Comparison of simultaneous signals obtained from a dual-field-of-view lidar and its application to noise reduction based on empirical mode decomposition

Wei Gong (龚 威)[1], Jun Li (李 俊)[1*], Feiyue Mao (毛飞跃)[1], and Jinye Zhang (张金业)[2]

[1]*State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China*

[2]*School of Science, Hubei University of Technology, Wuhan 430068, China*

[*]*Corresponding author: larkiner@gmail.com*

Received October 27, 2010; accepted December 17, 2010; posted online April 18, 2011

Although the empirical mode decomposition (EMD) method is an effective tool for noise reduction in lidar signals, evaluating the effectiveness of the denoising method is difficult. A dual-field-of-view lidar for observing atmospheric aerosols is described. The backscattering signals obtained from two channels have different signal-to-noise ratios (SNRs). The performance of noise reduction can be investigated by comparing the high SNR signal and the denoised low SNR signal without a simulation experiment. With this approach, the signal and noise are extracted to one intrinsic mode function (IMF) by the EMD-based denoising; thus, the threshold method is applied to the IMFs. Experimental results show that the improved threshold method can effectively perform noise reduction while preserving useful sudden-change information.

OCIS codes: 010.1100, 010.3640.

doi: 10.3788/COL201109.050101.

Lidar is widely used for observing atmospheric aerosols. As an active remote sensing instrument, it provides a high-spatial-resolution vertical profile of aerosol optical properties[1], but the effective range and data reliability are often limited by various noises. Performing a proper denoising method improves the quality of the obtained signals.

The measured lidar signal contains the laser backscattering signal from aerosol and various noises. It can be expressed simply as

$$V_{\mathrm{m}}(r) = V(r) + N_{\mathrm{b}}(r) + N_{\mathrm{e}}(r), \tag{1}$$

where $V_{\mathrm{m}}(r)$ is the actually measured signal, $V(r)$ is the backscattering signal from aerosol, $N_{\mathrm{b}}(r)$ denotes the noise caused by background light, and $N_{\mathrm{e}}(r)$ represents the noise caused by dark current and read out electronics. N_{b} and N_{e} can be statistically estimated by the signal obtained from an extremely far distance where the laser backscattering signal is negligible.

The power of the received signal typically falls with an increase in range, but noise is usually considered Gaussian white noise, which is stable with range. The signal-to-noise ratio (SNR) falls as the range increases, and the solution for the lidar equation becomes unstable and even fails because of the negative value produced by noise. Thus, the signal must be denoised before data retrieval for the aerosol properties.

Several signal analysis methods are widely adopted for noise reduction in the lidar signal. Most lidar systems employ multiple pulse averaging to enhance SNR. This method is a low pass filtering process at the cost of temporal resolution; the high-frequency backscattering signal is also smoothed. Wavelet analysis has been developed rapidly as an effective tool for noise reduction. Its main drawback is that the basis functions are fixed, and no such basis function has been proposed to correspond to the features of lidar signals.

Huang *et al.* introduced empirical mode decomposition (EMD) for analyzing signals from non-stationary and nonlinear processes[2]. The EMD method addresses completeness, orthogonality, locality, and adaptivity, which are necessary to describe non-stationary and nonlinear processes. The major advantage of EMD is posteriori adaptation because the basis functions are derived from the signal itself[2,3]. The signal can be decomposed into a series of intrinsic mode functions (IMFs) and the residual through the sifting process:

$$V_{\mathrm{m}}(r) = \sum_{j=1}^{n} \mathrm{IMF}_j(r) + R_n(r). \tag{2}$$

Each IMF satisfies two conditions: the number of

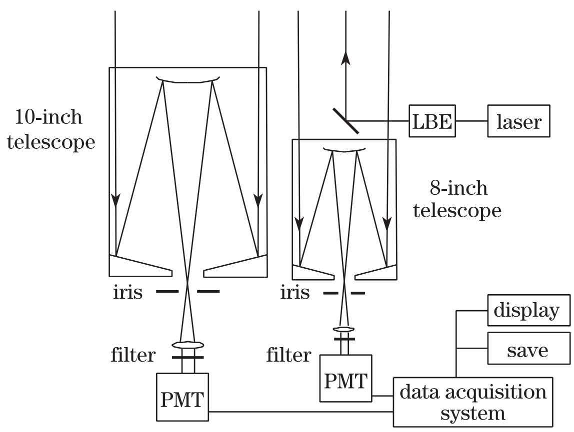

Fig. 1. Schematic of the DFL system. LBE: laser beam expander; PMT: photomultiplier tube.

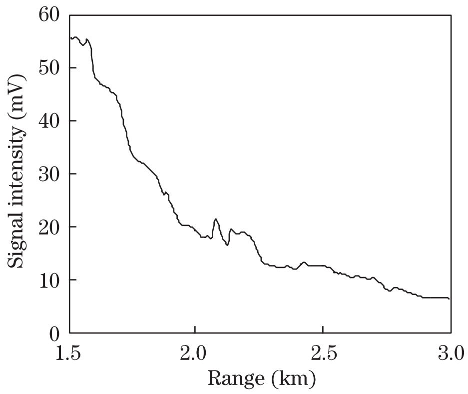

Fig. 2. Far-range channel signal.

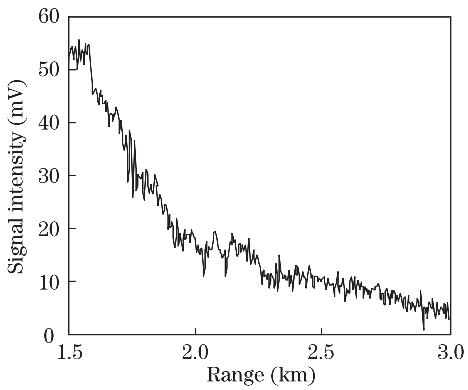

Fig. 3. Scaled near-range channel signal.

extrema and the number of zero crossings must either be equal or differ at most by one, and the mean of the upper and lower envelopes derived from local extrema is zero at any point[2,3]. This allows for physically meaningful instantaneous frequency and amplitude calculation through the Hilbert transform performed on the IMFs. Any IMF represents a simple oscillation mode. The low-order IMFs represent high-frequency oscillation components, whereas the high-order IMFs represent low-frequency oscillation components. The noise and signal are traditionally characterized by high and low frequencies, respectively, so that they can be distinguished by EMD. Although the EMD technique has been applied to various fields, and the theoretical base is empirical, some researches have shown that the EMD-based signal denoising method is effective in the analysis of a lidar signal[4].

The conventional EMD-based signal denoising method is achieved as follows[5].

Step 1: The signal is decomposed using the EMD method as shown in Eq. (2). Decomposition depends on the envelope calculations derived from local extrema. The mean of the derived envelopes is iteratively subtracted from the current signal until it is close enough to zero, and then the first IMF is extracted. After the first IMF is extracted and subtracted from the original signal, the procedure is repeated to obtain the subsequent IMFs. Finally, the lidar signal is decomposed into a series of IMFs and a trend.

Step 2: For reconstruction, in the time domain, the lower-order and higher-order IMFs represent the fine and coarse scales, respectively. We assume that low-order IMFs contain little value of the backscattering signal, and the conventional EMD-based signal denoising method is performed by obtaining the residual with the removal of some low-order IMFs.

After EMD-based noise reduction is applied to the actually measured signal, a suitable test must be performed to assess the effectiveness of the method. Simulation experiments are widely performed because the signals without noise are difficult to be estimated from the actual signals. However, the simulated signals cannot perfectly replace the actual signals because many effects in the real world are unexplained or extremely difficult to simulate. To solve this problem, we adopt an approach to compare the signals obtained from a dual-field-of-view lidar (DFL).

The DFL system was developed by the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University. Figure 1 shows the schematic of the DFL system. The lidar has two independent receiving channels to solve the problem of the dynamic range of lidar[6], one for the lower atmosphere (especially the planetary boundary layer) employing a coaxial system (near-range channel), and the other for higher altitudes using a biaxial system (far-range channel). The fields-of-view of the near-range and far-range channels are 10 and 1 mrad, respectively. The laser beam fully enters the fields-of-view of the near-range and far-range channels from distances of about 360 and 1000 m, respectively. The spatial resolution is 3.75 m relative to the 40 MHz sample frequency

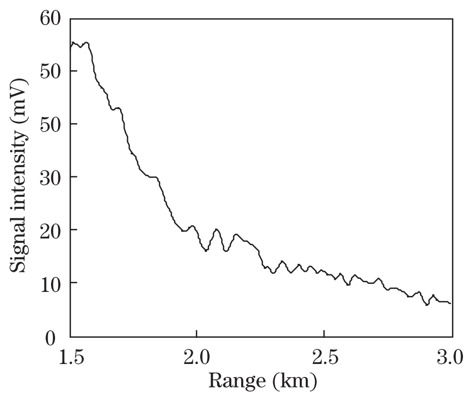

Fig. 4. Denoised near-range signal as residual 2.

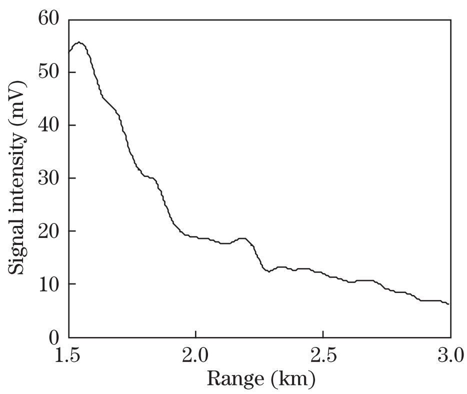

Fig. 5. Denoised near-range signal as residual 3.

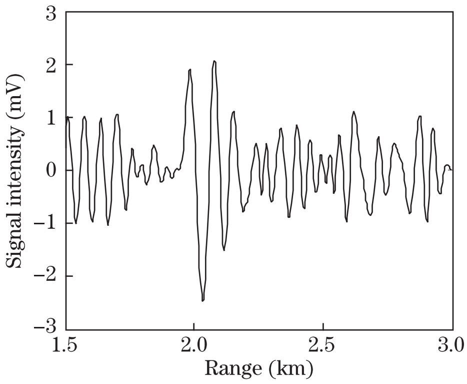

Fig. 6. IMF3 of near-range signal.

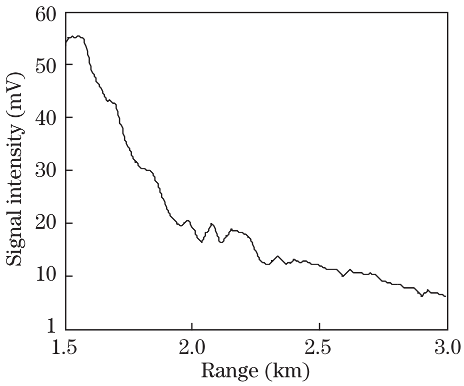

Fig. 7. Denoised result by the EMD-based threshold method.

of the data acquisition system.

The DFL employs simultaneous measurements performed by both channels. Figures 2 and 3 show the signals from 1.5 to 3 km, obtained from far-range and near-range channels at 20:10, Aug. 3, 2010. The comparison of these two simultaneous signals indicates similar useful signals and markedly different noise levels. The signals obtained from the two channels are similar because of simultaneous measurements and identical atmosphere altitudes, which is evidenced by long-term observing data. The noise intensity of the near-range channel is higher because of the large field-of-view and low efficiency of the optics and electronics. The larger field-of-view means more interference from the background light. The signal of the near-range channel is restricted to avoid saturation of the receiver. When the near-range signal is scaled to the far-range signal, the noise is also amplified by a factor that is typically more than 500. As a result, the standard deviation of the noise is 0.003%–0.02% of the signal in Fig. 2, but the value considerably increases to 1.7%–11% in Fig. 3. Thus, within the range of 1.5–3 km, the far-range signal can disregard noise contamination, but the near-range signal is buried in the noise. The comparison of these two signals provides a new approach to assessing the performance of the denoising method without simulation.

The result of the conventional method as residual 2 (Fig. 4), obtained by subtracting the first two IMFs, still contains some obvious high-frequency fluctuations. However, residual 3 (Fig. 5) shows the smoothed useful signal, especially the sudden change in signal at 2.1 km. This is because the backscattering signal and noise are both extracted to IMF3 (Fig. 6). The undulation at 2.1 km shows a structure similar to the far-range signal shown in Fig. 2, and the value of the peak beside 2.1 km shows a considerable difference with the partial IMF3 signal above 2.2 km. This shows the conflict between smoothing high-frequency noise and preserving the high-frequency signal. The conventional EMD-based signal denoising method is not always effective in the analysis of lidar signals because the backscattering signal and noise may be extracted to one IMF.

The threshold method can be adopted to improve performance[7,8]. The buried useful signal can be acquired by processing the soft threshold for the IMFs. The soft threshold can smooth the result and the threshold method shrinks the IMF as shown by the following functions:

$$\tau = \sigma\sqrt{2\,\log(L)}, \tag{3}$$

$$\sigma = \text{Median}(|\text{IMF}(r) - \text{Median}(\text{IMF}(r))|)/0.6745, \tag{4}$$

$$S(r) = \begin{cases} \text{IMF}(r) - \tau & \text{IMF}(r) \geq \tau \\ 0 & |\text{IMF}(r)| < \tau, \\ \text{IMF}(r) + \tau & \text{IMF}(r) \leq -\tau \end{cases} \tag{5}$$

where σ is the estimated noise level of the IMF, L is the length of the IMF, $S(r)$ denotes the buried useful signal that is extracted to IMF with noise, and τ represents the threshold value, determined as the conventional

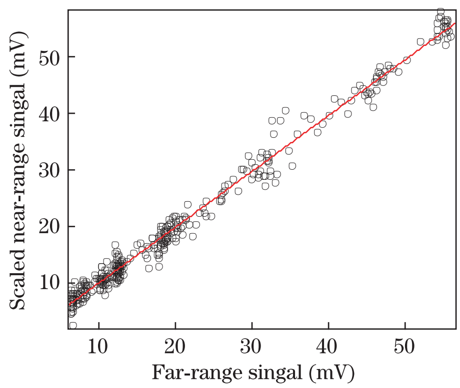

Fig. 8. Near-range signal versus far-range signal.

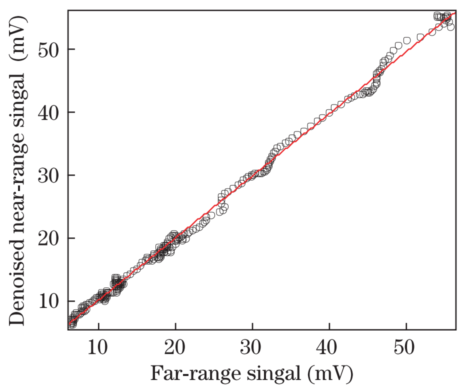

Fig. 9. Denoised near-range signal versus far-range signal.

Table 1. Results of the Linear Fitting and SNR

		Near	Residual 2	Residual 3	Wavelet	Threshold
20:10, Aug. 3	RMSE	1.541	0.7845	0.7589	0.797	0.6852
	R^2	0.9878	0.9968	0.9970	0.9967	0.9976
	SNR	24.0254	29.5414	29.6489	29.2647	30.4954
20:12, Aug. 3	RMSE	1.4830	0.6873	0.7938	0.7261	0.5978
	R^2	0.9870	0.9972	0.9963	0.9968	0.9979
	SNR	24.0292	30.5612	29.4353	30.1329	31.7869
20:07, Aug. 14	RMSE	0.9679	0.4925	1.454	0.4193	0.3938
	R^2	0.9942	0.9985	0.9861	0.9989	0.9990
	SNR	24.8741	30.7334	20.9294	32.1098	32.6644

threshold proposed by Donoho *et al.*[9] After $S(r)$ is obtained as the high-frequency useful signal, the denoising result can be modified by adding $S(r)$. This method can be called the EMD-based threshold method.

The EMD-based denoising method with soft threshold is applied to the near-range signal in Fig. 3. Usually, deciding whether the undulation of the actually obtained signal is the useful signal or noise is difficult. However, with the advantage of the DFL, the denoising result can be compared with the far-range signal. The result (Fig. 7) shows better performance than those of residuals 2 and 3.

To evaluate the performance of the EMD-based denoising method with soft threshold, the method was tested on the signals obtained at 20:10, Aug. 3, 2010, 20:12, Aug. 3, 2010, and 20:07, Aug. 14, 2010. The linear regressions of the near-range signal and denoising result are performed as a function of the far-range signals. The results of the linear fitting and SNR are shown in Table 1. The smaller root mean square error (RMSE) indicates better results. R^2 is a statistical measure of how well the regression line approximates the data points, and R^2=1.0 indicates that the regression line perfectly fits the data. The SNR is defined as

$$\text{SNR} = 10\log\frac{\sum V_{\text{far}}^2(r)}{\sum [V_{\text{de}}(r) - V_{\text{far}}(r)]^2}. \quad (6)$$

Firstly, the RMSE, R^2, and SNR values of residuals 2 and 3 show better performance than those of the original near-channel signal. This indicates that the conventional EMD-based noise reduction is effective for aerosol lidar signals. Secondly, residual 3 shows better results than residual 2 at 20:10, but residual 2 shows better results during the other two periods. This is because both the useful signal and noise are extracted to IMF3. The conventional method cannot effectively distinguish useful signals from noise. Residual 2 accepts these as useful signals, and residual 3 removes them as noise. Thirdly, after combining with the threshold method, the denoising results are markedly improved than those of residuals 2 and 3. The comparison of the near-range signal (Fig. 8) and the denoising result (Fig. 9) at 20:10, Aug. 3, 2010 also shows improvement after noise reduction, and the offset of linear fitting decreases. Finally, the EMD-based threshold method also performs better than the wavelet method. In particular, the wavelet method performs better than the conventional EMD-based method at 20:07, Aug. 14, 2010 and the EMD-based threshold method performs better than the wavelet method.

In conclusion, the comparison of signals obtained from the DFL provides a new approach for evaluating the performance of the denoising method without simulation. This approach avoids the drawbacks and limitations of simulation, and improves the reliability of the denoising results. The comparison shows that noise and high-frequency useful signals may be extracted to one IMF based on EMD. It decreases the performance of the conventional EMD-based denoising; thus, the threshold method is employed to obtain signals buried in the noise. On the basis of the analysis of received signals, the threshold method performs better than others, with promising results. Further work will be required to more effectively employ the EMD-based threshold method in more lidar systems.

This work was supported by the National "973" Program of China (Nos. 2009CB723905 and 2011CB707106) and the National Natural Science Foundation of China (Nos. 10978003 and 40871171).

References

1. W. Gong, J. Zhang, F. Mao, and J. Li, Chin. Opt. Lett. **8,** 533 (2010).
2. N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, Proc. R. Soc. Lond. A **454,** 903 (1998).
3. N. E. Huang and Z. Wu, Rev. Geophys. **46,** RG2006 (2008).
4. S. Wu, Z. Liu, and B. Liu, Opt. Commun. **267,** 137 (2006).
5. F. Zheng, D. Hua, and A. Zhou, Chinese J. Lasers (in Chinese) **36,** 1068 (2009).
6. J. Harms, Appl. Opt. **18,** 1559 (1979).
7. A. O. Boudraa, J. C. Cexus, and Z. Saidi, Int. J. Signal Process. **1,** 33 (2004).
8. Q. Li, G. Zhang, and Y. Liu, Spectrosc. Spect. Anal. (in Chinese) **29,** 142 (2009).
9. D. L. Donoho and I. M. Johnstone, Biometrika **81,** 425 (1994).