Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

Signal Processing 86 (2006) 2351–2363

SIGNAL PROCESSING

www.elsevier.com/locate/sigpro

Robust adaptive techniques for minimization of EOG artefacts from EEG signals

S. Puthusserypady[a,*,1], T. Ratnarajah[b,1]

[a]*Department of Electrical and Computer Engineering, National University of Singapore, 4 Engineering Drive 3, Singapore 117576, Singapore*
[b]*ECIT, Queen's University of Belfast, Queen's Road, Queen's Island, Belfast, BT3 9DT, UK*

Received 9 July 2003; received in revised form 7 June 2005; accepted 28 October 2005
Available online 5 December 2005

Abstract

In this paper, we propose the application of H^∞ techniques for minimization of electrooculogram (EOG) artefacts from corrupted electroencephalographic (EEG) signals. Two adaptive algorithms (*time-varying* and *exponentially-weighted*) based on the H^∞ principles are proposed. The idea of applying H^∞ techniques is motivated by the fact that they are robust to model uncertainties and lack of statistical information with respect to noise [B. Hassibi, A.H. Sayed, T. Kailath, Linear estimation in Krein spaces—Part 1: theory & Part II: applications, IEEE Trans. Automat. Control 41 (1996) 18–49]. Studies are performed on simulated as well as real recorded signals. Performance of the proposed techniques are then compared with the well-known least-mean square (LMS) and recursive least-square (RLS) algorithms. Improvements in the output signal-to-noise ratio (SNR) along with the time plots are used as criteria for comparing the performance of the algorithms. It is found that the proposed H^∞-based algorithms work slightly better than the RLS algorithm (especially when the input SNR is very low) and always outperform the LMS algorithm in minimizing the EOG artefacts from corrupted EEG signals.

Keywords: Electrooculogram (EOG) artefacts; Electroencephalogram (EEG); Robust adaptive filtering

1. Introduction

The brain generates electrical activity which can be recorded from the scalp by placing electrodes on the intact skull [1]. This recorded representation of brain electrical potentials is called the electroencephalogram (EEG). EEG is one of the commonly used non-invasive techniques for understanding the brain functions, abnormalities, injuries, etc. EEG mainly contains frequency-related activities. In addition to these frequency-related activities, EEG also contains specific transients or paroxysmal activities (called spikes or sharp waves or spike and wave activity, depending on their characteristics) which occur spontaneously [2].

Automation of EEG processing has been an intense research area in brain signal processing for several years. Compared to the manual processing

*Corresponding author. Tel.: +65 6874 2262; fax: +65 6779 1103.
E-mail addresses: elespk@nus.edu.sg (S. Puthusserypady), T.Ratnarajah@ieee.org (T. Ratnarajah).
[1]The authors would like to thank the anonymous reviewers for their valuable comments on the earlier version of this manuscript.

0165-1684/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2005.10.018

of EEG signals, automatic processing by computers is desirable in both clinical and experimental EEG analysis for the following reasons: (i) to obtain precise characterization and quantification, (ii) to avoid errors due to subjectivity and (iii) to minimize processing time. However, automated analysis of the EEG is bedeviled by the lack of reliable means for removing artefacts and, in particular, for distinguishing artefacts and pathological waves of cerebral origin.

Artefacts that are almost always present in EEG signals, are caused by both external and internal sources to the body. Of these, electrooculogram (EOG) generated by eye movements and/or blinks is found to be the most significant and common artefact. EOG signals spreads across the scalp to contaminate the EEG and is known as the ocular artefact (OA). It not only reduces the clinical usefulness of EEG signals but also make its analysis difficult. In some cases, it even makes the analysis impossible because of the similarity between artefacts and the pathological signals of interest. Therefore, effective minimization/rejection of artefacts from the collected data is essential in preparing the data for further analysis. Thus, the removal of EOG artefacts from contaminated EEG signals forms an important part of the analysis of EEG signals. Basically, there are three different approaches to the minimization/rejection of EOG artefacts from contaminated EEG recordings. They are: (i) rejection methods [3,4], (ii) eye fixation methods [5], and (iii) EOG subtraction methods [6–18]. Many other artefact minimization schemes based on independent component analysis (ICA), principal component analysis (PCA), and neural networks are also reported in various research publications [19–23].

EOG artefacts are generally high amplitude and low frequency in nature and affect mainly the low frequency region of EEG signals. Because of this spectral overlapping between the OA and some EEG signals, conventional low-pass filtering techniques cannot be applied for effective minimization of EOG from these corrupted EEG signals. Hence, we resort to adaptive noise cancellation (ANC) techniques which have many advantages over the fixed filtering schemes [24]. Related works have been reported in many publications [25–30] and the references therein.

The H^∞ approach was introduced in robust control theory on the hypothesis that the resulting minmax estimation techniques would be less sensitive to model uncertainties and parameter variations than conventional techniques [31]. These methods safeguard against the *worst-case* disturbances and therefore make no assumptions on the (statistical) nature of the signals. This replaces the method of modeling the disturbance signal as a random process with a given spectral density, a method that has led to the well-known linear quadratic regulator and Kalman filter. The performance of these classical estimation methods heavily depends on the validity of these model assumptions and hence limits their application.

In this work, the problem of minimizing EOG artefacts from EEG signals has been re-formulated to fit into the ANC framework. Two H^∞-based adaptive algorithms namely, the H^∞ *time-varying* (H^∞ TV) and H^∞ *exponentially weighted* (H^∞ EW) algorithms are proposed for effective minimization of EOG artefacts from contaminated EEG signals. Performance of these algorithms are compared with the popular adaptive filtering algorithms such as the least-mean square (LMS) and recursive least-square (RLS). In the following section, details of the H^∞ techniques for the EOG minimization problem are given.

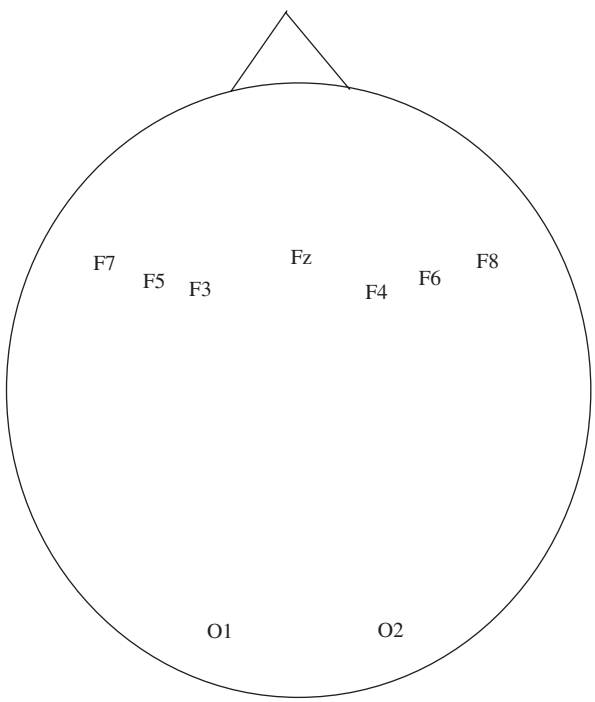

Fig. 1. Locations of the EEG electrodes.

2. Problem formulation

The schematic diagram of the EEG and EOG recording procedure is shown in Figs. 1 and 2, respectively. It is well known that an ANC scheme requires two inputs namely, the *primary input*, $y(n)$, and at least one *reference input*, $r(n)$ [24]. The signal scenario can be modeled as

$$y(n) = s(n) + z(n), \tag{1a}$$

$$r(n) = q(n) + u(n), \tag{1b}$$

where $s(n)$ is the signal of interest (EEG) and $z(n)$ is the artefact which corrupts the observations at the primary input sensor. $q(n)$ is the actual source (EOG) of this artefact which is correlated to $z(n)$ (i.e., $\mathscr{E}\{q(n)z(n)\} \neq 0$). We assume that $y(n)$ and $r(n)$ are correlated in some fashion whereas $s(n)$ and $r(n)$ are uncorrelated. $u(n)$ in the above equation is the measurement noise. The reference input $r(n)$ is filtered to produce an output (estimate) $\hat{z}(n)$ that is as close as $z(n)$. This estimate is to be subtracted from $y(n)$ to produce an estimate of the desired signal (estimate of EEG), $\hat{s}(n) = y(n) - \hat{z}(n)$. The schematic diagram of such an ANC scheme is shown in Fig. 3.

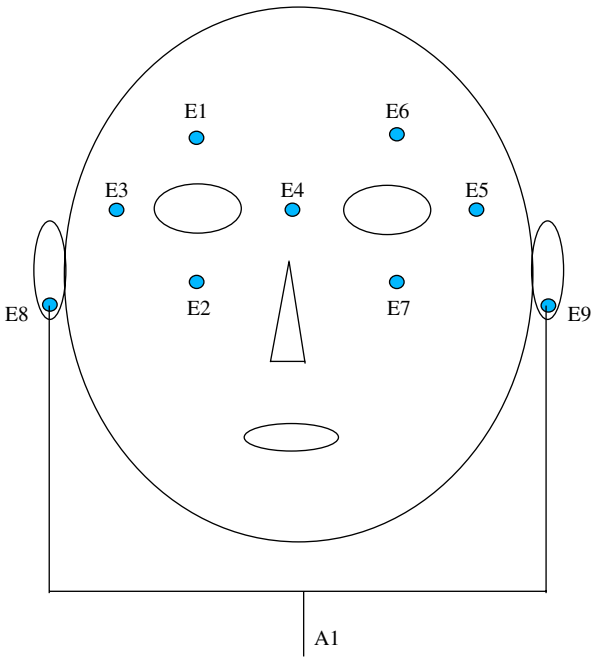

Fig. 2. Locations of the EOG electrodes.

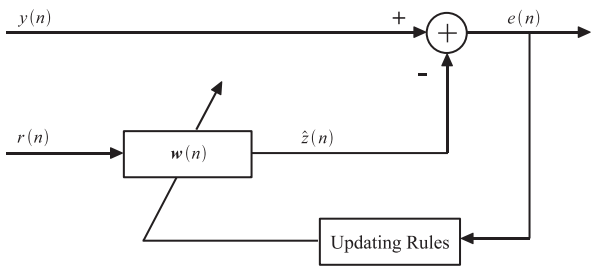

Fig. 3. Block diagram of the proposed noise minimization scheme.

3. H^∞ solution to artefact minimization problem

In this section, we present the development of adaptive algorithms based on the H^∞ approach for the estimation and minimization of EOG artefacts from contaminated EEG signals.

3.1. System model

It can be shown that the signal $y(n)$ at a particular time instant n obeys the following state-space model [1,31]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \Delta\mathbf{w}(n), \tag{2a}$$

$$y(n) = \mathbf{q}^{\mathrm{T}}(n)\mathbf{w}(n) + v(n), \tag{2b}$$

$$z(n) = \mathbf{q}^{\mathrm{T}}(n)\mathbf{w}(n) + v(n), \tag{2c}$$

$$v(n) = s(n) + v(n), \tag{2d}$$

where $\mathbf{w}(n) = [w_0, w_1, \ldots, w_{P-1}]^{\mathrm{T}}$ is the filter coefficient vector of size $P \times 1$, $\mathbf{q}(n) = [q(n), q(n-1), \ldots, q(n-P+1)]^{\mathrm{T}}$ is the artefact (actual) vector of size $P \times 1$ at the nth time instant and P is the order of the filter. In the above equation, $y(n)$ is the observed signal and $z(n)$ is the signal we intend to estimate. In this equation, $v(n)$ includes the EEG ($s(n)$) we intend to estimate and the model uncertainties ($v(n)$). $\Delta\mathbf{w}(n)$ represents the time variation in the filter coefficient vector. As it is unknown, we shall consider it as a disturbance.

Based on the above state-space model, three different H^∞ adaptive filter formulations are proposed for the purpose of coping with *time-variations*, *non-stationarity* and *model uncertainties*. They are:

(i) The time-variation problem.
(ii) The EW problem.
(iii) The finite-memory problem.

3.2. H^∞-based ANC schemes

A block diagram of the proposed noise minimization scheme for the linear model is shown in Fig. 3. Using the above state-space model (Eq. (2)), the objective is to estimate the unknown output signal $z(n)$ from the observations of $r(n)$. Let $\hat{z}(n) = \mathscr{F}(r(n),\ldots,r(n-P+1))$ denote the estimate of $z(n)$ given the observations $\{r(i); i = n,\ldots,(n-P+1)\}$. Since we are interested in estimating the output of the filter, we define the output estimation error as

$$\begin{aligned} e(n) &= y(n) - \hat{z}(n) \\ &= y(n) - \mathbf{r}^{\mathrm{T}}(n)\hat{\mathbf{w}}(n), \end{aligned} \tag{3}$$

i.e., as the difference between the uncorrupted output $(\mathbf{r}^{\mathrm{T}}(n)\mathbf{w}(n))$ and the output estimated $(\mathbf{r}^{\mathrm{T}}(n)\hat{\mathbf{w}}(n))$. Here, $\hat{\mathbf{w}}(n)$ is the weight vector at nth time instant. Moreover, the time variation in the filter coefficient vector (state vector) $\mathbf{w}(n)$,

$$\Delta\mathbf{w}(n) = \mathbf{w}(n+1) - \mathbf{w}(n) \tag{4}$$

is unknown, we shall consider it as a disturbance. Furthermore, define the weighted disturbances $\tilde{\mathbf{\Delta}}(n)$ and $\tilde{\mathbf{w}}(0)$ as follows:

$$\begin{aligned} \tilde{\mathbf{\Delta}}(n) &\equiv \mathbf{\Upsilon}_0^{-1/2}\Delta\mathbf{w}(n), \\ \tilde{\mathbf{w}}(0) &\equiv \mathbf{\Pi}_0^{-1/2}(\mathbf{w}(0) - \hat{\mathbf{w}}(0)), \end{aligned} \tag{5}$$

where $\hat{\mathbf{w}}(0)$ is an initial estimate of the state vector $\mathbf{w}(0)$; $\mathbf{\Pi}_0$ is a positive definite matrix reflecting a priori knowledge of how close $\mathbf{w}(0)$ is to $\hat{\mathbf{w}}(0)$; $\mathbf{\Upsilon}_0$ is a positive definite matrix that reflects a priori knowledge of how rapidly the state vector $\mathbf{w}(n)$ varies with time. For every choice of estimator $\mathscr{F}(\cdot)$, we will have a transfer operator $(\mathscr{T}_n(\mathscr{F}))$ from disturbances,

$$\begin{aligned} &\{\tilde{\mathbf{w}}(0), \lambda^{-k/2}v(k),\ldots,\lambda^{-(n-1)/2}v(n-1), \\ &\quad \tilde{\mathbf{\Delta}}(k),\ldots,\tilde{\mathbf{\Delta}}(n-1)\} \end{aligned}$$

to the output estimation error,

$$\{\lambda^{-k/2}e(k), \lambda^{-(k+1)/2}e(k+1),\ldots,\lambda^{-n/2}e(n)\},$$

where $0<\lambda<1$ is the forgetting factor that is chosen based on a priori knowledge of how fast the state matrix varies with time. The variable k in the above expressions is an arbitrary index to generalize the formulation of all the three algorithms mentioned at the end of this section. In the H^∞ framework, robustness is ensured by minimizing the maximum energy gain from the disturbances to the estimation errors. This leads to the following scenario [31].

Problem. Find an H^∞ optimal estimator $\hat{z}(n) = \mathscr{F}(r(n),\ldots,r(n-P+1))$ which satisfies the following:

$$\varepsilon^2 = \inf_{\mathscr{F}} \|\mathscr{T}_n(\mathscr{F})\|_\infty^2 = \inf_{\mathscr{F}} \sup_{\mathbf{w}(0),v(i),\tilde{\mathbf{\Delta}}(i)} \gamma(n), \tag{6}$$

where

$$\gamma(n) = \frac{\sum_{i=k}^{n} \lambda^{-i}e^2(i)}{\|\tilde{\mathbf{w}}(0)\|_2^2 + \sum_{i=k}^{n-1} \lambda^{-i}v^2(i) + \sum_{i=k}^{n-1} \tilde{\mathbf{\Delta}}^{\mathrm{T}}(i)\tilde{\mathbf{\Delta}}(i)}. \tag{7}$$

$\gamma(n)$ in Eq. (7) is the gain of the transfer operator which is the ratio of the energy of the estimation error to the energy of the disturbance. We shall assume, without loss of generality, that $\mathbf{\Pi}_0$ and $\mathbf{\Upsilon}_0$ have the special form $\mathbf{\Pi}_0 = \eta\mathbf{I}$ and $\mathbf{\Upsilon}_0 = \rho\mathbf{I}$, where η and ρ are positive constants. Note that for a filter that varies slowly with time, ρ will typically be very small.

The above formulation can be used to handle three different scenarios:

(i) The time-variation problem: $\lambda = 1$ and $k = 1$.
(ii) The EW problem: $0<\lambda<1$ and $k = 1$.
(iii) The finite-memory problem: $\lambda = 1$ and $k = n - L + 1$.

3.3. Solution 1: the TV algorithm

We have the following solution to the TV problem:

$$\mathbf{g}(n) = \frac{\tilde{\mathbf{P}}(n)\mathbf{r}(n)}{1 + \mathbf{r}^{\mathrm{T}}(n)\tilde{\mathbf{P}}(n)\mathbf{r}(n)}, \tag{8a}$$

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mathbf{g}(n)(y(n) - \mathbf{r}^{\mathrm{T}}(n)\hat{\mathbf{w}}(n)), \tag{8b}$$

$$\tilde{\mathbf{P}}^{-1}(n) = \mathbf{P}^{-1}(n) - \varepsilon_g^{-2}\mathbf{r}(n)\mathbf{r}^{\mathrm{T}}(n), \tag{8c}$$

$$\mathbf{P}(n+1) = [\mathbf{P}^{-1}(n) + (1 - \varepsilon_g^{-2})\mathbf{r}(n)\mathbf{r}^{\mathrm{T}}(n)]^{-1} + \mathbf{\Upsilon}_0, \tag{8d}$$

$$\mathbf{P}(0) = \eta\mathbf{I} = \mathbf{\Pi}_0. \tag{8e}$$

Here, $\mathbf{g}(n)$ is the gain factor, ε_g is a positive constant defined as in Eq. (6) and $\mathbf{P}(n)$ is the inverse of the correlation matrix. The correlation matrix is given by $\sum_{i=1}^{n} \mathbf{r}(i)\mathbf{r}^{\mathrm{T}}(i)$.

3.4. Solution 2: the EW algorithm

Alternatively, a forgetting factor λ is introduced and the EW algorithm is derived. In this manner the more recent data will be given more weighting than the earlier data samples. This will allow the algorithm to track the time variations of the underlying models. In particular, the prediction error and the disturbance energies are computed as

$$\sum_{i=1}^{n} \lambda^{-i} e^2(i) \tag{9}$$

and

$$\sum_{i=1}^{n} \lambda^{-i} v^2(i). \tag{10}$$

This forgetting factor λ is chosen on the basis of a priori knowledge of how fast the state vector varies with time (in the simulation examples below, we choose $\lambda = 0.9$). The EW problem is a special case of the general problem (Eq. (8)) with $0<\lambda<1$ and $k = 1$. The update equations are

$$\mathbf{g}(n) = \frac{\bar{\mathbf{P}}(n)\mathbf{r}(n)}{1 + \mathbf{r}^{\mathrm{T}}(n)\bar{\mathbf{P}}(n)\mathbf{r}(n)}, \tag{11a}$$

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{g}(n)(y(n) - \mathbf{r}^{\mathrm{T}}(n)\hat{\mathbf{w}}(n-1)), \tag{11b}$$

$$\bar{\mathbf{P}}^{-1}(n+1) = \lambda\bar{\mathbf{P}}^{-1}(n) + \lambda\mathbf{r}(n)\mathbf{r}^{\mathrm{T}}(n) - \varepsilon_{\lambda}^{-2}\mathbf{r}(n+1)\mathbf{r}^{\mathrm{T}}(n+1), \tag{11c}$$

$$\bar{\mathbf{P}}^{-1}(0) = \eta^{-1}\mathbf{I} - \varepsilon_{\lambda}^{-2}\mathbf{r}(1)\mathbf{r}^{\mathrm{T}}(1). \tag{11d}$$

Here, ε_{λ} is a positive constant defined as in Eq. (6).

3.5. Solution 3: the finite-memory algorithm

Another approach for dealing with time variations is the so-called sliding or finite-memory window. In this case, one considers data from a finite window of length L. Therefore as each new datum is observed, the least-recent datum is discarded so that we have memory of constant size L. Due to the fact that the old data is discarded, this method has the potential to cope with time variations in the underlying model (Eq. (2)). Thus the prediction error and disturbance energies are computed as

$$\sum_{i=n-L+1}^{n} e^2(i) \tag{12}$$

and

$$\sum_{i=n-L+1}^{n} v^2(i). \tag{13}$$

The finite-memory problem is a special case of the general problem (Eq. (8)) with $\lambda = 1$ and $k = n - L + 1$. The update equations are as follows:

$$\mathbf{g}^d(n) = \frac{\mathbf{P}^d(n)\mathbf{r}(n-L)}{-1 + \mathbf{r}^{\mathrm{T}}(n-L)\mathbf{P}^d(n)\mathbf{r}(n-L)}, \tag{14a}$$

$$\hat{\mathbf{w}}^d(n-1) = \hat{\mathbf{w}}(n-1) + \mathbf{g}^d(n)(y(n-L) - \mathbf{r}^{\mathrm{T}}(n-L)\hat{\mathbf{w}}(n-1)), \tag{14b}$$

$$(\mathbf{P}^d(n))^{-1} = \mathbf{P}^{-1}(n) - (1 - \varepsilon_L^{-2})\mathbf{r}(n-L)\mathbf{r}^{\mathrm{T}}(n-L), \tag{14c}$$

$$\mathbf{g}(n) = \frac{\mathbf{P}(n)\mathbf{r}(n)}{1 + \mathbf{r}^{\mathrm{T}}(n)\mathbf{P}(n)\mathbf{r}(n)}, \tag{14d}$$

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}^d(n-1) + \mathbf{g}(n)(y(n) - \mathbf{r}^{\mathrm{T}}(n)\hat{\mathbf{w}}^d(n-1)), \tag{14e}$$

$$\mathbf{P}^{-1}(n+1) = (\mathbf{P}^d(n))^{-1} + (1 - \varepsilon_L^{-2})\mathbf{r}(n+1)\mathbf{r}^{\mathrm{T}}(n+1). \tag{14f}$$

Once we have the estimate of $z(n)$ (i.e., $\mathbf{w}(n)$) using any of the three algorithms, the desired EEG signal can be estimated as

$$\hat{s}(n) = y(n) - \hat{z}(n) = y(n) - \mathbf{r}^{\mathrm{T}}(n)\hat{\mathbf{w}}(n). \tag{15}$$

The following two remarks will describe the other adaptive filtering approaches that are studied by researchers, see [26,31–33]. For brevity, we are considering only the TV model (Eq. (2)) with no explicit time variations of state vector such as the EW or finite memory.

Remark 1. *The H^{∞} approach*: An H^{∞}-optimal estimation strategy for time-invariant model is the well known LMS algorithm, where the robustness is ensured by minimizing the maximum (or worst-case) energy gain from the disturbances $\{\mu^{-1/2}\mathbf{w}, \{v(n)\}\}$, to the estimation error, $\{e(n)\}$ [32,33]. The LMS algorithm is given by

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mu\mathbf{r}(n)(y(n) - \mathbf{r}^{\mathrm{T}}(n)\hat{\mathbf{w}}(n-1)), \tag{16}$$

where μ is the learning rate.

Remark 2. *The H^2 approach*: Consider the time-invariant model and, in contrast to H^{∞} approach, suppose that $\mathbf{w}$ and $\{v(n)\}$ are zero-mean, uncorrelated

random variables with variances $\mu\mathbf{I}$ ($\mu>0$) and unity, respectively. Find an H^2-optimal estimation strategy that minimizes the sum of weighted-error squares:

$$\sum_{i=1}^{n} \lambda^{n-i} e^2(i) \tag{17}$$

for all n.

The solution to the above problem is the well-known EW RLS algorithm [34]. The update equations are as shown below:

$$\mathbf{g}(n) = \frac{\mathbf{P}(n)\mathbf{r}(n)}{\lambda + \mathbf{r}^{\mathrm{T}}(n)\mathbf{P}(n)\mathbf{r}(n)}, \tag{18a}$$

$$\bar{\mathbf{w}}(n) = \bar{\mathbf{w}}(n-1) + \mathbf{g}(n)(y(n) - \mathbf{r}^{\mathrm{T}}(n)\bar{\mathbf{w}}(n-1)), \tag{18b}$$

$$\mathbf{P}(n+1) = \lambda^{-1}\mathbf{P}(n) - \frac{\lambda^{-1}\mathbf{P}(n)\mathbf{r}(n)\mathbf{r}^{\mathrm{T}}(n)\mathbf{P}(n)}{\lambda + \mathbf{r}^{\mathrm{T}}(n)\mathbf{P}(n)\mathbf{r}(n)}, \tag{18c}$$

$$\mathbf{P}(0) = \eta\mathbf{I} = \mathbf{\Pi}_0. \tag{18d}$$

Here after, we refer this algorithm as RLS throughout the rest of this paper.

4. Results and discussion

Simulation studies are carried out to assess the performances of the proposed H^∞ adaptive-filtering algorithms in minimizing the EOG artefacts from corrupted EEG signals. These algorithms are then compared to the well-known LMS and RLS adaptive-filtering algorithms. This section is organized in three subsections. In the first subsection, the data details are explained. Results from the simulation studies and the real recorded signals are discussed in detail in the following two subsections.

4.1. Data description

The EEG and EOG data for the present study are obtained from the Biomedical Engineering Department of the Technical University of Eindhoven.[2] All the data are recorded on subjects aged between 18 and 23. Two sets of data, each containing multiple channel EEG and 4 channel EOG segments of 134 s duration, are used for the present study. The signals are digitized at a sampling rate of 256 Hz. EEG signals are filtered using a 100 tap linear phase finite impulse response (FIR) band pass filter with cut-off frequencies at 0.25 and 35 Hz. The EOG signals are band pass filtered using another 100 tap FIR filter with cut-off frequencies of 0.25 and 11.5 Hz.

[2] http://www.sps.ele.tue.nl/

4.2. Simulated EEG data

To compare the performance of the proposed H^∞ algorithms with the LMS and RLS algorithms, simulation studies are conducted using signal-to-noise ratio (SNR) improvement as the performance index.

4.2.1. Selecting the values of the filter order (P) and the convergence parameter (μ)

This study is conducted to determine the optimum filter order (P) for all the four proposed noise minimization algorithms and a suitable value for the step-size parameter (μ) for the LMS algorithm. Here, the EEG signal recorded from the electrode position O_1 (Fig. 1) is used as the *desired signal* ($s(n)$). The original EOG signal recorded between the electrodes E6 and A1 (Fig. 2) is used as the *reference signal* ($r(n)$). This signal is added to the desired signal at different signal to noise ratios (SNRs) to construct the *primary signal* ($y(n)$). Output SNR ($\mathrm{SNR}_{\mathrm{out}}$) has been used as the performance index for the quantitative comparison of the performance of the algorithms. The output SNR is calculated using

$$\mathrm{SNR}_{\mathrm{out}} = 10\log\left(\frac{\sum_{n=1}^{N} s^2(n)}{\sum_{n=1}^{N}(y(n)-\hat{s}(n))^2}\right), \tag{19}$$

where N is the total number of samples in the signal and $\hat{s}(n)$ is the estimate of the desired signal $s(n)$ using the ANC schemes.

For the LMS algorithm, the $\mathrm{SNR}_{\mathrm{out}}$ for different values of μ is calculated for filters orders (P) increasing from 1 to 14. Fig. 4 shows the $\mathrm{SNR}_{\mathrm{out}}$ curves corresponding to the input SNR of -20 dB. It can be seen that the output SNR curves reaches different steady-state values (corresponding to different μ values) for P values of 4 and beyond. For example, when $\mu = 2\times10^{-4}$, the output SNR reaches its steady-state value of about -4.33 dB (an improvement of 15.67 dB) for a P value of 4. Whereas if the μ value is less than 2×10^{-4}, the algorithm reaches its steady-state $\mathrm{SNR}_{\mathrm{out}}$ value (lower value) for a much larger filter order (say $P=6$ for $\mu = 1\times10^{-4}$). Corresponding performance curves for input SNR of 0 dB are shown in Fig. 5.

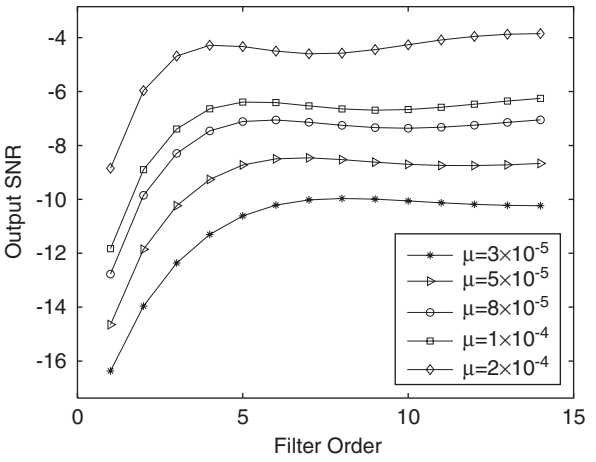

Fig. 4. Output SNR vs. the filter order (P) for LMS algorithm for different values of μ (input SNR $= -20$ dB).

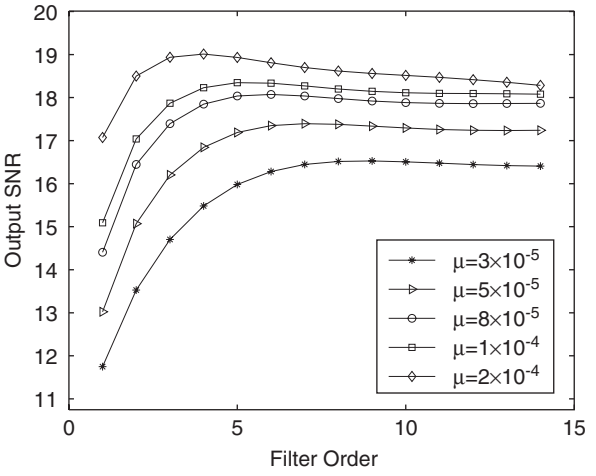

Fig. 5. Output SNR vs. the filter order (P) for LMS algorithm for different values of μ (input SNR $= 0$ dB).

Here also, it is observed that when $P = 4$ and $\mu = 2 \times 10^{-4}$, the output SNR is 18.9 dB (an improvement of 18.9 dB). From these observations, a suitable μ value of 2×10^{-4} and an optimal filter order of $P = 5$ has been selected for all the remaining simulation studies with the LMS algorithm.

4.2.2. Performance comparison

The $\mathrm{SNR}_{\mathrm{out}}$ values for the two H^{∞} algorithms and RLS algorithm are computed for varying P values from 1 to 14. The results are plotted along with that of the LMS algorithm (corresponding to the input SNR of -20 dB) in Fig. 6. The step-size parameter for the LMS algorithm is chosen as

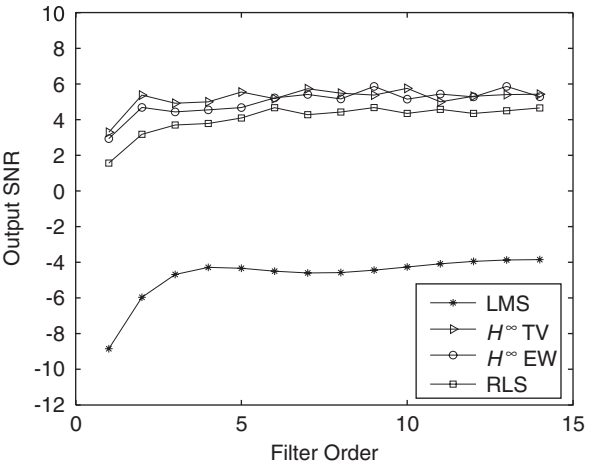

Fig. 6. Output SNR vs. the filter order (P) for all the proposed algorithms (input SNR $= -20$ dB).

Table 1
Output SNR for different adaptive filters

Filter	SNR $= -20$ dB	SNR $= -10$ dB	SNR $= 0$ dB
LMS	−4.33	7.61	18.90
H^{∞} EW	4.80	13.40	21.04
H^{∞} TV	5.01	13.45	21.40
RLS	4.07	13.24	21.02

2×10^{-4}. The parameters chosen for the H^{∞} TV algorithm are $\eta = 5 \times 10^{-3}$, $\rho = 10^{-5}$, and $\varepsilon_g = 1.5$. For the H^{∞} EW algorithm, the simulation parameters are $\lambda = 0.99$, $\varepsilon_{\lambda} = 1.5$ and $\eta = 5 \times 10^{-3}$. The parameter η for the RLS algorithm is chosen as 5×10^{-3}. From Fig. 6, it can be seen that the $\mathrm{SNR}_{\mathrm{out}}$ performance of the H^{∞} algorithms (SNR improvements of the order of 25 dB) is slightly better (≈ 1 dB) than the RLS algorithm. Their performance is much better (10 dB better) than the LMS algorithm (SNR improvement is of the order of 15 dB). It can also be noted that as the filter order is beyond 4 or 5, the output SNRs reaches to a steady-state value for all the algorithms considered. Accordingly, in all the following simulations, the filter order is chosen as $P = 5$ while keeping all the other parameters as mentioned above.

Table 1 shows the $\mathrm{SNR}_{\mathrm{out}}$ values of the proposed noise minimization schemes at different input SNRs. In the case of H^{∞} EW algorithm, the output SNR is 5.01 dB which is ≈ 25 dB improvement corresponding to the input SNR of -20 dB. The output SNR improvements corresponding to the

input SNRs of −10 and 0 dB are, respectively, ≈ 24 and ≈ 21 dB. With the use of H^{∞}TV algorithm, the SNR improvement is about 25 dB in the case of input SNR of −20 dB and is about 23 dB corresponding to −10 dB input SNR. For 0 dB input SNR, the output SNR improvement is ≈ 21 dB using this particular algorithm. It can be seen from the table that when using the H^{∞} algorithms, the improvement in the output SNRs is slightly better than that of the RLS algorithm at very low input SNRs (−20 dB) whereas at higher input SNR values, (say at −10 and 0 dB), their performance is almost similar. This is also clear from the two figures (Figs. 6 and 7). For the LMS algorithm, the corresponding SNR improvements are approximately 15, 17 and 19 dB, respectively, which are much less than the H^{∞} results. It is clear from the table that the H^{∞} methods outperforms the LMS algorithm consistently for all input SNR values ranging from −20 to 0 dB in minimizing the EOG artefacts from corrupted EEG signals. These results confirm the superiority of the H^{∞}-based algorithms in minimizing the EOG artefacts from contaminated EEG signals.

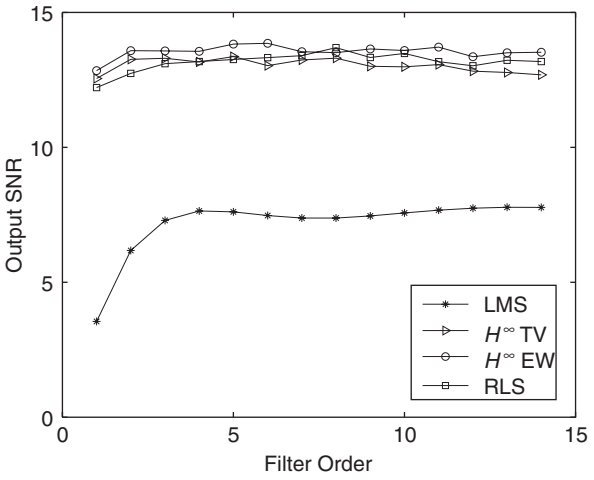

Fig. 7. Output SNR vs. the filter order (P) for all the proposed algorithms (input SNR = −10 dB).

Fig. 8 shows the desired ($s(n)$), reference ($r(n)$), and primary input ($y(n)$) signals along with the filtered signals using the proposed algorithms for an

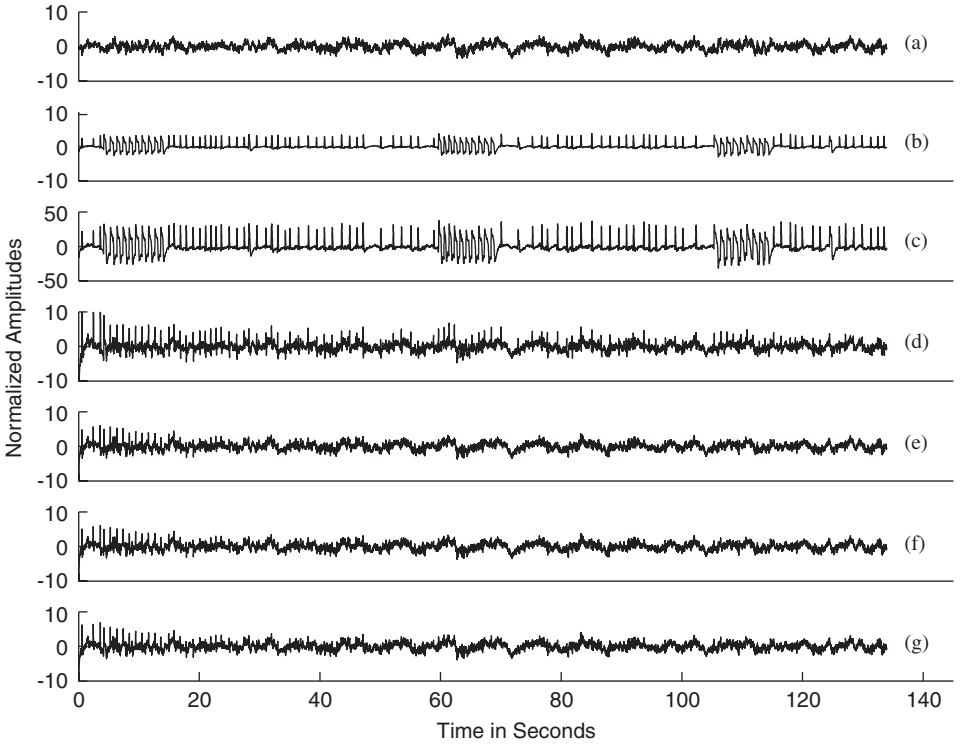

Fig. 8. Demonstration of EOG artefact minimization using the proposed algorithms: (a) desired signal; (b) corrupting EOG signal; (c) corrupted EEG signal (at −20 dB); (d) corrected EEG using LMS scheme; (e) corrected EEG using H^{∞} TV scheme; (f) corrected EEG using H^{∞} EW scheme; (g) corrected EEG using RLS scheme.

input SNR of −20 dB. Fig. 8(a) is the recorded and preprocessed EEG signal from O_1 electrode position on the scalp (desired signal). The preprocessed EOG signal (blinks) recorded between electrodes E6 and A1 is shown in Fig. 8(b). This signal is used as the reference input to the ANC schemes. The simulated primary input signal (at −20 dB) is shown in Fig. 8(c). The amplitude of the corrupting signal is so large that the EEG signal is not even visible to the naked eye. The estimated (desired) signals ($\hat{s}(n)$) using the four proposed artefact minimization schemes (LMS, H^∞ TV, H^∞ EW and RLS) are shown in Figs. 8(d)–(g), respectively. The corresponding output SNRs are, respectively −4.33, 5.01, 4.80 and 4.07 dB. From the visual inspection, it can be seen that except the LMS algorithm, all the other algorithms work better in estimating the desired signal from the contaminated EEG signal. In the case of the estimated output using the LMS algorithm, the residual EOG artefact is clearly visible. Corresponding results for the input SNR of 0 dB is shown in Fig. 9. Here, by visual inspection, the estimated EEG signals (Fig. 9(d)–(g)) look almost identical. But the output SNR values in Table 1 clearly show the superiority of the H^∞ algorithms to the LMS algorithm. The SNR_{out} value of the RLS algorithm is comparable to that of the H^∞ schemes. From these simulation studies, it can be clearly seen that the proposed H^∞ adaptive filtering algorithms perform slightly better than the RLS algorithm (especially at very low input SNRs) and much better (both in quantitative and qualitative terms) than the LMS algorithm in minimizing the EOG artefacts from corrupted EEG signals.

4.3. Real EEG data

For illustrating the performance of the proposed ANC schemes in minimizing the EOG artefacts from contaminated real EEG signals, EEG recorded from F3, F5, Fz, F4, and F6 electrode locations (Fig. 1) according to the expanded 10–20 international system are used in our study. Four different EOGs (2 horizontal EOGs and 2 vertical EOGs) are recorded by placing electrodes as shown in Fig. 2.

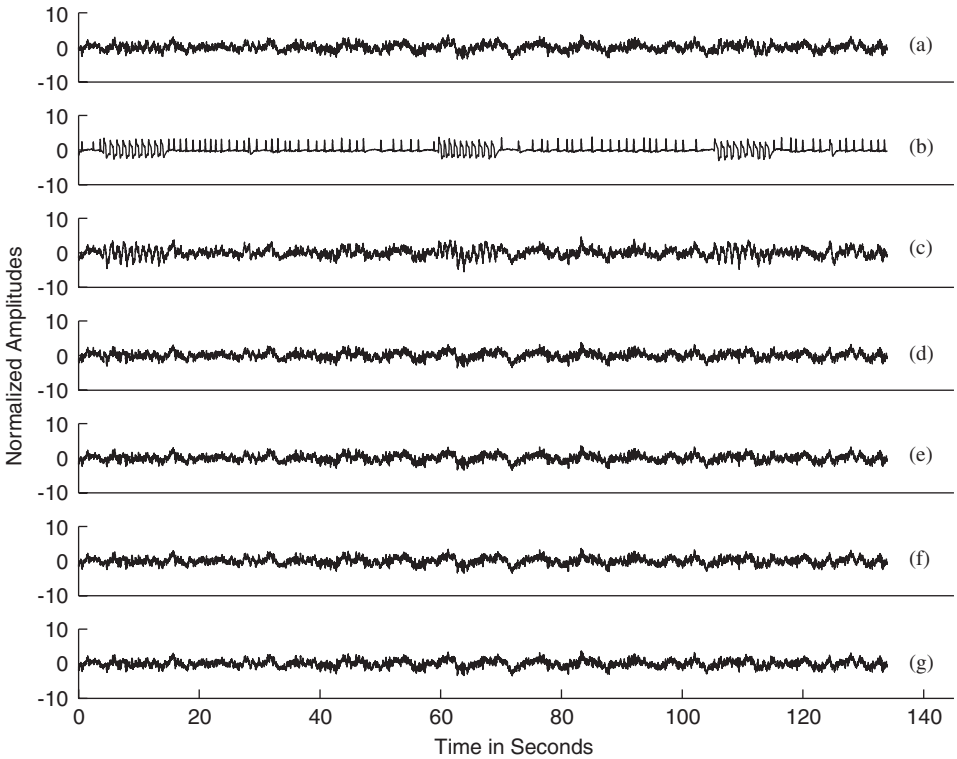

Fig. 9. Demonstration of EOG artefact minimization using the proposed algorithms: (a) desired signal; (b) corrupting EOG signal; (c) corrupted EEG signal (at 0 dB); (d) corrected EEG using LMS scheme; (e) corrected EEG using H^∞ TV scheme; (f) corrected EEG using H^∞ EW scheme; (g) corrected EEG using RLS scheme.

Horizontal left component is measured between electrodes E4 and E5 and the horizontal right component is measured between electrodes E3 and E4. Two vertical EOGs are measured from the left eye position. They are vertical "up" (between E6 and A1) and vertical "down" (between E7 and A1) components. In these studies, all the parameter values are kept as the same as in the case of studies with simulated data.

In the real experiment, since the desired signal is unknown, an appropriate measure defined in Eq. (20) is introduced to get a quantitative comparison of the performance of proposed algorithms in minimizing the EOG artefacts from corrupted EEG signals. It is defined as

$$R^2 = \frac{\sum_{n=1}^{N}(y(n) - \hat{s}(n))^2}{\sum_{n=1}^{N}\hat{s}^2(n)}. \quad (20)$$

Here, R^2 is the ratio of the power of the OA being removed from the primary signal to the power in the estimated EEG [8]. From this equation, it can be seen that the higher the value of R the better is the artifact minimization.

Since there are 4 EOGs that corrupts the EEG signals, simulations are conducted to determine the effect of number of reference inputs to the proposed ANC schemes. In each of these cases, the value of R is computed and is plotted against the varying filter order from 1 to 14. Fig. 10 shows the corresponding curves for LMS algorithm. It can be seen that the R values corresponding to the use of one and two reference EOGs (horizontal EOGs) are low indicating a poor EOG artefact minimization. This may be attributed to the fact that the two horizontal EOGs may be less correlated to the EEG at the Fz electrode location (ANC scheme works on the principle of correlation cancellation) and hence the numerator of Eq. (20) is small producing a low value of R. When three reference EOGs (two horizontal EOGs and one vertical EOG) are used, the R values are large (of the order of 2) for $P = 5$ and above. This is an indication of better artefact minimization compared to the use of two horizontal EOGs as reference inputs. When 4 reference EOGs are used, the performance index increases even more showing the effectiveness of more reference EOGs to the artefact minimization. The corresponding

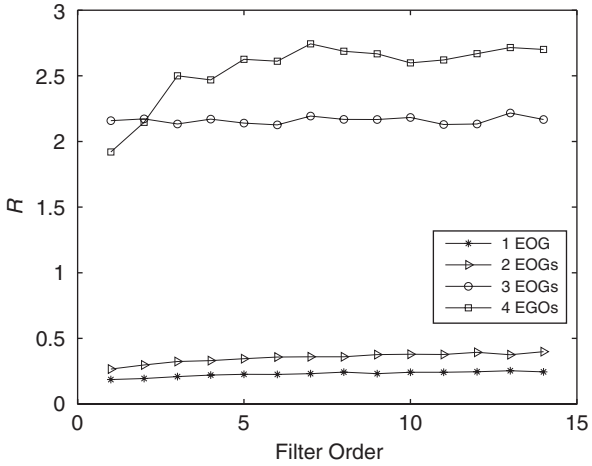

Fig. 11. R vs. the filter order (P) for H^∞ TV algorithm with multiple EOG reference inputs.

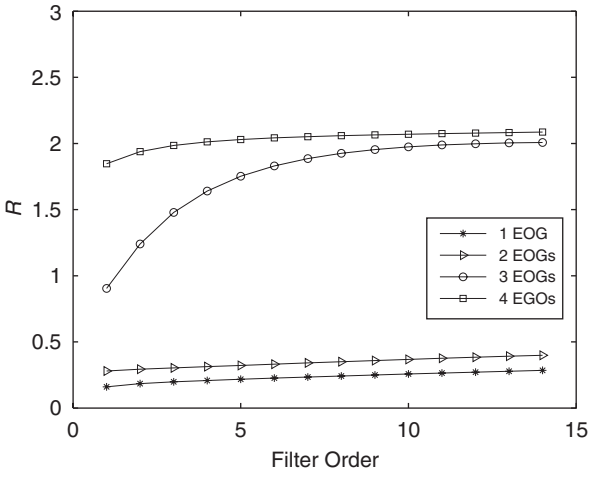

Fig. 10. R vs. the filter order (P) for LMS algorithm with multiple EOG reference inputs.

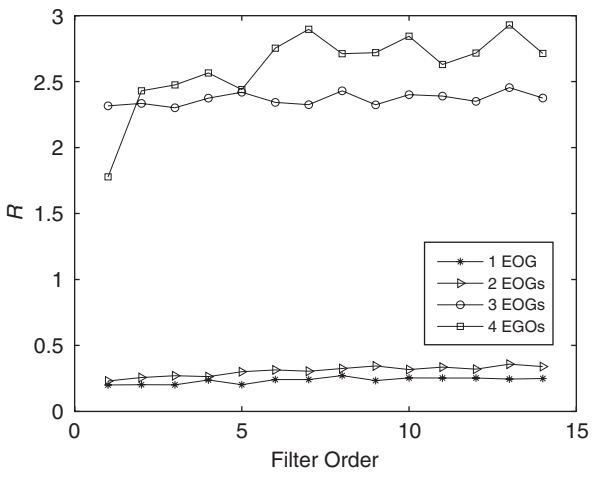

Fig. 12. R vs. the filter order (P) for H^∞ EW algorithm with multiple EOG reference inputs.

illustrations for the H^∞ and RLS filters are shown in Figs. 11–13. From these pictures, it is observed that the R values of the H^∞ filters are slightly higher than that of the RLS algorithm and are much better than the LMS filtering algorithm indicating a better EOG artifact minimization from corrupted EEG signals. Another important observation is that using more than one EOG reference inputs improve the performance of the proposed artifact minimization schemes.

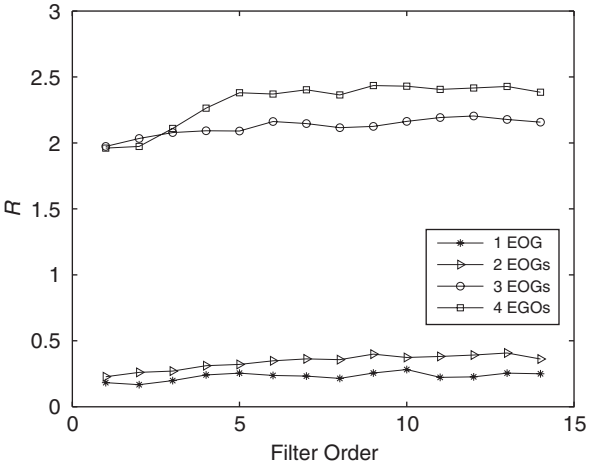

Fig. 13. R vs. the filter order (P) for RLS algorithm with multiple EOG reference inputs.

The representative results of the proposed artefact minimization algorithms are shown in Fig. 14. Fig. 14(a) shows the recorded EOG contaminated EEG signals from Fz position on the scalp. The EOG blink artefacts are clearly seen in this picture. The EOG minimized EEG signal using the LMS adaptive filtering algorithm is shown in Fig. 14(b), Here, the residual EOG artefacts are clearly visible showing the inferiority of the LMS algorithm, especially in the beginning. From Fig. 14(c) and (d), which corresponds to the outputs of the proposed H^∞ schemes, it can be seen that these algorithms minimize the EOG artefacts effectively from the recorded EEG signals compared to the LMS algorithm. The output (estimated EEG) of the RLS adaptive filtering scheme is shown in Fig. 14(e), which is comparable to the outputs of the proposed H^∞ adaptive filters.

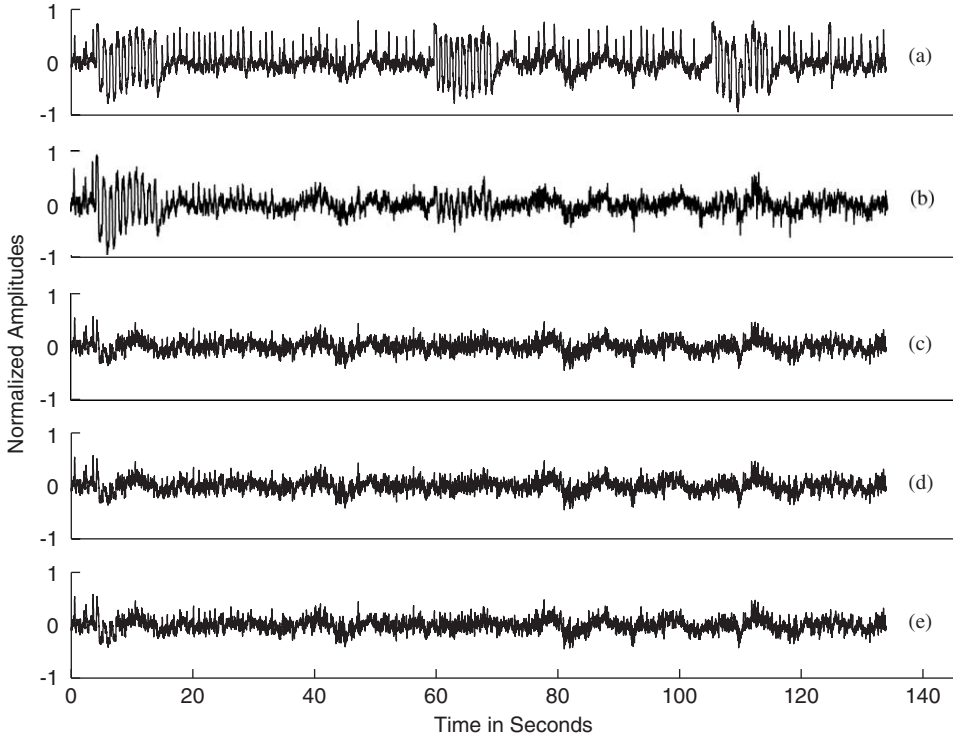

Fig. 14. Demonstration of EOG artefact minimization using the proposed algorithms (real EEG data): (a) EOG contaminated EEG signal recorded from Fz electrode location; (b) corrected EEG using LMS scheme; (c) corrected EEG using H^∞ TV scheme; (d) corrected EEG using H^∞ EW scheme; (e) corrected EEG using RLS scheme.

5. Conclusions

Signals from eye movements and blinks (EOGs) are of low frequency and several orders of magnitude larger than the brain-generated electrical potentials. These EOGs are one of the main sources of artefacts that affect the low frequency region of electroencephalographic (EEG) data. Effective minimization of artefacts from the collected data is an essential step in preparing the data for further analysis.

In this paper, we demonstrate the power and efficacy of H^∞ adaptive algorithms (H^∞ TV and H^∞ EW) for the minimization of EOG artefacts from corrupted EEG signals. The proposed algorithms are compared with the well-known LMS and RLS algorithms to measure its relative performance. We have evaluated the performance of these algorithms using simulated as well as real recorded signals. In the case of studies with simulated signals, output SNR improvement has been used as the measure for quantitative comparison along with time plots for qualitative comparison. In the case of real recorded signals, time plots as well as a performance index (R) defined in Eq. (20) are used as the measures for comparison. From all these extensive studies, we found that the H^∞-based algorithms performed slightly better (especially at very low input SNRs) than the RLS algorithm and convincingly outperform the LMS algorithm in minimizing the EOG artefacts from the corrupted EEG signals.

Acknowledgements

The authors of this paper express their gratitude to Prof. J.W.M. Bergmans and Dr. J.J.M. Kierkels of the Technical University of Eindhoven (TU/e) for providing the required data for successful completion of this work.

References

[1] B. Hassibi, A.H. Sayed, T. Kailath, Linear estimation in Krein spaces—Part I: theory & Part II: applications, IEEE Trans. Automat. Control 41 (1996) 18–49.
[2] J. Enderle, S. Blanchard, J. Bronzino, Introduction to Biomedical Engineering, Academic Press, New York, 2000.
[3] G. Gratton, M.G.H. Coles, E. Donchin, A new method for off-line removal of ocular artefacts, Electroenceph. Clin. Neurophysiol. 55 (1983) 468–484.
[4] R. Verleger, T. Gasser, J. Mocks, Correction of EOG artefacts in event related potentials of the EEG: aspects of reliability and validity, Psychophysiology 19 (1982) 472–480.
[5] R.P. Borda, J.J. Hablitz, Use of simple visual display to reduce eye movement artefacts in CNV recordings, Electroenceph. Clin. Neurophysiol. 34 (1973) 433–436.
[6] C. Fortgens, M.P. De Bruin, Removal of eye movement and ECG artefacts from non-cephalic reference EEG artefact, Electroenceph. Clin. Neurophysiol. 56 (1983) 90–96.
[7] B.W. Jervis, E.C. Ifeachor, E.M. Allen, The removal of ocular artefacts from electroencephalogram: a review, Med. Biol. Eng. Comput. 26 (1988) 2–12.
[8] B.W. Jervis, M. Thomlinson, C. Mair, J.M.L. Lopez, M.I.B. Garcia, Residual ocular artefact subsequent to ocular artefact removal from the electroencephalogram, IEE Proc. Sci. Meas. Technol. 146 (1999) 293–298.
[9] R.J. Croft, R.J. Barry, EOG correction: a new perspective, Electroenceph. Clin. Neurophysiol. 107 (1998) 387–394.
[10] R.J. Croft, R.J. Barry, EOG correction: a new aligned-artefact average solution, Electroenceph. Clin. Neurophysiol. 107 (1998) 395–401.
[11] G. Gratton, Dealing with artifacts: the EOG contamination of the event-related brain potentials, Behavior Res. Methods Instrum. Comput. 30 (1998) 44–53.
[12] R.J. Croft, R.J. Barray, Issues relating to the subtraction phase in EOG artifact correction of the EEG, Internat. J. Psychophysiol. 44 (2002) S.187–S.195.
[13] R.J. Croft, R.J. Barray, Removal of ocular artifacts from the EEG: a review, Neurophysiol. Clin. 30 (2000) S.5–S.19.
[14] G.L. Wallstrom, R.E. Kass, A. Miller, J.F. Cohn, N.A. Fox, Automatic correction of the ocular artifacts in the EEG: a comparison of regression-based and component-based methods, Internat. J. Psychophysiol. 53 (2004) 105–119.
[15] S.M. Haas, M.G. Frei, I. Osorio, B. Pasik-Duncan, J. Radel, EEG ocular artifact removal through ARMAX model system identification using extended least squares, Comm. Inform. Systems 3 (1) (2003) 19–40.
[16] T. Zikov, S. Bibian, G.A. Dumont, M. Huzmezan, A wavelet based de-noising technique for ocular artifact correction of the electroencephalogram, Private communication, 2004.
[17] M.J. Berryman, S. Messer, A. Allison, D. Abbott, Techniques for noise removal from EEG, EOG and airflow signals in sleep patients, Private communication, 2004.
[18] A. Schlogl, G. Pfurtscheller, EOG and ECG minimization based on regression analysis, Technical Report SIESTA-Task310 Artifact Detection and Denoising, 2004.
[19] R. Bogacz, U. Markowska-Kaczmar, A. Kozik, Blinking artifact recognition in EEG signals using artificial neural networks, Private communication, 2004.
[20] R. Vigario, Extraction of ocular artefacts from EEG using independent component analysis, Electroenceph. Clin. Neurophysiol. 103 (1997) 395–404.
[21] C.A. Joyce, I.F. Gorodnitsky, M. Kutas, Automatic removal of eye movement and blink artifacts from EEG data using blind component separation, Psychophysiology 41 (2003) 1–13.
[22] T.P. Jung, S. Makeig, C. Humphries, T.W. Lee, M.J. McKeown, V. Iragui, T.J. Sejnowski, Removing electroencephalographic artifacts by blind source separation, Psychophysiology 37 (2000) 163–178.
[23] E. Nezhadarya, M.B. Shamsollahi, EOG artifact removal from EEG using ICA and ARMAX modelling, Private communication, 2004.

[24] B. Widrow, et al., Adaptive noise cancellation: principles and applications, Proc. IEEE 63 (1975) 639–652.
[25] P.K. Sadasivan, D. Narayana Dutt, Development of Newton type adaptive algorithm for the minimization of EOG artefacts from noisy EEG signals, Signal Processing 62 (1997) 173–186.
[26] T. Ratnarajah, P.K. Sadasivan, An H^{∞} approach to adaptive minimization of EOG artefacts from EEG signals, in: Proceedings of the Eigth IEEE Workshop on DSP, Utah, USA, 1998.
[27] G.L. Wallstrom, R.E. Kass, A. Miller, J.F. Cohn, N.A. Fox, Correction of ocular artifacts in the EEG using Bayesian adaptive regression splines, Bayesian Statist., vol. 6, Springer, New York, 2002.
[28] P. He, G. Wilson, C. Russell, Removal of ocular artifacts from electroencephalogram by adaptive filtering, Med. Biol. Eng. Comput. 42 (2004) 407–412.
[29] A. Erfanian, B. Mahmoudi, Real-time ocular artifact suppression using recurrent neural network for electroencephalogram based brain-computer interface, Med. Biol. Eng. Comput. 43 (2) (2005) 296–305.
[30] S. Selvan, R. Srinivasan, Removal of ocular artefacts from EEG using an efficient neural network based adaptive filtering technique, IEEE Signal Processing Lett. 6 (12) (1999) 330–332.
[31] B. Hassibi, T. Kailath, H^{∞} adaptive filtering, in: Proceedings of ICASSP'95 1995, pp. 949–952.
[32] B. Hassibi, A.H. Sayed, T. Kailath, LMS is H^{∞} optimal, in: Proceedings of 32nd IEEE Conference on Decision and Control, 1993, pp. 74–80.
[33] B. Hassibi, A.H. Sayed, T. Kailath, H^{∞} optimality of the LMS algorithm, IEEE Trans. Signal Process 44 (2) (1996) 267–280.
[34] S. Haykin, Adaptive Filter Theory, third ed., Prentice-Hall, NJ, 1996.