Adaptive FIR Filter Use for Signal Noise Cancelling

M. Kolínová, A. Procházka, and M. Mudrová
Prague University of Chemical Technology
Department of Computing and Control Engineering
Technická 1905, 166 28 Prague 6, Czech Republic
Phone: 00420-2-2435 4198 ⋆ Fax: 00420-2-2431 1082
E-mail: Magdalena.Kolinova@vscht.cz, A.Prochazka@ieee.org,
Martina.Mudrova@vscht.cz

Abstract - Noise reduction, together with signal modelling, classification and prediction, forms a basic problem in many engineering, biomedical and econometric systems. This paper summarizes basic adaptive methods based upon the application of artificial neural networks simplified in many cases to an adaptive linear element only. Adaptive noise cancelling algorithms by such Finite Impulse Response (FIR) filters are studied in the first part of the paper both for the case that a reference signal is known and for the reference-free structures. Resulting signals are compared with those obtained by noise rejection in wavelet transform domain either based upon signal decomposition and reconstruction using properly chosen threshold levels for signal modification or incorporating adaptive FIR filtering. Signals preprocessed in this way are then used for their segmentation and classification to verify the effect of signal denoising. Methods presented in the paper are used for simulated signals at first and then applied to the processing of real signals representing physiological signals and technological time series.

1 Introduction

Signal modelling [12, 20], reconstruction [4] and noise reduction form fundamental problems in various applications including physiological signal analysis and both engineering and speech signal processing. The theory of non-adaptive noise reduction is based upon classical methods of digital filters and spectral analysis [18]. Further modern adaptive methods use both linear and non-linear structures of neural networks [3, 7, 8, 25]. This paper is restricted to application of the finite impulse response filter forming an adaptive linear element. While the use of this method based upon the known reference signal is described in many papers [26], it is not so common yet to apply this approach in the case that the reference signal is not known [7]. Here we deal with signal preprocessing, segmentation, feature extraction, and classification [17, 22]. Only one way of signal preprocessing is shown here (noise reduction), as well as its influence on signal classification. However, additive noise may be desirable in some cases, see [1] for this and for other ways of signal preprocessing. As for classification, the requirement is to classify the signal segments into a finite number of classes such that the average probability of misclassification is minimized. The solution to the classification problem may be obtained by neural networks using a Bayesian approach for signal segmentation [8].

In this paper, the mathematical analysis of the problem mentioned above is presented, together with the design of an adaptive processor and its coefficients optimization [6]. Results are then compared with those obtained by wavelet signal decomposition and reconstruction [16, 19, 21, 24].

2 Preprocessing of Real Signals

Signal preprocessing is fundamental in many applications of time series analysis, modelling, classification, and prediction. This paper is devoted to signal noise cancelling only to allow the following signal segmentation, feature extraction, and classification. Methods of signal denoising are studied both for simulated signals and for two different real time series. Their initial spectrum analysis is performed by the Discrete Fourier Transform method (DFT) defined by the relationship

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\, nk2\pi/N}\,, \quad k = 0, \ldots, N-1 \qquad (1)$$

for a length-N complex data sequence.

2.1 EEG Signals

The Electroencephalograph (EEG) signals are quite complex to analyze, and there is still a lot to be done in EEG signal processing. EEG signals are of chaotic nature; they have a dominant deterministic chaotic component and also a noise component. It is not clear whether the noise component is intrinsic in the EEG itself or whether it is due to the measurements [23].

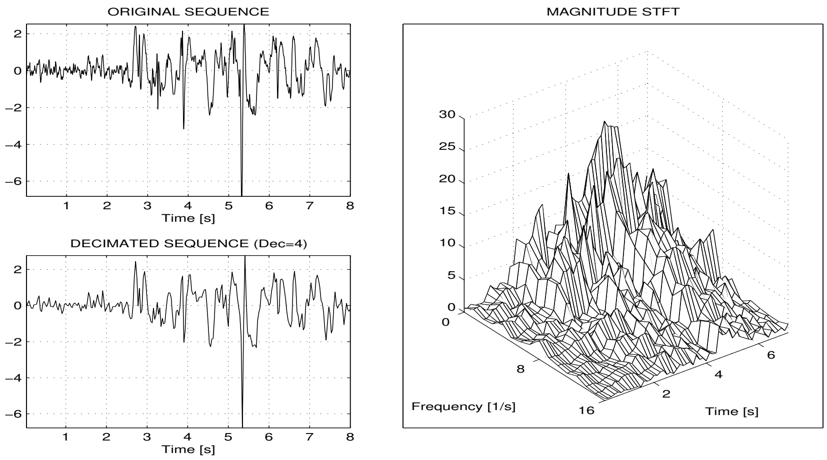

Figure 1: Evolution of frequency components of the EEG signal obtained by the STFT method. Number of values: 1024; decimation: 4; sampling frequency: 128 Hz; final frequency: 32 Hz

The original sampling of the EEG signal processed here was $128Hz$. The signal was denoised and decimated (by 4), and then it was processed using the Short Time Fourier Transform method (STFT) with the application of a sliding window of constant length (32 samples). A part of the signal is shown in Fig. 1.

2.2 Gas Consumption Signals

In Fig. 2, a part of the total normalized gas consumption in the Czech Republic is shown, measured with periodicity of two hours. Applying the DFT method to the data sequence, we find periodicity of one day and one week. As there is no significant evolution of frequency components, it is not necessary to apply the STFT method.

In the case of the following signal classification, it is necessary to find signal change-points in both cases. While this is relatively simple in this case, it is highly sensitive to the choice of the threshold level in biomedical signals.

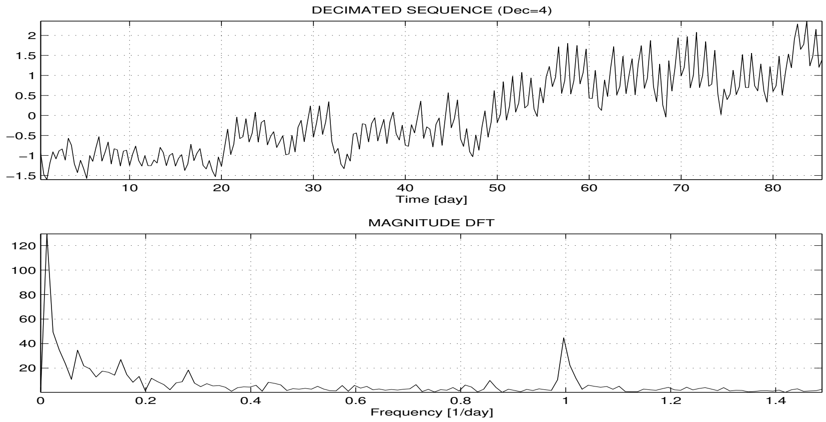

Figure 2: Results obtained by the STFT method applied to the gas consumption data set. Number of values: 512; decimation: 1; sampling

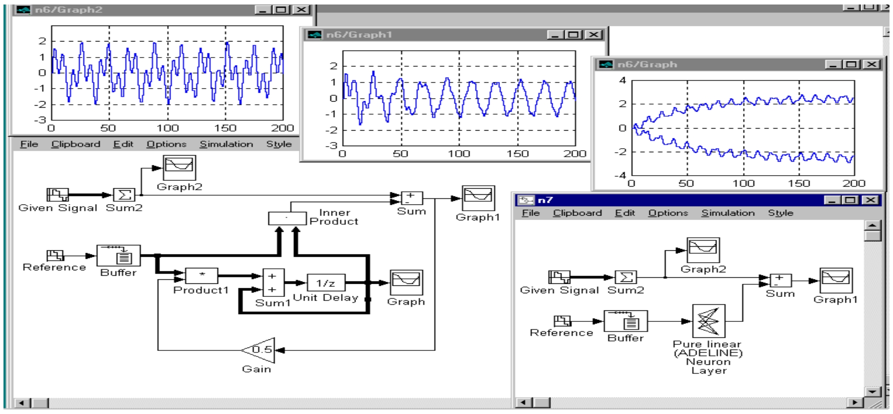

Figure 3: Application of the Widrow method to noise rejection and results obtained by using a linear adaptive element [14] with two coefficients to process the sum of two harmonic functions

3 Adaptive Noise Cancelling by FIR Filters

3.1 Reference Signal Use

The principle of the learning process of an adaptive element can be explained by using an algorithm presented in Fig. 3 for a given signal $\mathbf{Y} = [y_1, y_2, \cdots, y_N]'$ and for a reference signal having the form of a signal correlated with the observed signal noise component $\mathbf{G} = [g_1, g_2, \cdots, g_N]'$. Assuming an adaptive processor with its initial coefficients $\mathbf{W} = [w_1, w_2, \cdots w_R]$, it is possible to apply the mean square method to minimize the function

$$S = \sum_{n=1}^{N} (y_n - (w_1 g_n + w_2 g_{n-1} + \cdots + w_R g_{n-R+1}))^2 \tag{2}$$

The following gradient method applied to the last square error only results in the final iterative algorithm of the form

$$\mathbf{W}_n = \mathbf{W}_{n-1} - \alpha \frac{\partial S}{\partial \mathbf{W}_{n-1}} = \mathbf{W}_{n-1} + 2\,\alpha \mathbf{P}'_n\,(y_n - \mathbf{W}_{n-1} * \mathbf{P}_n) \tag{3}$$

for a chosen learning rate α and a pattern matrix $\mathbf{P}_n$ defined by the reference signal $\mathbf{G}$ as

$$\mathbf{P}_n = [g_n, g_{n-1}, \cdots, g_{n-R+1}]' \tag{4}$$

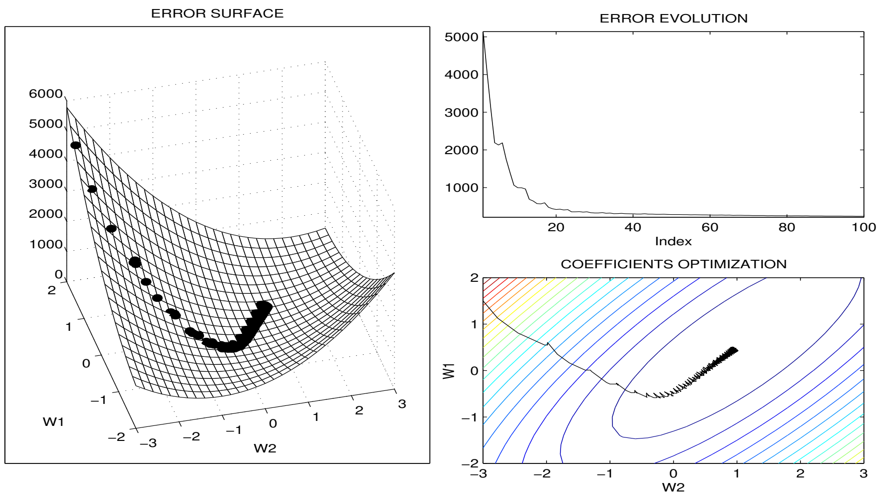

Figure 4: Optimization of the adaptive linear element with two coefficients for the given sequence $y(n) = sin(2\pi 0.05 n) + (0.5 r(n) + r(n-1))$ where $r(n) = sin(2\,\pi\,0.4\,n)$ and the learning rate $\alpha = 0.1$

An example of results of such an adaptive process is presented in Fig. 3, together with the final neural network structure that can be applied in the case that the learning process is complete and no further adaptation is necessary. Fig. 4 presents the optimization process for an artificial sequence and for a linear adaptive element with two coefficients. A similar principle can be used for real signal processing as well.

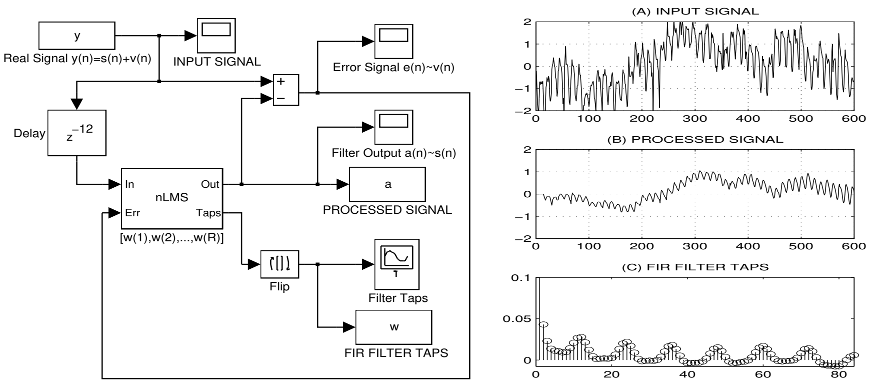

Figure 5: On the left, the principle of adaptive signal denoising using a reference signal identical to the delayed input signal is shown. The right part of the figure shows application of this principle to gas consumption signal processing: (A) the original signal, (B) the signal processed by a linear adaptive element with α being 0.01, and (C) filter taps

3.2 Noise Cancelling without a Reference Signal

In the case that the reference signal is not known, it is possible to use the delayed observed signal instead. While the output of the adaptive filter represents the noise estimate, in the original case it stands for the estimate of the observed time series for the case of noise cancellation without a reference signal [7]. The basic principle of this approach is given in Fig. 3.2 for an observed time series $\{y(n)\}$ composed of the original signal $\{s(n)\}$ and of an additive noise $\{v(n)\}$. In Fig. 3.2 (C) it can be noticed that the most important samples are those delayed by multiples of one day (12 samples).

The method presented above has been applied to processing of real signals, including both biomedical time series and engineering signals. Fig. 3.2 presents results of gas consumption in the Czech republic measured with the sampling period of two hours.

4 Wavelet Transform Domain Noise Rejection

Wavelet Transforms (WT) provide alternatives to the STFT method for the case of non-stationary signal analysis, and both STFT and WT result in decomposition of the signal into a two-dimensional function [2, 9, 10, 13, 15]. In the case of a continuous signal $x(t)$, the continuous WT coefficients are given by the relationship

$$C(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\; W(\frac{t-b}{a})\; dt \tag{5}$$

where $W(t)$ is the basic (mother) wavelet function, and values a and b stand for its translation and dilation, respectively. The discrete WT assumes evaluation of wavelet coefficients for discrete parameters of dilation $a = 2^m$ and translation $b = k\; 2^m$, using only the initial (mother) wavelet function $W(t)$ localized at a particular position.

Wavelet signal decomposition and reconstruction can be used very efficiently for noise cancelling in many applications. In the case we use certain, possibly modified decomposition levels, it becomes possible to extract signal components only. Using the threshold method introduced in [24], it is further possible to reject the noise and to enlarge the signal to the noise ratio. Wavelet coefficients are modified using a chosen thresholding level δ to determine new coefficients

$$\overline{C}(j,k) = \begin{cases} \text{sign } C(j,k)\; (\mid C(j,k) \mid -\delta) & \text{if} \mid C(j,k) \mid > \delta \\ 0 & \text{if} \mid C(j,k) \mid \leq \delta \end{cases}$$

Fig. 6 provides selected results for the processing of a real signal representing gas consumption mentioned above.

5 Noisy Signals Classification

Signal denoising is used in many cases as the initial step for signal decimation and prediction but the resulting signal can serve for signal classification as well. Fig. 7 presents results of the EEG signal segmentation [5, 11] and classification into three classes using a competitive neural network structure applied to features representing average signal frequency components in given frequency ranges. As a measure of classification results the variation of signal features related to the final class characteristics has been used. Typical results summarized in Table 1 show the effect of signal preprocessing. In the case of selecting other features based upon segments autoregressive modelling, signal preprocessing seems to be even more important.

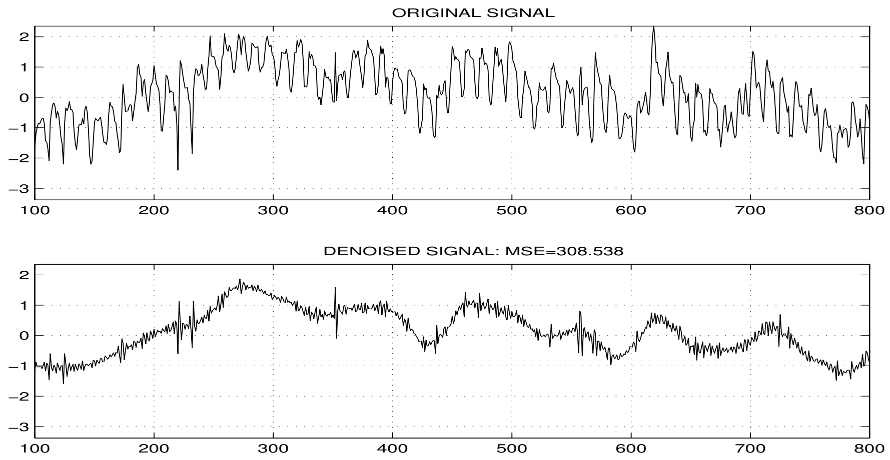

Figure 6: Signal denoising based upon the wavelet transform signal decomposition and reconstruction

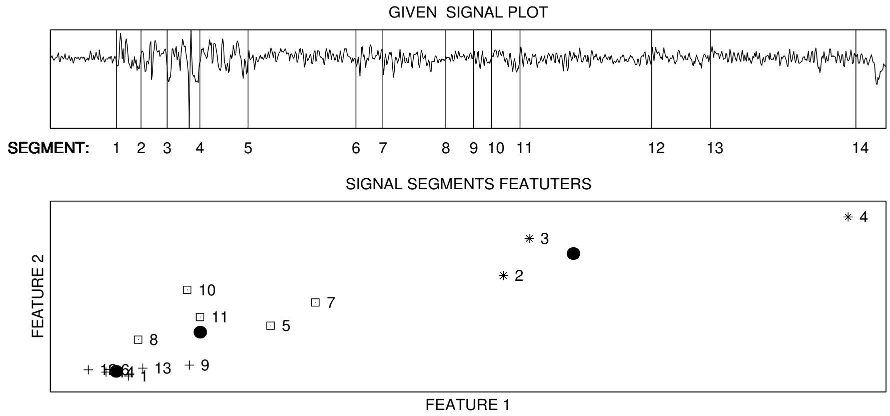

Figure 7: Results of the EEG signal segmentation and of its classification into three classes using features derived as average frequency components in given bands

Table 1: Variation of signal features in given signal classes

Signal	*Class A*	*Class B*	*Class C*
Original EEG signal	0.05	0.21	0.36
Denoised EEG signal	0.05	0.17	0.32

6 Conclusion

In this paper, use of the adaptive linear element to reduce the noise in an observed time series is presented, and possibilities of both the classical Widrow method and noise reduction without a reference signal are compared. It is possible to verify that in the case of the delayed process uncorrelated with the noise, the latter method can be used instead of the reference signal. An observed sequence can then be processed in the real time, even though segment-based noise reduction using for instance wavelet functions provides better results in most cases.

7 Acknowledgement

The paper has been supported by the Department of Computing and Control Engineering of the Prague University of Chemical Technology.

References

[1] Ch.M. Bishop. *Neural Networks for Pattern Recognition.* Clarendon Press, Oxford, 1995.

[2] A. Bruce, D. Donoho, and H.-Y. Gao. Wavelet Analysis. *IEEE Spectrum*, 33(10):26–35, October 1996.

[3] H. Demuth and M. Beale. *Neural Network Toolbox.* The MathWorks, Inc., Natick, Massachusetts 01760, January 1994.

[4] W. Etter. Restoration of Discrete-Time Signal Segments by Interpolation Based on the Left-Sided and Right-Sided Autoregressive Parameters. *IEEE Trans. on Signal Processing*, 44(5):1124–1135, May 1996.

[5] W.J. Fitzgerald, J.J.K. Ó Ruanaidh, and J.A. Yates. *Generalised Changepoint Detection.* University of Cambridge, 1994.

[6] M.T. Hagan and M.B. Menhaj. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, pages 989–993, November 1994.

[7] M.H. Hayes. *Statistical Digital Signal Processing and Modeling.* John Wiley and Sons, Inc., 1996.

[8] S. Haykin. *Neural Networks.* IEEE Press, New York, 1994.

[9] S. Hosur and A.H. Tewfik. Wavelet Transform Domain Adaptive FIR Filtering. *IEEE Trans. on Signal Processing*, 45(3):617–630, March 1997.

[10] Y.T. Chan. *Wavelet Basics.* Kluwer Academic Publishers, Boston, 1995.

[11] M. Lavielle. Optimal Segmentation of Random Processes. *IEEE Trans. on Signal Processing*, 46(5):1365–1373, May 1998.

[12] M.D. Macleod. Fast Nearly ML Estimation of the Parameters of Real or Complex Single Tones or Resolved Multiple Tones. *IEEE Trans. on Signal Processing*, 46(1):141–148, January 1998.

[13] J. Magarey and N.G. Kingsbury. Motion Estimation Using a Complex-Valued Wavelet Transform. *IEEE Trans. on Signal Processing*, 46(4):1069–1084, April 1998.

[14] The MathWorks, Inc., Natick, Massachusetts 01760. *Simulink, DSP Blockset*, May 1997.

[15] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi. *Wavelet Toolbox*. The MathWorks, Inc., Natick, Massachusetts 01760, March 1996.

[16] D.E. Newland. *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*. Longman Scientific & Technical, Essex, U.K., third edition, 1994.

[17] D.E. Newland. *Time-Frequency and Time-Scale Signal Analysis, In Signal Analysis and Prediction*. Birkhauser, Boston, 1998.

[18] J.G. Proakis and D.G. Manolakis. *Digital Signal Processing*. Prentice Hall, 1996.

[19] A. Procházka and V. Sýs. Time Series Prediction Using Genetically Trained Wavelet Networks. In *Neural Networks for Signal Processing 3 - Proceedings of the 1994 Workshop*, pages 195–203. IEEE, 1994.

[20] P.J.W. Rayner and W.J. Fitzgerald. The Bayesian Approach to Signal Modelling and Classification. In *Signal Analysis and Prediction I*, pages 65–75. ICT Press, 1997.

[21] O. Rioul and M. Vetterli. Wavelets and Signal Processing. *IEEE SP Magazine*, pages 14–38, October 1991.

[22] R. Setiono and H. Liu. Neural-Network Feature Selector. *IEEE Trans. on Neural Networks*, 8(3):654–662, May 1997.

[23] J. Sim and W.J. Fitzgerald. *Measuring the Depth of Anaesthesia Using Methods of Non-Linear Dynamics*. University of Cambridge, 1993.

[24] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1996.

[25] B. Widrow. 30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation. *Proceedings of the IEEE*, pages 27–53, September 1990.

[26] B. Widrow. Adaptive Signal Processing and Adaptive Neural Networks. *IEEE Press*, August 1992.