A New Adaptive Approach to Remove Baseline Wander from ECG Recordings Using Madeline Structure

J Mateo[1], C Sánchez[1], C Vayá[1], R Cervigon[1], JJ Rieta[2]

[1]Innovation in Bioengineering Research Group, University of Castilla-La Mancha, Cuenca, Spain
[2]Biomedical Synergy, Valencia University of Technology, Valencia, Spain

Abstract

Nowadays, there exist different approaches to cancel out noise effect and baseline drift in biomedical signals. However, none of them can be considered as completely satisfactory. In this work, an artificial neural network (ANN) based approach to cancel out baseline drift in electrocardiogram signals is presented. The system is based on a grown ANN allowing to optimize both the hidden layer number of nodes and the coefficient matrixes. These matrixes are optimized following the Widrow - Hoff Delta algorithm, offering much lower computational cost that the traditional back propagation algorithm.

The proposed methodology has been compared with traditional baseline reduction methods (FIR, Wavelet and Adaptive LMS filtering) making use of cross correlation, signal to interference ratio and signal to noise ratio indexes. Obtained results show that the ANN-based approach performs better, with respect to baseline drift reduction and signal distortion at filter output, than traditional methods.

1. Introduction

The Electrocardiogram (ECG) is a graphical representation of the electrical activity of the heart that offers information about the state of the cardiac muscle. With Filtering techniques applied in ECG it would be possible to improve the diagnosis of some diseases of the heart and diverse pa-thologies [1]. The bandwidth of the acquiring system is usually from the 0.05Hz to 100Hz with almost linear response, causing no distortion of the pulse waveform. However, distortion may arise from the movement of the subject respiration

The baseline variations take place due to several factors, like the movement of the patient during the acquisition of the electrocardiogram, breathing, and changes in the impedance of the electrodes. These variations suppose an interference of low frequency and amplitude that must be reduced in order to not change the last results. The frequency breathing components are usually below 0,8 Hz. Some researchers attempt to suppress baseline wander with high pass filter, which introduces nonlinear phase distortion and the key points displacement. To preserve phase information, the symmetric FIR digital filter is designed, but it does very little to attenuate low frequency baseline wander [2, 3]. Other authors [4, 5] have used adaptive filter to remove the ECG baseline wander. Similarly, adaptive filters and Wiener filters have been of limited value in the absence of prior knowledge of the physiological signal and its baseline drift [6, 7]. The use of Wavelets [8, 9] has been another possibility in order to eliminate baseline noise, obtaining quite acceptable results until a frequency of 0.2Hz. The method proposed has not been applied in the cancellation of baseline noise in ECG signals yet. This system has two important advantages: to provoke low signal distortion and to reduce baseline noise. Besides, it can be applied to a wide range of signals.

2. Materials

The electrocardiography treated signals validation requires a set of signals which will have to cover the pathologies, leads, etc. For this study, two types of signals were used: real recordings from the PhysioNet Database [10], and synthetic signals. The sampling frequency used is 1kHz.

Table 1. Signals used for the study

	N^o of Registers	Time (seg)
Synthetic	200	1049
Real+noise	550	106

550 recordings with different pathologies have been obtained from PhysioNet with different types of QRS morphologies.

Synthetic signals with different noises have been generated making use of the ECGSyn software [10]. White (myoelectric, thermal, etc.) and baseline noise are included in these registers.

3. Methods

3.1. Neural networks

The neuronal perceptron multilayer network (MLP) using the algorithm of backpropagation has been applied to diverse practical problems [11]. Perceptron multilayers method consists of at least three layers: A hidden inpuy layer, one or more hidden layers and an output layer.

A way to consider the optimal number of nodes in the hidden layer is to stop the training after a certain number of iterations and to determine how many signals were filtered with the present number of neurons used in the hidden layer. If the result of this test is not satisfactory it will add one or more neurons in the hidden layer to improve the performance of the network. In these cases, the network must be completely trained [11].

An alternative that seems more attractive is the development of increasing networks in which nodes are added in the hidden layer in systematic form during the learning process. With this idea, diverse structures have been proposed such as the increasing network cascade correlation [12], as well as neural networks [13, 14, 15].

3.2. Proposed system

The proposed system consists initially in a structure similar to the neural network ADALINE (ADAptive LINear Element) [16], which is used like initial structure because it is simple and easy to optimize using the algorithm of square minimums average, LMS [13]. It has had initially an input layer, one hidden layers and an exit layer, where they will be added neurons in the intermediate layer.

When the network has converged, if the operation obtained by the system is not the required one, a neuron is added in the hidden layer. In this case, the weights (w) that connect the input layer with the nodes of the intermediate layer, are congealed (these weights have been previously trained). The gains (b) that connect the hidden layer with the exit layer are adapted, as well as the weights (w) that connect the input layer with the neuron added in the hidden layer. The ANN is adapted using the Widrow - Hoff Delta algorithm which has obtained good results.

This new structure has got a special characteristic: it grows while it learns. It means that the neurons in the hidden layer are added one to one and their weights and gains are adapted. Besides, the weights of the input layer are conserved in the learning network. This mechanism, although sometimes could produce neural networks with a sub-optimal number of neurons in the hidden layer, makes it possible to estimate the size of the network. The proposed system with two neurons added in the hidden layer can be observed in figure 1.

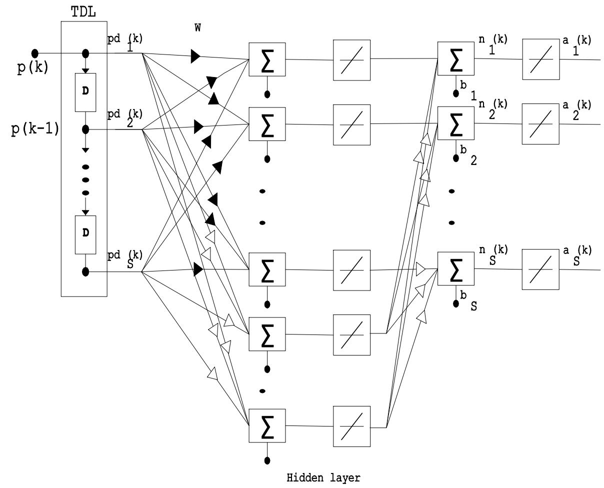

Figure 1. Proposed Neuronal Network with two neurons in the hidden layer. The black coefficients (W) are constants.

3.3. Learning algorithm using the Widrow-Hoff Delta rule

This network is a supervised learning network that needs to know the associated values in each input. The pairs of input/output are:

$$\{p_1, t_1\}, \{p_2, t_2\}, ..., \{p_Q, t_Q\} \quad (1)$$

where p_Q is the input to the network and t_Q is its corresponding wished exit, when a input p is presented to this network, the exit of the network is compared with the value of t (hoped exit) that is associated to him.

Adaptive LMS algorithm derives from the Widrow-Hoff Delta rule [17], a network Adaline, is deduced of the following way, according to the procedure described in Widrow [18, 19].

$$W(k+1) = W(K) + \alpha \frac{e(k)p(k)}{|p(k)|^2} \quad (2)$$

In which k shows the present iteration of the update process, $W(k+1)$ is the following value that will take the vector from weights and $W(k)$ is the present weights vector's value. Present error $e(k)$ is defined as the difference between wished answer $t(k)$ and the exit of network $a(k) = W^T(k)p(k)$ before the update:

$$e(k) = t(k) - W^T(k)p(k) \quad (3)$$

The variation of the error in each iteration is represented by:

$$\Delta e(k) = \Delta(t(k) - W^T(k)p(k)) = -p^T(k) * W(k) \quad (4)$$

The main characteristic of LMS algorithm is that safes the error and reduces the average quadratic error. In order to explain the quadratic mean error a network Adaline will be considered and an algorithm of approximated steps will be used, like Widrow and Hoff ; with this algorithm the function for the mean square error is:

$$e^2(k) = (t(k) - a(k))^2 \tag{5}$$

In equation (5), $t(k)$ shows the wanted exit in iteration k and $a(k)$ shows the exit of the proposed network; the square error has been replaced in iteration k, therefore in each iteration is had a gradient of the error of the following way:

$$\left[\nabla e^2(k)\right]_j = \frac{\partial e^2(k)}{\partial w_{i,j}} = 2e(k)\frac{\partial e(k)}{\partial w_{i,j}} \; para \; j = 1, 2, ...R \tag{6}$$

and

$$\left[\nabla e^2(k)\right]_{R+1} = \frac{\partial e^2(k)}{\partial b} = 2e(k)\frac{\partial e(k)}{\partial b} \tag{7}$$

The approach of $\nabla e(k)$ found in the equation (6) is replaced in the equation 2 that defines the process of update of weights for LMS algorithm; after the evaluation of partial derived the update process The weights (w) and gains (b) algorithm for this network Adaline are expressed:

$$w(k+1) = w(k) + 2\alpha e(k)p(k) \tag{8}$$

$$b(k+1) = b(k) + 2\alpha e(k) \tag{9}$$

4. Results

To evaluated the performance ofthe ANN filter, a clean pulse signal corrupted with different baselines is simulated. The obtained results of the ANN approach are compared with standard filtering techniques. The Butterworth high pass digital filter is a nonlinear phase filter, so the pulse waveform can be distorted. For this reason, a FIR filter using least squares error minimization has been selected as classical base line technique to compare with the proposed method. In this work, the biorthogonal 6.8 Wavelet family has been used, because it shows the better results. Table 2 lists the mean squared error (MSE) of the Wavelet and the ANN during different signal to noise ratio (SNR) of baseline's frequency is 0.1, 0.2, 0.4, 0.6 Hz respectively. FIR and LMS don't appear in the table because they have got bad results. Equation 10, shows MSE where x_{out} is the exit to the system and x the original signal without noise.

$$MSE = E\{\|x_{out} - x\|^2\} \tag{10}$$

When the baseline's frequency is less than 0.2 Hz, the Wavelets method is satisfied. The better cross correlation is obtained by the ANN, specially when the frequency of baseline is less than 0.6 Hz and bigger than 0.2 Hz.

Table 2. Comparison of errors of Wavelet and Neural Networks

Frequency	0.1 Hz			0.2 Hz		
SNR(dB)	0.8	-5.2	-8.7	0.8	-5.2	-8.7
Wavelet	0.01	0.03	0.3	0.1	0.03	4.9
ANN	0.1	0.3	0.5	0.2	0.3	0.5
Frequency	0.4 Hz			0.6 Hz		
SNR(dB)	0.8	-5.2	-8.7	6.8	0.8	-2.8
Wavelet	3.8	19	112	23	54	134
ANN	0.2	0.4	0.7	0.8	1.5	3.7

Table 3 shows the cross correlation results. In the case of synthetic signals, other paremeters can be measured. In this work, the Signal to Interference Ratio (SIR) has been selected to evaluate the performance of the compared methods. Equation 11, shows SIR expression where x_{in} shows the input to the system, x_{out} the exit and x the original registry without noise.

$$SIR = 20\log\left(\sqrt{\frac{E\{||x_{in} - x||^2\}}{E\{||x_{out} - x||^2\}}}\right) \tag{11}$$

Table 3. Obtained results of the cross correlation and SIR of baseline, average values

Methods	Synthetic	Real	SIR
FIR	$0,92 \pm 0.03$	$0,91 \pm 0.03$	10.8 ± 0.6
LMS	$0,63 \pm 0.32$	$0,60 \pm 0.35$	6.1 ± 2.34
Wavelet	$0,94 \pm 0.02$	$0,93 \pm 0.02$	14.6 ± 0.5
ANN	$0,97 \pm 0.02$	$0,96 \pm 0.02$	17.5 ± 0.4

If white noise (myoelectric, thermal, etc.) is added to baseline noise, the different techniques present more significant differences (Figure 2). But, in all cases, the ANN shows the best performance. The second method that approaches the data is Wavelet. The adaptive methods as LMS depend on the ECG and therefore its result is more variable. Methods FIR obtain intermediate values of baseline cancellation.

5. Conclusions

This paper introduces an approach for removing baseline in ECG signals using ANN. The paper illustrates the effectiveness of the approach by using examples with both simulated and measured ECG data. The current algorithm

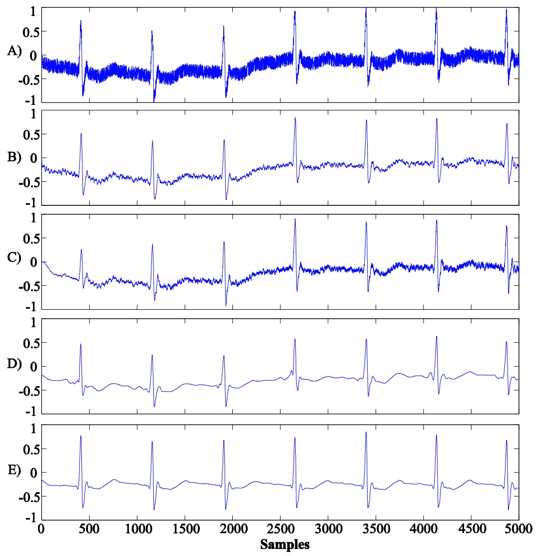

Figure 2. A. Input Signal. B. FIR Filter C. LMS Filter D. Wavelet Filter E. ANN

uses simultaneous perturbation which shows better results for noise reduction. The target application of our algorithm is preprocessing of ECG signals.

The signal cancellation depends on the convergence of the method in LMS methods. The results obtained are quite good with systems FIR and Wavelets. However, the System based on ANN, is the better to eliminate the baseline noise. It is possible to emphasize as well, that this last method, it is easier to implement and it provokes low signal distortion.

Acknowledgements

This work was partly funded by the project PAC-05-008-1 from Consejería de Educación de la Junta de Comunidades de Castilla-La Mancha, GV06/299 from Consellería de Empresa, Universidad y Ciencia de la Generalitat Valenciana and TEC2007–64884 from the Spanish Ministry of Science and Education.

References

[1] Sörnmo L, Laguna P. Bioelectrical Signal Processing in Cardiac an Neurological Applications. Elsevier Academic Press, 2005.

[2] Lian Y, HO P. ECG noise reduction using multiplier-free FIR digital filters. Proceedings of 2004 International Conference on Signal Processing 2004;2198–2201.

[3] Lian Y, Yu JH. The reduction of noises in ECG signal using a frequency response masking based FIR filter. IEEE International Workshop on Biomedical Circuits Systems 2004; 2(4):17–20.

[4] Laguna P, Jane R, Caminal P. Adaptive filtering of ECG baseline wander. Engineering in Medicine and Biology Society Proceedings of the Annual International Conference of the IEEE 1992;508–509.

[5] Laguna P, Thakor P, Caminal N. Adaptve baseline wander removal in the ECG: Comparative analysis with cubic spline technique. Computers in Cardiology 1992;143–146.

[6] Thakor N, Zhu Y. Application of adaptive filtering to ECG analysis: Noise cancellation and arrhythmia detection. IEEE Trans Biomed Eng 1991;38(8):785–794.

[7] Chiu C, Yeh S. A tentative approach based on wiener filter for the reduction of respiratory effect in pulse signals. Proc 19th Int Conf IEEE EMBS 1997;1394–1397.

[8] Nibhanupudi S. Signal Denoising Using Wavelets. Ph.D. thesis, University of Cincinnati, 2003.

[9] Xu L, Zhang D, Wang K. Wavelet-based cascaded adaptive filter for removing baseline drift in pulse waveforms. IEEE Trans Biomed Eng 2005;53(11):1973–1975.

[10] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000 (June 13);101(23):e215–e220. Circulation Electronic Pages: http://circ.ahajournals.org/cgi/content/full/101/23/e215.

[11] Haykin S. Neural networks: A comprehensive approach. In IEEE Computer Society Press. 1994; Piscataway, USA.

[12] Lehtokangas M. Fast initialization for cascade-correlation learning. IEEE Trans on Neural Networks 1999;10(2):410–414.

[13] Sanchez G, Toscano K, Nakano M, Perez H. A growing cell neural network structure with backpropagation learning algorithm. Telecommunications and Radio Engineering 2001;56(1):37–45.

[14] Hodge V. Hierarchical growing cell structures, trees GCS. IEEE Trans on Knowledge and Engineering 2001; 13(2):207–218.

[15] Schetinin V. A learning algorithm for evolving cascade neural networks. Neural Letters 2003;17(1):21–31.

[16] Hush DR, Horne BG. Progress in supervised neural networks. IEEE Signal Processing Magazine 1993;8–39.

[17] Hui S, Zak SH. The widrow-hoff algorithm for McCulloch-pits type neurons. IEEE Trans on Neural Networks 1994; 5(6):924–929.

[18] Widrow B, Lehr MA. 30 years of adaptive neural networks: Perceptrons, madeline and backpropagation. Proc of IEEE 1990;78:1415–1442.

[19] Wang ZQ, Manry MT, Schiano JL. LMS learning algorithms: Misconceptions and new results on convergence. IEEE Trans on Neural Networks January 2000;11(1):47–56.

Address for correspondence:

Jorge Mateo Sotos
Innovation in Bioengineering Research Group
Campus Universitario s/n, 16071 Cuenca (Spain)
E–mail: jorge.mateo@uclm.es