Suppression of the noise component of lidar signals based on discrete wavelet transform

A.D. Ershov

Institute of Atmospheric Optics,
Siberian Branch of the Russian Academy of Sciences, Tomsk

Received June 11, 2003

The procedure of noise suppression in lidar signals based on wavelet transform is considered. Multiresolution analysis is briefly reviewed. It is shown that the discrete wavelet transform can successfully be used in such practical applications as noise suppression, while the signal structure in this case is kept unchanged.

Introduction

At laser sensing of the atmosphere, the signal continuously experiences various distortions caused by both the instrumental noise and the noise due to the atmospheric background. Noise of different type at the receiver input restricts the lidar capabilities. One of the ways to improve the accuracy and reliability of lidar measurements is application of signal processing algorithms allowing the increase of the signal-to-noise ratio.

The investigations carried out in this field gave a rather wide set of different data processing algorithms. Thus, an algorithm using Markov filtration apparatus is described in Refs. 1 and 2. References 3 to 5 present the methods for separation of the signal from the background of noise through application of adaptive least-squares (LS) algorithms. But recently (in the mid-1980s) arisen signal filtering algorithms that are currently being intensely developed based on wavelet transform are capable of providing satisfactory results even with the noise level higher than 100% (Ref. 6). The aim of this paper is to study the applicability of the methods of wavelet analysis to isolation of the signal from the background of noise in lidar data.

Analysis of signals in the terms of wavelet transform is divided into two types: continuous and discrete (dyadic). The continuous wavelet transform is largely used for analysis of transient processes, detection of sharp and hidden variations in a signal, study of nonstationarities, etc.[7,8] The discrete wavelet transform (DWT) has gained its widest utility in applied problems, for example, signal compression and noise suppression, because it can be carried out efficiently and without extra consumption of computer memory with the properly selected basis functions (the vector of coefficients recursively substitutes for the vector of the initial values). Since DWT is based on the ideas of multiresolution analysis (MRA), consider them first.

1. Orthogonal multiresolution analysis

Orthogonal MRA in the Hilbert state $L^2(R)$ is a series of closed nonoverlapping subspaces $V_j \subset L^2(R)$, $j \in Z$, whose incorporation gives, in the limit, $L^2(R)$ [Refs. 9–11]:

$$V_j \subset V_{j+1} \ , \ \bigcup_{j\in Z} V_j = L^2(R) \ , \ \bigcap_{j\in Z} V_j = 0 \ ;$$

for any function $s(t) \in V_j$, its extended version belong to the space

$$V_{j+1} \ (s(t) \in V_j \Leftrightarrow s(2t) \ V_{j+1});$$

there is such a function $\varphi(t)$ that

$$\{\varphi(t-k)\}_{k\in Z} \ , \ \left(\int_{-\infty}^{\infty} \varphi(t)\mathrm{d}t \neq 0 \right)$$

form an orthonormal basis of the space V_0. The function $\varphi(t)$ is called the generating scaling function or the scaling function. Since $\varphi_{0,k}(t)$ form the orthonormal basis of the space V_0, the functions $\varphi_{j,k}(t) = 2^{j/2}\varphi(2^j t - k)$ form the orthonormal basis of the space V_j (the normalized factor $2^{j/2}$ appears, because at the change of the variable $t \to 2t$ the norm of the function decreases by $\sqrt{2}$ times, which should be compensated for).

Thus, the functions $\varphi(t-k)$ of the basis V_1 result from the basis V_0 through simple double shrinkage of its elements $\varphi(2t-k)$. Since any element of V_0 belongs also to V_1, $\varphi(t)$ can be resolved in terms of $\varphi(2t-k)$ [Refs. 10–12, 14]:

$$\varphi(t) = \sqrt{2}\sum_{k\in K} h_k \varphi(2t-k), \qquad (1)$$

where $h_k \in R$, $k \in K$, $K \subset Z$.

0235-6880/03/10 856-07 $02.00

Equation (1) is called the scaling relationship for the scaling function, and the coefficients h_k are called the coefficients of the scaling function or the scaling filter.

In the majority of cases, there is no explicit equation for $\varphi(t)$ [Ref. 12]. However, there is a fast algorithm, which uses the detailing equation for the scaling function $\varphi(t)$ at points of the dyadic grid ($t = 2^{-j}k$, $j, k \in Z$) [Ref. 13]. For every pair of subspaces $V_j \subset V_{j+1}$ of the MRA, there is a subspace W_j, being the complement of V_j up to V_{j+1}, such that

$$V_j \perp W_j,\ V_{j+1} = V_j \oplus W_j,$$

that is, every element of V_{j+1} can be written unambiguously as a sum of the elements W_j and V_j. The subspaces W_j are called refining or detailing in the meaning that they contain the detailed information needed for transition from the approximation with the resolution j to the approximation with the resolution $j + 1$. For the space W_0, the generating wavelet $\psi(t)$ is defined so that the series $\{\psi(t - k)\}_{k\in Z}$ forms the orthonormal basis in W_0. The functions

$$\psi_{j,k}(t) = 2^{j/2}\,\psi(2^j t - k)$$

are called wavelets. The scaling relationship for wavelets is determined as:

$$\psi(t) = \sqrt{2}\sum_{k\in K} g_k \varphi(2t - k), \qquad (2)$$

where $gk = (-1)^k h_{1-k}$.

It follows from Eq. (2) that wavelets are fully determined by the scaling functions (1). Now the signal $s(t)$ can be represented as a linear combination[14]:

$$s(t) = \sum_{k=1}^{\infty} c_{j0,k}(k)\varphi_{j0,k}(t) + \sum_{j=j_0}^{\infty}\sum_{k=1}^{\infty} d_j(k)\,\psi_{j,k}(t). \qquad (3)$$

The coefficients in Eq. (3) are determined from the condition of orthogonality of the basis functions:

$$c_{j0}(k) = \int_{-\infty}^{\infty} s(t)\varphi_{j0,k}(t)\mathrm{d}t\,, \qquad (4)$$

$$d_j(k) = \int_{-\infty}^{\infty} s(t)\psi_{j,k}(t)\mathrm{d}t\,. \qquad (5)$$

The coefficients $c_{j0}(k)$ and $d_j(k)$ are DWT coefficients (spectrum of the signal $s(t)$ in the bases of scaling functions and wavelets). The terms in Eq. (3) can be considered as representation of the signal $s(t)$ in two projections: projection onto the space V_{j0} (first term) and projection onto the orthogonal complement of V_{j0} up to $L^2(R)$ (second term). The space structure is such that the projection of the signal onto the first space is "coarse" (low-frequency), while the second one is the high-frequency (it contains the detailed information about the signal that was lost at projection onto V_{j0}).

Figure 1 shows the schematic layout of DWT. In every first unit, time readouts of the signal are processed, while every next unit processes the corresponding coefficients c_j. As a result, we obtain $j_m - j_0 - 1$ vectors of wavelet coefficients d_j and one vector c_j.

Equations (1) and (2) allow obtaining the algorithms for fast calculation of $c_j(k)$ and $d_j(k)$. For an arbitrary scaling coefficient j, Eqs. (1) and (2) have the form:

$$\varphi_{j,k}(t) = 2^{(j+1)/2}\sum_n h(n)\,\varphi(2^{j+1}t - 2k - n) =$$

$$= 2^{(j+1)/2}\sum_n h(n-2k)\,\varphi(2^{j+1}t - n), \qquad (6)$$

$$\psi_{j,k}(t) = 2^{(j+1)/2}\sum_n g(n-2k)\varphi(2^{j+1}t - n)\,. \qquad (7)$$

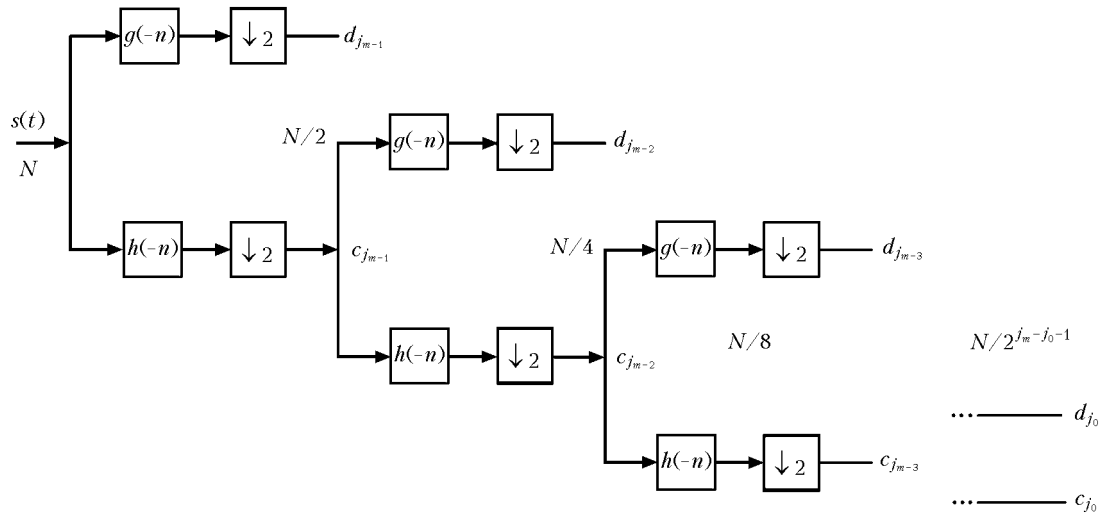

Fig. 1. Schematic layout of the direct discrete wavelet transform. The decimation procedure is shown by arrows.

Substituting Eqs. (6) and (7) into Eqs. (4) and (5), we obtain the recurrence equations for calculating $c_j(k)$ and $d_j(k)$ (from "fine" to "coarse" scale):

$$c_j(k) = \sum_n h(n-2k)c_{j+1}(n);$$
$$d_j(k) = \sum_n g(n-2k)c_{j+1}(n). \quad (8)$$

Here we face a problem on the first step of calculation of the coefficients c_j at some value of the scale $j = j_m$. For continuous signals, the highest level of resolution in the general case is equal to $+\infty$. Therefore, the coefficients of signal representation with some finite resolution j_m are usually taken as the first step, and the higher j_m, the higher the accuracy of representation. The coefficient c_j of the initial representation is determined from the equation:

$$c_{j_m}(k) = \int_{-\infty}^{\infty} s(t)\varphi_{j_m,k}(t)\mathrm{d}t.$$

In the case of a discrete signal, the components of the signal itself are taken as the coefficients $c_j(k)$ of the level of resolution j_m, that is, $c_{jm}(k) = s(k)$. The iteration procedures (6) and (7) are completed at $j = j_0$, which is chosen based on the signal duration.

2. Noise filtering procedure

Classification of signals into all possible types is difficult, but wavelets turn out a powerful tool for rather simple characterization of a wide class of signals.[15] Regardless of the nature of the information processed, for noise suppression it is necessary to take into account a number of problems, for example, to check for validity of the hypothesis on the Gauss character of the statistics and the spectral composition of the noise component, selection of the type of a function for threshold processing along with the criterion for calculation of the threshold itself, as well as standard problems of wavelet analysis having the general character: selection of the most appropriate wavelet basis, establishment of the needed depth of data decomposition, etc.

Appearance of a number of methodological paradigms of noise suppression, as well as software and hardware solutions based on them has led to the significant progress in solution of the problem of noise suppression as one of the important problems of information processing.[16]

In the terms of wavelet analysis, the noise filtering procedure consists of three stages (Donoho–Johnston paradigm)[17–19]:

(a) *Signal decomposition*

The type of wavelet and the number of the levels of resolution j_m are selected. DWT of the signal $s(t)$ to the level j_m is calculated.

(b) *Selection of threshold for wavelet coefficients*

For any level from j_0 to j_m, the threshold ε_j is selected and the coefficients are modified by a certain rule.

(c) *Signal reconstruction*

Inverse DWT is carried out with the use of modified wavelet coefficients by Eq. (3).

From the statistical point of view, such a technique is nonparametric estimation of the regression model of the signal with the use of the orthogonal basis.[20] This method works well with many real data, including signals having nonstationary noise characteristics, but the best results are obtained for rather smooth signals, that is, the signals, in whose decomposition only a small number of detailing coefficients differ markedly from zero. The optimal values of the parameters of decomposition depend on the signal characteristics and should be selected experimentally.[21] Here we can recommend the following:

(1) The depth of decomposition determines the scale of rejected details: the larger its value, the large signal variations will be filtered out. At rather large values of this parameter ($j_m > 7$), not only noise is suppressed, but also the signal is smoothed (its peaks are cut off).

(2) The order of the wavelet determines the smoothness of the reconstructed signal: the smaller the wavelet order, the more pronounced are signal peaks, and vice versa – for wavelets of high orders the signal peaks are smoothed. The order of the wavelet becomes an important factor at inverse DWT, when it is necessary to smooth errors caused by rejection of small wavelet coefficients.[22]

In practice, first, some initial level of decomposition (usually equal to three) is selected and the wavelet filtering is performed. Then the level of decomposition is increased until the best result is achieved, that is, the decomposition details contain a noise-like component, and the approximation rather well describes the initial signal.

Consider the operation of the noise filtering procedure for problems of laser sensing. The lidar echo signal (test signal) calculated by the lidar equation in the single scattering approximation for the 5-km long horizontal path[23,24] is taken as an input signal. The profile of the aerosol scattering coefficient is set homogeneous along the path ($\sigma_{aer} = 0.2$ km^{-1}, $\sigma_{mol} = 0.012$ km^{-1}) with imposed three aerosol inhomogeneities ($\sigma_{aer1} = 0.4$ km^{-1}, $\sigma_{aer2} = 0.6$ km^{-1}, $\sigma_{aer3} = 0.6$ km^{-1}). The echo signal obtained is supplemented with white noise so that the signal-to-noise ratio does not exceed unity already at the distance about 3 km.

To perform DWT, it is first necessary to choose the type of wavelet. Here there are no clear criteria. The similarity of the signal (as a whole, rather than some its part) to the transformation function usually serves as a parameter decisive for this choice, but this is a rather subjective factor. To facilitate this procedure, it is possible to use the criterion of minimum entropy.[22] Entropy of a signal with respect to the wavelet basis

reflects the number of significant terms in the decomposition (3) and is determined by the value of

$$\exp\left(-\sum_{j,k}\left|d_{j,k}\right|^2 \log\left|d_{j,k}\right|^2\right).$$

The smaller this value, the more optimal is the basis of analysis for this signal. Usually, the higher is the order of the wavelet, the lower is entropy. For the lidar echo signal, the wavelet of tenth order from the family of orthogonal wavelets with a compact carrier — Symlet wavelet — was taken. The scaling and wavelet functions of the selected basis are shown in Fig. 2.

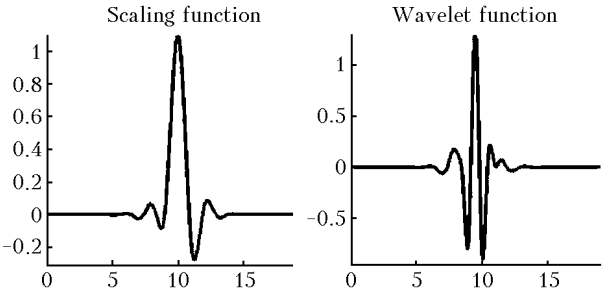

Fig. 2. Plot of the scaling function and Symlet wavelet of the 10th order.

This family of wavelets is described in a more detail in Refs. 9, 11, and 25. The detailing coefficients $d_{j,k}$ of the five-level DWT of the lidar signal are shown in Fig. 3*a*. The histograms of the distribution of $d_{j,k}$ are depicted in Fig. 3*b*.

It is seem from Fig. 3 that the values of $d_{j,k}$ are mostly close to zero, and the variance of the detailing coefficients decreases with the increase of the level of decomposition of the function ($d_{j0} \to d_{j4}$) and the histogram acquires more and more Gaussian form.

The next stage of signal filtering is selection of the threshold and the rules for modifying the coefficients. Selection of the threshold ε is a significant problem in the noise suppression procedure. If the value of the threshold ε is low enough, then it leads to partial conservation of the noise component in the detailing coefficients and only insignificant increase of the signal-to-noise ratio. At the same time, large threshold values can cause the loss of the coefficients bearing information about the behavior of the profile of the optical atmospheric characteristics. The optimal value of ε is such that the highest possible signal-to-noise ratio is provided with the lowest distortion of the reconstructed signal.

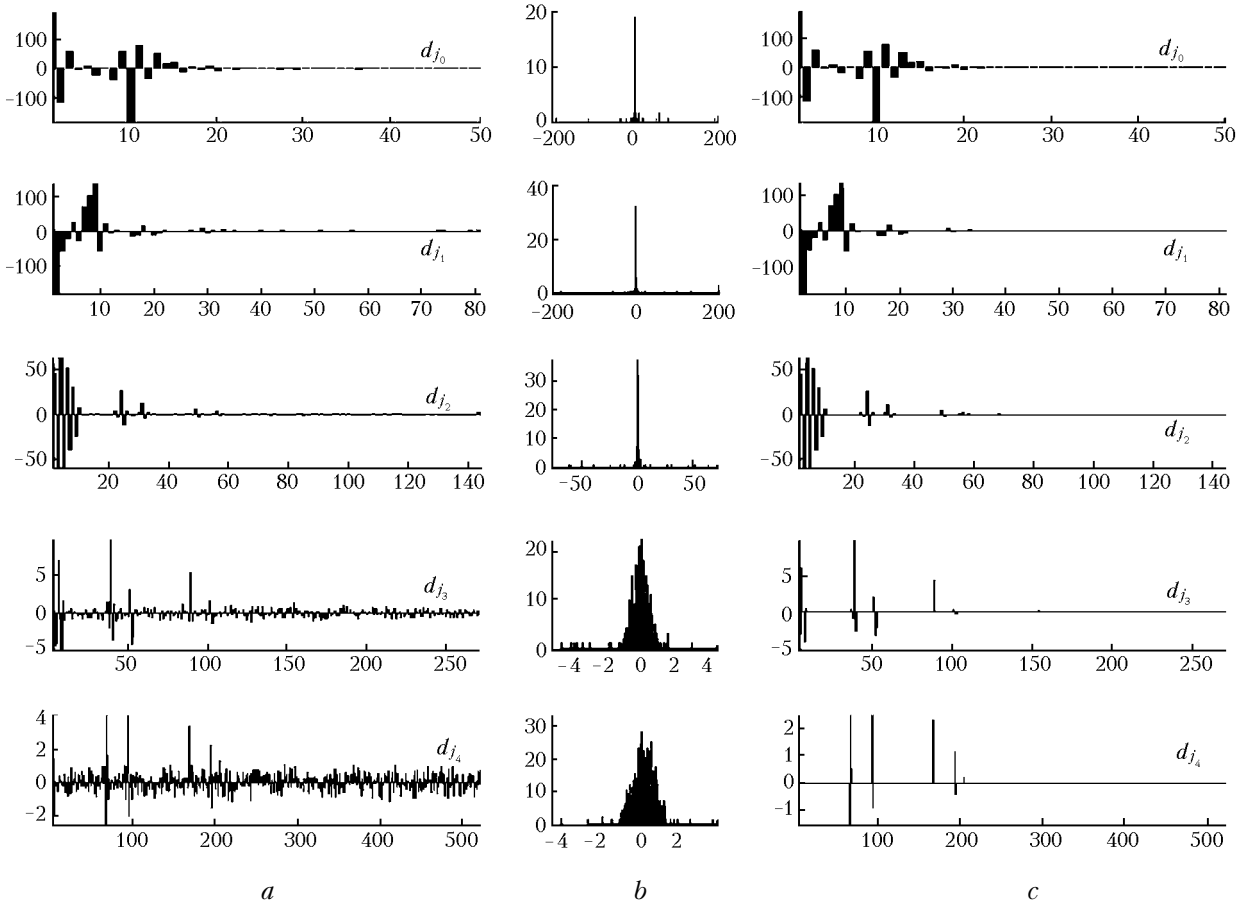

Fig. 3. Coefficients of DWT detailization of a test signal at the levels $j_0 - j_{m-4}$; $d_{j_0} - d_{j_{m-4}}$ before the procedure of finding the threshold (*a*); histograms of the distribution of $d_{j_0} - d_{j_{m-4}}$ (*b*); $d_{j_0} - d_{j_{m-4}}$ after the threshold finding procedure (*c*).

The value of ε is estimated by certain criteria, which are quite numerous. The detailed information on them can be found in Refs. 17, 26, and 27. The most widely used criteria implemented in the majority of computer mathematics are:

(1) The "universal" Donoho–Johnston criterion:

$$\varepsilon_j = \sigma\sqrt{2\log(n_j)}\,,$$

where n_j is the length of realization of the wavelet decomposition of the level j; σ is the variance of noise. The variance can be estimated using a rather smooth part of the initial signal.

(2) The Stein risk assessment criterion[28] minimizing the square loss function

$$R_j(\varepsilon) = \frac{1}{n_j}\left\| d_j^* - d_j \right\|^2$$

for the chosen model of noise, where $R_j(\varepsilon)$ is the square loss function, d_j^* are detailing coefficients of the filtered signal, d_j are the detailing coefficients of the signal hypothetically without an additive noise.

(3) The minimax criterion. The filtered signal is approximated by the regression model, and ε is selected so that it realizes the minimum of the maximum root-mean-square error obtained for worst function in this set.

Modification of $d_{j,k}$ is performed by the two main rules[16,17,20,26]:

– hard thresholding:

$$T_{\mathrm{h}}(d_{j,k}) = \begin{cases} 0, & \text{if} \quad d_{j,k} < \varepsilon_j \\ d_{j,k}, & \text{if} \quad d_{j,k} \geq \varepsilon_j \end{cases},$$

and

– soft thresholding:

$$T_{\mathrm{s}}(d_{j,k}) = \mathrm{sign}(d_{j,k})\max(0, |d_{j,k}| - \varepsilon_j)\,.$$

In the problems of noise suppression, it is better to use the method of soft thresholding, while in the problems of signal compression the rule of hard thresholding is used more often. Hard thresholding has two disadvantages that decrease its value for noise suppression. The first of them consists in that conservation of detailing coefficients exceeding some preset threshold value assumes conservation of noise present in them. Another disadvantage is connected with the appearance of parasitic harmonics in the resulting signal due to artificial introduction of lacunas formed from zeroed coefficients into the series.[16] The general equation for the procedure of threshold processing can be written in the following form:

$$s(t) = \sum_{k=1}^{\infty} c_{j0,k}(k)\varphi_{j0,k}(t) + \sum_{j=j_0}^{\infty}\sum_{k=1}^{\infty} T_{j,k}\{d_j(k)\}\,\psi_{j,k}(t). \quad (9)$$

In Eq. (9) the threshold value in T can be taken common for the whole level of decomposition (global thresholding) or it is possible to use ε varying from one level to another (local thresholding) and depending on the number of the detailing coefficients (microlocal thresholding). Local versions of thresholding are preferred in noise suppression problems, because they have higher adaptability to the initial data as compared to the global threshold processing. Let us illustrate the peculiarities of different criteria of threshold processing using a test signal.

Figure 4*a* depicts the profile of the reconstructed lidar echo signal corrected for range square after application of soft local thresholding to the detailing coefficients $d_{j,k}$ (Fig. 3*a*) by the criteria described above. The detailing coefficients $d_{j,k}$ after the thresholding procedure (Stein criterion) are depicted in Fig. 3*c*.

It is seen from Fig. 4*a* that the use of the threshold ε calculated by the Stein criterion yields better results as compared to that obtained using the universal and minimax criteria. In this case, the noise in the lidar signal is suppressed without considerable distortion of the echo signal itself. In the range from 3 to 4 km, the mean variation of the signal amplitude decreases from 190% in the initial data to 12% in the echo signal reconstructed after thresholding.

Now let us illustrate the capability of noise suppression using a lidar echo signal due to Raman scattering by nitrogen molecules recorded in the photon counting mode. The noise component recorded by the photon counter consists of elastic scattering pulses coming through side lobes of the interference filter, pulses caused by the sky radiation, and dark pulses of the photocathode. Since the recorded signal was weak, the relative rms error already at the distance of 5 km achieved 40%, and at 10 km it exceeded 80%. Figure 4*b* shows the initial (curve *1*) and reconstructed (curves *2* and *3*) lidar signals.

The threshold processing of detailing coefficients of the lidar signal was performed by the rule of soft thresholding. The highest values of the threshold ε are obtained when using the estimate of the universal Donoho–Johnston criterion. In this case, it is seen from Fig. 4*b* (curve *2*) that even in the case of application of the maximum values of ε parasitic harmonics are still present in the lidar signal. In this case, to completely remove the noise component, it is necessary to selectively increase the threshold for the levels j_m, j_{m-1} (Fig. 4*b*, curve *3*).

Conclusions

Applying the noise suppression procedure based on DWT to lidar signals, it is possible to achieve a rather deep suppression of noise with the signal structure kept unchanged. The main disadvantage is that calculation of the threshold ε by the existing criteria not always gives satisfactory results and it is necessary to vary the threshold for obtaining acceptable results.

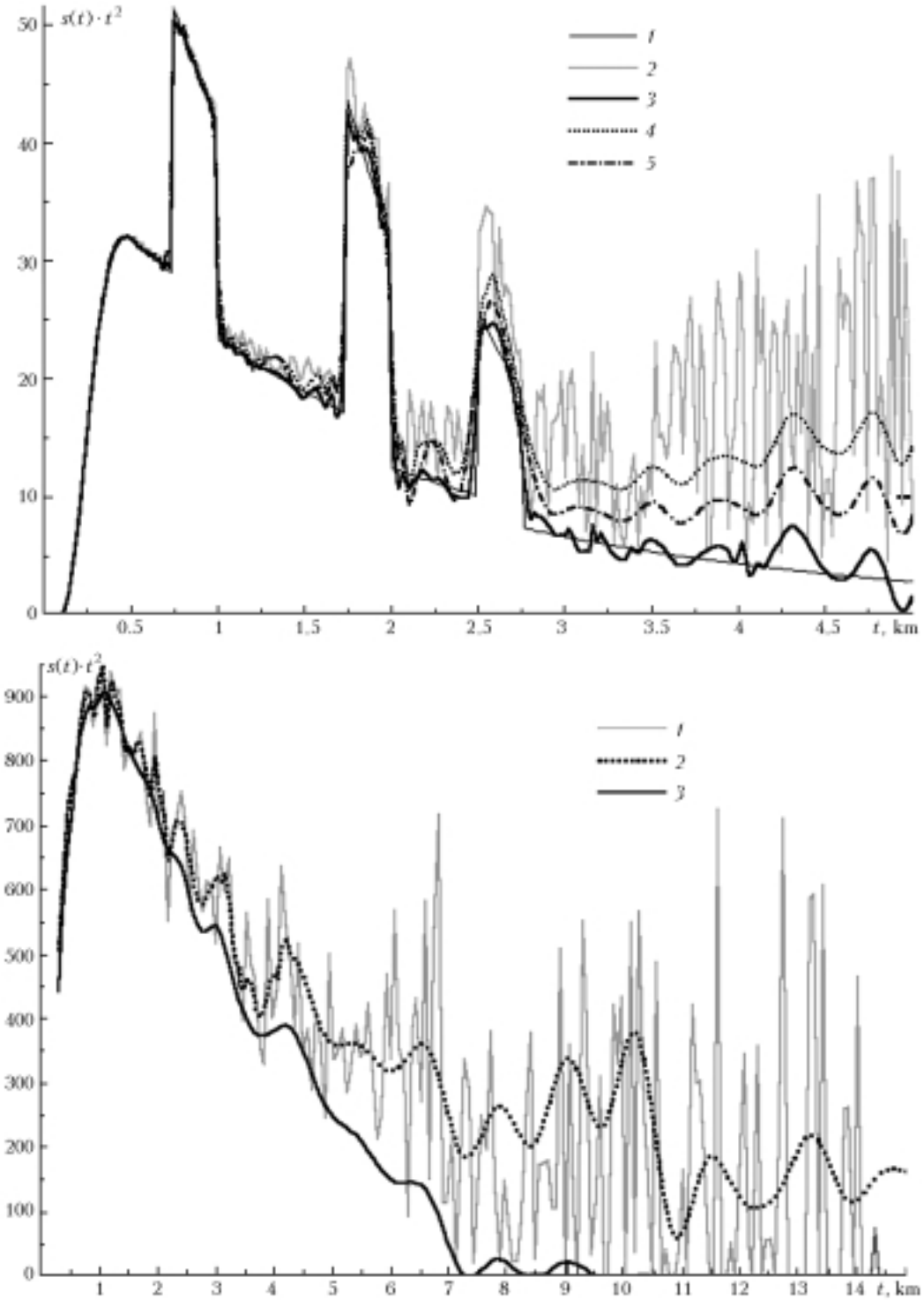

Fig. 4. Signal and its reconstruction by the tenth-order symlet: test signal (*a*): initial signal (*1*), noisy signal (*2*); signal reconstructed using the thresholding procedure by the Stein (*3*), Donoho–Johnston (*4*), and minimax (*5*) criterion; Raman lidar signal (*b*): initial signal (*1*); reconstructed signal ["soft" thresholding with the estimate ε by the Donoho–Johnston criterion] (*2*); the same but ε for $d_{j4} - d_{j3}$ is increased threefold (*3*).

Acknowledgments

This work was supported, in part, by RFBR Grant No. 02–05–64486, INTAS–01–0239, and CRDF RG2–2357–TO–02 Grants.

References

1. A.I. Isakova and G.M. Igonin, in: *Abstracts of Reports at the First Inter-Republic Symp. on Atmospheric and Ocean Optics*, Tomsk (1994), Part 2, pp. 126–127.

2. G.N. Glazov, G.M. Igonin, and D.M. Leshchinskii, Opt. Atm. **1**, No. 5, 71–76 (1988).
3. S.N. Volkov and A.I. Nadeev, Atmos. Oceanic Opt. **8**, No. 5, 624–626 (1995).
4. S.N. Volkov, B.V. Kaul', V.A. Shapranov, and D.I. Shelefontyuk, Atmos. Oceanic Opt. **13**, No. 8, 702–706 (2000).
5. J.-L. Zarader, A. Dabas, P.H. Flamant, B. Gas, and O. Adam, IEEE Transaction on Geoscience and Remote Sensing **37**, No. 6, 2678–2691 (1999).
6. M. Golub, http://www.osp.ru/school/1999/03/05.htm
7. V. Geppener, A. Lanne, and D. Chernichenko, http://www.chipinfo.ru/literature/chipnews/200006/2.html
8. S.A. Terekhov, *Wavelets and Neural Networks*, http://alife.narod.ru/lectures/wavelets2001.
9. V.P. D'yakonov, *Wavelets. From Theory to Practice* (Solon-R, Moscow, 2002), 440 pp.
10. A.V. Pereberin, Vychislitel'nye Metody i Programmirovanie **2**, No. 2, 133–158 (2001).
11. S.G. Mallat, Trans. Am. Math. Soc. **315**, No. 1, 69–87 (1989).
12. B. Jawerth and W. Sweldens, "*An overview of wavelet based multiresolution analyses*," The Wavelet Digest V. 2, Issue 2 maxwell.math.sc.edu/pub/wavelet/papers//overview.ps
13. I. Daubechies, Comm. Pure Appl. Math. **41**, 909–996 (1988).
14. L.V. Novikov, Nauchnoe Priborostroenie **10**, No. 3, 57–64 (2000).
15. V. Spiridonov, Computerra, No. 8 (1988).
16. K.A. Alekseev, http://www.matlab.ru/wavelet/book5/index.asp.
17. D.L. Donoho, IEEE Trans. on Inform. Theory **41**, No. 3, 613–627 (1995).
18. D.L. Donoho and I.M. Johnston, Bernoulli, No. 1, 39–62 (1996).
19. V. Geppener, A. Lanne, and D. Chernichenko, http://www.chipinfo.ru/literature/chipnews/200007/16.html
20. A. Kiselev, www.basegroup.ru/filtration/wavelet_applications.htm
21. A. Kiselev, www. basegroup.ru/filtration//wavelet_utils.htm
22. I.M. Dremin, O.V. Ivanov, and V.A. Nechitailo, Usp. Fiz. Nauk **171**, No. 5, 465–501 (2001).
23. Signals and Noise in Lidar (Radio i Svyaz', Moscow, 1985), 264 pp.
24. Yu.S. Balin, I.V. Samokhvalov, and V.S. Shamanaev, in: Laser Sensing of the Atmosphere (Nauka, Moscow, 1976), pp. 118–121.
25. A. Cohen, I. Daubechies, and J. Feauveau, Comm. Pure Appl. Math. **45**, 485–560 (1992).
26. D.L. Donoho and I.M. Johnston, Biometrika **81**, No. 3, 425–455 (1994).
27. A. Antoniadis and J. Fan, J. of the American Statistical Association **96**, No. 455, 939–965 (1999).
28. C. Stein, Annals of Statistics **9**, 1135–1151 (1981).