Neurocomputing 103 (2013) 222–231

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Artifacts removal in EEG signal using a new neural network enhanced adaptive filter

Aysa Jafarifarmand, Mohammad Ali Badamchizadeh *

Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

ARTICLE INFO

Article history:
Received 26 January 2012
Received in revised form
18 July 2012
Accepted 26 September 2012
Communicated by J. Zhang
Available online 2 November 2012

Keywords:
EEG
Artifact
Adaptive FLN–RBFN-based filter
Adaptive noise cancellation

ABSTRACT

EEG signal is an important clinical tool for diagnosing, monitoring, and managing neurological disorders. This signal is often affected by a variety of large signal contaminations or artifacts, which reduce its clinical usefulness. In this paper, a new adaptive FLN–RBFN-based filter is proposed to cancel the three most serious contaminants, i.e. ocular, muscular and cardiac artifacts from EEG signal. The basic method used in this paper for the elimination of artifacts is adaptive noise cancellation (ANC). The results demonstrate the effectiveness of the proposed technique in extracting the desired EEG component from contaminated EEG signal.

1. Introduction

Electroencephalogram (EEG) is the noninvasive measurement of the electrical activity of the brain obtained by means of some electrodes placed on the scalp over multiple areas of the brain. EEG, which is an important clinical tool for diagnosing, monitoring, and managing neurological disorders, is widely affected by a variety of large signal contaminations or artifacts, which reduce its clinical usefulness.

Artifacts are undesired signals with non-cerebral origins recorded by EEG electrodes. Based on their origin, artifacts can be divided into two types:

- *Biological artifacts*: mostly caused by electrooculogram (EOG) or ocular activities, electromyogram (EMG) or muscular activities, and electrocardiogram (ECG) or cardiac activities;
- *External artifacts*: caused due to technical factors such as line-interference, leads, and electrodes.

While the influence of artifacts of technical origin can be reduced to a large degree by improving technology and paying extra attention to the attachment of electrodes to the body surface, it is impossible to avoid the influence of artifacts of biological origin. Accordingly, most algorithms developed for EEG artifact processing are intended for the reduction of biological artifacts [1].

The most appropriate approach introduced to handle the artifacts' problem is the artifact removal method, that is, the process of identifying and removing artifacts from brain signals as well as keeping related neurological phenomenon intact. To remove artifacts from EEG signals, different techniques have been introduced.

An ocular artifact removal method based on wavelet transform has been proposed in [2], and in [3–5], wavelet transform has been applied to remove EOG artifacts from EEG signal. The idea of wavelet based denoising methods relies on the assumption that the signal magnitudes dominate the magnitudes of the noise in a wavelet representation. Basically, filtering is performed by comparing each wavelet coefficient to a predetermined threshold and setting it to zero if its magnitude is less than the threshold. On this basis selecting the threshold level plays a crucial role in wavelet based denoising methods. The main difficulty in EEG artifact removal problem is selecting the threshold level so that it should not remove the original EEG signal coefficients and not to keep the artifact signals as original ones.

Independent component analysis (ICA) is the other technique used to remove EEG artifact. In [6], ICA has been used to separate the EOG artifacts from EEG signals, and in [7], authors have used the ICA algorithm to remove the ocular and muscular artifacts from EEG signal. This method requires offline analysis and processing of data collected from a sufficiently large number of channels, and its success largely depends on correct identification of the noise components [8].

* Corresponding author.
E-mail addresses: a.jafari@tabrizu.ac.ir (A. Jafarifarmand), mbadamchi@tabrizu.ac.ir (M.A. Badamchizadeh).

0925-2312/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.09.024

To remove the artifacts from EEG signal, many regression-based techniques have also been proposed. Two nonlinear regression methods have been compared with the linear regression in [9], and in [10] a regression based method has been proposed to remove different kinds of ocular artifacts from EEG signal. In all the regression-based approaches, calibration trials are first conducted to determine the transfer coefficients between the EOG, EMG or ECG channels and each of the EEG channels. These coefficients are later used in the correction phase to estimate the artifactual component in the EEG recording for removal by subtraction [8], and therefore these methods are not suitable for real-time implementation. Also, during the performance, the true coefficients may be off from the adjusted value, and they cannot be calibrated again automatically [11].

Recently, adaptive filtering has achieved widespread applications and successes in many areas. According to time varying signals arising from the human body, adaptive filtering has been an appropriate method in EEG artifact removal. Although adaptive linear filters are the most widely used among the various adaptive filters, their performance is not satisfactory for dealing with nonlinear problems. Since the artifact signals pass through the nonlinear path from their origin to the place of EEG electrodes on the scalp [12], linear adaptive filters are not appropriate to remove these artifacts from EEG signals.

As a result, due to this nonlinearity, the adaptive filters should be used that have a high capacity to deal with nonlinear problems. With the development of neural networks (NN) and fuzzy interference systems (FIS) and based on the fact that both of them are global nonlinear approximators, many new approaches of designing adaptive nonlinear filters for the purpose of noise cancellation have been proposed [13].

Thanks to the powerful learning and generalization abilities, neural networks have become an attractive approach in adaptive signal processing. However, it is not easy to determine the structure of neural networks, because the internal layers of neural networks are not clear to users [13]. The internal nodes and weights have no individual conceptual meaning and therefore cannot directly correspond to a rule, consequently the meaning associated with internal nodes and weights of neural networks is not easily interpreted by users.

Fuzzy systems can be understood by users because the rule base is constructed by linguistic IF–THEN rules. However, the learning capacity of fuzzy systems is less than that of neural networks [14].

A promising approach of reaping both the benefits of NN and FIS and solving their respective problems is to combine them into an integrated system termed fuzzy neural networks [13].

In [15], a hybrid soft computing technique called adaptive neuro-fuzzy inference system (ANFIS) has been proposed to estimate the interference and separate the EEG signal from its EOG, ECG and EMG artifacts; the proposed method successfully removes the artifacts and extracts the desired EEG signal. Also in [16,17], ANFIS has been applied to extract the original bioelectrical signals from the contaminated measured ones. Er et al. have proposed an adaptive radial basis function networks (RBFN) based filter for the adaptive noise cancellation (ANC) problem in [18,19], which implements Takagi–Sugeno–Kang (TSK) fuzzy systems functionally. The proposed filter has the salient features that no predetermination of RBF neurons is needed and the structure learning algorithm is automatic as well as the parameters learning algorithm. It is also shown that the filter successfully cancels noise with a parsimonious structure and performs better than ANFIS.

TSK-type neuro-fuzzy systems consist of a type of fuzzy statements in which the consequence part of each fuzzy rule is a linear combination of input variables. Since the traditional TSK-type neuro-fuzzy system does not take full advantage of the mapping capabilities that may be offered by the consequent part, introducing a nonlinear function, especially a neural structure, to the consequent part of the fuzzy rules seems to be appropriate to improve the accuracy of functional approximation.

In the present paper, a new adaptive FLN–RBFN-based filter is proposed to remove ocular, muscular and cardiac artifacts from EEG signals using the ANC method. In the proposed filter, the functional link neural network (FLN) is applied to the consequent part of the fuzzy rules, in the RBFN-based filter proposed in [18,19], and thus the consequent part of the proposed model is a nonlinear combination of input variables that increases its universal nonlinear approximation ability considerably. FLN is a single-layer neural structure capable of forming arbitrarily complex decision regions by generating nonlinear decision boundaries with nonlinear functional expansion [14]. In addition, it has all the features of the mentioned RBFN-based filter, the proposed adaptive filter performs much better and cancels the artifacts efficiently.

This paper is organized as follows. In Section 2, the concept of ANC is explained. Section 3 describes the structure of the suggested adaptive filter. Section 4 presents the learning algorithms. Section 5 presents the results of simulations and discussion to show the performance of the proposed method. Finally, Section 6 shows the conclusion.

2. Adaptive noise cancellation

The basic method used in this paper for elimination of artifacts is ANC. It is a process by which the interference signal can be filtered out by identifying a model between a measurable noise source and the corresponding immeasurable interference [15]. The method uses a noisy signal as primary input and a reference input that consists of noise correlated in some unknown way with the primary noise. It is also assumed that the desired clean signal is uncorrelated with noise source and interference signals. By adaptively filtering and subtracting the reference input from the primary input, the output of the adaptive filter will be the error signal, which acts as a feedback to the adaptive filter [20]. With this setup, the adaptive filter will be able to cancel the noise and obtain an estimate of the less noisy signal.

Fig. 1 shows the basic concept of adaptive noise cancellation using the proposed adaptive FLN–RBFN-based filter. The primary input of the canceler is the measurable EEG signal, $q(k)$, which is naturally contaminated by artifacts. In other words, the artifact signal $n(k)$ (noise source signal) goes through unknown nonlinear dynamics f (the route in human body from the artifact generating source to each EEG electrode on the scalp) and generates a distorted noise $d(k)$, which is then added to clean EEG signal $s(k)$, in the place of electrodes to form $q(k)=s(k)+d(k)$, which is the very signal measured by the EEG electrodes. The reference input to the canceler, which is the input to the adaptive filter in

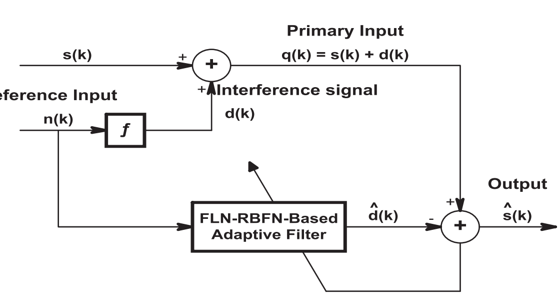

Fig. 1. Schematic diagram of ANC.

fact, is the noise source signal, $n(k)$, measured directly from the artifact generating origin i.e. EOG for ocular artifacts, EMG for muscular artifacts and ECG for cardiac artifacts. The noise $n(k)$ is filtered to produce $\hat{d}(k)$ which is as close a replica as possible of interference signal, $d(k)$, the distorted and delayed version of $n(k)$. The aim is to retrieve clean EEG, $s(k)$, from the measured EEG signal, $q(k)$, that is procurable by estimating the $\hat{d}(k)$ using the adaptive filter, which is then subtracted from measured $q(k)$ to produce the output of the system, $\hat{s}(k)$, that would be close to the required signal $s(k)$.

In the noise canceling application, the objective is to produce an error signal (output) that is a best fit in the least squares sense to the signal $s(k)$. This is accomplished by feeding back the system output to the adaptive filter to minimize the error signal until it reaches the value $\hat{s}(k)=s(k)$. As a result, the output of the noise canceler $\hat{s}(k)$ is the estimation of corrected or clean EEG

$$\hat{s}(k)=s(k)+d(k)-\hat{d}(k) \tag{1}$$

Squaring both sides

$$\hat{s}(k)^2=s(k)^2+[d(k)-\hat{d}(k)]^2+2s(k)[d(k)-\hat{d}(k)] \tag{2}$$

Taking the expected value for both sides, and assuming that $s(k)$ is uncorrelated with $n(k)$ and $d(k)$ and thus with $\hat{d}(k)$, which is well accepted for EEG signal and its artifacts:

$$\begin{aligned} E[\hat{s}(k)^2] &= E[s(k)^2]+E[[d(k)-\hat{d}(k)]^2]+2E[[s(k)[d(k)-\hat{d}(k)]] \\ &= E[s(k)^2]+E[[d(k)-\hat{d}(k)]^2] \end{aligned} \tag{3}$$

Since the signal power $[s^2(k)]$ is a specified value, the minimum output power will be given by

$$\min E[\hat{s}(k)^2]=E[s(k)^2]+\min E[[d(k)-\hat{d}(k)]^2] \tag{4}$$

Therefore, when the filter is adjusted so that $E[\hat{s}^2(k)]$ is minimized, $E[[d(k)-\hat{d}(k)]^2]$ is also minimized. The filter output $\hat{d}(k)$ is then the best least squares estimate of the interference signal $d(k)$, and with considering

$$E[\hat{s}(k)-s(k)]^2=E[d(k)-\hat{d}(k)]^2 \tag{5}$$

the output signal $\hat{s}(k)$ will also be the best least squares estimate of the clean EEG signal $s(k)$

$$\hat{s}(k)\approx s(k) \tag{6}$$

The basic idea is to use the least square algorithms to develop an adaptive filter that can be used in ANC applications [20].

It also should be noted that it may take some time for artifact signal to get the EEG electrodes on the scalp from its origin, which means a delay between the artifact generation time and the time it contaminates the EEG signal. Therefore, to consider this delay possibility a tapped delay line (TDL) can be used for the noise source signal. Fig. 2 shows the structure that the noise source signal passes through the TDL before going into the adaptive filter.

The noise source signal passes through $r-1$ delays. The output of the TDL, which is the input of the adaptive filter, is an r-dimensional vector, $X(k)=\begin{bmatrix} x_1(k) & x_2(k) & \cdots & x_r(k)\end{bmatrix}^T$, made up of the noise source signal at the current time and the previous ones.

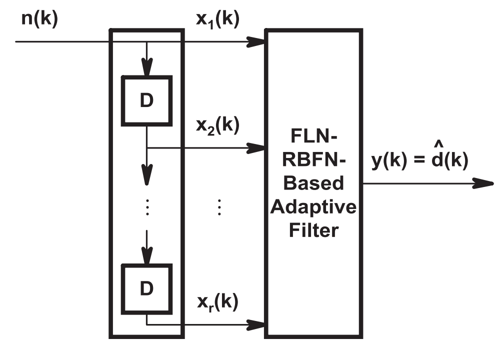

Fig. 2. Applying tapped delay line to the input of the adaptive filter.

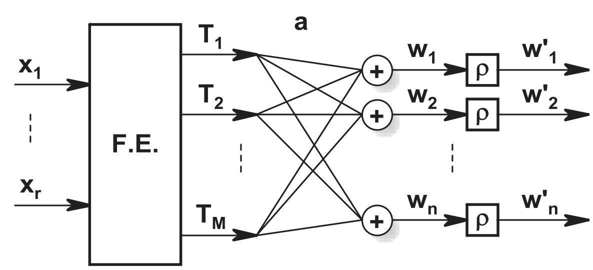

Fig. 3. Structure of an FLN.

3. Structure of the adaptive FLN–RBFN-based filter

This section details the structure of FLN and the structure of the proposed FLN–RBFN-based adaptive filter. In the proposed filter, FLN is applied to increase the dimensions of the input space in the consequent part of the fuzzy rules. Accordingly, the input representation is enhanced using the functional expansion and the consequent part of the fuzzy rules is a nonlinear combination of input variables. As a result, the local properties of the consequent part enable the proposed filter to approximate a nonlinear combination of input variables more effectively.

3.1. Functional-link neural network

FLNs are higher-order neural networks without hidden units introduced by Klassen and Pao in 1988. Despite their linear nature, FLNs can capture non-linear input–output relationships, provided that they are fed with an adequate set of polynomial inputs, or the functions might be a subset of a complete set of orthogonal basis functions, which are constructed out of the original input attributes [21]. In other words, FLN is essentially a single-layer perceptron network, in which the input space is enhanced using suitable orthogonal polynomials and linear separability is achieved in the extended space. In FLN, a set of basis functions, T, and a fixed number of weight parameters α are used to approximate a continuous function. With a suitable set of basis functions, the problem is then to find the weight parameters, α, that provide the best possible approximation of the desired function on the set of input–output examples [22].

The structure of an FLN is shown in Fig. 3. The FLN consists of M basis functions $T_1, T_2, \cdots, T_M$ with the following input–output relationship for jth output:

$$w_j=\sum_{i=1}^{M}\alpha_{ji}T_i, \tag{7}$$

$$w'_j=\rho(w_j) \tag{8}$$

where function $\rho(\cdot)$ is an activation function such as $\tanh(\cdot)$, w'_j is the jth output of the FLN and w_j denotes the jth local output of the FLN structure that provides the weight of the jth RBF neuron or the consequent part of the jth fuzzy rule in the proposed adaptive FLN–RBFN-based filter; $j=1,2,\ldots,n$ and n is the number of RBF neurons or fuzzy rules which will be described in the next section.

Variety of orthogonal polynomials for functional expansion, such as Legendre, Chebyshev and trigonometric polynomials are commonly used. In this paper, based on the investigations conducted and also considering their powerful non-linear approximation capacity [21], Chebyshev orthogonal polynomials have been selected as functional expansion.

The Chebyshev polynomial basis functions are given by

$$Ch_0(x)=1$$
$$Ch_1(x)=x$$
$$Ch_2(x)=2x^2-1$$
and
$$Ch_{m+1}(x)=2xCh_m(x)-Ch_{m-1}(x) \tag{9}$$

Vector T of the basis functions shown in Fig. 3 obtained by using Chebyshev functions is given as follows:

$$T=\begin{bmatrix} T_1 & T_2 & \ldots & T_M \end{bmatrix}^T = \begin{bmatrix} 1 \\ Ch_1(x_1) \\ Ch_2(x_1) \\ \vdots \\ Ch_1(x_r) \\ Ch_2(x_r) \\ \vdots \end{bmatrix} \tag{10}$$

In the FLN–RBFN-based filter, the links to the local outputs do not exist in the initial state, and the number of local outputs is consistent with the number of fuzzy rules (RBF neurons) generated by the learning algorithms. Section 4 details the learning algorithms.

3.2. FLN–RBFN-based adaptive filter

The proposed adaptive FLN–RBFN-based filter is depicted in Fig. 4. The proposed filter implements the fuzzy inference system that the consequent part of each fuzzy rule corresponds to a functional link neural network. For input vector $X=\begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix}^T$, jth fuzzy IF–THEN rule would be in the following form:

Rule j: IF x_1 is A_{1j} and x_2 is A_{2j} … and x_r is A_{rj}THEN

$$\hat{y}_j=\sum_{i=1}^{M}\alpha_{ij}T_i = \alpha_{1j}T_1+\alpha_{2j}T_2+\cdots+\alpha_{Mj}T_M \tag{11}$$

where $\hat{y}_j$ is the jth local output of the proposed filter; A_{lj} for

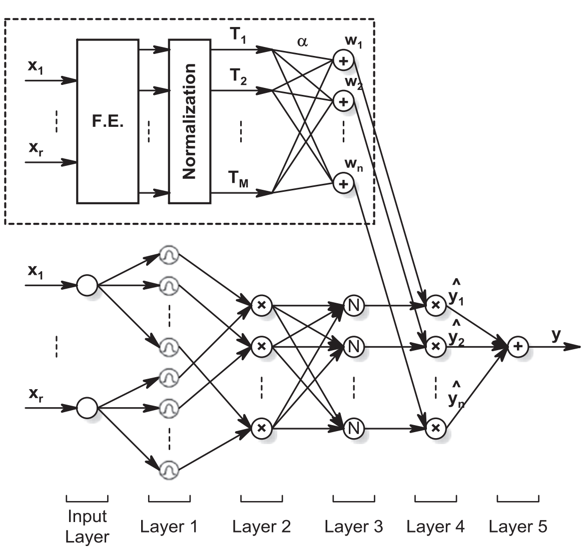

Fig. 4. Structure of proposed adaptive FLN–RBFN-based filter.

$l=1,2,\ldots,r$ is the linguistic term of the precondition part with Gaussian membership function, α_{ij} is the link weight of the local output, T_is are the basis functions (Chebyshev polynomials as mentioned) of input variables, which are normalized before being used for the proposed filter, and M is the number of basis functions.

The functions of the nodes in each of the five layers are described as follows:

Layer 1: each node in this layer acts as a one-dimensional membership function (MF) which is considered a Gaussian function of the following form:

$$\mu_{lj}(x_l)=\exp\left[-\frac{(x_l-c_{lj})^2}{\sigma_j^2}\right] \tag{12}$$

where

$l=1,2,\ldots,r$, r is the number of input variables;

$j=1,2,\ldots,n$, n is the number of membership functions;

μ_{lj} is the jth membership function of the lth input, variable x_l, c_{lj} is the center of the jth MF of x_l and σ_j is the width of the jth MF.

Layer 2: each node in this layer with the linked nodes of the previous layer constitutes an RBF neuron, and also each node in this layer performs the precondition part of a fuzzy rule. The number of rules in this system is exactly the number of RBF neurons. For the jth rule R_j, its output is

$$\phi_j=\exp\left[-\frac{\sum_{l=1}^{r}(x_l-c_{lj})^2}{\sigma_j^2}\right] = \exp\left[-\frac{\|X-C_j\|^2}{\sigma_j^2}\right] \tag{13}$$

where $C_j=\begin{bmatrix} c_{1j} & \cdots & c_{lj}\end{bmatrix}^T$ is the center of the jth RBF neuron.

Layer 3: nodes in this layer are called normalized nodes. The output of the normalized nodes is given by

$$\psi_j=\frac{\phi_j}{\sum_{z=1}^{n}\phi_z} = \frac{\exp\left[-\frac{\|X-C_j\|^2}{\sigma_j^2}\right]}{\sum_{z=1}^{n}\exp\left[-\frac{\|X-C_z\|^2}{\sigma_z^2}\right]} \tag{14}$$

Layer 4: in this layer, the outputs of the Layer 3 are weighted by the local outputs of the FLN.

Layer 5: each node in this layer represents an output variable which is the summation of incoming signals from Layer 4. Its output is given by

$$y(X)=\sum_{j=1}^{n}w_j\psi_j = \frac{\sum_{j=1}^{n}w_j\exp\left[-\frac{\|X-C_j\|^2}{\sigma_j^2}\right]}{\sum_{j=1}^{n}\exp\left[-\frac{\|X-C_j\|^2}{\sigma_j^2}\right]} \tag{15}$$

where y is the value of the output variable and w_j is essentially the consequent part of the jth fuzzy rule, which comes from local output of FLN.

$$w_j=\alpha_{1j}T_1+\alpha_{2j}T_2+\ldots+\alpha_{Mj}T_M \tag{16}$$

where $j=1,2,\ldots,n$, the parameters $\alpha_{j1},\alpha_{j2},\ldots,\alpha_{jM}$ are the linear parameters in the consequent part of the jth fuzzy rule and $T_1,T_2,\ldots,T_M$ are the normalized Chebyshev polynomials of inputs.

4. Learning procedure of the adaptive FLN–RBFN-based filter

This section presents a learning algorithm for constructing the proposed adaptive filter in both offline and online mode. The proposed learning algorithm comprises structure identification phase and parameter determination phase. As an outstanding feature of the proposed adaptive filter, there are no nodes in the adaptive filter initially and the nodes are created automatically as learning proceeds. The rest of this section details the structure identification phase and the parameters determination phase firstly and then the learning algorithms are described.

4.1. Structure identification and parameter determination

Structure identification includes the generation of RBF neurons and elimination of redundant RBF neurons (fuzzy rules). With the generation of a new RBF neuron, the input space is automatically partitioned into the receptive fields of the corresponding RBF neurons [18,19]. Parameter determination is the process of allocating membership functions or adjusting the parameters in the precondition part and also computing the linear parameters in the consequent part.

4.1.1. Generation of RBF neurons

To make sure that every pattern can be represented in sufficient match degree by at least one fuzzy rule (RBF neuron), the minimum-firing-strength (MFS) criterion is used to generate new rules. The main idea of the MFS criterion is as follows: for any input in the operating range, there exists at least one RBF neuron (fuzzy rule) so that the match degree (firing strength) is greater than a predefined constant, that is, the value of MFS [18]. To meet MFS criterion, for any newly arrived pattern, $X(k)$, the firing strengths of all existing RBF neurons are calculated using Eq. (13) and a new rule is generated if no one exceeds the criterion of rule generation, F_{gen}. $F_{gen}(F_{gen} \in [0,1])$ is a pre-specified threshold that increases during the learning process. Initially, it is set small to achieve rough but global representation and then, it gradually increases for fine learning. It is given by

$$F_{gen} = \min\left[F_{\min}\delta^{-i}, \quad F_{\max}\right] \tag{17}$$

where $\delta \in (0,1)$ is the decay constant, and i is the number of existing RBF neurons [18].

4.1.2. Allocation and adjustment of RBF units (precondition parameters)

For any new generated RBF neuron, the initial mean and width values need to be determined for the new membership function. These parameters, which are essentially the precondition parameters of the fuzzy rules, are then updated and adjusted according to the newer incoming patterns.

For the first generated neuron, the parameters are set as

$$\begin{aligned} C_1 &= X(1) \\ \sigma_1 &= \sigma_{\text{int}} \end{aligned} \tag{18}$$

where $X(1)$ is the first input and σ_{int} is a prespecified constant.

Since the generalization of RBF neurons is dependent on the appropriate choice of the width value, the width must be selected so carefully to ensure proper and sufficient degree of overlapping. Hence, due to the fact that the centers will be adjusted automatically according to the next input patterns, the centers are initially allocated as follows:

$$C_{new} = X(k) \tag{19}$$

and following the main idea of the MFS criterion, to ensure sufficient match degree for any pattern in the input space, the width of the newly generated RBF neuron is computed as follows:

$$\sigma_{new} = \frac{\max\{\|X(k)-C_a\|, \|X(k)-C_b\|\}}{\sqrt{\ln(1/F_{gen})}} \tag{20}$$

where C_a and C_b are the two nearest neighboring centers to the newly arrived pattern, in the sense of Euclidean distance. After the center and width values are allocated by the aforementioned method, the next arrived pattern which is represented by the newly generated RBF neuron or the existing RBF neurons will meet the MFS criterion so that the match degree (firing strength) will be greater than F_{gen} [13].

In cases where there is no need to generate a new RBF neuron, that is, the existing RBF units can represent the new incoming pattern in sufficient degree to meet the MFS criterion, the centers of the existing membership functions are adjusted by a self-organizing algorithm. The main idea is to set the centers of the membership functions so that only the regions of the input space containing data to be covered. This adaptive formulation runs independently for each input linguistic variable as follows [18]:

$$\begin{aligned} \|x_l(k)-c_{closest}\| &= \min\|x_l(k)-c_{lj}\| \\ c_{adjusted} &= c_{closest} + \beta \times [x_l - c_{closest}] \\ c_{closest} &= c_{adjusted} \end{aligned} \tag{21}$$

where learning rate β is monotonically decreasing and $j = 1,2,\cdots,n$, n is the number of existing rules.

4.1.3. Determination of linear parameters

To optimize the linear parameters of the proposed adaptive FLN–RBFN-based filter so that the main objective of the adaptive noise cancellation system is achieved, the following cost function is defined:

$$E = \sum e^2(k) = \sum [q(k)-y(k)]^2 \tag{22}$$

where $q(k)$ is the desired output for the adaptive filter and $y(k)=y(X(k))$.

According to Eq. (16), the linear parameters of each fuzzy rule are

$$\mathrm{A}_j = \begin{bmatrix} \alpha_{1j} & \alpha_{2j} & \dots & \alpha_{Mj} \end{bmatrix} \tag{23}$$

and the linear parameters of all fuzzy rules can be written as follows:

$$A = \begin{bmatrix} A_1 & A_2 & \dots & A_n \end{bmatrix}^T \tag{24}$$

Eq. (15) for each incoming pattern can be rewritten in the following form:

$$y(k) = \sum_{j=1}^{n} w_j\psi_j = \sum_{j=1}^{n} \mathrm{A}_j T\psi_j \tag{25}$$

where T is the basic functions vector introduced in Eq. (10). Assuming that $\theta_j = T\psi_j$

$$y(k) = \sum_{j=1}^{n} A_j\theta_j = \mathrm{A}^T\theta \tag{26}$$

As a result, for a special N pieces group of incoming patterns it can be written as

$$Y = \Theta \mathrm{A} \tag{27}$$

where

$$Y = \begin{bmatrix} y(k) & y(k+1) & \cdots & y(k+N-1) \end{bmatrix}^T$$

and

$$\Theta = \begin{bmatrix} \theta(k) & \theta(k+1) & \cdots & \theta(k+N-1) \end{bmatrix}^T$$

Eqs. (22) and (27) show that the problem of determining the linear parameters is essentially a linear least square problem and is feasible to be solved by some linear methods such as least square error (LSE) and recursive least square (RLS).

In the proposed method in this paper, both RLS and LSE algorithms are applied as is detailed in the next section.

Eq. (22) is rewritten as

$$E = \|Q - Y\| \tag{28}$$

To find an optimal coefficient vector A^* using the LSE algorithm it is approximated

$$Q = \Theta A^* \tag{29}$$

The optimal $A^*$is given by

$$A^* = \Theta^+ Q \tag{30}$$

where $\Theta^+$is the pseudo-inverse of Θ.

$$\Theta^+ = \left(\Theta^T \Theta\right)^{-1} \Theta^T \tag{31}$$

Also, the modification of the linear parameters using the RLS algorithm with forgetting factor λ is given by

$$\begin{aligned} A^*(i+1) &= A^*(i) + S_{i+1}\theta(i+1)\left(q(i+1) - \theta(i+1)^T A^*(i)\right) S_{i+1} \\ &= \frac{1}{\lambda}\left[S_i - \frac{S_i\theta(i+1)\theta(i+1)^T S_i}{\lambda + \theta(i+1)^T S_i \theta(i+1)}\right] \end{aligned} \tag{32}$$

where $S_0 = \gamma I$ that γ is a large positive constant and I is identity matrix of rank Mn.

4.1.4. Eliminating the redundant neurons

Sometimes, an RBF neuron may be generated initially, but eventually contributes little to the system. Therefore, a pruning technique, which is the error reduction ratio (ERR) method applied in [18], is used here to simplify the system structure of the FLN–RBFN-based filter.

Eqs. (27) and (28) can be rewritten as

$$Q = \Theta A + E \tag{33}$$

Error reduction ratio (ERR) is an indication of the significance of each RBF neuron towards the reduction in the total mean squared error. To calculate the individual contribution of each neuron, Θ is transformed into a set of orthogonal basis vectors. Matrix Θ can be decomposed into

$$\Theta = HR \tag{34}$$

where R is an $nM \times nM$ upper triangular matrix and H is an $N \times nM$ matrix with orthogonal columns h_i. It is worth mentioning that the space spanned by the set of orthogonal basis vectors h_i, is the same space spanned by the set of the column vectors of matrix Θ [23].

Substituting Eq. (34) into Eq. (33), it yields

$$Q = HRA + E = HG + E \tag{35}$$

The LSE solution of G is given by

$$G = \left(H^T H\right)^{-1} H^T Q \tag{36}$$

or

$$g_i = \frac{{h_i}^T Q}{{h_i}^T h_i} \tag{37}$$

Because h_i and h_j are orthogonal for $i \neq j$, the energy of Q is

$$Q^T Q = \sum_{i=1}^{nM} g_i^2 h_i^T h_i + E^T E \tag{38}$$

Therefore, the ERR due to h_i can be considered as

$$err_i = \frac{g_i^2 h_i^T h_i}{Q^T Q} \tag{39}$$

Substituting g_i by Eq. (37) yields

$$err_i = \frac{\left(h_i^T Q\right)^2}{h_i^T h_i Q^T Q} \tag{40}$$

In order to measure the contribution of each RBF neuron, the ERR matrix $\Delta = \left[\begin{array}{cccc} \delta_1 & \delta_2 & \cdots & \delta_n \end{array}\right]$ is constructed by the elements obtained from Eq. (40) in the order that $\Delta \in \Re^{M \times n}$ and the jth column of Δ is the ERR corresponding to the jth RBF unit, where $j = 1,\ldots,n$ [18].

Ultimately η_j is defined, which represents the significance of the jth RBF neuron.

$$\eta_j = \sqrt{\frac{\delta_j^T \delta_j}{M}} \tag{41}$$

and then the RBF neurons, for which $\eta_j < k_{err}$, are eliminated to simplify the structure of the adaptive filter. k_{err} is a pre-specified threshold [18].

4.2. Learning algorithm

Although generally the multichannel EEG distribution is considered as multivariate Gaussian, the mean and covariance properties generally change from segment to segment. Therefore EEGs are considered stationary only within short intervals, i.e. quasi-stationarity. Normal spontaneous activity is essentially stationary. This Gaussian assumption holds during a normal brain condition, but during mental and physical activities this assumption is not valid. When considering long time periods, several factors make it necessary to treat the EEG signal as a non-stationary, stochastic process. Some examples of non-stationarity of the EEG signals can be observed during the change in alertness and wakefulness, during eye blinking and during the transitions between various ictal states. Yet another factor is the intermittent occurrence of transient waveforms such as epileptic spikes or pulse-shaped artifacts [1,12].

Offline learning algorithm has more simple computations, but as it is done based on all incoming data the same, the EEG signal is assumed stationary indeed. Although this assumption is a bit far from reality, as mentioned, the algorithm works well in artifact removal and the improvement in results is significantly considerable. Online learning algorithm, however has more complicated computations, is useful in online applications and also as it is updated based on the new incoming patterns in each cycle, well meets the non-stationarity of the EEG signal.

4.2.1. Offline learning algorithm

The offline learning algorithm for the proposed adaptive filter is summarized as follows. The linear weights determination in this algorithm is done by standard LSE method based on all input–output patterns the same.

i. For first incoming pattern, generate a new RBF neuron.
ii. For the kth new incoming pattern, check the MFS criterion.
iii. If met, adjust the existed neurons.
iv. Else, generate a new RBF neuron.

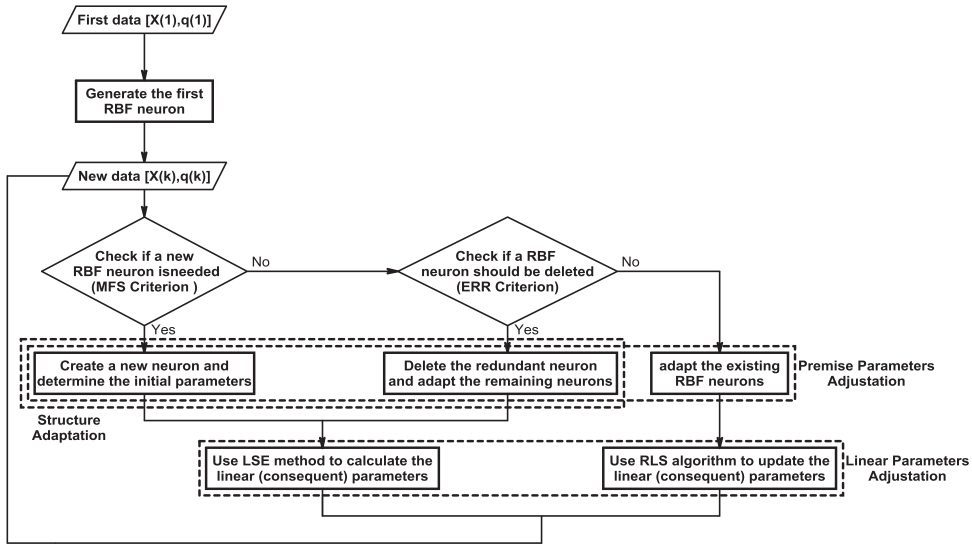

Fig. 5. Flow diagram of the online learning algorithm for adaptive FLN–RBFN-based filter.

v. Check the ERR criterion and delete the redundant neurons.
vi. Determine the linear weights using LSE.

4.2.2. Online learning algorithm

Fig. 5 presents flow diagram of the online learning scheme for the adaptive FLN–RBFN-based filter. Both structure identification phase and parameter determination phase are depicted in the diagram. At linear parameters modification level, in the case where no neurons have been added to or deleted from the hidden layer, the modification of the weights is performed using the RLS algorithm with forgetting factor λ. If the structure of the hidden layer has been modified either by adding or deleting a node, then the connection weights need to be recalculated. This calculation is based on a time window, where a specified number of past input–output data are stored. The new connection weights are obtained using the LSE on the time window. As a result, in both cases the influence of the older data points is weakened and more importance is given to the new data points.

5. Results and discussion

Real signals taken from the database available in the physionet website [24] were used to perform simulations. For demonstrating the proposed method, a real measured EEG signal, which is contaminated to artifacts by its nature, and artifact source signals i.e. EOG, EMG and ECG measured simultaneously with the EEG signal are taken. The signals are collected from a subject in the least muscular activities mode. ANC method using the proposed adaptive FLN–RBFN-based filter is performed to remove the ocular, muscular and cardiac artifacts from EEG signal. In all simulations, the primary input of the ANC system (the target signal for training the adaptive filter) is the measured EEG signal that is naturally contaminated by artifacts and two reference inputs (the inputs of the adaptive filter) are considered include the artifact source signal measured directly from the artifact generation origin and its delayed version i.e. $r=2$.

5.1. EEG artifact removal

In this section, the results of ocular, muscular and cardiac artifacts removal from EEG signals using the proposed adaptive FLN–RBFN-based filter are discussed.

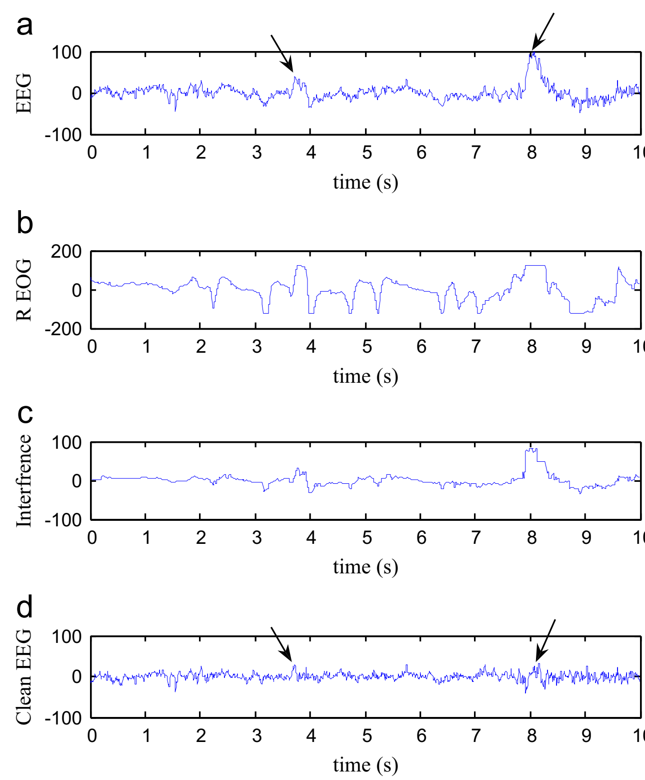

Fig. 6. EOG artifact cancellation: (a) EEG record contaminated with EOG artifacts; (b) real EOG; (c) estimated ocular interference; and (d) cleaned EEG (output signal from the ANC system).

Fig. 6 indicates the performance of the proposed filter in eliminating the ocular artifact. Fig. 6a and b shows the measured EEG signal and directly measured EOG signal respectively, the estimated ocular interference signal using the proposed adaptive filter is shown in Fig. 6c, and ultimately Fig. 6d illustrates the clean EEG resulting from subtracting the estimated interference signal from the measured EEG as shown in the ANC system. Considering the parts marked by arrows that indicate the severe presence of ocular artifacts, it is obvious that the proposed filter

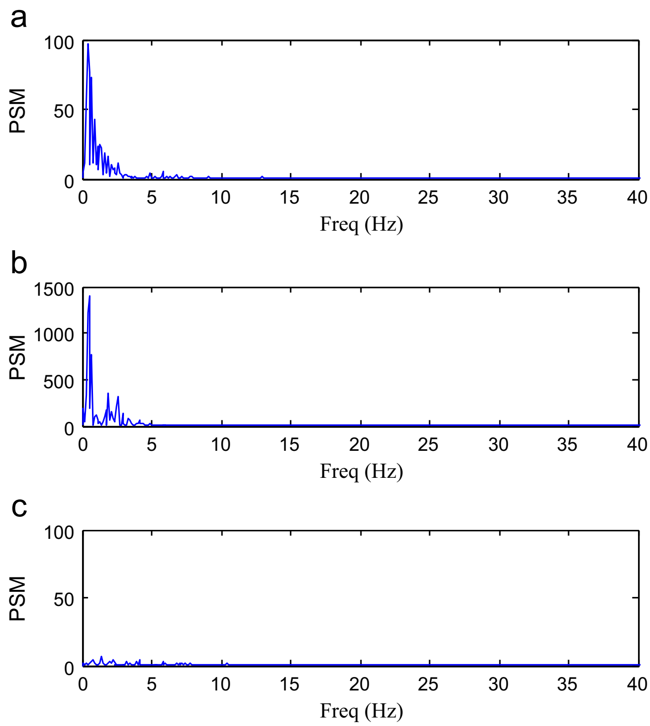

Fig. 7. EOG artifact cancellation in frequency domain: (a) PSM of measured EEG; (b) PSM of real EOG; and (c) PSM of the clean EEG.

has removed the artifacts successfully and has desirable performance in this area.

The result is also achieved in frequency domain using power spectrum magnitude (PSM) criterion. Fig. 7a shows the power spectrum of the EEG signal contaminated with EOG, the power spectrum of EOG signal is shown in Fig. 7b and the power spectrum of cleaned EEG is illustrated in Fig. 7c. Note that the absence of low frequencies of EOG in output signal indicates the removal of EOG artifact.

The performance of the proposed filter in eliminating the EMG and ECG artifacts from EEG signal are also illustrated in Figs. 8 and 9 respectively. As expected, since the subject is in relaxing mode with the least muscular activity and also has the healthy heart in the selected sample data, the contamination resulting from EMG and ECG artifacts is not so severe. It is obvious from Figs. 8 and 9 that the proposed adaptive filter has been able to detect and remove even these little artifacts properly. To illustrate the performance more clearly, the correlation criterion is also used to indicate the level of artifact removal. The results are shown in Tables 1 and 2 respectively. The considerable decrease in the amount of correlation between the artifact source signal and clean EEG compared to the contaminated EEG in both cases reveals the desirable performance of the proposed adaptive filter.

To demonstrate the superiority of proposed filter, performance of adaptive FLN–RBFN-based filter was compared to the adaptive neuro-fuzzy inference system (ANFIS) [15–17] and the adaptive RBFN-based filter [18,19]. In order to compare the results, the mean square error (*MSE*) criterion is applied as follows:

$$\mathrm{MSE} = \frac{\sum_{k=1}^{p} (q(k)-y(k))^2}{p} \tag{42}$$

where, $q(k)$ and $y(k)$ are the desired and real output of the adaptive filter respectively, and p is the number of patterns. Since it has the most severe contamination effect, the comparison is done for ocular artifact removal from EEG signal. The results shown in Table 3 indicate that the proposed adaptive filter has removed the artifact more successfully and the least value of *MSE* is obtained by using the least number of neurons among three compared filters. The results represent the superiority of the proposed adaptive filter to other mentioned ones.

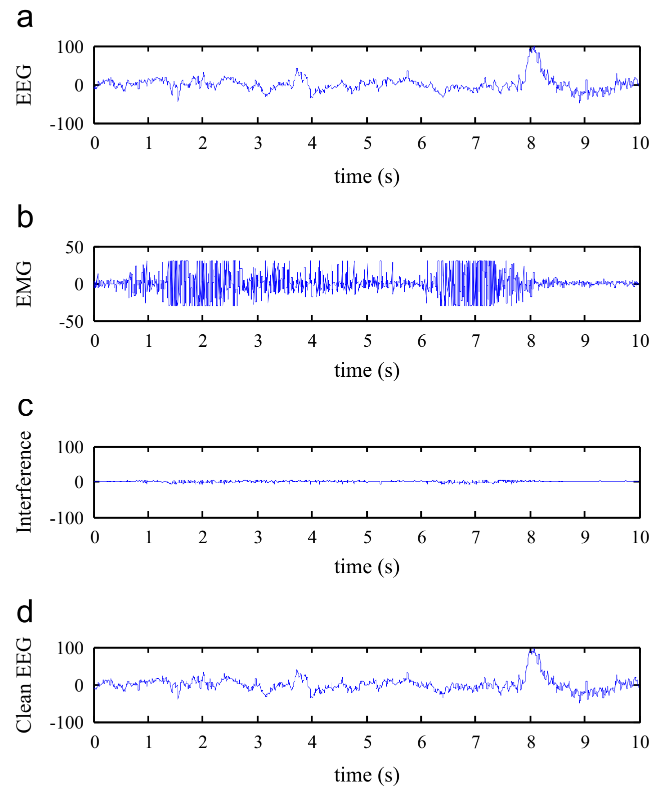

Fig. 8. EMG artifact cancellation: (a) EEG record; (b) real EMG; (c) estimated ocular interference; and (d) cleaned EEG.

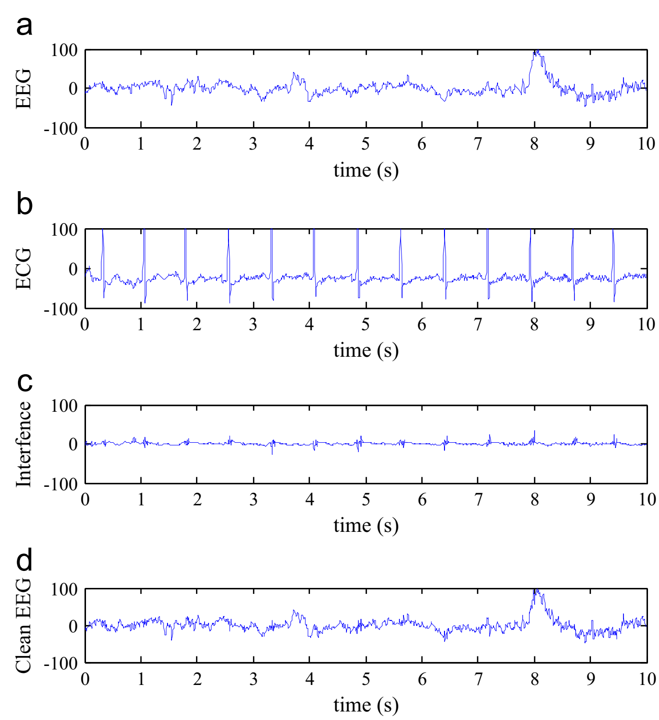

Fig. 9. ECG artifact cancellation: (a) EEG record; (b) real ECG; (c) estimated ocular interference; and (d) cleaned EEG.

Table 1
Correlation between EEG signal and EMG interference.

Mode	Correlation
Before muscular artifact removal	7.81×10^{-4}
After muscular artifact removal	2.62×10^{-16}

Table 2
Correlation between EEG signal and ECG interference.

Mode	Correlation
Before cardiac artifact removal	-0.04
After cardiac artifact removal	2.13×10^{-13}

Table 3
Comparison of proposed filter with two older filters.

Adaptive filter	Number of neurons	*MSE*
ANFIS	9	90.8377
RBFN-based	9	90.6095
FLN–RBFN-based	8	80.0320

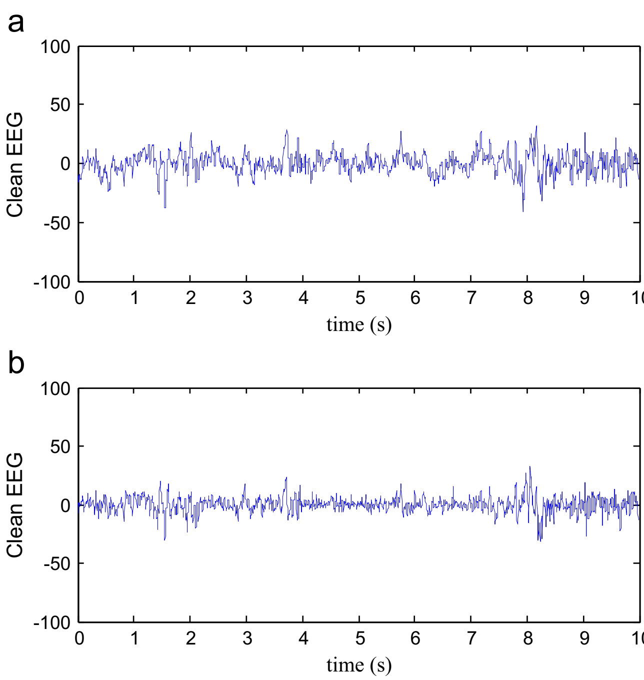

Fig. 10. Performance of online and offline learning algorithms in EOG artifact removal from EEG Signal: (a) offline algorithm, and (b) online algorithm.

5.2. Comparison of online and offline learning algorithm

As observed from the results obtained in previous sections the proposed adaptive filter can eliminate the artifacts properly using the offline algorithm and performs better than ANFIS and adaptive RBFN-based filter. However, while in many cases the offline algorithm is adequate and preferable because of its more simple computations than the online algorithm, in addition to the necessity of using online algorithm for online applications, it is shown that the online learning algorithm performs better in non-stationarities. To do the comparison, since the ocular activity has caused a non-stationarity in EEG signal, both offline and online learning algorithms have been applied to remove the ocular artifact from EEG signal using the proposed adaptive filter. Using the online learning algorithm $MSE = 69.6312$ is achieved using maximum 4 neurons in each training level; whereas, for the offline algorithm $MSE = 80.032$ was obtained using 8 neurons.

The resulted clean EEG signals using both algorithms are also shown in Fig. 10. The smoother and cleaner signal obtained, proves the better behavior of the online learning algorithm against non-stationarities.

6. Conclusion

In this paper, a new adaptive FLN–RBFN-based filter is proposed to remove ocular, muscular and cardiac artifacts from EEG signal using ANC principle. The proposed filter uses a Chebyshev FLN to the consequent part of the fuzzy rules constructed by RBF neurons. Simulation results show that the proposed adaptive filter performs acceptable and removes artifacts from EEG signals successfully. Both online and offline learning algorithms are presented for the proposed filter each of which is suitable in certain conditions and also for specific applications. The advantage of the offline algorithm is its more simple computations than the online algorithm, but this algorithm cannot be applied for online applications and also does not act suitable for non-stationary conditions. Therefore, presented offline algorithm is more preferable for the offline applications with low level of non-stationarity. Online learning algorithm is applied in online applications and also removes the artifacts successfully in non-stationary situations such as severe ocular artifacts.

The performance of the proposed adaptive filter in removing the artifacts from EEG signals has also been compared to ANFIS and adaptive RBFN-based filter. The results confirm the superiority of the proposed filter.

References

[1] L. Sornmo, P. Laguna, Bioelectrical Signal Processing in Cardiac and Neurological Applications, Elsevier Academic Press, Amsterdam, 2005.
[2] T. Zikov, S. Bibian, G.A. Dumont, M. Huzmezan, C.R. Ries, A wavelet based de-nosing technique for ocular artifact correction of the electroencephalogram, in: Proceedings of the 24th Annual Conferenceof the IEEE Engineering in Medicine and Biology Society, 1, 2002, pp. 98–105.
[3] V. Krishnaveni, S. Jayaraman, N. Malmurugan, A. Kandaswamy, K. Ramadoss, Non adaptive thresholding methods for correcting ocular artifacts in EEG, Acad. Open Internet J. 13 (2004).
[4] V. Krishnaveni, S. Jayaraman, S. Aravind, V. Hariharasudhan, K. Ramadoss, Automatic identification and removal of ocular artifacts from EEG using wavelet transform, Meas. Sci. Rev. 6 (4) (2006) 45–57.
[5] P. Senthil Kumar, R. Arumuganathan, K. Sivakumar, C. Vimal, Removal of ocular artifacts in the EEG through wavelet transform without using an EOG reference channel, Int. J. Open Probl. Comput. Sci. Math. 1 (3) (2008) 188–200.
[6] T.-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, T.J. Sejnowski, Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects, Clin. Neurophysiol. 111 (2000) 1745–1758.
[7] T.-P. Jung, S. Makeig, C. Humpheries, T.-V. Lee, M.J. Mckeown, V. Iragui, T.J. Sejnowski, Removing electroencephalographic artifacts by blind source separation, Psychophysiology 37 (2000) 163–178.
[8] A. Garcés Correa, E. Laciar, H.D. Patiño, M.E. Valentinuzzi, Artifact removal from EEG signals using adaptive filters in cascade, J. Phys.: Conf. Ser. 90 (2007) 1–10.
[9] Z. Sakkaff, A. Nanayakkara, Removal of ocular artifacts from EEG signals in brain computer interface, in: Proceedings of the Technical Sessions 24, 2008, pp. 51–57.
[10] A. Schlogl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, G. Pfurtscheller, A. Fully, Automated correction method of EOG artifacts in EEG recordings, Clin. Neurophysiol. 118 (2007) 98–104.
[11] P. He, G. Wilson, C. Russell, Removal of ocular artifacts from electroencephalogram by adaptive filtering, Med. Biol. Eng. Comput. 42 (3) (2004) 407–412.
[12] S. Sanei, J.A. Chambers, EEG Signal Processing. John Wiley & Sons Ltd. Chichester England, 2007.

[13] M.J. Er, Z. Li, Adaptive noise cancellation using online self-improved fuzzy filter with applications to multimedia processing, Intell. Multimedia Process. Soft Comput. (2004) 389–414.
[14] C.-H. Chen, C.-J. Lin, C.-T. Lin, A functional-link based neuro-fuzzy network for nonlinear system control, IEEE Trans. Fuzzy Syst. 16 (5) (2008) 1362–1378.
[15] C.K.S. Vijilal, P. Kanagasabapathy, S. Johnson, V. Ewards, Artifacts removal in EEG signal using adaptive neuro-fuzzy inference system, in: Proceedings of the IEEE International Conference on Signal Processing, Communications and Networking, 2007 pp. 589–591.
[16] C.K.S. Vijila1, C.E.S. Kumar, Interference cancellation in EMG signal using ANFIS, Int. J. Recent Trends Eng. 2 (5) (2009) 244–248.
[17] S. Saranya, S. SujaPriyadharsini, A. Novel, Hybrid soft computing technique for extracting fetal ECG from maternal ECG signal, Int. J. Comput. Appl. 3 (8) (2010) 7–13.
[18] Z. Li, M.J. Er, Y. Gao, An adaptive RBFN-based filter for adaptive noise cancellation, in: Proceedings of the 42nd IEEE Conference on Decision and Control, 6, 2003, pp. 6175–6180.
[19] F. Liu, M.J. Er, A learning algorithm for constructing fuzzy neural networks with application to regression problems, in: Proceedings of the International Conference on Information Science and Technology, 2011, pp. 318–322.
[20] M.A. Abdulmagid, D.J. Krusienski, S. Pal, W.K. Jenkins, Principles of adaptive noise cancelling, Annu. Res. J. 2 (2004) 52–59.
[21] S. Dehuri, S.-B. Cho, A comprehensive survey on functional link neural networks and an adaptive PSO–BP learning for CFLNN, Neural Comput. Appl. 19 (2) (2010) 187–205.
[22] G. Panda, D.P. Das, Functional link artificial neural network for active control of nonlinear noise processes, in: Proceedings of the International Workshop on Acoustic Echo and Noise Control, 2003, pp. 163–166.
[23] S. Chen, C.F.N. Cowan, P.M. Grant, Orthogonal least squares learning algorithm for radial basis function networks, IEEE Trans. Neural Networks 2 (2) (1991) 302–309.
[24] 〈www.physionet.org〉.

Mohammad Ali Badamchizadeh was born in Tabriz, Iran, in December 1975. He received the B.S. Degree in Electrical Engineering from University of Tabriz in 1998 and the M.Sc. Degree in Control Engineering from University of Tabriz in 2001. He received the Ph.D. Degree in Control Engineering from University of Tabriz in 2007. He is now an Assistant Professor in the Faculty of Electrical and Computer Engineering at University of Tabriz. His research interests include adaptive control, hybrid dynamical systems, stability of systems, time delay systems.

Aysa Jafarifarmand was born in Tabriz, Iran, in 1986. She received the B.S. Degree in Power Electrical Engineering from University of Tabriz in 2008. She received the M.Sc. Degree in Control Engineering from University of Tabriz in 2011. She is currently a Ph.D. candidate of Control Engineering at University of Tabriz. Her research is centered on artificial intelligence and brain-computer interface.