Yao WANG
Ming J ZUO
Yaguo LEI
Xianfeng FAN

POPRAWA APROKSYMACJI LOKALNEJ ŚREDNIEJ W ROZKŁADZIE NA MODY EMPIRYCZNE DLA CELÓW DETEKCJI USZKODZEŃ PRZEKŁADNI

IMPROVEMENT OF LOCAL MEAN APPROXIMATION IN EMPIRICAL MODE DECOMPOSITION FOR GEAR FAULT DETECTION

Rozkład na mody empiryczne (EMD) to adaptacyjna metoda przetwarzania sygnału w połączonej dziedzinie czasu i częstotliwości, która jest całkowicie sterowana przez same dane. Metody interpolacji funkcjami sklejanymi trzeciego stopnia (cubic spline interpolation) używa się do aproksymacji średniej lokalnej w procesie przesiewu EMD. Niniejsza praca bada podejścia do poprawy aproksymacji średniej lokalnej w celu otrzymania lepszych charakterystyk EMD. Do aproksymacji średniej wartości obwiedni (envelope mean approximation) zastosowano metodę zmodyfikowanej monotonicznej interpolacji Hermite'a funkcjami sklejanymi (modified monotone piecewise Hermite interpolation, MMPHI), jako że wykazuje ona przewagę nad metodą funkcji sklejanych trzeciego stopnia. Zbadano również jeden z typów bezpośredniej aproksymacji lokalnej średniej, tzw. podejście okienkowanej średniej lokalnej (windowed local mean, WLM), i pokazano jego zalety w wykrywaniu impulsów.

Słowa kluczowe: *Rozkład na mody empiryczne, interpolacja Hermite'a, funkcja sklejana 3-ego stopnia, aproksymacja średniej wartości obwiedni, okienkowana średnia lokalna, detekcja uszkodzeń przekładni.*

Empirical mode decomposition (EMD) is an adaptive time-frequency domain signal processing method that is completely driven by data itself. The cubic spline interpolation method has been used to approximate the local mean in the sifting process of EMD. This study explores approaches to improve local mean approximation to obtain better EMD performance. A modified monotone piecewise Hermite interpolation (MMPHI) method is applied to envelope mean approximation, because it demonstrates advantages over the cubic spline method. A type of direct approximation of the local mean, i.e., the windowed local mean (WLM) approach, is also investigated and its merit in identifying impulses is demonstrated.

Keywords: *Empirical mode decomposition, hermite interpolation, cubic spline, envelope mean approximation, windowed local mean, gear fault detection.*

1. Introduction

Research in machine condition monitoring and maintenance [6, 10-12] requires a comprehensive understanding of advanced signal processing methods. Signal processing plays a very important role in fault detection of machinery since it bridges the gap between the collected physical signals and the signatures of faulty conditions.

To analyze signals adaptively, a new approach was recently introduced [7] where a signal was written as a finite sum of intrinsic mode functions (IMFs). The decomposition method to obtain IMFs is called empirical mode decomposition (EMD). Huang *et al.* proposed a general approach which requires two steps in analyzing the data [7]: (1) decompose the data into a number of IMFs, thus expanding the data on a basis derived from itself; (2) apply Hilbert transform to the decomposed IMFs and construct an energy-frequency-time distribution, designated as the Hilbert spectrum. In other words, this method uses the instantaneous frequency and energy rather than the global frequency and energy defined by Fourier spectral analysis. Physically, the necessary conditions for defining a meaningful instantaneous frequency are that the functions are symmetric with respect to the local zero mean, and have the same number of zero crossing as the number of extrema. As a result, an IMF is defined by two conditions [7]: (1) in the whole data set, the number of extrema and the number of zero crossings must either be equal or differ at most by one; (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. Unfortunately, at any given time, the data may involve more than one oscillatory mode so that we have to decompose the data into IMFs.

A systematic way of extracting oscillatory modes is called a sifting process. By virtue of the IMF definition, the decomposition method can simply use the envelopes defined by the local maxima and the local minima separately. Once the extrema are identified, all the local maxima, $E_{max}(t)$, are connected by a cubic spline line as an upper envelope, and all the local minima, $E_{min}(t)$, are connected by a cubic spline line as a lower envelope. Their mean is denoted as m_{11}, and the difference between a data $x(t)$ and m_{11} is the first component, h_{11}, i.e.,

$$x(t) - m_{11} = h_{11} \quad (1)$$

Here, m_{jk} and h_{jk} represent variables obtained at the jth decomposition level and the kth iteration operation, respectively. Ideally, h_{11} should be an IMF, because the construction of h_{11} described above seems to have been made to satisfy all the requirements of IMF. In reality, however, the cubic spline interpolation can generate new extrema, and shift or exaggerate existing ones. The purposes of the sifting process are to eliminate riding waves and make the wave profiles more symmetric. Toward this end, the sifting process has to be repeated a number of times. In the next sifting process, h_{11} will be treated as a new signal, then

$$h_{11} - m_{12} = h_{12} \quad (2)$$

where m_{12} is the mean of the upper and the lower envelopes of h_{11}. One can repeat this sifting procedure k times, until h_{1k} is an IMF, that is

$$h_{1(k-1)} - m_{1k} = h_{1k} \quad (3)$$

and the result, the first IMF from the signal, is denoted as

$$c_1 = h_{1k} \quad (4)$$

To guarantee that the IMFs retain enough physical sense of both amplitude and frequency modulation, Huang *et al.* gave a sifting process stop criterion [7] by limiting the size of the standard deviation, *SD*, computed from two consecutive sifting results as

$$SD = \sum_{t=0}^{T} \left[\frac{\left| h_{1(k-1)}(t) - h_{1k}(t) \right|^2}{h_{1(k-1)}^2(t)} \right] \quad (5)$$

A typical value for *SD* can be set between 0.2 and 0.3 [1].

Overall, c_1 should contain the finest scale or the shortest period component of $x(t)$. One can separate c_1 from the rest of the data by

$$x(t) - c_1 = r_1 \quad (6)$$

Since residue r_1 may still contain information of longer period components, it is treated as a new original data and subjected to the same sifting process as described above. Eventually, we obtain $c_1, c_2, \ldots, c_n$.

The whole sifting process can be stopped by any of the following predetermined criteria: when component c_n becomes so small that it is less than the predetermined value of substantial significance, or when residue r_n becomes a monotonic function from which no more IMFs can be extracted. Finally we have decomposed the original signal $x(t)$ into

$$x(t) = \sum_{i=1}^{n} c_i + r_n \quad (7)$$

Thus, the data has been decomposed into n empirical modes, and a residue which can be either the mean trend or a constant.

Although the effectiveness of this technique has been demonstrated in many applications (for example, see [2, 13, 18]), EMD itself is not perfect. Even the authors of the original paper have stated in [7], "At any rate, improving the spline fitting is absolutely necessary." Possible improvements of the envelope mean approximation will be investigated in this paper. Section 2 points out problems of two reported interpolation methods used for envelope mean approximation. An alternative interpolation, i.e., the modified monotone piecewise Hermite interpolation (MMPHI), will be investigated in Section 3. Section 4 proposes a direct approximation using windowed local mean (WLM) method which has better performance in identifying impulses. Conclusions are given at the end.

2. Issues of reported interpolation methods in envelope mean approximation

Envelope mean defined by Huang *et al.* is used to force local symmetry instead of local mean [7]. This is a necessary approximation to avoid the definition of a local averaging time scale for non-stationary signal analysis. This type of approximation is called an envelope mean approximation. The cubic spline interpolation was used to construct the upper and lower envelopes based on extrema. This interpolation is a type of Hermite cubic interpolation that requires continuity of both the first and the second derivatives at knots (extrema) [8]. A "Not-a-Knot" cubic spline interpolation is actually used in [7] that provides the third derivative continuity at the second to the first point and the second to the last point to obtain a unique solution.

For a given data set, however, we may not know any features of its upper or lower envelopes. As a result, we prefer to have the interpolation to represent the shapes of the envelopes as they are, i.e., to avoid the imposition of any additional details that are not confirmed by the data [1]. Either undershooting or overshooting may occur due to the interpolation [7]. This is because the quality of approximation of the local mean affects the iteration times of obtaining IMFs. A bad approximation will slow down the speed of decomposition. In addition, the purpose of the decomposition is to identify IMFs before conducting Hilbert transform or other subsequent analysis. Low quality of the approximation will cause more serious asymmetry of IMFs that makes the results from subsequent processes less meaningful. Huang *et al.* suggested using high-order spline in [7].

3. Application of the MMPHI to envelope mean approximation

Monotonicity is one of the most important shape properties to preserve in order to have a good-looking interpolation [9]. The monotone piecewise Hermite interpolation is a modified version of the piecewise Hermite cubic interpolation that preserves the monotonicity of the data [4].

3.1. Modified Monotone Piecewise Hermite Interpolation (MMPHI)

The piecewise Hermite cubic interpolation is a type of interpolation that puts constraints on only the first derivative at known knots, i.e.,

$$f'(x_i) = p'_{i-1}(x_i) = p'_i(x_i) = d_i \quad (8)$$

where x_i, f, p_i, and d_i are knot i, interpolating function, piecewise cubic function in the interval $[x_i, x_{i+1}]$, and the derivative value at knot i, respectively. This interpolation is a general base and is open to provision of various conditions that could give a unique solution. Studies have been reported to provide alternative conditions for the piecewise Hermite interpolation to maintain data monotonicity. Fritsch et al. derived necessary and sufficient conditions for a cubic function to be monotonic in an interval and proposed a new algorithm [3, 4]. A function, G, is constructed such that

$$d_i = G(\Delta_{i-1}, \Delta_i), \qquad i = 2, \cdots, n-1 \quad (9)$$

where $\Delta_i = (g_{i+1} - g_i)/h_i$ is the slope of the line segment joining the data to be interpolated and $h_i = x_{i+1} - x_i$. A specific form of the G function is given in [4]. The key difference between the modified monotone piecewise Hermite interpolation (MMPHI) and the cubic spline interpolation is that MMPHI releases the constraints of the continuity of the second derivative, while adding constraints related to monotonicity to have a unique piecewise cubic interpolating function.

3.2. Motivation of the study

The purpose of the interpolation in EMD is to show the behaviors of the envelopes of the data. If an obvious feature is present, the interpolating function is expected to represent it in a suitable way [1]. Since monotonicity is a key shape property of a curve, we want to preserve monotonicity of the data in our study. The monotone piecewise Hermite interpolation is not the only method that considers monotonicity. Wolberg [17] described methods that minimize the second derivative discontinuity while the algorithm of [4] guarantees continuity of the first derivative. But Wolberg's methods involve linear and quadratic programming which consume a great deal of time. The linear piecewise interpolation is the simplest way to keep monotonicity; but, with this approach, a smooth curve's upper and lower envelopes, and therefore its envelope mean, can become too sharp, losing its smoothness. Since the interpolation of EMD is based on extrema which do not provide any information on the continuity of the second derivative, only continuity of the first derivative is considered for our application. Therefore, our motivation for applying MMPHI to the envelope mean approximation is to maintain monotonicity of the data and guarantees continuity of the first derivative without increasing computational complexity.

3.3. Using MMPHI for envelope mean approximation

To see a more accurate decomposition of an original signal, we introduce MMPHI to the envelope mean approximation instead of the cubic spline or high-order spline interpolations. This is because MMPHI has less undershooting or overshooting problems so that the local mean being approximated is more accurate. We also expect MMPHI to have an advantage over the cubic spline interpolation with regard to calculation time. It has been reported that doing one MMPHI interpolation needs $O(3n)$ arithmetic operations while doing one cubic spline interpolation needs $O(11n)$ arithmetic operations [16]. Obviously, the total calculation time of a type of EMD is also dependent on the number of iterations of the sifting process and the number of IMFs obtained. With a more accurate envelope mean approximation, we will have a fewer number of iterations. Therefore, the total calculation time is expected to be shortened as well.

Before testing our proposed methods, two issues of EMD should be considered: stopping criteria and end point swings. After a comparison of reported approaches, we choose the partial criterion [15] and the mirroring approach [19] to deal with these two issues, respectively [16]. Now, we can compare the proposed EMD that constructs envelope means by MMPHI to those by cubic spline interpolation and by high-order spline interpolation (HOSI). Here the fourth order is used. Two types of simulated signals are used for comparisons. The first one is a combination of multiple sinusoid waves with different frequencies and amplitudes plus a global trend, as given below,

$$x(t) = 0.5\sin(2\pi 100t) + 2\sin(2\pi 20t) + \\ + \sum_{\Omega=6,10,15,25} \sin(2\pi 50t/\Omega) + 0.0005t \qquad (10)$$

This signal will help us assess each method's decomposition capability. This signal is analyzed using Matlab 7.0 on a personal computer with Intel Core 2 Duo CPU 2.67 GHz. Figs. 1 (a), (b), (c) show the IMFs of the 3000-point data by MMPHI, cubic spline, and HOSI, respectively. A preliminary assessment is given based on the visual observation of the IMFs.

It can be seen from Fig. 1 that (1) MMPHI generates more IMFs than the other two methods; (2) the first three sinusoidal components are decomposed in all approaches; (3) although the cubic spline interpolation distorts the rest of the components a little worse than does MMPHI, the difference is not very obvious through visual observation; and (4) HOSI has different frequencies mixed up in IMF4. The averaged mean square error (MSE) is used as a performance indicator in this study. MSEs between

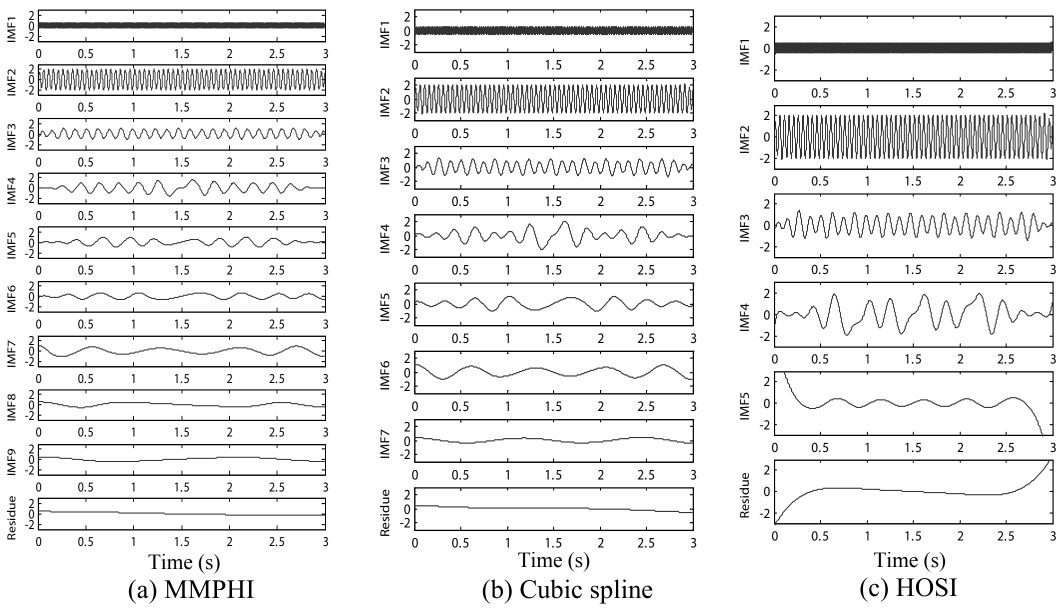


Fig. 1. Decomposition of the 3000-point signal of x(t) in Eq. (10)

decomposed IMFs and the corresponding frequency components are calculated by

$$MSE_i = \frac{1}{L}\sum_{j=1}^{L}[IMF_i(t_j) - x_i(t_j)]^2, \quad i = 1, 2, \cdots, m \quad (11)$$

where L, m, $IMF_i(t_j)$,and $x_i(t_j)$ are the data length, the number of known components, IMF i, and the corresponding signal component, respectively. The averaged MSE of the decomposition is expressed as

$$MSE_{avg} = \frac{1}{m}\sum_{i=1}^{m} MSE_i \quad (12)$$

The calculation time and MSE_{avg} for the data are shown in Table 1. We can see that MMPHI is slightly slower than the cubic spline algorithm. We also see that the proposed MMPHI approach is better than the cubic spline and the HOSI algorithm in terms of the averaged MSE because its MSE_{avg} is the smallest one in Table 1.

Tab. 1. Comparison of the performance of the proposed method and
other reported methods on the 3000-point signal of x(t)in Eq.
(10).

Interpolation methods:	Cubic spline	MMPHI	HOSI
Calculation time (s):	1.3440	1.4533	1.5084
MSE_{avg}:	0.2414	**0.1530**	0.3126

Another simulated signal is a combination of a periodic impulse signal and a chirp signal. Each impulse can be expressed as

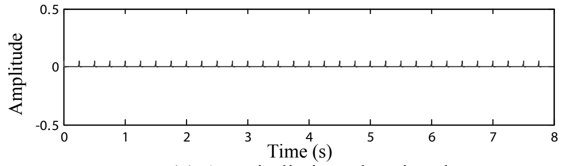
$$y_i(t) = 0.1e^{-100t}\sin(1000t) \quad (13)$$

The time interval between every two adjacent impulses is 0.25 seconds. The chirp signal can be expressed

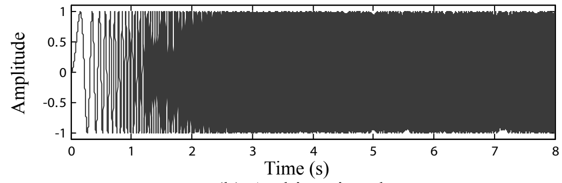
$$y_c(t) = \sin(100\pi t^2) \quad (14)$$

The signal is shown in Fig. 2. The mixing ratio between the impulses and the chirp signal is 1:1. It is hard to identify impulses in the mixed signal in Fig. 2 (c). Because periodic impulses may represent fault signatures of a machine, this simulated signal is used to validate the proposed method to detect periodic impulses.

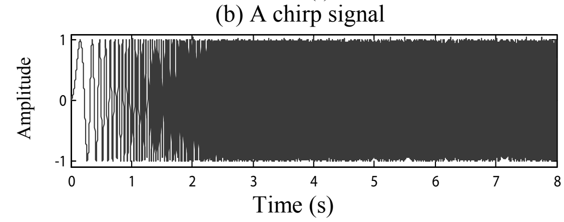
We tested a data series with 3000 data points and set the time interval of any two adjacent points at 0.001 seconds. Fig. 3 shows the IMFs of the 3000-point data. It can be seen that there is not much difference based on visual observation among the three decompositions. The chirp signal stays mainly in the first IMF and none of the three approaches give clear information of the impulses. The averaged MSE is still used as an indicator of accuracy but this time $m = 2$ in Eq. (12). The calculation time and the MSE_{avg} for the 3000 point data series are shown in Table 2. The



(a) A periodic impulse signal

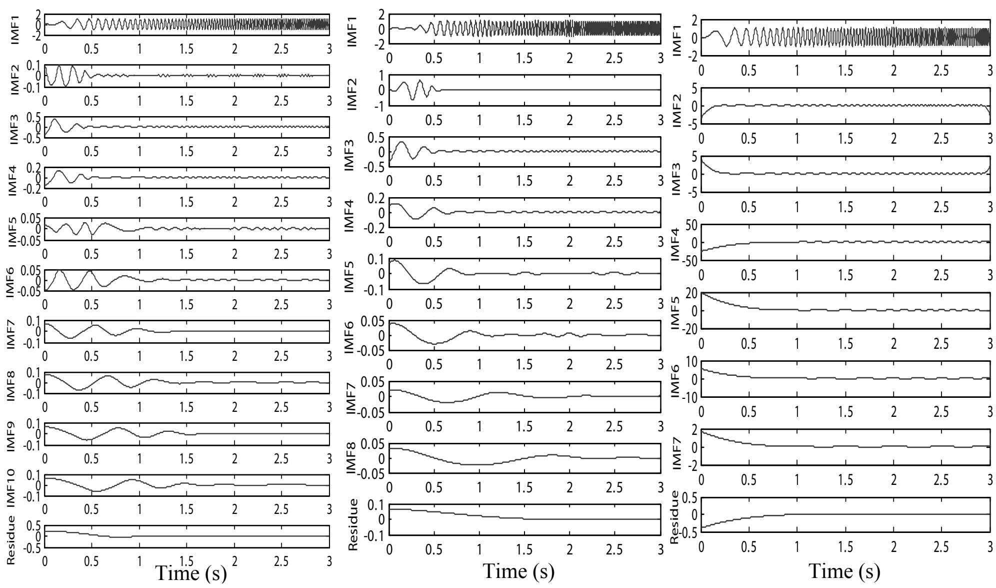


(b) A chirp signal



(c) The combination of signals in Figs. 2 (a) and (b)

Fig. 2. A periodic impulse signal, a chirp signal, and their combination



(a) MMPHI (b) Cubic spline (c) HOPI

Fig. 3. Decomposition of the 3000-point signal in Fig. 2 (c)

proposed MMPHI approach is the best with regard to calculation time and accuracy. The HOPI approach is time consuming.

Tab. 2. Comparison of the performance of the proposed method and other reported methods on the 3000-point signal in Fig. 2 (c).

Interpolation methods:	Cubic spline	MMPHI	HOSI
Calculation time (s):	10.073	3.6877	264.6530
MSE_{avg}:	0.0338	0.0099	0.1080

4. Direct approximation using Windowed Local Mean (WLM)

It should be noted that all the three interpolation approaches investigated in Section 3 cannot identify impulses clearly (See Fig. 3). Essentially, they all tend to use envelope means to approximate local means. The true local mean is supposed to be the mean of a signal so it should reflect the global trend of the signal. We also expect that some regional signatures of the original signal should appear on the curve of the local mean. These two requirements will ensure that the difference between the original signal and the local mean converges on an IMF. Based on the algorithm of the envelope mean approximation, values for the points between extrema are really not important if they do not exceed their neighboring extrema; therefore, for two signals with the same extrema positions and values but quite different points between the extrema, the means obtained by the envelope mean approach will be the same. This is proof that envelope means do not approximate local means very well. Therefore, direct approximation that does not require calculations of envelopes is adapted as an alternative to envelope mean approximation in this section. Gai *et al.* introduced a method for approximating local means called Local Mean Mode Decomposition (LMMD) [5]. LMMD uses a time-varying filter to calculate the means. But LMMD still uses the cubic spline interpolation which would result in the same shortcoming encountered with the envelope mean approximation and no effort was done to support the statement that LMMD is faster than EMD. We will introduce an approximation using WLM and show its advantage in the following subsections.

4.1. Windowed local mean (WLM)

RÄosler [14] considered the mean of a signal within a window and called this the local mean with respect to that window. We termed this type of direct approximation as windowed local mean (WLM). For a continuous function, $y = f(x)$, the expression of its windowed local mean is

$$m_\delta[f](x) = \frac{1}{\delta}\int_{x_\delta} f(\tau)d\tau \qquad (15)$$

where $x_\delta = [x - \delta/2, x + \delta/2]$, and δ is the width of the integration window. The purpose of giving such an expression is not for using it for EMD but just to introduce the WLM to interpret the concept of "local mean" mentioned in [7].

Our purpose is to find an improved direct approximation with a better capacity for identifying impulses. Such an approximation should consider the relatively macro view of a signal but the consideration has to be restricted to be within relatively micro windows to avoid losing local signatures. We can see that this method does have such a nice feature. An expression of the WLM for discrete data points is not given in [14] but is needed to deal with real signals and to process them on a computer. We define here for a set of data, x_i, $i = 1,2,\cdots,n$, the discrete form of its WLM as

$$m_\delta(i) = \frac{1}{\delta+1}\sum_{j=i-\delta/2}^{i+\delta/2} x_j, \quad i = 1,2,\cdots,n \qquad (16)$$

where $\delta+1$ is the number of data points in the window centered at data point x_i. Using the method proposed, WLM is calculated at each point of the data set. The width of the summation window is centered at this point. Thus, all points contribute to the approximation of the local mean and no interpolation is required. We propose to apply the discrete form of the windowed local mean to EMD to directly approximate the local mean. We expect it to have a better capacity for identifying impulses without sacrificing extra decomposition capacity or adding calculation time.

4.2. Selection of window width and end point extension

To have an integrated algorithm, two issues need to be discussed. The first one is the selection of the window width. Apparently, the selection of the proper width of the summation window is absolutely crucial to the effectiveness of WLM because it determines the relative relationship between the local and the global perspectives. It has been given in [16] that a proper width for a signal with multiple frequency components is exactly equal to the shortest period, or the period of the highest frequency component. But this statement needs to be verified by testing more types of signals. For a given signal, the period of the highest frequency component is difficult to determine. Alternatively, [16] uses the interval between two neighboring maxima or two neighboring minima to estimate the shortest period. If the values of the shortest periods are not the same, [16] calculates the average length of all intervals between two neighboring maxima and between two neighboring minima first and then uses this average value to calculate windowed local mean at the current level. The decomposition process will not stop until the residual is a trend signal or the number of the decomposed IMFs has reached a preset value.

As shown in Eq. (16), WLM is calculated as a summation at each data point. When the distance between a center point and its nearest end point is less than half the width of the window, no more data values can be provided beyond the end point and the summation cannot be completed. Therefore, we have to have the ends of the data extended to guarantee that a summation can be conducted at every data point. The data should be extended at least half of the window width at the first and last point. A description of the extension is explained in [16] using segments of cosine waves.

4.3. Comparisons on simulated data

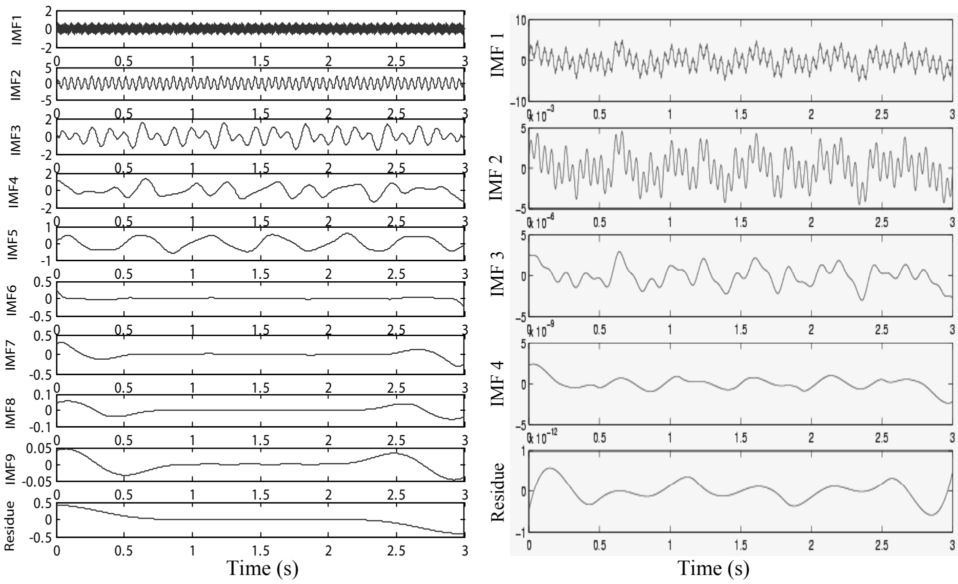
In this section, we are going to use the same simulated signals as used in Section 3 to test the proposed WLM method and compare it to the method which uses the reported direct-mean approximation (LMMD [5]). For comparison on the combination of sinusoid waves, the calculation time for the proposed approach and the LMMD on the 3000 point multiple sinusoid combinations is shown in Table 3. We can see that the proposed method does not have an advantage with regard to calculation time. This is because at every level of decomposition, every data point participates in the calculation of the windowed summation.

Tab. 3. Comparison of the calculation time for WLM and LMMD methods on the 3000-point x(t) in Eq. (10).

Direct-mean approximation:	LMMD	WLM
Calculation time (s):	0.1400s	4.0800s

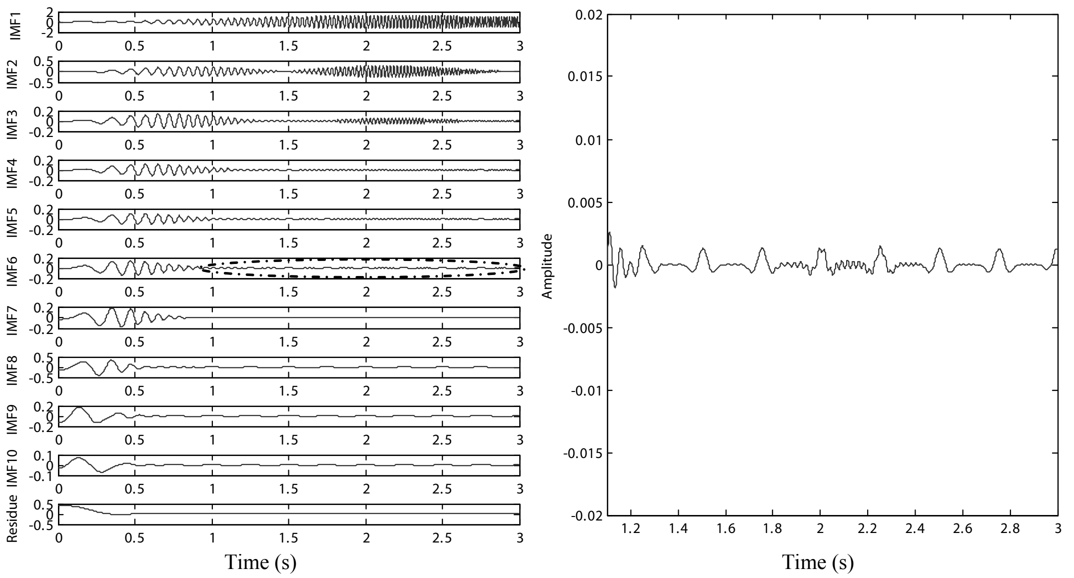
As the number of extrema decreases with decomposition, the width of the summation window becomes larger so that time consumption increases as the decomposition nears its conclusion, with regard to decomposition performance. However, we will show that the proposed approach has its advantages over LMMD. In Fig. 4 (b), it can be seen that there is a serious problem with LMMD. Only four IMFs have been generated because components are mixed up. The proposed approach does not have this problem. Although the IMFs beyond IMF3 are distorted to some extent, at least the frequencies of the first three IMFs are clear and they reflect the designed components within the simulated signal.

We applied the two methods to the combination signal of the chirp and impulse signals and the calculation time of the 3000 point data series is shown in Table 4. The proposed approach is slower than LMMD for the reason mentioned above but the time consumption is not that large. From visual observations of the decompositions shown in Fig. 5, it can be seen that some impulses are present in IMFs4-10. The marked region in IMF6 is enlarged and displayed in Fig. 5 (b) where most of the periodical impulses are clearly shown and it is easy to identify that the time interval between adjacent impulses is just 0.25 seconds as designed. These observations demonstrate the potential use of the



(a) WLM (b) LMMD

Fig. 4. Decomposition of the 3000-point x(t) in Eq. (10) using windowed local mean and LMMD



(a) Windowed local mean (b) The enlarged region as marked in IMF6 (Fig. 5(a))

Fig. 5. Decomposition of the 3000-point signal in Fig. 2 (c) by WLM method

WLM approach for identification of impulses. On the other hand, for the LMMD approach as shown in Fig. 6, the information on impulses is too blurred to reveal their intervals even from enlarged IMFs, which is our real concern when this type of simulated signal is tested, because impulses may represent faults being monitored. This shows that the LMMD approach is not as effective as the WLM approach.

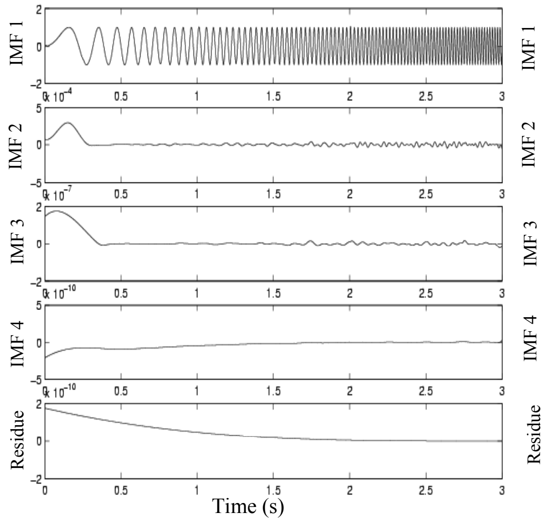


Fig. 6. Decomposition of the 3000-point data signal in Fig. 2 (c) by LMMD

Tab. 4. Comparison of the calculation times by WLM and LMMD methods on the 3000-point signal in Fig. 2 (c).

Direct-mean approximation:	LMMD	WLM
Calculation time (s):	0.1090	4.4690

Based on the comparison with the LMMD method, the proposed WLM approach is not as fast as LMMD. The LMMD method, however, does not pass the basic decomposition capability test because it uses only one average value to represent points between two extrema and it ignores local features. The proposed approach using WLM shows more useful information than does LMMD method. It identifies impulses that are hidden in chirp signals without losing its basic decomposition capability. This is due to the fact that the WLM approach takes every data point into consideration and captures more local features than does LMMD.

5. Conclusions and Discussion

We have discussed improvements to two types of approximation to the local mean in EMD, i.e. the envelope mean approximation and the direct mean approximation. Based on the work in this paper, the following conclusions have been reached.

(1) The MMPHI approach has advantages over the other two envelope mean methods, i.e., the cubic spline and the HOSI methods, with regard to calculation time and accuracy. When a visual observation is not able to detect much difference, the averaged MSE is used to help with the performance assessment.

(2) The WLM approximation is better than the LMMD method, with regard to its capability to identify impulses that are hidden in other signals. This merit can be obtained without sacrificing too much calculation time and without losing its basic decomposition ability.

(3) In applying EMD to processing signals, we suggest using the MMPHI approach to have a quick look at what basic frequency components are contained in the raw data. After that the WLM approximation can be used to detect impulses and any characteristic frequency that may exist due to faulty conditions in the system being monitored.

It should be noted that some interesting issues may be focused on in future research. The relationship between data length and MMPHI capability needs further investigation. The selection of a proper window width is crucial to the effectiveness of decomposition in the WLM approximation. We used a single average width value in our decomposition iteration. The average width may not work well for data with dynamic frequencies so that we may use varying widths along the length of the data to capture the local features more precisely. Last but not least, visual observation is not enough to assess how well impulse identification is being performed. A reasonable indicator of the accuracy of a decomposition needs to be defined, especially when real signals are analyzed. The validation of the proposed methods for the detection of variant gear fault modes besides tooth missing needs to be conducted.

****

This study was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

****

6. References

1. Butland J. A method for interpolating reasonable-shaped curves through any data. North-wood Hills Middlesex UK, 1980; Computer Graphics 80.
2. Cheng J S, Yu D J, Tang J S, and Yang Y. Local rub-impact fault diagnosis of the rotor systems based on EMD. Mechanism and Machine Theory 2009; 44: 784-791.
3. Fritsch F N, Butland J. A method for constructing local monotone piecewise cubic interpolants. SIAM J. Scientific and Statistical Computing 1984; 5: 300-304.
4. Fritsch F N, Carlso R E. Monotone piecewise cubic interpolation. SIAM Journal on Numerical Analysis 1980; 17: 238-246.
5. Gai Q, Ma X J, Zhang H Y, Zou Y K. Processing time-varying signals by a new method. Beijing; 2001 Radar, 2001 CIE International Conference on.
6. Huang H Z, Liu Z J, Murthy D N P. Optimal reliability, warranty and price for new products. IIE Transactions 2007; 39: 819-827.

7. Huang N E, Shen Z, Long S R, Wu M C, Shih H. H, Zheng Q, Yen N C, Tung C C, Liu H H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Pro. Royal Society 1998; 454: 903-995.
8. Kahaner D, Moler C, Nash S. Numerical Methods and Software. Prentice Hall, Englewood Cliffs: New Jersey, 1989.
9. Levesley J, Anderson I J, Mason J C. Algorithms for Approximation IV, University of Huddersfileld, 2002, pp. 24-35.
10. Liu Z J, Chen W, Huang H Z, Yang B. A diagnostics design decision model for products under warranty. International Journal of Production Economics 2007; 109: 230-240.
11. Liu Y, Huang H Z. Comment on ''A framework to practical predictive maintenance modeling for multi-state systems'' by Tan C.M. and Raghavan N. [Reliab Eng Syst Saf 2008; 93(8): 1138–50]. Reliability Engineering and System Safety 2009; 94: 776-780.
12. Peng W, Huang H Z, Zhang X, Liu Y, Li Y. Reliability based optimal preventive maintenance policy of series-parallel systems. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2009; (2): 4-7.
13. Rai V K, Mohanty A R. Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert–Huang transform. Mechanical Systems and Signal Processing 2007; 21: 2607-2615.
14. RÄ*o*sler H. A study on the empirical mode decomposition. M.Sc. Thesis, Mathematics and Computer Sciences, University of Amsterdam, Amsterdam, Netherlands, 2002.
15. Rilling G, Flandrin P, Goncalves P. On empirical mode decomposition and its algorithms. Grado Italy 2003; IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing.
16. Wang Y. Improvements of approximation of the local mean in the process of empirical mode decomposition. M.Sc. Thesis, Mechanical Engineering, University of Alberta, Canada, 2006.
17. Wolberg G, Alfy I. An energy-minimization framework for monotonic cubic spline interpolation. Journal of Computational and Applied Mathematics 2002; 143: 145-188.
18. Wu F J, Qu L S. An improved method for restraining the end effect in empirical mode decomposition and its applications to the fault diagnosis of large rotating machinery. Journal of Sound and Vibration 2008; 314: 586-602.
19. Zhao J P. Improvement of the mirror extending in the empirical mode decomposition method and the technology for eliminating frequency mixing. High Technology Letters 2002; 8: 40-47.

Yao WANG, M.Sc.
Prof. Ming J. ZUO, Ph.D.
Yaguo LEI, Ph.D.
Xianfeng FAN, Ph.D.

Department of Mechanical Engineering
University of Alberta
Edmonton, Alberta, T6G 2G8, Canada
e-mail: Ming.zuo@ualberta.ca