Real time ocular and facial muscle artifacts removal from EEG signals using LMS adaptive algorithm

Saeid Mehrkanoon, mehrkanoon@perdana.um.edu.my
Mahmoud Moghavvemi, mahmoud@um.edu.my
Hossein Fariborzi fariborzi@perdana.um.edu.my
Dept. of Electrical Engineering University Malaya
Kuala Lumpur, Malaysia

***Abstract*- The EEG signal is most useful for clinical diagnosis and in biomedical research. ElectroOculoGram (EOG), ElectroMyoGram (EMG) artifact are produced by eye movement and facial muscle movement respectively. An adaptive filtering method is proposed to remove these artifacts signals from EEG signals. Proposed method uses horizontal EOG (HEOG), vertical EOG (VEOG), and EMG signals as three reference digital filter inputs. The real-time artifact removal is implemented by multi-channel Least Mean Square algorithm. The resulting EEG signals display an accurate and artifact free feature.**

Key words **real-time -adaptive filter, Least Mean Square, Finite Impulse Response, EEG, EMG, EOG, noise cancellation**

I. Introduction

During the eye movement (saccade,blink,etc) the electric field around the eye generates a signal known as Electro-occulogram (EOG). EOG signals have two dimensions, vertical and horizontal. Facial muscles movements generate an electrical large amplitude signals known as Electromyogram (EMG). EOG and EMG signals appear in the Electro-encephalogram (EEG) signal as noise or artifact. Pure EEG signal is generated by brain activities. Artifact signals cause serious problems in pure EEG signals identification and interpretation for clinical diagnosis applications. Activities in each part of the brain make specific signals. If the artifacts contaminate the EEG data, identifying the brain activities would be limited. Therefore the artifacts should be removed by a proper method until the pure EEG signals and data are extracted.

To overcome these problems many regression-based techniques have been proposed, including simple time-domain regression [1], multiple-leg time-domain regression [2] and regression in the frequency domain [3]. The regression methods need calibration trials for determining the transfer coefficients between EOG, EMG and each of the EEG channels; they are not suitable for real-time implementation. During the performance of regression based method the calibration factors maybe off from the adjusted value, and they can’t be calibration again automatically. The tap weights of adaptive filter are initialized once, and the taps follow the optimizing tap weight process for readjustment.

The filter will track the variations of the input usually; adaptive filters are implemented as off-line application. In the stationary environment the adaptive LMS, RLS [4] algorithms are employed. In non-stationary environment Kalman filter [5] is used with adaptive algorithm. All of adaptive algorithms are applied for functional system optimizing and variable coefficients adjustment. Further attention is to use them in real-time application, but there are risks to implement them, because an adaptive algorithm can be applied for real-time application which should be stable and convergence with auto tracking capability. The RLS adaptive algorithm was implemented in off-line artifact removal. The RLS algorithm uses a tapped delay line or transversal structure. It relies on the matrix inversion lemma, which results in lack of numerical robustness and excessive numerical complexity [6-9]. A simple way to minimize the error signal is using the least mean square. LMS algorithm simplifies the computation by estimating the gradient from instantaneous values of correlation matrix of the tap inputs and the cross-correlation vector between the desired response and the tap weights. [10-13]

II. Signal Separation Beased on Adaptive Learning Algorithm

A. *Least Mean Square Adaptive Algorithm (LMS)*

Recall the Wiener filter problem [5] $\{x_k\},\{d_k\}$ which jointed wide sense stationary to find W minimizing $\mathrm{E}[{e_k}^2]$.

$$e_k = d_k - y_k = d_k - \sum_{i=1}^{M-1}(w_i x_{k-i}) = d_k - X^{k^T} W^k \qquad (1)$$

Where e_k , $d_{k,} y_k$ denote an error signal, desired and filter output signals, respectively.

$$X^k = \begin{pmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-M+1} \end{pmatrix} \qquad (2) \qquad\qquad W^k = \begin{pmatrix} w_0^k \\ w_1^k \\ \vdots \\ w_{M-1}^k \end{pmatrix} \qquad (3)$$

The superscript denotes absolute time, and the subscript denotes time or a vector index.

The solution can be found by setting the gradient $= 0$.

$$
\begin{aligned}
\nabla^{k} &= \frac{\partial}{\partial W}(\mathrm{E}[e_{k}^{2}]) \\
&= \mathrm{E}[2e_{k}(-X^{k})] \\
&= \mathrm{E}[-2(d_{k} - X^{k^{T}}W_{k})X^{k}] \\
&= -(2\mathrm{E}[d_{k}X^{k}]) + (\mathrm{E}[X^{k^{T}}])W \\
&= -2\mathrm{P} + 2RW \\
&\Rightarrow (W_{opt} = R^{-1}\mathrm{P})
\end{aligned} \tag{4}
$$

Where E denotes the mathematic expectation, the W_{opt} can be found iteratively using a gradient descent technique.

$$W^{k+1} = W^{k} - \mu\nabla^{k} \tag{5}$$

In fact, in the practice we don't know R and P exactly, and in an adaptive context. They maybe slow varying with time. To find the approximation of Wiener filter, approximating the gradient is useful to do.

$$
\begin{aligned}
\nabla^{k} &= \frac{\partial}{\partial W}(\mathrm{E}[e_{k}^{2}]) \\
&= \frac{\partial}{\partial W}(e_{k}^{2}) = 2e_{k}\frac{\partial}{\partial W}(d_{k} - W^{k^{T}}X^{k}) \\
&= 2e_{k}(-X^{k}) = -(2e_{k}X^{k})
\end{aligned} \tag{6}
$$

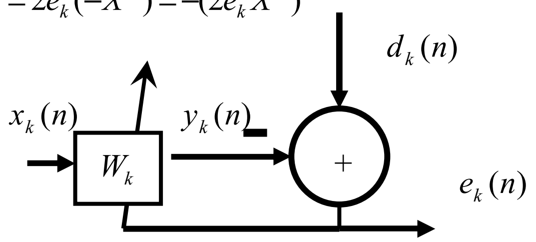

Fig.1 Adaptive filter block diagram

The FIR filter W_k is shown by fig.1.

The relationship between filter input and output is defined by:

$$
\begin{aligned}
y_{k}(n) &= w_{0}(n)x_{k}(n) + w_{1}(n)x_{k}(n-1) + \ldots \\
&+ w_{M-1}x_{k}(n-M+1) \\
&= \sum_{i=1}^{M-1} w_{k}(n)x_{k}(n-k) = w^{T}(n)x_{k}(n)
\end{aligned} \tag{7}
$$

The filter parameters are changed by:

$$w(n+1) = w(n) + \mu x_{k}(n)e(n)$$
(8)

Where μ denotes the adaptation step size, and $x(n)$ is filter input.

B. NLMS ALGORITHM

The NLMS algorithm is normalized LMS algorithm. To make NLMS algorithm the parameter vector is modified at time (n) from $W(n)$ to $W(n+1)$. [14-15]

Fulfilling the constraint obtained as:

$$d(n) = W^{T}(n+1)X(n) \tag{9}$$

The "least modification" of $W(n)$ based on least Euclidian norm of difference is defined as:

$$W^{T}(n+1) - W(n) = \delta w(n+1) \tag{10}$$

Minimizing (10):

$$\| \delta w(n+1) \|^{2} = \sum_{k=1}^{M-1}(w_{k}(n+1) - w_{k}(n))^{2} \tag{11}$$

Under the constraint

$$d(n) = w^{k}(n+1)x(n) \tag{12}$$

The solution can be obtained by Lagrange multipliers method:

$$
\begin{aligned}
&J(w(n+1,\lambda)) = \\
&\sum_{k=0}^{M-1}(w_{k}(n+1) - w_{k}(n))^{2} + \lambda\left(d(n) - \sum_{i=0}^{M-1} w_{i}(n+1)x(n-i)\right)
\end{aligned} \tag{13}
$$

To obtain the minimum of $J(w(n+1),\lambda)$ we check

$$J(w(n+1),\lambda) = \| \delta w(n+1) \|^{2} + \lambda(d(n) - w^{k}(n+1)x(n)) \tag{14}$$

The zeros of criterion partial derivatives:

$$\frac{\partial J(w(n+1),\lambda)}{\partial w_{j}(n+1)} = 0$$

$$\frac{\partial}{\partial w_{j}(n+1)}\left[\sum_{k=0}^{M-1}(w_{k}(n+1) - w_{k}(n))^{2} + \lambda\left(d(n) - \sum_{i=0}^{M-1} w_{i}(n+1)x(n-i)\right)\right] = 0 \tag{15}$$

$$2\left(w_{j}(n+1) - w_{j}(n)\right) - \lambda.x(n-j) = 0$$

$$\Rightarrow w_{j}(n+1) = w_{j}(n) + \frac{1}{2}\lambda.x(n-j) \tag{16}$$

Where λ will result from:

$$
\begin{aligned}
d(n) &= \sum_{i=0}^{M-1} w_{i}(n+1)x(n-i) \\
&= \sum_{i=0}^{M-1}\left(w_{i}(n) + \frac{1}{2}\lambda.x(n-i)\right)x(n-i) \\
&= \sum_{i=0}^{M-1} w_{i}(n)x(n-i) + \frac{1}{2}\lambda\sum_{i=0}^{M-1}\left(x(n-i)\right)^{2}
\end{aligned} \tag{17}
$$

$$
\begin{aligned}
\Rightarrow \lambda &= \frac{2\left(d(n) - \sum_{i=0}^{M-1} w_{i}(n)x(n-i)\right)}{\sum_{i=0}^{M-1}\left(x(n-i)\right)^{2}} \\
&= \frac{2e(n)}{\sum_{i=0}^{M-1}\left(x(n-i)\right)^{2}}
\end{aligned} \tag{18}
$$

Thus the minimum point of the criterion $J\left(w(n+1),\lambda\right)$ will be obtained by using the adaptation equation:

$$w_j(n+1) = w_j(n) + \frac{2e(n)}{\sum_{i=0}^{M-1}\left(x(n-i)\right)^2}.x(n-j) \qquad (19)$$

To make a flexible adaptation for adaptive processing, one constant $\tilde{\mu}$ is defined to control the step size. It will be introduced by:

$$w_j(n+1) = w_j(n) + \tilde{\mu}\frac{1}{\sum_{i=1}^{M-1}\left(x(n-i)\right)^2} e(n)x(n-j) \qquad (20)$$

$$= w_j(n) + \frac{\tilde{\mu}}{\| x(n) \|^2} e(n)x(n-j)$$

The possible numerical difficulties will happen if $\| x(n) \|$ is very close to zero, so to overcome this problem a constant $'a > 0'$ is used:

$$w_j(n+1) = w_j(n) + \frac{\tilde{\mu}}{a + \| x(n) \|^2} e(n)x(n-j) \qquad (21)$$

III. Removing EOG, EMG Artifacts by Adaptive Learning Algorithm

In this paper description of noise cancellation method based on real-time adaptive filtering to remove EOG and EMG artifacts from EEG signals is proposed. This method is particularly suitable, because it does not require calibration trials, and it is stable and has capability of tracking the multi signal variations. For this reason one is able to cancel EOG, EMG signals as real-time recording the EEG signals.

Fig.2 shows the block diagram of the artifacts canceller used in this application. The primary artifacts in the corrupted EEG signal known as HEOG, VEOG, and EMG. The corrupted EEG signal $S(n)$ is obtained by (for example point of F8 from skull).

The signal $S(n)$ is modeled by pure EEG signal and artifact components $B(n), C(n)$. Signal $S(n)$ is uncorrelated signal with $B(n)$, and $C(n)$.

B1(n), B2(n), and C1(n) are three reference digital filter input signals consist of HEOG, VEOG, and EMG signals respectively. $W_h(k), W_v(k)$, and $W_m(k)$ represent three FIR filters. Each of FIR filter has enough tap weights to product desired signals. The tap weights based on adaptive learning algorithm help to remove any unwanted signal that comes to the input filter.

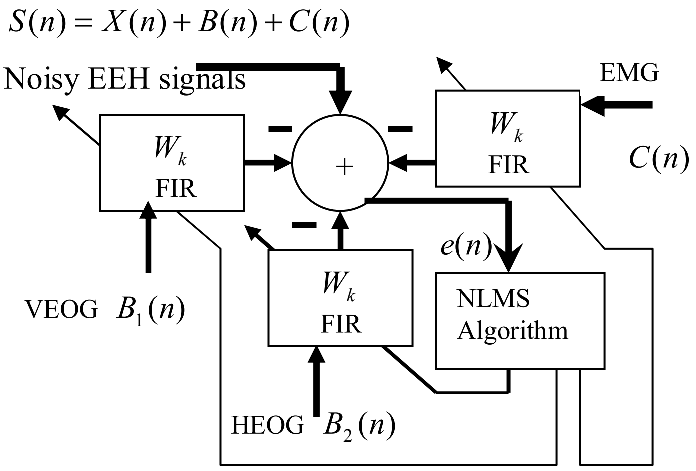

Fig2. Block diagram of adaptive EOG, EMG artifact removal

The output signal of sums of six signals $\hat{B}_1(n), \hat{B}_2(n), \hat{C}_1(n), X(n), B(n), C(n)$ is $e(n)$ which known error signal. The results of minimization of the error signal display clean EEG signal.

$$e(n) = X(n) + B(n) + C(n) - \hat{B}_1(n) - \hat{B}_2(n) - \hat{C}_1(n) \qquad (22)$$

The signals $\hat{B}_1(n), \hat{B}_2(n), \hat{C}_1(n)$ are output signals for FIR filters $W_h(n), W_v(n), W_m(n)$ respectively. Our assumption is that the signal $X(n)$ is a zero mean stationary. So, the signals $X(n)$ and $B(n), C(n)$ are uncorrelated with each other as:

$$\begin{aligned} \mathrm{E}[X(n)B(n-k)] &= 0 \\ \mathrm{E}[X(n)C(n-k)] &= 0 \end{aligned} \qquad \text{For all } k \qquad (23)$$

But the signals $B(n), C(n)$ and $B_1(n), B_2(n), C_1(n)$ are correlated with each other as:

$$\mathrm{E}[B(n), B_1(n-k)] = p_1(k) \qquad \text{For all } k \qquad (24)$$

$$\mathrm{E}[B(n), B_2(n-k)] = p_2(k) \qquad \text{For all } k \qquad (25)$$

$$\mathrm{E}[C(n), C_1(n-k)] = p_3(k) \qquad \text{For all } k \qquad (26)$$

$$n = 1,2,\ldots\ldots, M$$

Where E denotes mathematic expectation.

The $p_1(k), p_2(k), p_3(k)$ are unknown cross-correlation for dilation k. The expectation of signal $e(n)$ is as close to signal $X(n)$. To make the 'best', minimization of mean square of error signal e(n) defined by:

$$\xi(n) = \min\ \mathrm{E}[e^2(n)] \tag{27}$$
$$= \min \mathrm{E}[(X(n)+B(n)+C(n)-(\hat{B}_1(n)+\hat{B}_2(n)+\hat{C}_1(n)))^2]$$

Where $\xi(n)$ denotes minimized error value by sample n^{th}.

As we wish to best estimate $\hat{B}_1(n), \hat{B}_2(n)$ and $\hat{C}_1(n)$ the signal power $\mathrm{E}[e^2(n)]$ will be unaffected as the weights are being adjusted to minimize $\mathrm{E}[e^2(n)]$. From the optimization equation (21), the FIR filter tap-weights $W_j(n)$ are adjusted by minimization of right side of equation (22), and the right side is also minimized by minimizing the mean square of signal $e(n)$. Error minimization is our aim during the iteration for optimizing the filter coefficients. So the real-time noise canceller based on adaptive LMS algorithm generates the signal $e(n)$ which is close to EEG signal $X(n)$. Development of algorithm for the case of three reference filters inputs as real-time processing at time t_n is presented as the following tap weights optimizing equations. According to the samples:

$$S(n), B_1(n), B_2(n), C_1(n),\ e(n)\quad n = 1,2,....,N \tag{28}$$

The NLMS algorithm is computing the M-by-1 tap weight vectors. The tap weights of implemented FIR filter are updated by following formulas:

$$\hat{W}_h(n+1) = W_h(n) + \frac{\mu'}{\| B1(n) \|^2} B1(n)e^*(n) \tag{29}$$

$$\hat{W}_v(n+1) = W_v(n) + \frac{\mu'}{\| B2(n) \|^2} B2(n)e^*(n) \tag{30}$$

$$\hat{W}_m(n+1) = W_m(n) + \frac{\mu'}{\| C1(n) \|^2} C1(n)e^*(n) \tag{31}$$

Where μ' is adaptation step for NLMS algorithm and it is also dimensionless. The produced signals $B_1(n), B_2(n), C_1(n)$ which come from scalp are normalized by the squared Euclidian norm of the tap-input vectors.

IV. SIMULATION STUDY

This simulation study is based on MATLAB programming environment. Extensive simulation studies were performed to test the capabilities and stabilities of multi-channel adaptive LMS algorithm application in real-time artifact removal. In this simulation study the sinusoidal signal instead of EEG signal is used. The signal is recorded by sampling rate of $16KHz$. The total samples of sinusoidal signal are 48000. The 1500 samples out of overall 48000 samples have been selected to appear the resolution of collected signal in the Fig3. Fig.3A represents the sinusoidal signal, 3B), 3C), 3D show each digital filter output signals and 3E) display corrupted sinusoidal signal. Extracted sinusoidal signal is presented by fig3F) based on implemented multi-channel NLMS algorithm to noise or "artifact" removal. The three artifact signals are defined by definition of three types of random noise signals. The artifact signals are generated simultaneously with samples of sinusoidal signal. According to the algorithm the pure real original signal will be extracted by subtraction of adaptive filter output and sinusoidal signal. The FIR filter which is used in this simulation study has eight adjustable tap-weights. The iteration of FIR filter is primary error and reference samples signals by using 200 samples. So, the NLMS algorithm is trained by primary 200 samples of error and sinusoidal signals with learning step size 0.015. For this reason, the algorithm needs 200 samples periodically until 48000th for training. The tap-weights are readjusted by error samples based on equations (29-31). No artifact signal is seen as long duration samples as 1500 samples in the 'extracted clean signal' by the following Fig3F. The result is that adaptive algorithm could remove the unwanted noises and it also extracts the original pure signal.

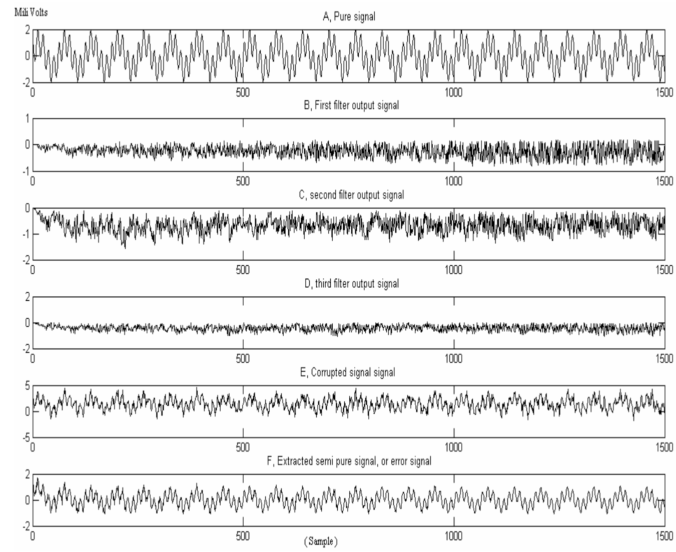

Fig.3 A) pure signal, B) first filter output C) second filter output D) error signal E) corrupted signal F) clean signal

V. Experiments setup

All the EEG, VEOG, HEOG, EMG signals are recorded simultaneously by EEG machine. The signals were transferred to processor. The processor extracts the pure EEG signals based on NLMS algorithm simultaneously as signal recording from the scalp and skull.

Fig4 presents the block diagram of simultaneously signal recording and artifact removal system based on EEG machine and high performance computer.

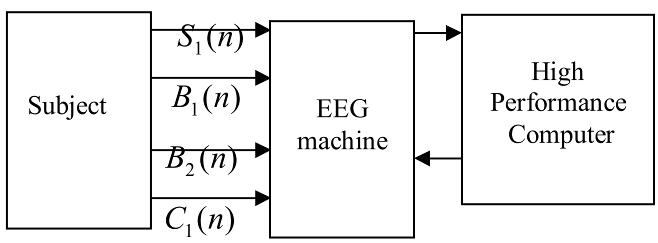

Fig.4 Block Diagram of EEG Signal Decomposition by High performance Computer

Fig4. The signals $S(t), B_1(t), B_2(t), C_1(t)$ indicate corrupted EEG signal, artifact signals HEOG, VEOG, and EMG respectively. The EEG signals are recorded by sampling rate of 256 Hz per each EEG channel by Ag/AgCI scalp electrodes from positions F3, Fz, F4, Pz, Cz, C4, and Fp2. those positions are defined by 10-20 electrodes system based on the electrode placement system by Fig5. A Pair of electrodes is placed above (positive terminal) and below (negative terminal) the right eye to record VEOG. And another pair of electrodes were placed at the right and left temple to record HEOG. Another pair of electrodes is placed at the protuberances of right and left cheek to record EMG. All recorded EEG channels are referenced to the left earlobe.

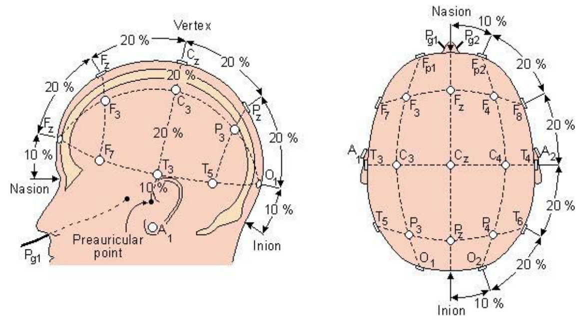

Fig5. Standard Locations of the EEG electrodes - (10-20 Standard)

Fig 6a presents VEOG signals, Fig 6b as HEOG signals. Each of above signals consists of 4800 samples with magnitude range between -250 and 500 micro-volts and -250 and 50 micro-volts respectively. Fig.6c and Fig.6d present noisy EEG signals and corrected EEG signals, respectively. As Fig6d there is no EOG and EMG artifacts effect in clean EEG signals. The VEOG signal illustrates several positive going pulses corresponding to eye blinks, and the HEOG signal also shows several pulses between samples 1800 and 4800.

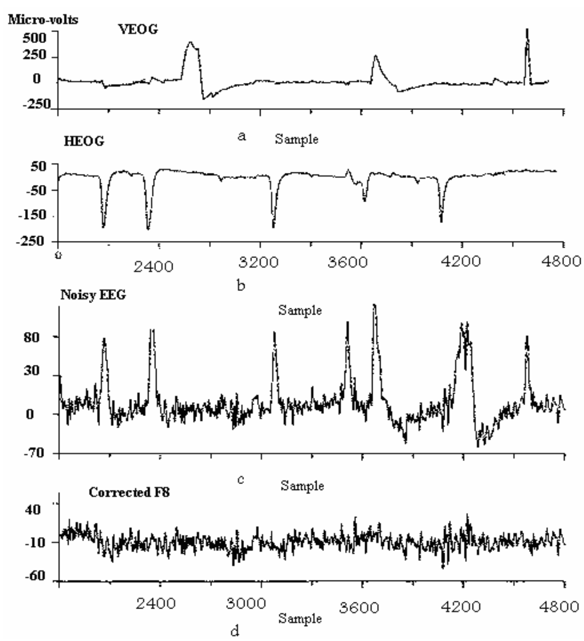

Fig.6 a) VEOG Signal b) HEOG Signal c) Noisy EEG Signal d) Corrected EEG signal

There is no EOG and EMG artifacts effect in clean EEG signals. The vertical EOG signal illustrates several positive going pulses corresponding to eye blinks. The HEOG signals also shows several pulses between samples 1800 and 4800. The artifacts can be identified easily by original EOG signal processing.

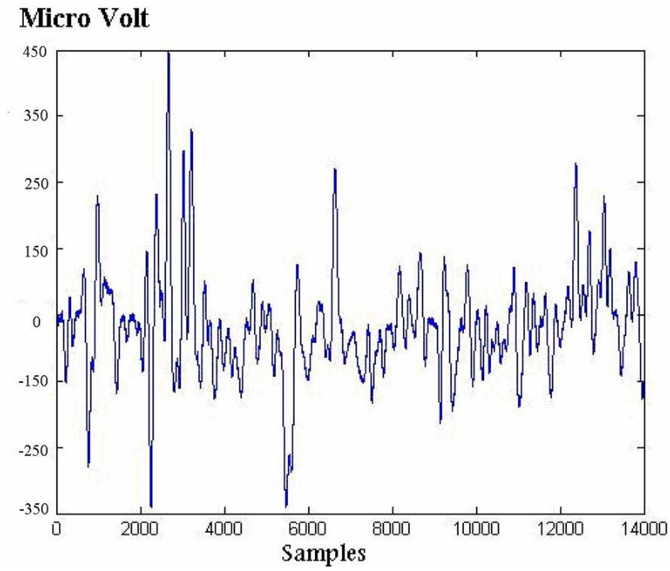

Fig.7 represents the EMG artifact including head movement and chewing artifact and blinking artifact respectively.

Fig.7 illustrates EMG artifact signals including head movement artifact, facial muscle movement. It is a part of recorded EMG signals in the terms of 8, 10, and 20 seconds. The number of samples is 2048, 2560s in terms of recording for head movement, cheek movement artifact signals. Maximum and minimum magnitudes of head movement artifact signal are between 300 and -250 micro-volts, respectively. The magnitudes of facial muscle movement are between ± 150 micro-volts in amplitude.

Fig.8 illustrates portion of measured EEG signals and artifact free EEG signals which have been obtained by implementing the EEG machine and high performance computer. All the signals magnitudes are Micro-volts in the Fig.8a, 8b

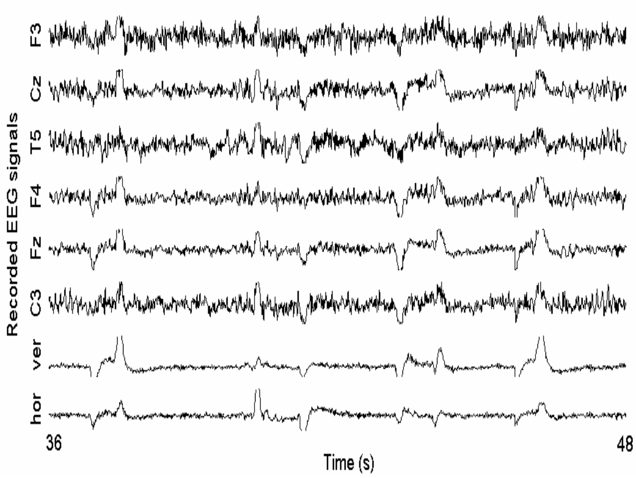

Fig8 a) Portion of measured EEG signals

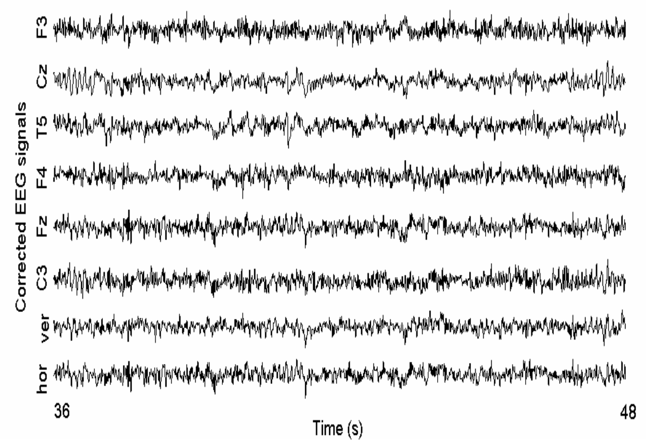

Fig8 b) Artifact-free EEG signals

VI. Conclusion

In this paper, a method for real-time multi-channel artifact removal based on adaptive learning algorithm is proposed and investigated. Three different lengths of digital FIR filter which provided the proper estimation of three artifacts VEOG, HEOG, EMG are used. The results of real-time artifact removal algorithm satisfied the expected protocol and design. The digital filters which tracked different types of noises and artifacts in time and frequency variant were presented. Presented digital filters simultaneously computed the anti-noise. The true and pure EEG signals with separated VEOG, HEOG, and EMG artifacts were obtained. We also have applied to classify the pure extracted EEG signal for brain cognitive state.

Reference

[1] R.J. Croft, R.J. Barry, "Removal of ocular artifact from the EEG: a review," *Neurophysiologie Clinique*, vol. 30, pp. 5-19, 2000.

[2] G. Gratton, M.G. Coles, and E. Donchin,.A new method for off-line removal of ocular artifact., Electroencephalography and clinical neurophysiology, vol.55, pp. 484-486, 1993

[3] Dimitris G. Manolakis, statistical and adaptive signal processing, MC Graw Hill, 2000 *Clinique*, vol. 30, pp. 5-19, 2000.

[4] P. He, G. Wilson, and C. Russell, .Removal of ocular artifacts from electro-encephalogram by adaptive filtering.. Medical & Biological Engineering & Computing, vol. 42, pp. 407-412, 2004.

[5] Sorenson, H. W. 1970. "Least-Squares estimation: from Gauss to Kalman," *IEEE Spectrum*, vol. 7, pp. 63-68, July 1970.

[6] P. He, G. Wilson, and C. Russell, "Removal of ocular artifacts from electro-encephalogram by adaptive filtering." Medical & Biological Engineering & Computing, vol. 42, pp. 407-412, 2004.

[7] T. Gasser, L. Sroka, and J. Möcks, "The transfer of EOG activity into the EEG from eyes open and closed," *Electroencephalography and Clinical Neurophysiology*, vol. 61, pp. 181-193, 1985.

[8] JUNG, T. P., ET all: 'Removal of eye activity artifacts from visual event-related potential in normal and clinical subjects', Clin. Neurophysiol., 111, pp. 1745–1758

[9] KENEMAN, J. L. (1991): 'removal of the ocular artifact from the EEG: A comparision of time and frequency domain methods with simulated and real data', Psychophysiology , 28, pp.115-121

[10] N. Saito and T. Sone, Influence of Modeling Error on Noise Reduction Performance of Active Noise Control Systems using Filtered-X LMS Algorithm, J. Acoust. Soc. Jpn. (E), Vol. 17, No. 4, 1996, pp. 195- 202.

[11] Sorenson, H. W. 1970. "Least-Squares estimation: from Gauss to Kalman,"*IEEE Spectrum*, vol. 7, pp. 63-68, July 1970.

[12] Douglas S.C. and Meng T.H.Y "Normalized Data Nonlinearities for LMS Adaptation", IEEE Transactions on Signal Processing, 42(6):1352-1365, 1994.

[13] Arenas-García, J., Figueiras-Vidal, A., and Sayed, A., "Mean-Square Performance of a Convex Combination of Two Adaptive Filters," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1078-1090, March 2006.

[14] N. J. Bershad, "Analysis of the Normalized LMS Algorithm with Gaussian Inputs," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP- 34, no. 4, pp. 793-806, April 1986.

[15] J. Homer, ''Detection guided NLMS estimation of sparsely parametrized channels'', *IEEE Trans. on Circuits and Systems II*, Vol. 47, pp. 1437-1442, Dec. 2000