2010 3rd International Conference on Computer and Electrical Engineering (ICCEE 2010)
IPCSIT vol. 53 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V53.No.1.29

Random Noise Attenuation Based on Support Vector Regression and Adaptive Wiener Filtering

Xiaoying Deng[1+] and Yong Luo[2]

[1]Department of Electronic Engineering, Beijing Institute of Technology, Beijing, China
[2]Military Representative Office of PLA, Twelfth Research Institute of CEC, Beijing, China

Abstract. For suppressing the random noise of the seismic data, we propose a composite method compounded of the 1-D least squares support vector regression (LS-SVR) and the 2-D adaptive Wiener filtering. The former can efficiently suppress the strong noise and recover the desired seismic wavelets trace by trace from time domain, but it does not consider the space-domain information; the latter can make full use of the correlations of multi-trace data from time-space domain, but for the case of the strong noise it works less well. The combination of two methods has complementary advantages. By simulation experiments, we find that compared with the results obtained by using the 1-D LS-SVR and the 2-D adaptive Wiener filtering separately, the results obtained by using the combined method have much higher signal-to-noise ratio (SNR) and better waveform quality.

Keywords: Support vector regression, adaptive Wiener filtering, random noise reduction

1. Introduction

For the raw seismic data, existences of random noise degrade the record not to identify the useful information. Noise reduction is a very important task. All kinds of the denoising approaches have been proposed and applied to the practical processing. These methods can be classified into two main types: time-domain methods and transform-domain methods. In the former, typical methods include stacking [1], polynomial fitting [2], median filtering [3], singular value decomposition [4], Wiener filtering [5], etc. In the latter, typical methods include wavelet transform [6], seislet transform [7], curvelet transform [8], K-L transform [9], f-x prediction filtering [10], etc. All of these methods have their own characteristics and applicable scopes.

Support vector machine (SVM) is a new powerful learning machine based on the statistical learning theory [11]. The SVM has been widely applied to classification, regression, prediction, recognition, etc. Some researchers have used the SVM to eliminate the noise, such as [12], [13]. We have proposed to use the least squares support vector regression (LS-SVR) based on the Ricker wavelet kernel for random noise reduction of the seismic data [14]. It is a 1-D denoising method. However, we know that the seismic data have the strong correlations in time-space domain. Theoretically, a 2-D method may work better. In order to remedy the insufficiency of 1-D LS-SVR, we propose a composite method made up of the 1-D LS-SVR and the adaptive Wiener filtering in this paper. And we compare the combined method with the individual methods by experiments.

[+] Corresponding author.
E-mail address: xydeng@bit.edu.cn.

2. Theoretical Basis

Firstly, we briefly introduce the framework of the LS-SVR and the adaptive Wiener filtering for signal processing, respectively. Then give the composite scheme.

2.1 LS-SVR

Given a training data set of N points $\{t_i, x_i\}_{i=1}^{l}$, where x_i corresponds to the sampling value of the signal at the time t_i, the optimization problem of the LS-SVR [15] can be formulated as

$$\min\left\{\frac{1}{2}\|\boldsymbol{\omega}\|^2+\frac{\gamma}{2}\sum_{i=1}^{l}e_i^2\right\}$$
$$\text{s. t. } x_i=\boldsymbol{\omega}\cdot\varphi(t_i)+b+e_i\,(i=1,\cdots,l) \tag{1}$$

where $\varphi(t_i)$ is a function that maps the input space into a higher dimensional feature space, $\boldsymbol{\omega}$ is the weight vector, e_i is the error variable, γ is a preset regularization parameter, and b is a bias term. The Lagrangian function of the above problem is

$$L(\boldsymbol{\omega},e,b,\alpha)=\frac{1}{2}\|\boldsymbol{\omega}\|^2+\frac{\gamma}{2}\sum_{i=1}^{l}e_i^2-\sum_{i=1}^{l}\alpha_i\left[\boldsymbol{\omega}\cdot\varphi(t_i)+b+e_i-x_i\right] \tag{2}$$

with Lagrange multipliers $\alpha_i(i=1,2,\cdots,l)\in R$. Using the optimality conditions of problem (2), we obtain

$$b=\frac{\boldsymbol{1}^T(\boldsymbol{A}^{-1})^T\boldsymbol{X}}{\boldsymbol{1}^T\boldsymbol{A}^{-1}\boldsymbol{1}},\quad \boldsymbol{\alpha}=\boldsymbol{A}^{-1}(\boldsymbol{X}-\boldsymbol{1}b), \tag{3}$$

where $\boldsymbol{A}=\boldsymbol{\Omega}+\gamma^{-1}\boldsymbol{E}, \Omega_{ij}=\varphi(t_i)\cdot\varphi(t_j)=K(t_i,t_j), K(t_i,t_j)$ is a kernel function that satisfies Mercer's condition [16], $\boldsymbol{E}$ is a unit matrix, $\boldsymbol{X}=[x_1,x_2,\cdots,x_l]^T$, $\boldsymbol{1}=[1,1,\cdots,1]^T$, and $\boldsymbol{\alpha}=[\alpha_1,\alpha_2,\cdots,\alpha_l]^T$. Then the regression function can be expressed as

$$f(t)=\sum_{i=1}^{l}\alpha_i K(t_i,t)+b\,. \tag{4}$$

Next we compute the regression outputs on all training data by using the regression function (4). We have

$$\begin{aligned}\boldsymbol{Y}&=\left[f(t_1),f(t_2),\cdots,f(t_l)\right]^T\\&=\boldsymbol{\Omega\alpha}+b\boldsymbol{1}=\boldsymbol{\Omega}\left[\boldsymbol{A}^{-1}(\boldsymbol{X}-\boldsymbol{1}b)\right]+\boldsymbol{1}b\\&=\left[\boldsymbol{\Omega A}^{-1}-(\boldsymbol{\Omega A}^{-1}-\boldsymbol{E})\boldsymbol{1}\frac{\boldsymbol{1}^T(\boldsymbol{A}^{-1})^T}{\boldsymbol{1}^T\boldsymbol{A}^{-1}\boldsymbol{1}}\right]\boldsymbol{X}.\end{aligned} \tag{5}$$

Let $\boldsymbol{M}=\left[\boldsymbol{\Omega A}^{-1}-(\boldsymbol{\Omega A}^{-1}-\boldsymbol{E})\boldsymbol{1}\frac{\boldsymbol{1}^T(\boldsymbol{A}^{-1})^T}{\boldsymbol{1}^T\boldsymbol{A}^{-1}\boldsymbol{1}}\right]_{l\times l}$, where $\boldsymbol{M}$ is the filtering matrix which depends only on the kernel matrix $\boldsymbol{\Omega}$. Then we may directly get the filtering output $\boldsymbol{Y}=\boldsymbol{MX}$ from input $\boldsymbol{X}=[x_1,x_2,\cdots,x_l]^T$. In this case, we may precompute the filtering matrix $\boldsymbol{M}$ by selecting appropriate parameters according to the input signal to be processed. Then obtain the output by using a multiplication by the input signal.

2.2 Adaptive wiener filtering

The adaptive Wiener filtering [17] is a two-dimensional noise-removal filtering method. It uses a pixelwise adaptive processing method based on statistics estimated from a local neighborhood of each pixel $a(n_1,n_2)$ of an image. Firstly the adaptive Wiener filtering estimates the local mean and variance around each pixel

$$\begin{aligned}u&=\frac{1}{NM}\sum_{n_1,n_2\in S}a(n_1,n_2),\\\sigma^2&=\frac{1}{NM}\sum_{n_1,n_2\in S}a^2(n_1,n_2)-\mu^2,\end{aligned} \tag{6}$$

where S is the N-by-M local neighborhood of each pixel in the image. Then creates a pixelwise output using these estimates

$$b(n_1,n_2)=\mu+\frac{\sigma^2-v^2}{\sigma^2}\left[a(n_1,n_2)-\mu\right], \tag{7}$$

where v^2 is the noise variance. If the noise variance is not known, use the average of all the local estimated variances.

The adaptive Wiener filtering is a type of linear filter. And this approach, tailoring itself to the local image variance, often produces better results than linear filtering. It is more selective than a comparable linear filter, preserving edges and other high-frequency parts of an image.

2.3 Composite method

For seismic data, in terms of single trace, every-trace data contains information representing different underground layers. From a single-trace point of view, the 1-D LS-SVR based on the Ricker wavelet kernel [14] can reduce the strong random noise and enhance the desired seismic wavelets trace by trace. But this method only uses the time-domain information, not considering the space-domain information. On the other hand, from a multi-trace point of view, the seismic data have very strong correlation and continuity. Where the variance is large, the 2-D adaptive Wiener filtering performs little smoothing, and where the variance is small, it performs more smoothing. So when the additive noise is very strong, the adaptive Wiener filtering works less well. From the above, we combine the two methods to expect to achieve better denoising performance.

3. Simulation Experiments

In this section, we investigate the performance of the proposed method and compare with the single 1-D LS-SVR method and 2-D adaptive Wiener filtering. Here for the 1-D LS-SVR, the kernel parameter of the Ricker wavelet kernel is set as 30 and the regularization parameter is set as 1. For the 2-D Wiener filtering, the window size is set as 3-by-3. For the composite method, the same parameters as the above are selected.

According to the parameters listed in Table I, we synthesize a seismic record as shown in Figure 1a. The record includes 60 traces, 3 reflection events and geophone interval 50m. By adding the Gaussian white noise with different levels trace by trace on the initial seismic record, we obtain noisy records with different SNRs. By using the above two methods and the composite method, we list the SNRs before and after denoising and mean square errors (MSE) as shown in Table II. From Table II, compared with the results obtained by using the LS-SVR and Wiener filtering separately, the combined method can obtained much higher SNR (4dB or so higher than LS-SVR, and 7dB or so higher than Wiener filtering) and lower MSE.

Figure 1 shows an example on random noise suppression corresponding to the third case in Table II, i.e. the line with SNR before denoising -0.57 dB. From Figure 1b, the strong random noise nearly masks the useful events, and the seismic wavelets are distorted so as to not be recognized. But after denoisng by using different methods, the three events have became visible, especially the result obtained by using the

Table 1 Parameters Used to Generate Synthetic Record

Event	t_0 (s)	v (m/s)	f (Hz)	A (V)
1	1	1800	32	1
2	1.5	2200	30	0.9
3	1.58	2250	28	0.8

Table 2 Comparisons for SNR and MSE

SNR before denoising (dB)	SNR after denoising (dB)			MSE(V^2)		
	LS-SVR	*Wiener*	*combined*	*LS-SVR*	*Wiener*	*combined*
4.18	14.20	11.48	18.19	0.006	0.012	0.0027
1.18	11.45	8.18	15.02	0.012	0.025	0.0053
-0.57	9.93	6.53	13.12	0.018	0.038	0.0081
-1.82	8.90	5.53	11.78	0.024	0.051	0.011

composite method as shown in Figure 1e. And the noise remained in Figure 1e is the weakest. Comparably, the adaptive Wiener filtering works too badly, and the result in Figure 1c has remained much random noise.

Next we randomly select a trace, such as trace 15, and detailedly compare the recovered seismic wavelets. Figure 2 draws the corresponding 15th trace in Figure 1, respectively. We can see that the three seismic wavelets in Figure 2e are most approximate to the initial signals, and the noise energy remained is little.

In addition, from the spectra of trace 15 as shown in Figure 3, the spectrum of the result obtained by using the LS-SVR basically reduces the low-frequency and high-frequency noise, but it does nothing in the dominant-frequency band [10, 70] Hz. However, the spectrum of the result obtained by using the composite method not only reduces low-frequency and high-frequency noise, but also adjusts those frequency components in the dominant-frequency band.

From the above, the performance of the combined method greatly excels the single LS-SVR and the single Wiener filtering.

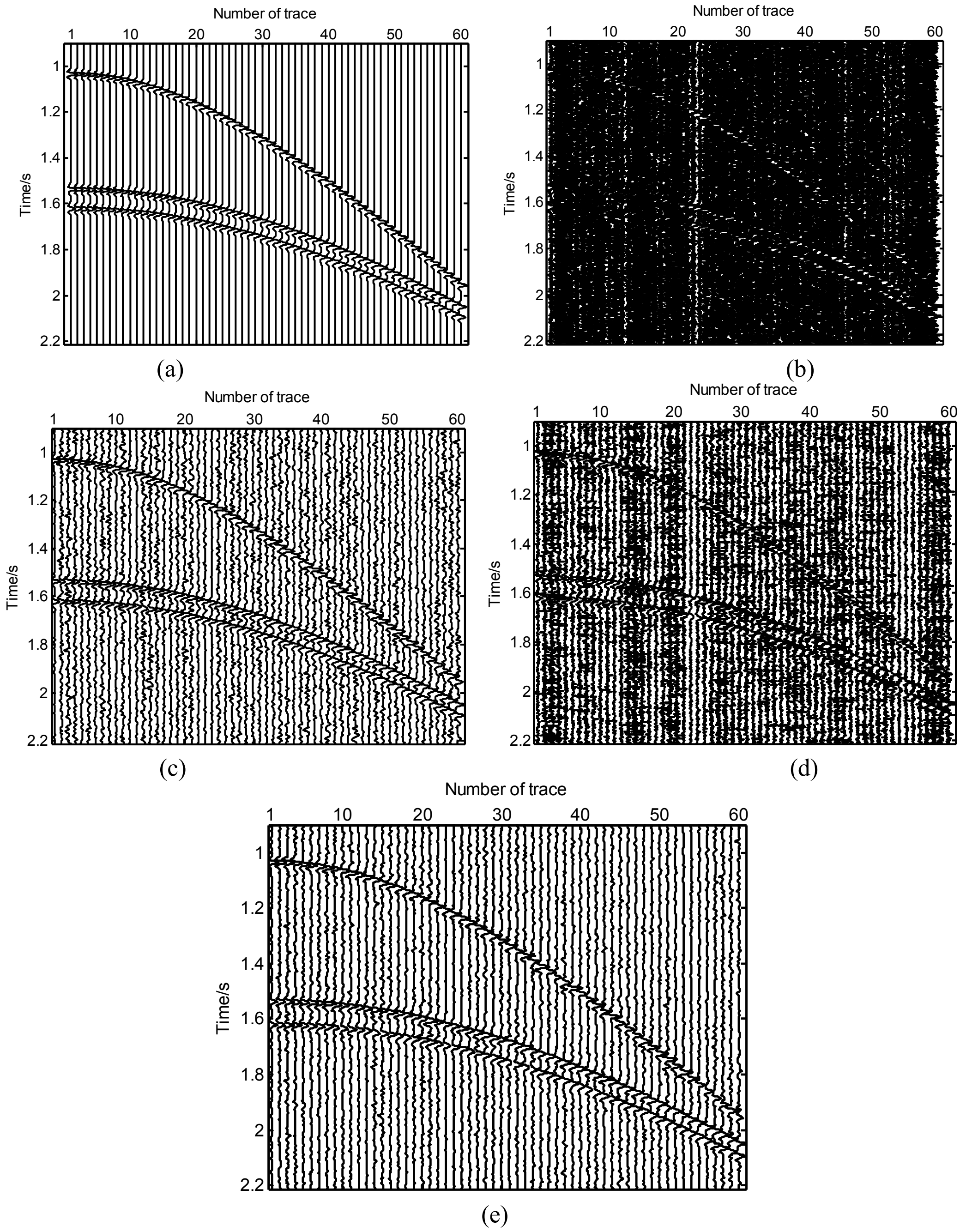

Figure 1. An example on random noise suppression: (a) initial record; (b) noisy record; (c) result obtained by using the 1-D LS-SVR; (d) result obtained by using the adaptive Wiener filtering; (e) result obtained by using the composite method.

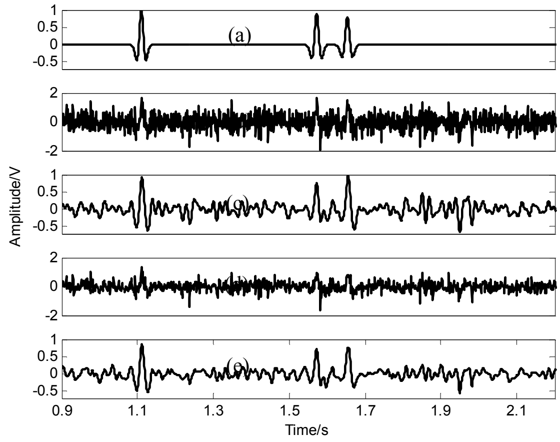

Figure 2. Trace 15 corresponding to Figure 1: (a) initial trace; (b) noisy trace; (c) result obtained by using the 1-D LS-SVR; (d) result obtained by using the adaptive Wiener filtering; (e) result obtained by using the composite method.

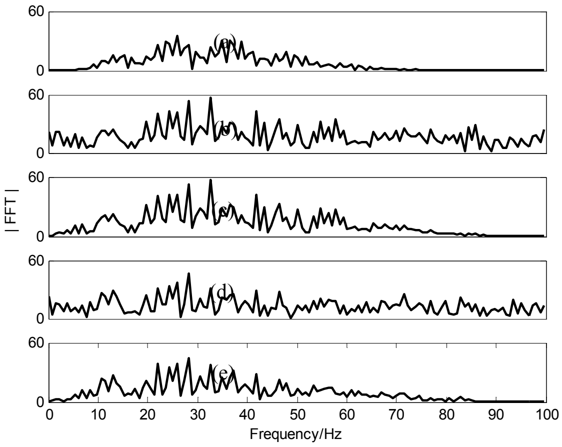

Figure 3. Spectra of Trace 15 corresponding to Figure 2: a) initial trace; (b) noisy trace; (c) result obtained by using the 1-D LS-SVR; (d) result obtained by using the adaptive Wiener filtering; (e) result obtained by using the composite method.

4. Discussion and Conclusion

The LS-SVR deals with the seismic data trace by trace without the limit of the good continuity of events. It can recover the seismic wavelets well. The adaptive Wiener filtering essentially belongs to a smoothing method, and works best when the noise is constant-power additive noise, such as Gaussian noise. But when the noise is very strong, its performance will decline rapidly. So we combine the above two methods. Firstly by using the LS-SVR, the SNR of seismic data is improved. Then by using the Wiener filtering, the noise in time-space domain is removed, and the SNR is improved further. Note that we should first apply the LS-SVR to the seismic data, and then apply the Wiener filtering, but not vice versa.

By the simulation experiments in this paper, the combined method can greatly improve the SNR of seismic data, and works much better than the 1-D LS-SVR and 2-D adaptive Wiener filtering used separately. So the combined method is an efficient denoising method, and may be applied to the random noise removal of the seismic data, which will lay a foundation for the succeeding processing and identification.

5. Acknowledgment

This work was supported by the national natural science foundation of China under Grant No. 40804022.

6. References

[1] R. G. Anderson, and G. A. Mcmechan, “Weighted stacking of seismic data using amplitude-decay rates and noise amplitudes,” Geophysical Prospecting, vol. 38, pp. 365–380, May 1990.

[2] S. Yu, X. Cai, and Y. Su, “Seismic signal enhancement by polynomial fitting,” Applied Geophysics, vol. 1, pp. 57-65, January 1989.

[3] Y. Liu, C. Liu, and D.Wang, “A 1D time-varying median filter for seismic random, spike-like noise elimination,” Geophysics, vol.74, pp. V17–V24, January 2009.

[4] W. Lu, “Adaptive noise attenuation of seismic images based on singular value decomposition and texture direction detection,” Journal of Geophysics and Engineering, vol.3, pp. 28-34, January 2006.

[5] J. S. Lim, Two-Dimensional Signal and Image Processing, Englewood Cliffs, NJ, Prentice Hall, pp. 536-540, 1990.

[6] S. Y. Cao, and X. P. Chen, “The second-generation wavelet transform and its application in denoising of seismic data,” Applied geophysics, vol. 2, pp. 70–74, June 2005.

[7] S. Fomel, and Y. Liu, “Seismic data analysis with one-dimensional seislet frame,” Expanded Abstracts, 78th Annual International Meeting, SEG, 2008, pp. 2581–2585.

[8] R. Neelamani, , A. I. Baumstein, D. G. Gillard, M. T. Hadidi, and W. I. Soroka, “Coherent and random noise attenuation using the curvelet transform,” The Leading Edge, vol. 27, pp. 240–248, 2008.

[9] I. F. Jones, and S. Levy, “Signal-to-noise ratio enhancement in multichannel seismic data via the Karhunen-Loeve transform,” Geophysical Prospecting, vol. 35, pp. 12-32, 1987.

[10] L. L. Canales, “Random noise reduction,” Expanded Abstracts, 54th Annual International Meeting, SEG, 1984, pp. 525–527.

[11] V. Vapnik, The Nature of Statistical Learning Theory, New York: Springer-Verlag, 1995.

[12] B.Y. Sun, D.S. Huang, and H.T. Fang, “Lidar signal denoising using least-squares support vector Machine,” IEEE Signal process. lett. vol. 12, pp. 101–104, February 2005.

[13] H. Cheng, J.W. Tian, J. Liu, and Q.Z. Yu, “Wavelet domain image denoising via support vector regression,” Electronics lett. vol. 40, pp. 1479–1480, November 2004.

[14] X. Deng, D.Yang, and J. Xie, “Noise reduction by support vector regression with Ricker wavelet kernel,” Journal of Geophysics Engineering, vol.6, pp.177–188, May 2009.

[15] J.A.K. Suykens and J. Vandewalle, “Least squares support vector machine classifiers,” Neural Process. Lett. pp.293–300, June 1999.

[16] J. Mercer, “Functions of positive and negative type and their connection with the theory of integral equations,” Philosophical Trans. of the Royal Society of London Series A 209, pp. 415–446, October 1909.

[17] The MathWorks Inc. MATLAB, The language of Technical Computing, Image Processing Toolbox, Version 7.0.0.19920 (R14), 2004.