Raman spectroscopy de-noising based on EEMD combined with VS-LMS algorithm*

YU Xiao (俞潇)[1†], XU Liang (许亮)[2†], MO Jia-qing (莫家庆)[1], and LÜ Xiao-yi (吕小毅)[1,3]**

1. School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China

2. School of Electrical Engineering, Tianjin University of Technology, Tianjin 300384, China

3. Institute of Health and Environmental Medicine, Academy of Military Medical Science, Tianjin 300050, China

(Received 14 October 2015)

This paper proposes a novel de-noising algorithm based on ensemble empirical mode decomposition (EEMD) and the variable step size least mean square (VS-LMS) adaptive filter. The noise of the high frequency part of spectrum will be removed through EEMD, and then the VS-LMS algorithm is utilized for overall de-noising. The EEMD combined with VS-LMS algorithm can not only preserve the detail and envelope of the effective signal, but also improve the system stability. When the method is used on pure R6G, the signal-to-noise ratio (*SNR*) of Raman spectrum is lower than 10 dB. The de-noising superiority of the proposed method in Raman spectrum can be verified by three evaluation standards of *SNR*, root mean square error (*RMSE*) and the correlation coefficient ρ.

Document code: A **Article ID:** 1673-1905(2016)01-0016-4

DOI 10.1007/s11801-016-5200-9

The main methods of de-noising for Raman spectrum include wavelet[1], ensemble empirical mode decomposition (EEMD) and least mean square (LMS) adaptive filter techniques. Wavelet transform requires prior knowledge, such as the selection of wavelet function and the decomposed level of wavelet[1-4]. Empirical mode decomposition (EMD)[5] is a new kind of adaptive signal processing method, which is suitable for nonlinear and non-stationary signals. One important drawback of the EMD is modal mixing, which distorts the signal. Focusing on the mode mixing problem caused by EMD, the EEMD algorithm was proposed by Wu et al[6] in order to overcome the modal aliasing of EMD. A non-stationary signal produced using EEMD can produce a number of smooth intrinsic mode functions (IMFs) in turn. However, when signal-to-noise ratio (*SNR*) is low and the system is influenced by abnormal events, EEMD of high frequency IMFs will suffer different degrees of white noise. LMS adaptive filter does not need prior knowledge, which has made it be the standard for the adaptive filter algorithm. However, the steady-state error and the convergence rate inevitably come into conflict with each other, so the de-noising results can be hit or missed[7].

Research results show that the de-noising effect of LMS adaptive filter is much better than that of wavelet or EEMD[8]. However, there is still room to improve the resolution of the peaks and troughs. The LMS adaptive filter is a training network, and it may cause some instability problems for nonlinear and non-stationary signal filtering. For this reason, this paper proposes that EEMD can be combined with the VS-LMS algorithm to be used for Raman spectrum de-nosing. The algorithm overcomes the instability of the network and improves the resolution of the spectrum.

The general principle diagram of the adaptive filter is shown in Fig.1. The output signal can be produced by digital filter with adjustable parameters after the signal gets into the system, and the error signal can be got by comparing the output signal with the expected signal. The adaptive algorithm is used to adjust the parameters of the filter, in order to achieve the smallest error signal.

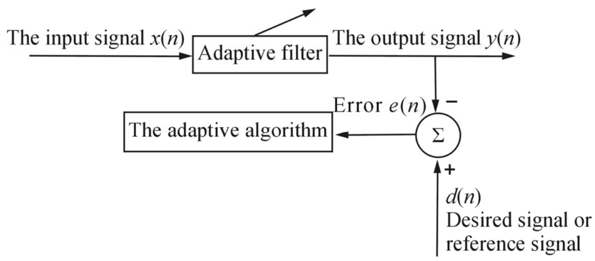

Fig.1 The principle diagram of the LMS adaptive filter

* This work has been supported by the National Natural Science Foundation of China (No.61308120), the Doctor Startup Project of Xinjiang University (No.BS120122), the Young Talents Project in Xinjiang Uygur Autonomous Region (No.2013731003), and the Xinjiang Science and Technology Project (Nos.201412107 and 2014211B003).

† These authors contributed equally to this work and should be considered co-first authors.

** E-mail: xiaoz813@163.com

The specific algorithm about LMS adaptive filter has been given in Refs.[9] and [10].

Raman spectra of non-stationary and nonlinear signals are similar. The LMS filter is a training network, and it is more suitable for periodic signal than other methods. When the noise intensity is large in the Raman spectrum, it becomes difficult to control the size of the step, and the problems of waveform disorder and system instability appear. For this reason, the Raman spectrum containing noise should firstly be treated with EEMD to filter out the bulk of the noise. The high-frequency component contains a lot of white noise, so it can serve as a filter. Then the signal is reconstructed. The resulting signal contains a small amount of white noise. However, after EEMD filtering, the resolution of peaks in Raman spectrum is not high, which affects the analysis of the characteristics of the sample. Then the VS-LMS adaptive filter can be used to achieve a better de-noising effect. The VS-LMS algorithm is obtained by improving the LMS algorithm. The step size factor in the VS-LMS algorithm is established as

$$mu = \frac{mu_0}{1+\dfrac{n}{5\,000}}, \tag{1}$$

where mu_0 and mu are the first and the nth step size factors, respectively.

This method can prevent the contradiction between the convergence rate and the steady-state error of LMS algorithm. When the VS-LMS adaptive filter begins to work, because of the large error, a fast convergence is realized by using the large step size. In the continuous learning process, it is required to improve the stability and reduce the imbalance, and the length of each step also decreases, which produces a better de-noising effect and improves the stability of the network. The flow chart of the proposed EEMD combined with VS-LMS algorithm is shown in Fig.2.

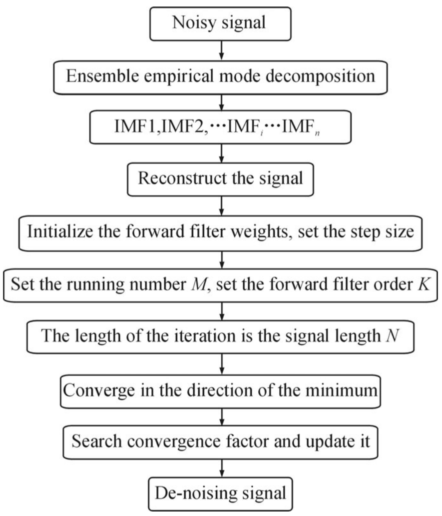

Fig.2 Flow chart of the EEMD combined with VS-LMS algorithm

The Raman spectrum of pure R6G was recorded by RENISHAW Raman spectrometer made in German BRUKER company. The Raman spectrum of pure R6G as shown in Fig.3 can be obtained in the frequency range from 578.560 76 cm^{-1} to 3 997.854 88 cm^{-1} at 532 nm continuous laser excitation with output power of 20 mW, and the data acquisition time is 5 s.

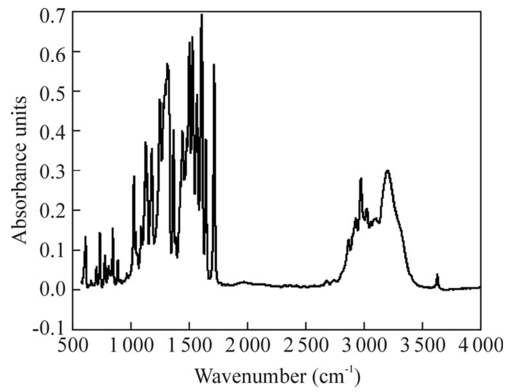

Fig.3 The Raman spectrum of pure R6G

Adding the white noise with the intensity of 0.08 to the Raman spectrum of pure R6G, the Raman spectrum containing noise is obtained as shown in Fig.4, and then the *SNR* of Raman spectrum is got to be 6.103 3 dB.

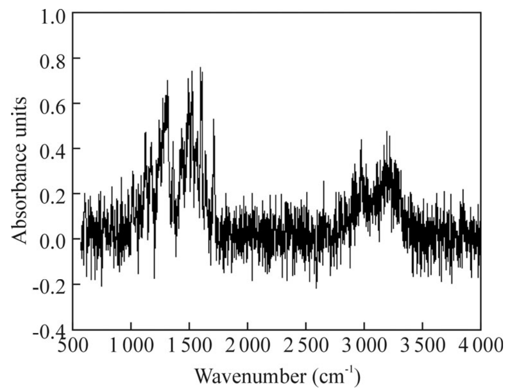

Fig.4 The Raman spectrum containing noise

We use wavelet, EEMD, LMS adaptive filter and EEMD combined with VS-LMS algorithm for de-noising processing, respectively. The three evaluation standards of *SNR*, root mean square error (*RMSE*) and correlation coefficient ρ can prove the de-noising advantages of EEMD combined with VS-LMS algorithm in Raman spectrum.

Different decomposition levels, threshold selection methods and wavelet basis functions based on Raman spectrum are selected and compared for many times. Finally, sym8 is selected as the wavelet basis function, 3-layer decomposition is used, the rigrsure method is adopted to determine the threshold value, and the soft threshold is used to de-noise the spectrum. The Raman spectrum after wavelet sym8 de-noising is shown in Fig.5.

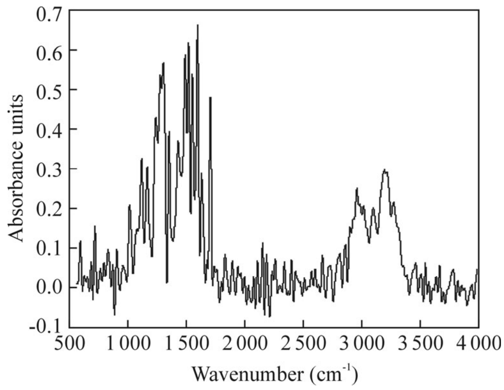

Fig.5 The Raman spectrum after wavelet sym8 de-noising

Fig.5 shows that the spectrum resolution is improved through wavelet de-noising. However, it is still not satisfactory. Compared with Fig.3, the Raman spectrum in Fig.5 still has some noise, especially in the ranges of 1 700—2 500 cm^{-1} and 3 500—4 000 cm^{-1}.

Using EEMD adaptive decomposition, the EEMD diagrams are obtained as shown in Fig.6.

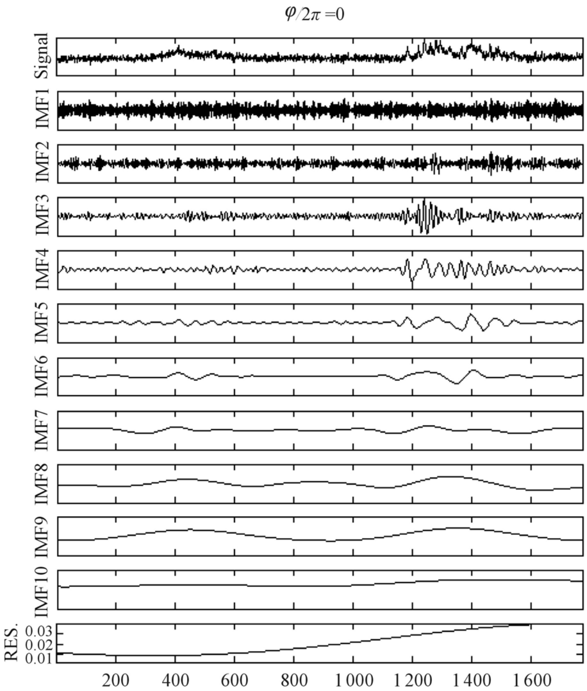

Fig.6 EEMD diagrams

In Fig.6, IMF1 to IMF10 are the IMFs of the signals containing noise. In theory, they are single frequency components from high frequency to low frequency. The last diagram shows the trend of the spectral envelope. IMF1 is white noise, and IMF2—IMF10 are the effective signal components. However, it can be seen from Fig.6 that the modal mixing phenomenon obviously appears in IMF3—IMF6. It can be found from experimental results that the reconstruction signals from IMF2 to IMF10 contain a lot of noise, which suggests that the IMF1 and IMF2 also contain a lot of noise. Therefore, the reconstructed signals cannot obtain better de-noising results. After a series of experiments, we find that the reconstructed signal from IMF3 is the best, which is shown in Fig.7.

Fig.7 shows that the spectrum resolution is improved through EEMD de-noising. But some details are still submerged by noise. It still has some noise, especially in the ranges of 1 700—2 500 cm^{-1} and 3 500—4 000 cm^{-1}. Compared with the results in Fig.3, the main peaks are interfered by noise in the range of 3 000—3 300 cm^{-1}.

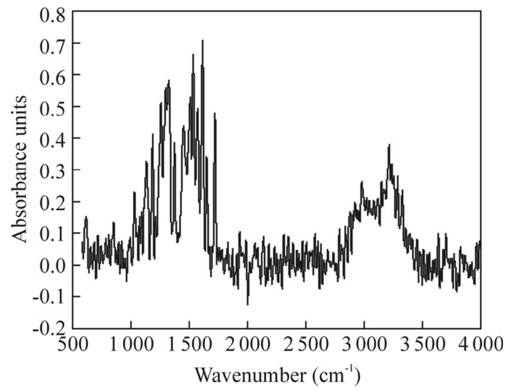

Fig.7 The Raman spectrum after EEMD de-noising from IMF3

After a series of experiments, the LMS adaptive filter is designed. Step size factor is mu=0.56, the filter order is K=12, and the running number is 30. The number of iterations is from 1 to 1 774. Initialize the forward filter weights as W=[0.164 2, 0.134 1, 0.052 9, −0.062 4, −0.158 6, −0.193 2, −0.155 5, −0.059 9, 0.058 4, 0.122 9, 0.110 6].

The Raman spectrum after LMS adaptive filter de-noising is shown in Fig.8.

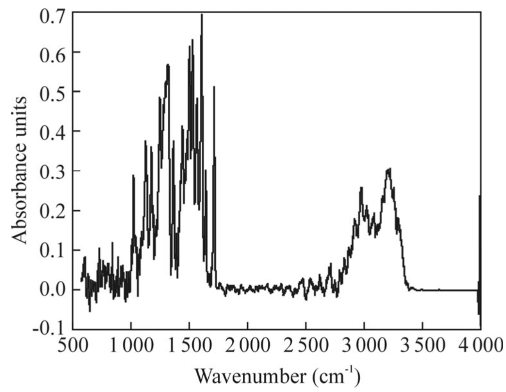

Fig.8 The Raman spectrum after LMS adaptive filter de-noising

Fig.8 shows that the details are more clear than those in Figs.5 and 7. But there is still some noise in the range of 1 700—2 500 cm^{-1}, and several peaks still have a small amount of noise in the range of 1 200— 1 700 cm^{-1}. Consequently, the resolution in Fig.8 is still not ideal.

Through a series of experiments, the EEMD combined with VS-LMS adaptive filter is designed. Firstly, choose EEMD to de-noise the Raman spectrum containing noise. The intensity of white noise is 0.04, the number of EMD decompositions is 50, and the maximum number of decompositions in each time is 300. 10 components of IMF1—IMF10 are obtained, then we filter IMF1 and IMF2, and finally the reconstructed signal is obtained. Secondly, the related parameters of VS-LMS adaptive filter are designed as follows: the order number is K=12, the running number is M=30, the first step length factor is mu_0=0.56, the number of iterations is from 1 to 1 774, and mu is the nth step length factor. Initialize the forward filter weights as W=[0.164 2, 0.134 1, 0.052 9, −0.062 4, −0.158 6, −0.193 2, −0.155 5, −0.059 9, 0.058 4, 0.122 9, 0.110 6].

The Raman spectrum after EEMD combined with VS-LMS adaptive filter de-noising is shown in Fig.9. It can be seen that the Raman spectrum has more clear details. Compared with the result in Fig.8, several peaks and smooth places in Fig.9 have no noise, especially in the range of 1 200—2 500 cm^{-1}. The spectrum in Fig.9 is more close to the original one. In the process of experiment, the filter stability is enhanced significantly.

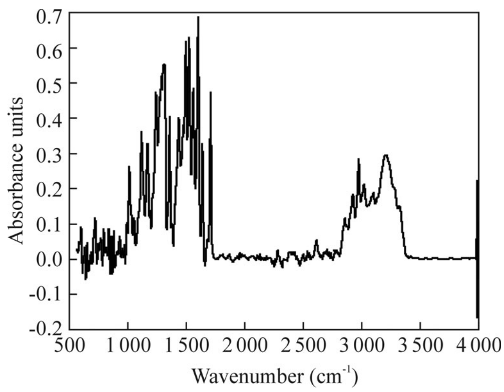

Fig.9 The Raman spectrum after EEMD combined with VS-LMS adaptive filter de-noising

The evaluation standards are shown in Tab.1. The computer simulation results show that the EEMD combined with VS-LMS algorithm has better performance than the wavelet sym8 algorithm, the EEMD algorithm and the LMS adaptive filter algorithm. Compared with those of LMS filter, the *SNR* of EEMD combined with VS-LMS algorithm is improved by 12.8%, *RMSE* is down to 0.019 9, and the correlation coefficient of the waveform is also increased to 0.988 1.

Tab.1 Evaluation standards of four de-noising methods

	Noisy signal	Sym8	EEMD	LMS	EEMD-VSLMS
SNR	6.103 3	13.851 1	12.551 5	15.548 7	17.539 4
RMSE	0.074 3	0.030 4	0.035 4	0.025 0	0.019 9
ρ	0.859 3	0.971 4	0.963 9	0.980 9	0.988 1

The main signal and noise of Raman spectrum are overlapping. The EEMD combined with VS-LMS adaptive filter is used for de-noising of the Raman spectrum containing noise, and the de-noising result is compared with those of wavelet transform, EEMD adaptive decomposition and LMS adaptive filter methods. The wavelet transform can improve the resolution of the Raman spectrum when *SNR* is less than 10 dB. However, its de-noising effect is not satisfactory. The Raman spectrum after wavelet de-noising is compared with the original one. Some details are submerged in the noise, and a lot of peaks and troughs are visible after de-noising, which do not exist before. When spectrum's *SNR* is less than 10 dB, the EEMD de-noising is highly stable, and it preserves the wave characteristics. After de-noising, the resolution of the peak in spectrum needs to be further improved. The de-noising effect of the LMS adaptive filter is remarkable, and the evaluation standards are higher than those of the wavelet and EEMD methods. However, some noise still appears in smooth parts of the signal, the network becomes unstable and the algorithm does not converge, because it is not easy to control the step size factor. The new method of EEMD combined with the VS-LMS algorithm has remarkable de-noising effect. The evaluation standards are higher than those of wavelet, EEMD and LMS adaptive filter methods. The network is stable in the process of de-noising, and this algorithm preserves the characteristics of the original spectrum. After de-noising with the EEMD combined with VS-LMS algorithm, the pure R6G Raman spectrum can indicate the features of the sample accurately.

References

[1] Gao Zong-li, Ye Wei-lin, Zheng Chuan-tao and Wang Yi-ding, Optoelectronics Letters **10**, 299 (2014).

[2] Wang Jian-feng, Zhang Zai-xuan, Dai Bi-zhi, Li Chen-xia and Insoo S. Kim, Optoelectronics Letters **2**, 195 (2006).

[3] Yin Hong, Chen Zeng-qiang and Yuan Zhu-zhi, Optoelectronics Letters **2**, 369 (2006).

[4] Chen Yong, Wang Kun, Liu Huan-lin, Chen Li-juan and He Ming-ling, Journal of Optoelectronics·Laser **24**, 1053 (2013). (in Chinese)

[5] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung and Henry H. Liu, Royal Society of London Proceedings **454**, 903 (1998).

[6] Zhaohua Wu and Norden E. Huang, Advances in Adaptive Data Analysis **1**, 1 (2009).

[7] Dariusz Bismor, Variable Size LMS Algorithm with Fast Convergence Capabilities for Active Noise Control, 18th International Congress on Sound and Vibration 1, 282 (2011).

[8] Zhang Yuan-yuan, Li Shun-ming, Hu Yi-xian, Jiang Xing-xing and Guo Hai-dong, Journal of Vibration and Shock **32**, 61 (2013).

[9] Ma Guo-dong, Yan Shu-tian, He Cheng-zhu and Yang Chen, Electronic Design Engineering **22**, 43 (2014).

[10] Roberto Cristi, Modern Digital Signal Processing, America: CL-Engineering, 2003.