

A Research on Local Mean in Empirical Mode Decomposition

Yong-Ping Huang, Xue-Yao Li, and Ru-Bo Zhang

National Laboratory on Machine Perception, Peking University, Beijing, P.R. China
College of Computer Science and Technology, Harbin Engineering University,
Harbin 150001, P.R. China
{hwangyungping, lixueyao, zrb}@hrbeu.edu.cn

Abstract. The conventional Empirical mode decomposition (EMD) method uses envelope mean interpolated by cubic spline fitting, which is sensitive to extrema. An efficient method for finding the local mean is generated by using support vector regression machines. The analysis results indicate that the proposed algorithm has higher performance in the ability of frequency separation, insensitive to the sampling frequency and can eliminate mode mixing in small-amplitude sine waves intermittence.

Keywords: Empirical Mode Decomposition, Support Vector Regression Machines, Mode Mixing, Local Mean, Hilbert-Huang Transform.

1 Introduction

Empirical mode decomposition (EMD) introduced by Huang [1], is a method for decomposing non-linear and multi-component signals. The intrinsic mode functions (IMFs) resulting from EMD are obtained in the sifting process, which requires that a local mean should be defined [2]. So the critical part of EMD method is how to find the local mean. The conventional EMD proposed by Huang use envelope mean, the mean of upper and lower envelopes which interpolate all the local maxima or minima by the cubic spline, as local mean, but it is not the real one. In addition, the cubic spline fitting suffers overshoot or undershoot.

To find the real local mean, at present there are some solution approaches, such as the stationary-point EMD method (SPEMD) introduced by Zhong [3], a new envelope algorithm using subsection power function proposed by Qin [4], the B-spline approach for empirical mode decomposition adopted by Chen [5]. However, all these approaches use curve interpolation, which is sensitive to extrema, calculated by a parameter interpolation equation. This major and serious disadvantage will lead some difficulties when extrema are untrue or corrupted by intermittency.

Therefore, it is necessary to find a better method that is insensitive to extrema and has a robust performance. For the insensitive ability and robustness in interpolation, we introduce SVRM to estimate the local mean in sifting process, and the EMD algorithm based on SVRM is proposed.

2 The Local Mean Based on SVRM

In sifting process, SVRM is trained by local extrema (include local minima and maxima) of time series, then predicts directly the local mean. For the simplicity and the high generalization performance of SVRM, we choose the RBF kernel function and linear \mathcal{E} -insensitive loss function in SVRM. It is well known that SVRM generalization performance depends on a good setting of parameters [6].

For the adaptive and computation of analytical selection method, we adopt the analytical approach to choose the parameters. In the case of a signal $y=f(x)$, the extrema are $(ex_1, ey_1), \dots, (ex_i, ey_i), \dots, (ex_n, ey_n)$, where ex_i and ey_i are the coordinates and the corresponding y value of extrema, n is the number of extrema. Note that the extrema should be sorted by coordinate order for computing the information between adjacent extreme. The analytical parameter selections used in sifting process are given as follows, the regularization parameter C :

$$C = \max\left(\overline{ey} + 3\delta_{ey}, \left|\overline{ey} - 3\delta_{ey}\right|\right) \quad (1)$$

here \overline{ey} and δ_{ey} are the mean and the standard deviation of the y value of extrema. In addition, the following RBF kernel parameter γ :

$$\gamma = \max\left(\overline{dis_ex} + \delta_{dis_ex}, \left|\overline{dis_ex} - \delta_{dis_ex}\right|\right) \quad (2)$$

where $\overline{dis_ex}$ and δ_{dis_ex} are the mean and the standard deviation of the coordinate distances between adjacent extrema. The parameter of insensitive loss function \mathcal{E} is:

$$\mathcal{E} = \frac{\overline{dis_ey}}{2} \quad (3)$$

where $\overline{dis_ey}$ is the mean of the y value distances between adjacent extrema.

3 Experiments

To illustrate the behavior of the local mean based on SVRM, we used the conventional EMD introduced by Huang [1] to compare the algorithm using the local mean based on SVRM (hereinafter EMD-SVRM for short). First, we generated a synthetic signal of two superimposed tones:

$$f(x) = a_1 \sin(2\pi f_1 x) + a_2 \sin(2\pi f_2 x), x \in [0, 10\pi] \quad (4)$$

where both f_1 and f_2 are the frequency of sine tones, respectively. The local mean based on SVRM and the envelope mean are shown in Fig.1. The fluctuation in the SVRM-based local mean is considerable, which implies that the sifting process can separate the tones. As proposed in Ref [10], the conventional EMD can not accurately separate two superimposed tones when the frequency proportion $0.5 < f_2/f_1 < 2.0$. So

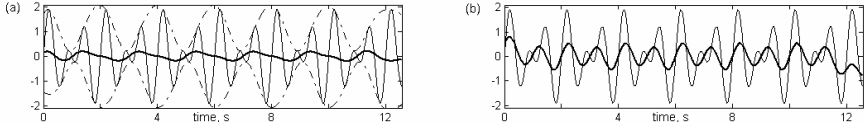


Fig. 1. Illustration of the envelope mean and the local mean using SVRM: (a) the signal in thin solid line, with the upper and lower envelopes in dot-dashed lines and the envelope mean in thick solid line; (b) the signal in thin solid line and the local mean based on SVRM in thick solid line. The signal tested is of the form in Eq.4 with $f_1 = 1$ and $f_2 = 1.5$ with unit amplitude.

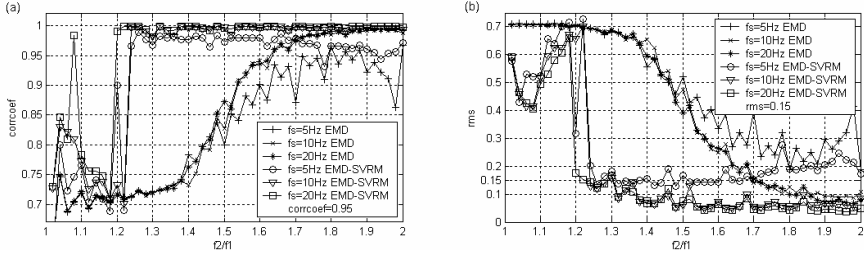


Fig. 2. Comparison of the *corrcoef* and the *rms* between the first IMF and the actual high frequency component: (a) calculated results of *corrcoef*; (b) calculated results of *rms*

we compare the performance in different sampling frequencies and different frequency proportions $f_2/f_1 \in (1.0, 2.0]$.

Fig.2 presents calculated results of the correlation coefficient (*corrcoef*) and the root mean square error (*rms*) between the first IMF and the actual high frequency component. No substantial differences in the results of the second IMF, we omit them for the sake of brevity. The measured results show the significant improvement of the proposed algorithm including separate components that are similar in frequency that would be inseparable with conventional one. And the EMD-SVRM algorithm is less insensitive to sampling frequency for the good performance when sampling frequency $fs=5\text{Hz}$, while the performance of conventional EMD is unstable and big fluctuation.

In this test, we consider the small-amplitude and high-frequency sine waves intermittent in large-amplitude tones. The signal is given as following:

$$y = \sin(8\pi x) + 0.03\sin(200\pi x) \left[u\left(x - \frac{85}{300}\right) - u\left(x - \frac{100}{300}\right) \right] + 0.03\sin(200\pi x) \left[u\left(x - \frac{165}{300}\right) - u\left(x - \frac{180}{300}\right) \right] \quad (5)$$

where $x \in [0, 1]$ and the sample frequency fs is 300Hz. Using the conventional EMD, we will obtain the first three components which contain seriously mixed modes especially the first IMF, as shown in Fig.3.a. And with the EMD-SVRM algorithm, the result is given in Fig.3.b. Clearly, the small-amplitude sine wave intermittent was separated from the large-amplitude tones and the amplitude of the third IMF component is very small. The test shows the EMD-SVRM algorithm can eliminate the mode mixing caused by intermittency and do not need the intermittency criterion introduced by Huang [1].

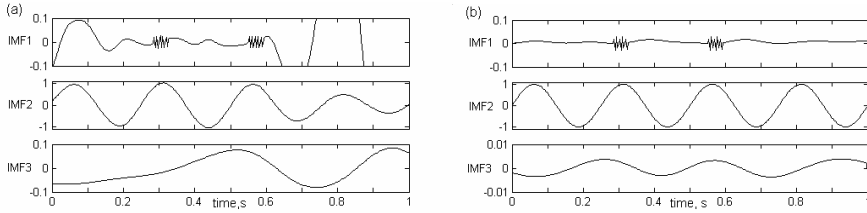


Fig. 3. Comparison of the IMF components between the conventional EMD and the EMD-SVRM for the intermittency tones: (a) the first three IMF components from conventional EMD algorithm; (b) the first three IMF components from EMD-SVRM approach

4 Conclusions

The local mean based on SVRM is designed to overcome the deficiency of sensitivity to extrema, and adopt the generalization performance of SVRM. Compared with conventional one, our approach shows significant improvements over frequency separation and good performance of eliminating mode mixing in intermittency. Future work include finding better model selection especially the analytic parameter selection and using the ν -SVR instead the \mathcal{E} -SVR may get better performance.

Acknowledgements. This work was supported by the China National Science Foundation (Grant No.60475016) and the Foundational Research Fund of Harbin Engineering University (HEUF 04092).

References

1. Huang, N.E., Shen, Z.: A new view of nonlinear water waves: the Hilbert spectrum. *A. Rev. Fluid Mech*, Vol. 31. (1999) 417-457
2. Deering, R., Kaiser, J.F.: The use of a masking signal to improve empirical mode decomposition. *ICASSP2005*, Vol. 4. (2005) 485-488
3. Zhong, Y.M.: Research on theoretic evidence and realization of directly-mean EMD method. *Chinese Journal of mechanical engineering*, Vol. 17. (2004) 399-404
4. Qin, S.R., Zhong, Y.M.: A new envelope algorithm of Hilbert–Huang Transform. *Mechanical Systems and Signal Processing*, Vol. 20. (2006) 1941–1952
5. Chen, Q.H., Huang, N.E., Riemenschneider, S.: A B-spline approach for empirical mode decompositions. *Advances in Computational Mathematics*, Vol. 24. (2006) 171–195
6. Cherkassy, V., Ma, Y.Q.: Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks*, Vol. 17. (2004) 113-126