Neural Network Based Canceller for Powerline Interference in ECG Signals

J Mateo[1], C Sánchez[1], A Torres[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

Power line interference may severely corrupt a biomedical recording. Notch Filters and adaptive cancellers have been suggested to suppress this interference. In this paper, an improved adaptive canceller for the reduction of the fundamental Power Line Interference component in electrocardiogram (ECG) recordings is proposed. A comparison is made between the performance of our method and a narrow and a wide Notch Filter and Notch Adaptive Filter in suppressing the fundamental power line interference component. For this purpose, a real ECG signal is corrupted by an artificial power line interference signal. The cleaned signal after applying all methods is compared with the original ECG signal. Results indicate that power line interference of ECG are removed effectively by this new method. Interference elimination can be performed continuously and rapidly even if the situations of interference are changing with time or frequency. In the worst conditions 48.5Hz and 51.5Hz (BW 1.5Hz), ANN obtained results show the efficiency (CCC$=0.96\pm0.02$ *SIR*$=17.3\pm0.4$*) in comparison with the classical technique with the best performance (CCC*$=0.91\pm0.03$ *SIR*$=13.2\pm0.6$*). The method is easy to implement and it is applicable not only to ECG but also other biomedical signals.*

1. Introduction

Electrocardiogram (ECG) is a method which has been based on recording the heart electrical activity. ECG register is a non-stationary signal which includes valuable clinical information. However, this information is being often corrupted by noise. In addition, ECG signals have also been interfered by 50/60 Hz power line. This comes from the feeding lines of measurement systems, despite having installed a grounding protection, a shielding and amplifier incorrect design [1]. There is a single previous step in order to eliminate this kind of interference in ECG registers. This step implies to maintain the signal characteristics for diagnosis. When ECG is being recorded, frequency is being modified by power line around the nominal 50Hz or 60Hz (specifications are 50/60Hz $\pm 3\%$). When this frequency is being changed, then there will not be an accurate answer to this filter. For this purpose, an adaptive filter will be required.

The 50/60Hz Notch Filter (NF) rejects a narrow frequency band around 50/60Hz. These methods are easy to implement at a computational low cost. Moreover, these are generating an undesirable signal modification. The interferences are being eliminated but some important frequency components of ECG signal are also removed [2]. First of all, adaptive filtering has been proposed by Widrow [3]. This method does not disturb the ECG frequency spectrum but this requires a reference signal, which adds to the complexity of hardware and software [4, 5, 6]. In order to solve this problem, different adaptive structures have been studied for removing power line interference [7, 8, 9, 10, 11, 12]. These methods are also able, on the one hand, to decrease the ECG noise. But, on the other, they change the original signal. For this reason, a new method has been created and developed to reduce the signal modification and to decrease power line interference only in one step. The system which is being proposed, is based on a growing Artificial Neural Network (ANN). This ANN has been chosen mainly because its adaptability to the nonlinear and time-varying features of the ECG signal. The proposed ANN can be trained to filter out the power line interference. In addition, the structure allows to optimize both the number of neurons by which the hidden layer is made up of and the coefficients matrices (weights and bias). The matrices are optimized according to the Widrow - Hoff Delta algorithm [13]. This system has three important advantages, these being: First of all, the power line interference and white noise are reduced; second, a low modification of the signal is being caused; and third, this system can be applied to a wide range of biomedical signals.

2. Materials

In this study, two types of signals have been used. These have been referred to either real recordings from the PhysioNet Database [14] or synthetic signals. 450 recordings with different pathologies have been obtained as a result from PhysioNet with different types of QRS morphologies and 200 synthetic signals with different noises have been generated by means of using the ECGSyn software [14]. White (myoelectric, thermal, etc.) and power line interference noise are included in these registers. The sampling frequency used is 1kHz.

3. Methods

The Multi-Layer Perceptron (MLP) is a neural network which Back-Propagation algorithm has been used. MLP has been applied to the resolution of various problems [15]. At least, MLP is based on three layers, these being: An input layer, one or more hidden layers and an output layer. There are two phases in the method in order to estimate the optimal number of nodes which have been located in the hidden layer. The first phase will consist on stopping the training after a certain number of iterations. The second phase will determine the noise quantity which has been properly removed from the registers with the current number of neurons which have been used in the hidden layer. If the result of this test is not satisfactory, one or more neurons will be added in the hidden layer in order to improve the performance of the proposed network. However, in these cases the network has to be completely trained [15].

3.1. Proposed system

The proposed system initially consists on a simple structure which is similar to the neural network [16] ADALINE (ADAptive LINear Element) as Figure 1 shows (where w is the weight of this neuron and b the bia). This network structure has initially an input layer, one hidden layer (made up of 30 neurons) and an ouput layer.

This new structure has been provided with a special characteristic: it is growing while it is learning. By which is meant that the neurons, which embed the intermediate layer, have been added one by one and their weights (w) and their bias (b) have also been adapted. However, the input layer weights, which had been previously adapted, remain stable in order to conserve the learning process. Sometimes, this mechanism could produce neural networks with a sub-optimal number of neurons in the hidden layer. Thus, this mechanism allows how to estimate the network size in order to develop a concrete work. But each time that a new neuron has been added to the hidden layer, this network would not need to be completely trained.

3.2. Training neural network

A single hidden layer has been used as a network prototype. In Figure 1, $P = (p_i)$ is the input vector, $Y = (y_i)$ is the output vector (clean signal), $W = (w_{ij})$ is the matrix of weights between the input layer and the hidden layer, $b = (b_{ij})$ is the matrix of the bia between the hidden layer and the output layer. The characteristic of this hidden layer

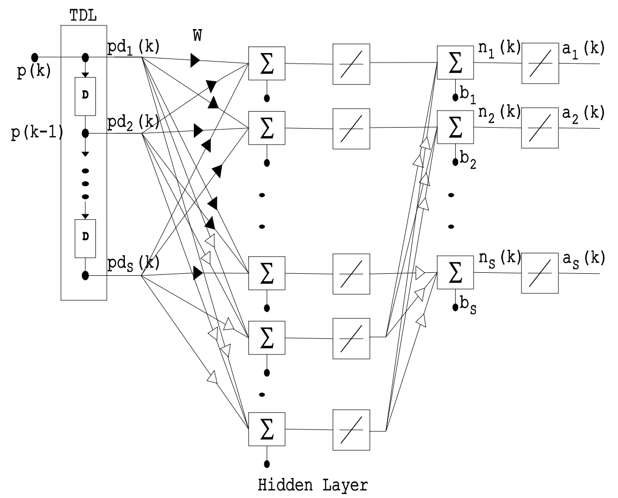

Figure 1. Proposed Neural Network with two neurons in the hidden layer. The black coefficients are constants.

is sigmoidal. $t_i(k)$ represents the expected output of the pth set of the input vector. The initialization procedure has been utilised to select which kind of network would be the most initialized for training. The Widrow-Hoff Delta Rule has been used to improve the network synaptic weights and bias in order to minimize the error function. The weights (W) and bias (b) have been chosen by a uniform distribution of random numbers from the interval $[-1, 1]$. These values have been obtained experimentally by means of adapting the network with a single pass.

3.3. Initialization process

This network has been built by a hidden layer network which uses a sigmoidal activation function. There have been created a number of candidate networks (H) which contains a number of hidden neurons in order to initiate the network learning. The value of H has been decided by means of a test. For each candidate network, the sum of absolute values of covariances have been calculated from Equation 1.

$$F_j = \frac{1}{N} \sum_{k=1}^{M-1} \left| \sum_{p=1}^{N} (y_{j,p} - \overline{y_j})(e_{k,p} - \overline{e_k}) \right|, \quad j = 1, ..., H \tag{1}$$

where $y_{j,p}$ is the output of the jth candidate network for the pth training pattern. The parameter $\overline{y_j}$ is the mean of the jth hidden unit outputs, $e_{k,p}$ is the output error at the kth output unit for the pth training pattern and $\overline{e_k}$ is the mean of the output errors at the kth output unit. Then, the network with the maximum covariance F_j is selected as the most promisingly network to be initialized. An optimum value $H = 30$ has been obtained.

3.4. Learning algorithm using the Widrow-Hoff Delta rule

The proposed network (Adaline) has been based on a supervised learning which needs get to to know the associated values in each input. These pairs of input/output are:

$$\{p_1, t_1\}, \{p_2, t_2\}, ..., \{p_Q, t_Q\} \tag{2}$$

where p_Q is the network input and t_Q is its corresponding expected output, when p input is presented to the proposed network, the network exit is compared with t_{value} (the expected output) which has been associated to it. Our proposed network (Adaline) has been deduced by the following way, according to the procedure which has been described in Widrow [13]. The process to update the weights is being given by Equation 3.

$$W(k+1) = W(K) + \alpha \frac{e(k)p(k)}{|p(k)|^2} \tag{3}$$

where k shows the current iteration of the updated process, $W(k+1)$ is the following value which will be assigned to the vector weights, and $W(k)$ is the current value of the vector weights. The current error $e(k)$ has been defined as the difference between the expected output $t(k)$ and the network exit $a(k) = W^T(k)p(k)$ before its updating process:

$$e(k) = t(k) - W^T(k)p(k) \tag{4}$$

In each iteration, the variation of error has been represented by:

$$\Delta e(k) = \Delta(t(k) - W^T(k)p(k)) = -p^T(k) * W(k) \tag{5}$$

In order to explain our Mean Squared Error (MSE), another kind of algorithm will be applied. This algorithm has been named Approximated Descendant Steps, whose method has been based on the Widrow and Hoff Delta Rule. By means of this algorithm, the instant gradient has been calculated in each iteration instead of the true gradient. For this reason, equation 6 is proving MSE Function, this being:

$$e^2(k) = (t(k) - y(k))^2 \tag{6}$$

In Equation 6, $t(k)$ represents the expected output in the iteration k and $y(k)$ shows the exit of the network. Our updating process for weights and bias has been defined in equations 7 and 8. The best results have been obtained when 12 neurons were added to the hidden layer. When more than 12 neurons are added, there is no improvement of both the computational load and the noise reduction.

$$w(k+1) = w(k) + 2\alpha e(k)p(k) \tag{7}$$

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

4. Results

In this section, our system has been compared with the standard filtering techniques (NF, Notch Adaptive Filter (NAF)), which had been introduced previously. The equation MSE($\widehat{\theta}$)=E$\left\{(\widehat{\theta} - \theta)^2\right\}$ has been used to calculate MSE in order to verify learning of the neural network. $\widehat{\theta}$ shows the exit to the system and θ shows the clean signal.

The power line was a simulated 50Hz sinusoide. In addition, the signal frequency has been modified between 48.5 Hz and 51.5 Hz. The following table shows the obtained MSE for different bandwidths (0.5Hz and 1.5Hz) and the selected techniques; NF, NAF and ANN.

Table 1. Average values of MSE for ECG recordings. Bandwitdh of 0.5Hz and 1.5Hz

BW	0.5 Hz			1.5 Hz		
Method	NF	NAF	ANN	NF	NAF	ANN
48.5 hz	47	23.4	6.8	41.8	31.6	15.5
49 Hz	46	22.7	6.2	38.5	30.4	13.9
49.5 Hz	34	21.8	5.9	20	29.5	12.5
50 Hz	3	1.32	1.3	3.5	2.1	1.5
50.5 Hz	33.5	21.7	5.6	19.4	29.2	12.4
51 Hz	46.5	21.9	6	39	30.6	14.1
51.5 Hz	47.5	23.6	6.5	41.05	21.9	16.2

Table 1 indicates that the MSE for NF is minimum at 50Hz but becomes significant as the power line wanders away from 50Hz. Error is relevant, even when bandwidth has been increased to 1Hz. The MSE for a ANN is quite small and remains nearly constant even when frequency line varies around 50Hz. Only at 51Hz the MSE becomes relevant as bandwidth is increased above 1 Hz. For NAF the MSE value is bigger than ANN for all frequencies, the difference is the smallest for 50Hz.

In addition, the Interference Ratio Signal (SIR) has also been computed by means of Equation 9, where x_{in} shows the input signal to the system; x_{out} the output and x the original recording without noise. The results of cross correlation coefficient (CCC) and SIR are shown in Table 2. As it can be observed, ANN achieves higher values than traditional methods reflecting an improved performance.

$$SIR = 20 \log \left(\sqrt{\frac{E\{||x_{in} - x||^2\}}{E\{||x_{out} - x||^2\}}} \right) \tag{9}$$

In Figure 2, the results which have been obtained by the NF, NAF and ANN method, have been represented. Where ANN has been achieved to cancel the power line interference, other systems can not reduce this interference completely.

Table 2. Average values of CCC and SIR for ECG recordings

Method	NF	NAF	ANN
CCC	0.89 ± 0.3	0.92 ± 0.3	0.97 ± 0.2
SIR	13.8 ± 1.3	15.4 ± 1.4	18.5 ± 1.2

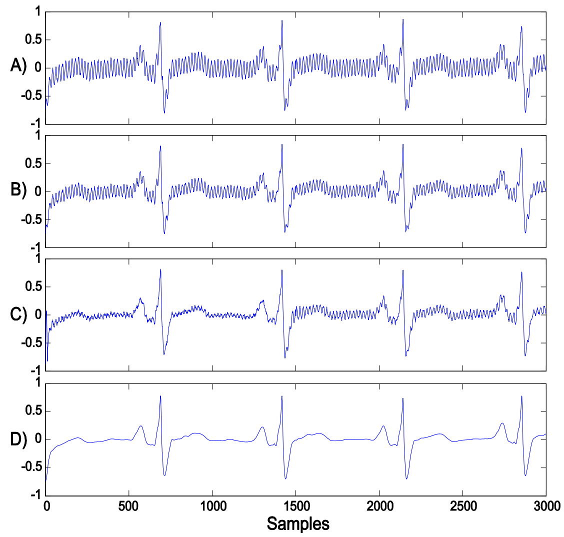

Figure 2. A. Input Signal (Real ECG). B. Notch Filter output signal C. Notch Filter Adaptive output signal D. Neural Network output signal. The Power line frequency is 50Hz with BW of 1Hz.

5. Conclusions

This paper proves how the proposed growing ANN has been used to remove white noise and power line interference from ECG data in one step. Throughout all the stages, our ANN has been adapted by means of using the Widrow - Hoff Delta Algorithm, which has been improved in order to achieve our target. By means of this improvement, ANN has obtained bigger values of CCC and SIR than the other methods. This ANN system achieves to correct the changes produced by noise and to reduced significantly interferences in the ECG signals. As a way of conclusion, suffice is to say that the neural network-based approach obtains both more noise (power line and white) reduction and low modification of the signal results in comparison with systems which had been based on NF and NAF methods.

Acknowledgements

This work was partly sponsored by Junta de Comunidades de Castilla-La Mancha, Patronato Universitario Cardenal Gil de Albornoz from Cuenca and TEC2007–64884 from the Ministry of Science and Education (Government of Spain).

References

[1] Huhta JC, Webster JG. 60-hz interference in electrocardiography. IEEE Trans Biome Eng 1973;20(2):91–101.

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

[3] Widrow B, Glover J, McCool JM, Kaunitz J, Williams CS, Hearn R, Zeidler JR, Dong E, Goodlin R. Adaptive noise cancelling: Principles and applications. Proc IEEE 1975; 63:1692–1716.

[4] Mitov I. A method for reduction of power line interference in the ECG. Medical Engineering Physics 2004; 26(10):879–887.

[5] Leski JM, Henzel N. ECG baseline wander and powerline interference reduction using nonlinear filter bank. Signals Processing 2005;85(4):781–793.

[6] Hejjel L. Supression of power-line interference by analog notch filtering in the ECG signal for heart rate variability analysis: To do or not to do. Diagnostic and Medical Technology 2004;.

[7] Ziarini AK, Konrad A. A nolinear adaptive method of elimination of power line interference in ECG signals. IEEE Trans Biomed Eng 2002;49(6):540–547.

[8] Bor-Shyh L, Nor-Shing L, Wan-Chi L, Fok-Ching C, Yue-Der L. Removing residual power-line interference using WHT adaptive filter. En 24rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, volumen 1. 2002; 155–156.

[9] Martens S, Mischi M, Oei SG, Bergmans JWM. An improved adaptive power line interference canceller for electrocardiography. IEEE Trans Biomed Eng 2006; 53(11):2220–2231.

[10] Nimunkar, Arnit J, Tompkins WJ. Emd-based 60-hz noise filtering of the ecg. En 29rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2007; 1904–1907.

[11] Wu Y, Rangayyan RM, Zhouc Y, Ngd SC. Filtering electrocardiographic signals using an unbiased and normalized adaptive noise reduction system. doi.10.1016/j.medengphy.2008.03.004. Medical Eng Physics 2008;10.

[12] Lin YD, Hu YH. Power-line interference detection and suppression in ecg signal processing. IEEE Trans Biomed Eng 2008;55(1):354–357.

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

[14] 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.

[15] Haykin S. Neural Networks: A Comprehensive Approach. 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

[16] Hush R, Horne BG. Progress in supervised neural networks. IEEE Signal Processing Magazine 1993;10(1):8–39.