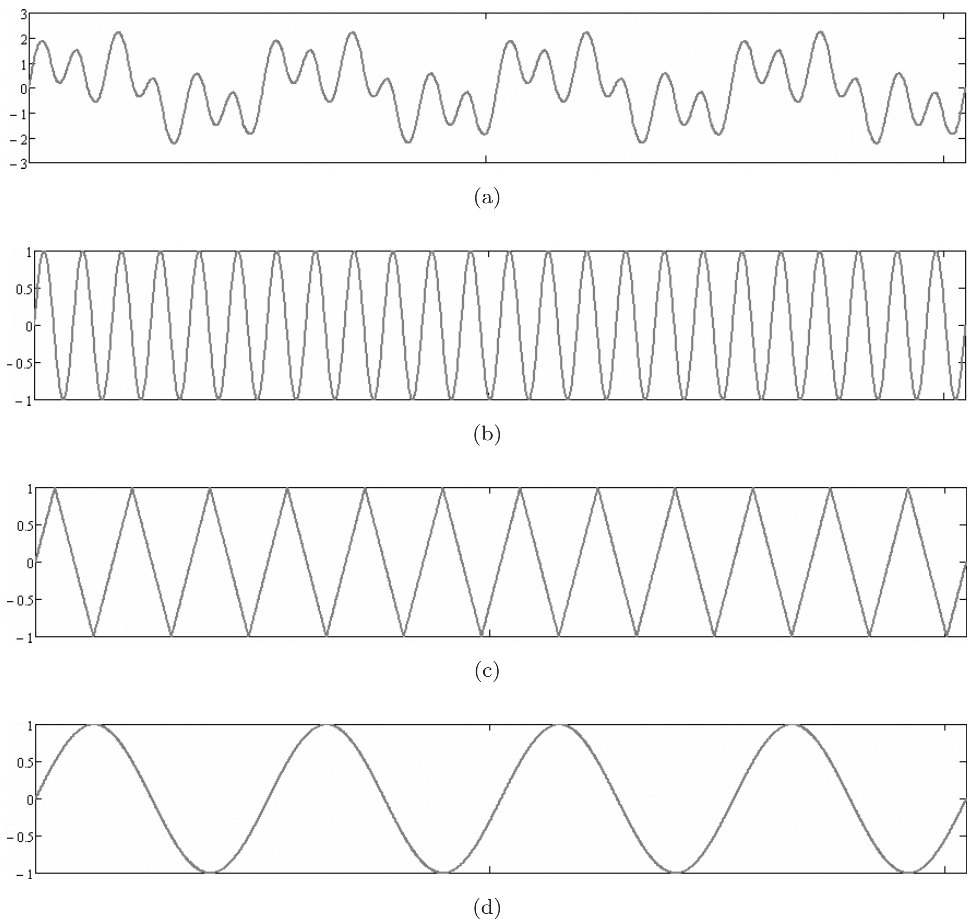


Fig. 10. (a) The testing signal consisted of a high and low frequency sinusoidal waves with an intermediate frequency component of saw-tooth function (b) the high frequency component at 24 Hz (c) saw teeth component at 12 Hz (d) low frequency part at 4 Hz.

The resulting IMFs are given in Figs. 11(a)–11(c). Clearly, the performance of EMD and SEMD are comparable for all the components, except the end effect is more severe for the linear envelope case. The comparison of the different approaches on the saw-tooth function is enlightening. None of the method actually gives the exact saw-tooth original functional form. The SEMD had obviously smoothed the sharp corners off, so is the cubic spline based EMD method. They all serve as approximation to the similar degree.

It should be pointed out that the most important achievement here is not how well the IMFs agree with the truth here. Rather, it is impressive that one could use a straight line envelope mean *ab initio*, and still have the result converges to the truth. Thus, we have proved the uniqueness of the smoothing procedure for EMD under a weak limitation of spline fitting.

(3) Length-of-Day data

Finally, we will use a set of data from a natural phenomenon, the Length-of-Day (LOD) covering 1965 to 1985. This set of data represents the true rotating speed of the Earth. Unlike artificial testing, the natural signal contains unknown noise, non-stationary and nonlinear variation produced by the interactions among the Earth, the Moon and the Sun. The LOD is well known in EEMD testing and has been studied by many of scientists for its physical meanings [Chao, 1989; Huang *et al.*, 2003; Wu and Huang, 2009]. The total signal consisted of the rotation speed of the Earth influenced by the moon and the Sun. Several hidden intermittence oscillations probably caused by natural events have made the IMFs suffering the mode mixed problem. In spite of mixing problems, this experiment shows SEMD and EMD generate the same quantity of IMF and similar Hilbert spectra as shown in Figs. 12 and 13. Because of the smoothing process invoked in SEMD, the resulting Hilbert spectrum from it is quite stable. The results from EMD and SEMD are very similar in IMFs, but the Hilbert spectrum from EMD contains noticeably more fluctuations. To examine the details, we expand a section of the result covering only 1966–1967 in Fig. 14. This component represents the monthly tidal cycle caused by the Moon; therefore, there are around 13 cycles per year. The differences between EMD and SEMD in IMFs are small. But the instantaneous frequency values are quite different: The frequency values from SEMD are near constant and smooth; while, the values from EMD show large intra-wave fluctuations. The fact that SEMD gives a near constant frequency without intra-wave modulation in the Hilbert spectrum of the monthly tide, as shown in Fig. 14, is an intriguing result. This is very similar to the result produced by Olhede and Walden (2003), when they used the wavelet packet as the decomposition tool on the chirp signal from a bat. As the wavelet is a linear tool, the decomposed components can represent only the linear properties. In all natural signals, it is well known that when one lumps all the Fourier components into a single amplitude modulated function as in the monthly tide here, there should be frequency modulations as well. Indeed, the amplitude modulated monthly tide signal contains many Fourier terms (see,

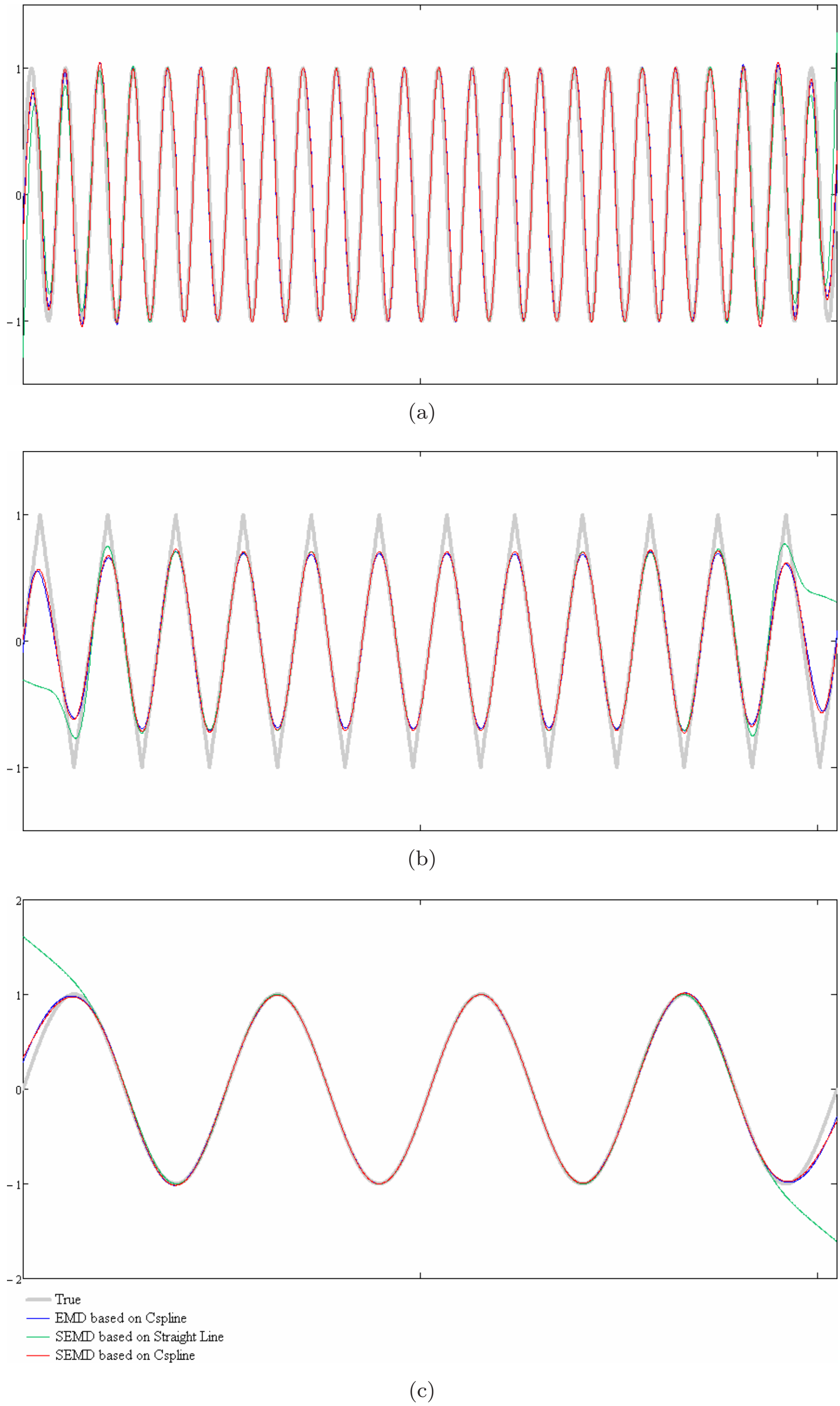


Fig. 11. Comparison between EMD and SEMD based on Cubic spline and straight line (a) The 1st sinusoidal component at 24 Hz (b) The 2nd saw teeth component at 12 Hz (c) The 3rd sinusoidal component at 4 Hz.

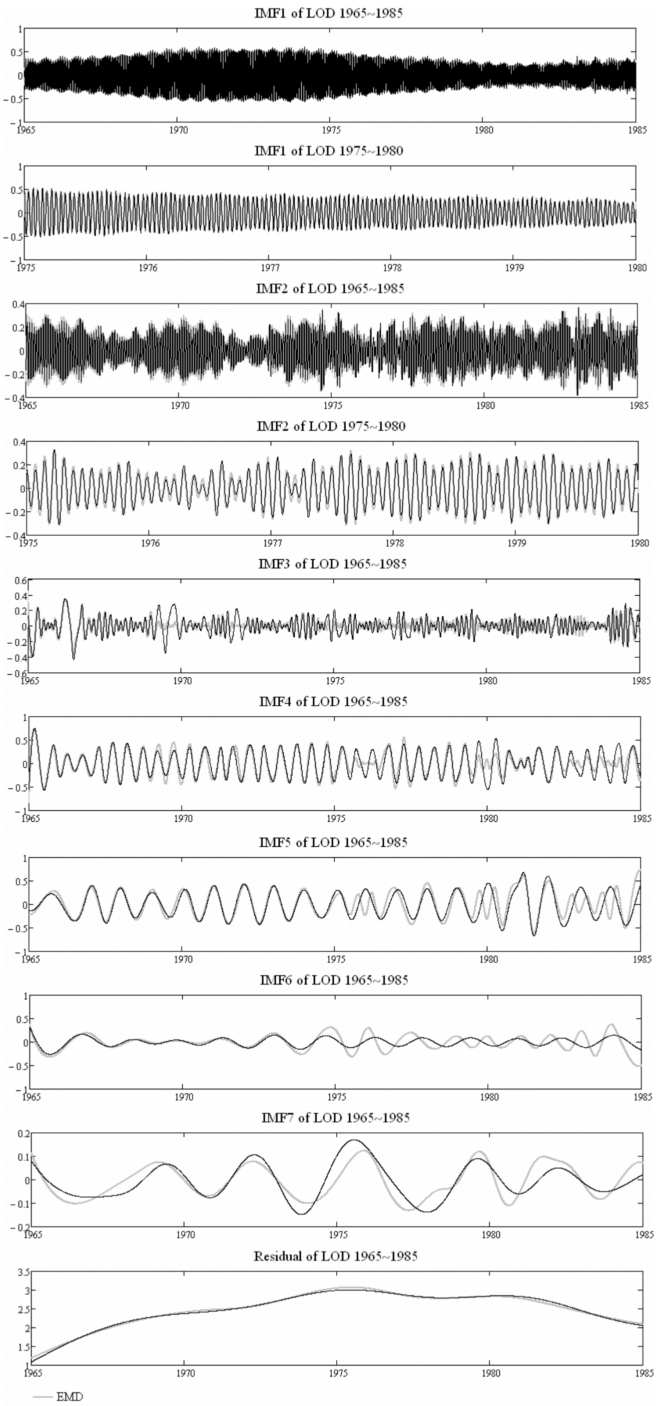


Fig. 12. The comparison of LOD-IMF decomposed by SEMD and EMD from 1965 to 1985.

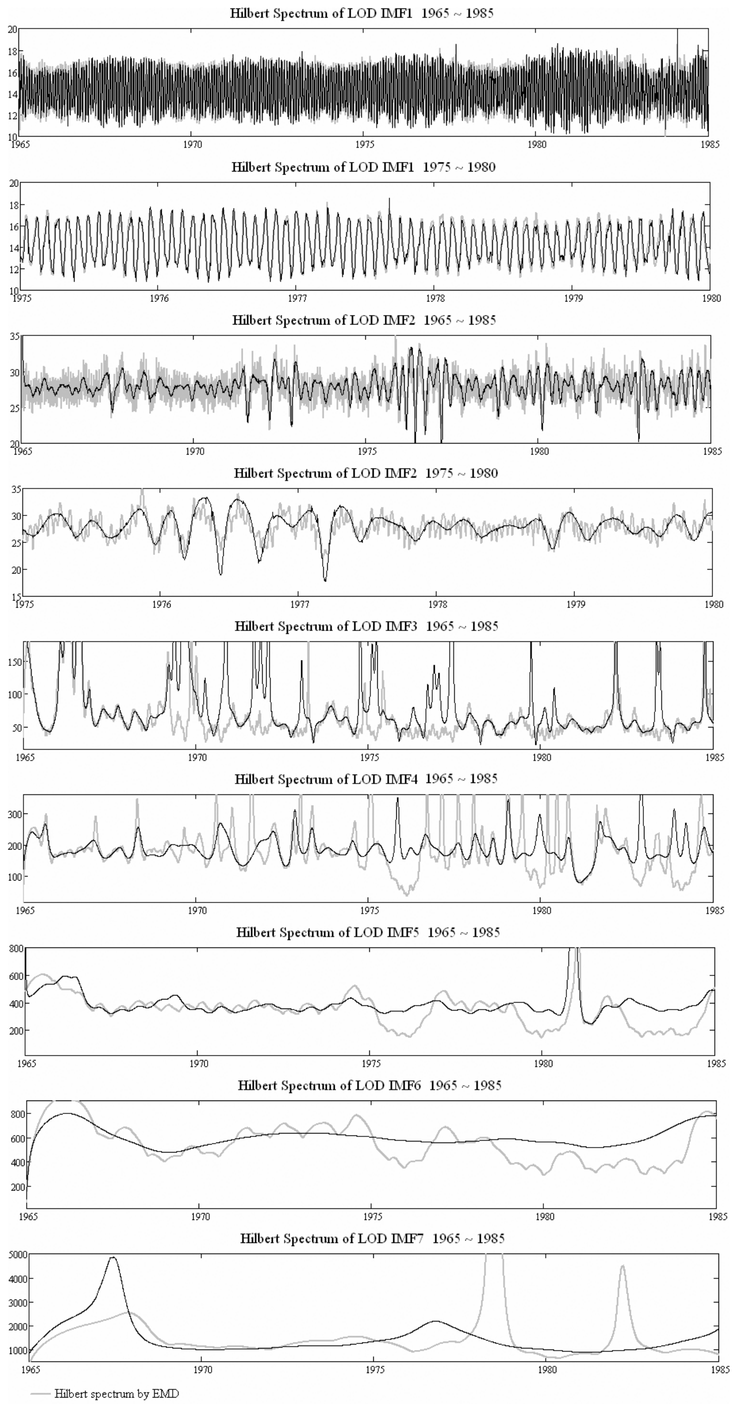


Fig. 13. The Hilbert spectrum of LOD-IMF decomposed by SEMD and EMD from 1965 to 1985.

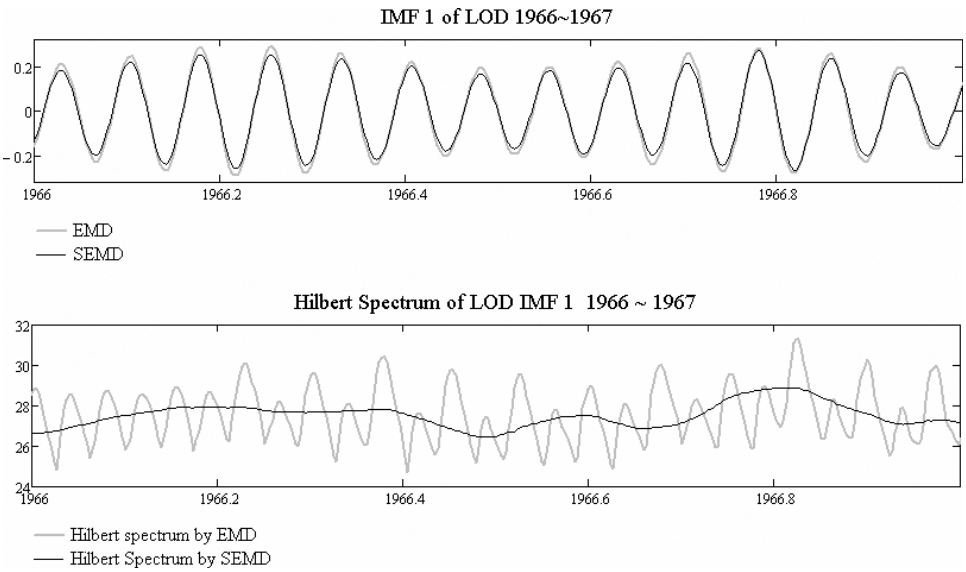


Fig. 14. The Hilbert spectrum of LOD-IMF1 decomposed by SEMD and EMD from 1966 to 1967.

for example, Cartwright, 1999). Yet, the frequency modulation is absent from the result produced by SEMD as shown in Fig. 14. This effect would be studied in more details in the future.

4. Conclusion

In most applications of HHT so far, IMFs had attracted more users than the associated Hilbert Spectral analysis. As a result, people had put more emphasis on the form of the IMFs, and the amplitude modulation in IMF has evolved as the standard for testing the results. Based on our present study, it should be pointed out that when frequency is of concern, one should be very cautious to the amplitude comparison method only. Our experiments here explain that the difference between EMD and SEMD in IMFs might be small, but the difference in the instantaneous frequency might be large. Because the smoothing is a reliable patch to eliminate the effect of chosen cubic spline as basis for the decomposition, it could produce smooth IMFs as demonstrated clearly. Whether the instantaneous frequency values produced are also correct should be investigated further as discussed in the case of LOD above. A crucial feature of SEMD is this: It really opens the window to make of chosen spline of EMD uncritical. The uniform convergence of different initial spline functions to the same final IMF sets indicates that SEMD has established the uniqueness of EMD method. Thus SEMD has made a great advances and valuable contribution to the theoretical foundation of empirical mode decomposition method.

Having made this statement, we should also point out the lingering doubt on the theoretical side: Could SEMD really give the true answer precisely? The answer is definitely no, as indicated by the saw-tooth example. Like EMD, SEMD is also an approximation; it is a super smooth one. This raises another question: is smoothness a desirable quality? We need to study this question in more details in the future. As shown in Eqs. (1) and (2), cubic spline based EMD always gives IMFs in terms of cubic spline functions. The same should also be true for the SEMD: all IMFs would consist of the smoothing function as given in Eq. (8), which is essentially the Fourier expansion of the true functions. The completeness could be proved if we have infinite many terms. With finite terms as used here, the smoothing might be an over-powering factor.

From the definition of instantaneous frequency, the EMD is one of optimal solution to solve physical problems especially in wave motion. In the past, this phenomenon is hardly described completely by filters or any basis based decomposition such as STFT and wavelet. Those conventional tools create mathematical artifacts in high frequency harmonics for nonlinear processes. If the treated signal is not understood in advance, the analyzed results would be meaningless physically. This is why we need an adaptive and non-basis analyzer to prevent from information missed in unfamiliar signals.

EMD is an adaptive tool to separate an input into finite frequency and amplitude modulations. Generally speaking, EMD is a kind of numerical solution rather than an exact solution in decomposition. After decomposition, all IMFs admit well behaved Hilbert transformation with more information in power and frequency variations. As claimed by Huang *et al.* (1998), EMD could be applied to nonlinear signal without prior basis. Hypothetically, these spline basis used in EMD are assumed to be averaged out or vanished when the sifting is applied recursively. The final answer is always an approximation. Whatever errors, produced in the computation steps, would be more obvious in frequency modulation rather than amplitude modulation. This is reason why the Hilbert spectrum in each IMF is noisier. Certainly, computed errors could be ignored if you are only interest in the application of amplitude variation. Even so, the refined IMF is necessary to prevent from accumulated errors which may cause serious distortion in low frequency modes.

SEMD is also spline based, but with the added novel smoothing procedure it is no longer dependent on the specific spline functional form. The envelopes used in sifting could be constructed by Cubic Spline passing through all extrema of the input data. These envelopes could also be replaced by piecewise straight lines as initial guess too. During sifting, the smoothed window is given adaptively by the minimal spacing between extrema which represents the highest digital frequency of the signal. The part of smoothed mean is removed in each sifting can be well controlled by damped ratio of the chosen frequency. These smoothed effects can be proven by Fourier analysis theoretically. There is one nagging problem: the end effect. The EMD and SEMD all decompose the signal into the IMFs from higher to lower frequency irreversibly. Any high frequency ripple caused by poor estimation at ends would contaminate the lower frequency components. In the smoothing procedure, we need to extend the data outside the range in order to obtain a mean for the points near the boundary. To minimize the end effects, we use mirror boundary to replace the repeated boundary and make sure the extended boundary is differentiable at ends and contains the same bandwidth with treated signal. To sum up, the concept of SEMD originally come from theoretical analysis, but the implement of SEMD is still quite empirical.

SEMD is an attempt to eliminate the drawback of EMD. The smoothing concept in SEMD is tested out in both artificial signals and natural dataset. In this paper, SEMD could be seen as an adaptive filter and constrained to admit Hilbert transformation. Theoretically, the stronger smoothing one chooses the more efforts and time one would have to use. Practically, the accuracy of results would meet its limit because of numerical computation. In spite of the time consuming sifting, SEMD almost spends similar computational time to EMD to meet required stoppage criteria in our experiments. Stoppage criterion in SEMD could be asked as strict as possible to avoid under decomposition. The sifting number in SEMD is advised less than 30 times for computational efficiency. However, SEMD sometimes still suffers from over decomposition. The tolerance of over decomposition is independent on chosen spline as initial guess in each sifting, but better chosen spline would save smoothing time and efforts. The rule of optimal decomposition is still studied further; so far, the cubic spline seems to be the optimal among all the spline functions tried so far.

Because of smoothing approach without prior basis and uniform convergence, the Hilbert spectrum generated by SEMD looks similar to the optimization of EMD in our experiments. It seems that the smoothing is a more robust and accurate approach to implement the decomposition. And more importantly, the smoothing EMD has made a big step toward establishing the uniqueness of the empirical mode decomposition method.

Acknowledgments

N. E. Huang was supported by a grant from Federal Highway Administration, DTFH61-08-00028, and the grants NSC 98-2627-B-008-004 (Biology) and NSC 98-2611-M-008-004 (Geophysical) from the National Science Council, and finally a grant from NCU 965941 that have made the conclusion of this study possible. He is also supported by a KT Lee endowed Chair at NCU.

References

Bhuiyan, S. M. A., Attoh-Okine, N. O., Barner, K. E., Ayenu-Prah, A. Y. and Adhami, R. R. (2009). Bidimensional empirical mode decomposition using various interpretation techniques. *Adv. Adaptive Data Anal.*, **1**: 309–338.

Chao, B. F. (1989). Length-of-day variations caused by El Niño-southern oscillation and quasi-biennial oscillation. *Science*, **243**: 923–925.

Cartwright, D. (1999). *Tides, A Scientific History.* p. 292, Cambridge University Press, Cambridge, UK.

Chen, Q., Huang, N., Riemenschneider, S. and Xu, Y. (2006). A B-spline approach for empirical mode decompositions. *Adv. Computational Math.*, **24**: 171–195.

Deering, R. and Kaiser, J. F. (2005). The use of masking signal to improve empirical mode decomposition. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing — Proceedings*, pp. IV485–IV488.

Hou, T. Y., Yan, M. P. and Wu, Z. (2009). A variant of the EMD method for multi-scale data. *Adv. Adaptive Data Anal.*, **1**: 483–516.

Huang, N. E., Wu, M. L., Long, S. R., Shen, S. S., Qu, W. D., Gloersen, P. and Fan, K. L. (2003). A confidence limit for the empirical mode decomposition and hilbert spectral analysis. *Proc. Roy. Soc. Lond.*, **A459**: 2317–2345.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis. *Proc. Roy. Soc. Lond.*, **A454**: 903–995.

Huang, N. E. and Attoh-Okine, N. O. (2005). *The Hilbert-Huang Transform in Engineering*, p. 313, CRC Press, Boca Raton.

Huang, N. E., Shen, Z. and Long, S. R. (1999). A new view of nonlinear water waves: The Hilbert spectrum. *Ann. Rev. Fluid Mech.*, **31**: 417–457.

Huang, N. E., Wu, Z., Long, S. R., Arnold, K. C., Chen, X. and Blank, K. (2009). On instantaneous frequency. *Adv. Adaptive Data Anal.*, **1**: 177–229.

Lin, L., Wang, Y. and Zhou, H. (2009). Iterative filtering as an alternative algorithm for empirical mode decomposition. *Adv. Adaptive Data Anal.*, **1**: 543–560.

Liu, Z., Wang, H. and Peng, S. (2004). Texture classification through directional empirical mode decomposition. *Proceedings of the 17th IEEE International Conference on Pattern Recognition* (ICPR '04), Vol. 4, pp. 803–806, Cambridge, UK, August 2004.

Nunes, J. C., Bouaoune, Y., Delechelle, E., Niang, O. and Bunel, P. (2003a). Image analysis by bidmensional empirical mode decomposition. *Image and Vision Computing*, **21**: 1019–1026.

Nunes, J. C., Niang, O., Bouaoune, Y., Delechelle, E. and Bunel, P. (2003b): Bidimensional empirical mode decomposition modified for texture analysis, *Image Analysis, Proceedings*, **2749**: 171–177.

Nunes, J. C., S. Guyot and E. Delechelle, (2005). Texture analysis based on local analysis of the bidimensional empirical mode decomposition. *Machine Vision and Applications*, 16, pp. 177–188.

Nunes, J. C. and Delechelle, E. (2009). Empirical mode decomposition applications on signal and image processing. *Adv. Adaptive Data Anal.*, **1**: 125–175.

Olhede, S. and Walden, A. T. (2004). The Hilbert spectrum via wavelet projections. *Proc. Roy. Soc. Lond.*, **A460**: 955–975.

Qin, S. R. (2006). A new envelope algorithm of Hilbert–Huang transform. *Mechanical Systems and Signal Processing*, **20**: 1941–1952.

Riemanschneider, S., Liu, B., Xu, Y. and Huang, N. E. (2005). B-Spline based empirical mode decomposition. *Hilbert-Huang Transform and Its Applications*, eds. N. E. Huang and S. S. P. Shen, World Scientific, pp. 27–55.

Wang, G., Chen, X. Y., Qiao, F. L., Wu, Z. and Huang, N. E. (2010). On intrinsic mode function. *Adv. Adaptive Data Anal.*, **2**: 277–293.

Wu, Z. and Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise assisted data analysis method. *Adv. Adaptive Data Anal.*, **1**: 1–41.

Wu, Z., Huang, N. E. and Chen, X. (2009). The multi-dimensional ensemble empirical mode decomposition method. *Adv. Adaptive Data Anal.*, **1**: 339–372.

Xu, Guanlei, Wang, Xiaotong, Xu and Xiaogang (2007). Neighborhood limited empirical mode decomposition and application in image processing. *Proceedings of the 4th International Conference on Image and Graphics*, ICIG 2007, 149–154.

Yuan, Y., Jin, M., Song, P. and Zhang, J. (2009). Empiric and dynamic detection of the sea bottom topography from synthetic aperture radar image. *Adv. Adaptive Data Anal.*, **1**: 243–263.