Proceedings of the 24th IASTED International Multi-Conference
BIOMEDICAL ENGINEERING
February 15-17, 2006, Innsbruck, Austria

EXTRACTION AND SEPARATION OF HIGH-AMPLITUDE ARTIFACTS IN ELECTROENCEPHALOGRAMS FROM EPILEPTIC PATIENTS

A. R. Teixeira, A.M. Tomé
DETUA/IEETA
University of Aveiro
Aveiro 3810-193 Portugal
email: ana@ieeta.pt

E.W. Lang, P. Gruber
Institute of Biophysics
University of Regensburg
D-93040 Regensburg, Germany
email: elmar.lang@biologie.uni-regensburg.de

A.Martins da Silva
HGSA and ICBAS/IBMC
University of Porto
4099-001 Porto, Portugal
email:ams@icbas.up.pt

ABSTRACT
Electroencephalogram (EEG) recordings are often distorted by high-amplitude artifacts which hamper its correct visual inspection. In this work we present a method which can be applied separately to each channel to extract high-amplitude components. The method is called local singular spectrum analysis (SSA) and is a principal component analysis in clusters formed after embedding the signals in their time-delayed coordinates. The extracted signal can be subtracted from the original channel resulting in a corrected EEG version. The algorithm is applied to real EEG segments containing paroxysmal epileptiform activity contaminated by artifactual activity. The extracted artifact as well as the corrected EEG will be presented.

KEY WORDS
Singular spectrum analysis, artifacts removal, Electroencephalogram, Electrooculogram.

1 Introduction

Electroencephalograhic recordings can be distorted by electric signals from eye movements, eye blinking, muscle activity, head movements, heart beats and line noise. These artifacts constitute a problem for the interpretation of the EEG signals because in many cases they constitute the most prominent signal. In particular, the ocular activity creates artifacts in Electroencephalogram(EEG) signals, especially those recorded from frontal channels. When eyes move (blinking or other movements) the electric field around the eyes changes, producing an electrical signal known as Electrooculogram (EOG). As the signal propagates over the scalp, it also appears in the recorded Electroencephalogram (EEG). The large amplitude artifactual (EOG) signal is more visible in the frontal channel than in the derivation of the occipital region. In some studies, like single-trial event-related potentials (ERP), data from frontal electrodes located near the eyes are often discarded in the subsequent analysis and it is also common to reject all segments of signals (epochs) which are contaminated. It is common practice to exclude, manually or semi-automatically, all epochs which contain signals larger than a given arbitrary threshold. This is a tedious work and it results in a substantial data loss and a probably incomplete clinical analysis. In long-term EEG monitoring, like epileptic recording sessions, high-amplitude EOG artifacts often mask the initial onset of focal seizures compromising the visual interpretation. With the availability of digital EEG recordings it has become a desirable procedure trying to correct the data with other techniques than only simple linear digital filtering. The primary goal of these correction methods is to remove artifacts without distorting the underlying brain signals of interest. A variety of automatic procedures have been proposed to correct or remove ocular artifacts from EEG recordings. Some techniques are based on regression analysis, principal component analysis, and more recently independent component analysis(ICA) [12], [8] and [7] or blind source separation [6] or adaptive filtering techniques [4]. The traditional method is regression analysis which basically consists in the subtraction of the scaled EOG channel (or horizontal and vertical EOG recording channels) from the EEG signal. The most recent works use independent component analysis: [7] used the informax algorithm, [11] and [14] applied the joint approximative diagonalization algorithm (JADE) of eigen-matrices, in [6] an approximate joint diagonalization of time-delayed correlation matrices (SOBI) was used while in [12] the fast fixed point algorithm (FastICA) has been applied. In all the works but one [11], the EOG channels are included in the processed data set of signals. In one of the works [12] it was argued that it can be possible to achieve the computation of the independent components without the inclusion of EOG recordings. One important issue in ICA methods is the identification of the components related with the ocular artifacts. Hence it is needed to eliminate the components related with the artifacts in order to reconstruct the data without artifacts. Most of the works did not give any emphasis to the identification task which seems to be achieved in a visual/manual manner. In spite of the number of available works in the field, the performance of the different methods is not easy to compare as they use different databases, different measures and goals. In [8] the authors conclude that ICA performed better than PCA, where some remnants of the Electrooculogram (EOG) were still visible in the corrected data. In [13] the performance of regression methods, principal component analysis and an independent component analysis method were compared using real and simulated data and the authors conclude that ICA distorts

the power of the EEG signal in the range of $5Hz$ to $20Hz$.

In this work we will present a method based on singular spectrum analysis [3] to extract artifacts from EEG recordings, in particular the high amplitude signals resulting from eye and blinking movements. The proposed method is applied in parallel to a chosen set of channels in order to extract the prominent artifact. This signal is subtracted from the original and a corrected version of the EEG is obtained. Results of real EEG segments, containing paroxysmal epileptic activity contaminated with artifacts will be shown before and after the correction by the method. The removed artifacts are also presented.

2 Denoising and Singular Spectrum Analysis

In many signal processing applications the sensor signals are contaminated with noise which is assumed to be additive and non-correlated with source signals. The general purpose of SSA analysis is the decomposition of a time series into additive components that can be interpreted as "trends" , "oscillatory" and "noise" components. The SSA strategy is widely used in climatic, meteorologic and geophysics data analysis [3]. The SSA analysis techniques rely on the embedding of the sensor signal in the high-dimensional space of its time-delayed coordinates. The multidimensional signal obtained is then projected onto the most significant directions computed using singular value decomposition (SVD) or principal component analysis (PCA). The embedding strategy in conjunction with the FastICA algorithm was used in [5] to decompose the EEG into distinct components.

In this work we propose a clustering step before the PCA (or SVD), hence the multidimensional data is then projected locally, i.e. in each cluster the data is projected in k directions associated to the most significant eigenvalues [10]. The choice of the number of components is based on the application of the MDL (minimum description length) criterion [9].

In the next sections we will describe the main steps of the basic SSA analysis, then the steps of our modified version of the singular spectrum analysis are also described. Finally, the MDL principle to choose the number of significant eigenvalues is presented.

2.1 Embedding and SSA analysis

Embedding can be regarded as a mapping that transfers a one-dimensional time series $x = (x[0], x[1], ..., x[N-1])$ to a multidimensional sequence of lagged vectors. Let M be an integer (window length) with $M < N$. The embedding procedure forms $L = N - M + 1$ lagged vectors, $\mathbf{x}_l, l = 1...L$, that constitute the columns of the trajectory matrix

$$\mathbf{X} = \begin{bmatrix} x[M-1] & x[M] & \dots & x[N-1] \\ x[M-2] & x[M-1] & \dots & x[N-2] \\ x[M-3] & x[M-2] & \dots & x[N-3] \\ \vdots & \vdots & \ddots & \vdots \\ x[0] & x[1] & \dots & x[N-M] \end{bmatrix} \quad (1)$$

Note that the matrix has identical entries along its diagonals. The multidimensional vectors may be centered in the time-delayed space by computing

$$\mathbf{X}_c = \mathbf{X}(\mathbf{I} - \frac{1}{L}\mathbf{j}_L\mathbf{j}_L^T) \quad (2)$$

where $\mathbf{j}_L$ forms a vector of ones with dimension $L \times 1$, and $\mathbf{I}$ represents an $L \times L$ identity matrix. With the centered trajectory matrix the eigendecomposition of the $M \times M$ covariance matrix $\mathbf{S} =< \mathbf{X}_c\mathbf{X}_c^T >$ is computed. In SSA analysis, the strategy to choose the eigenvectors of $\mathbf{S}$ to project and reconstruct the multidimensional signal is called grouping [3]. The choice of eigenvectors depends on the goal of the analysis. The denoising can be achieved by first projecting the multidimensional signal into the space spanned by the $k < M$ eigenvectors corresponding to the most significant eigenvalues and then reconstruct the denoised signals. Considering the matrix $\mathbf{U}$ with the k eigenvectors in its columns, the denoised multidimensional signal is obtained via

$$\hat{\mathbf{X}} = \mathbf{U}\mathbf{U}^T\mathbf{X}_c + \mathbf{X}\frac{1}{L}\mathbf{j}_L\mathbf{j}_L^T \quad (3)$$

Notice that it cannot be expected that $\hat{\mathbf{X}}$ possesses identical elements along each descending diagonal like in the case of $\mathbf{X}$ (eqn. 1). This can be accomplished, however, by replacing the entries in each diagonal by their average along the diagonal to form the matrix $\hat{\mathbf{X}}_d$. Then the extracted one-dimensional signal is obtained by reverting the embedding ,i.e, by taking a sample from each diagonal of the matrix $\hat{\mathbf{X}}_d$ forming the sequence $\hat{x}[n]$.

2.2 Local SSA

The method is explained in detail in [10] but basically it introduces a clustering step in the SSA technique and the MDL criterion to choose the components in each cluster. For convenience we review the steps of all the procedure:

- After embedding the multidimensional signal $\mathbf{x}_l, l = 1...L$, i.e. the columns of the trajectory matrix are clustered using a clustering algorithm (like k-means [1]). After the clustering the set of indices of the columns of $\mathbf{X}$ is subdivided in q disjoint subsets $c_1, c_2, ...c_q$. Thus sub-trajectory matrices $\mathbf{X}^{(c_i)}$ are formed with those columns of the matrix $\mathbf{X}$ which belong to the subset of indices c_i.

- A covariance matrix is computed in each cluster and the data is projected using the eigenvectors corresponding to the k largest eigenvalues. Applying the

MDL criterion (explained next section), the number of significant directions may be different in each cluster. Then the reconstructed sub-trajectory matrices $\hat{\mathbf{X}}^{(c_i)}$ are obtained for each cluster and a processing very similar to the one described by equation 3 follows.

- The clustering is reverted, i.e. each column of the extracted sub-trajectory matrix, $\hat{\mathbf{X}}^{c_i}$, will be assigned to a column of $\hat{\mathbf{X}}$ according to the contents of subset c_i.
- the one-dimensional signal, $\hat{x}[n]$, is obtained by reverting the embedding, i.e. by averaging the corresponding descending diagonals of $\hat{\mathbf{X}}$ to form a sample of the sequence $\hat{x}[n]$.

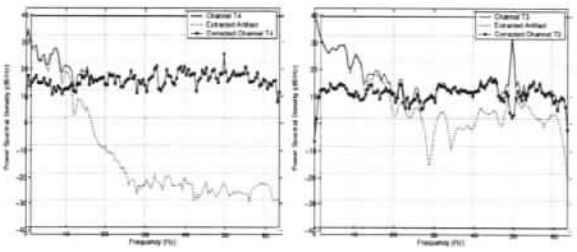
Usually the time series is considered a superposition of components. For instance, if the signal is assumed to be distorted with a non-correlated additive gaussian noise, the following model is considered: $x[n] = y[n] + r[n]$, where $r[n]$ represents a Gausssian noise. The embedding step turns the sequence into a multidimensional vector and the denoising problem, i.e, the reduction of the noise level has to be seen as a nonlinear operation. The clustering step contributes to achieve a non-linear processing with a linear method (the PCA) by choosing directions of maximum variance in sub-groups of vectors in the multidimensional delayed-coordinates.

3 The parameters of the Local SSA

The implementation of the algorithm is achieved by following the steps described above after assigning the following parameters: the embedding dimension (M) and the number of clusters (q) to group the columns of the trajectory matrix. A third parameter can be chosen by the user or may be automatically assigned using an MDL criterion: the number of significant directions k to project and reconstruct the multidimensional vectors.

3.1 Embedding Dimension and Number of Clusters

In SSA applications, the embedding dimension is also discussed and it is recommended to choose it approximately half of the segment length (N)[3]. It is also recommended that it should be close to the periodicity to be extracted from the time series [3]. A similar strategy is suggested in [5] where a lower bound is recommended which is related to the resolution in frequency which each column of the trajectory matrix should represent, i.e, $M > f_s/f_r$ where f_s is the sampling frequency and f_r is the minimal frequency that has to be extracted.
In local SSA the number of clusters has to be assigned also by the user and it was experimentally verified that the number of samples N constitutes the natural maximum threshold once the PCA (and the MDL) performance in each cluster is related with the number of sample vectors to perform the estimates. In particular the cardinality of each cluster cannot be lower than the embedding dimension. However it was experimentally verified that for the same embedding dimension, increasing the number of clusters achieves less distortion in low frequency bands (theta and low-alpha) [10]. Then a simple heuristic is used to create an automatic procedure. The number of clusters starts with a maximum value q_{max} checking afterwards if all clusters end up with a cardinality higher than M and the number of directions in clusters are $k < (M/2)$. If both criteria are not met then the number q of clusters is decreased and the process is repeated until a reliable decomposition in each cluster is achieved.



(a) Channel T4 (b) Channel T3

Figure 1. Power Spectral Density (dB/Hz) versus Frequency (Hz) of 2 channels of segment 1: original, extracted artifact and corrected signals.

3.2 MDL criterion

The determination of the number of significant directions is based on the application of a maximum likelihood estimation of the parameter vector of the covariance matrix of each cluster. Considering ordered eigenvalues, i.e. $\lambda_1 > \lambda_2 > \ldots > \lambda_M$, the k largest eigenvalues of the covariance matrix are sought which minimize the following expression [9]

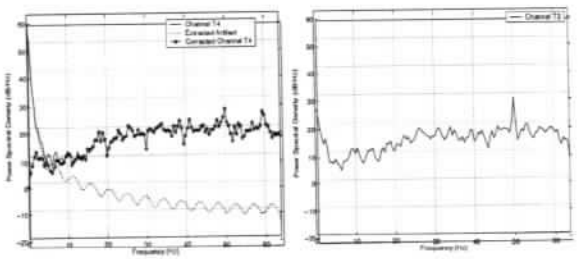
$$MDL(k) = \alpha(k) + \beta(k) \tag{4}$$

where $\alpha(k)$ is related with the log-likelihood function of an estimate of the parameters of the covariance matrix. It is calculated according to

$$\alpha(k) = -N(M-k)\ln\left[\frac{\prod_{i=k+1}^{M}\lambda_i^{1/(M-k)}}{\frac{1}{M-k}\sum_{i=k+1}^{M}\lambda_i}\right] \tag{5}$$

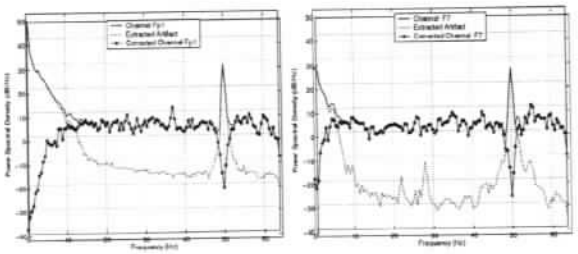
and $\beta(k)$ is a penalty term that represents the complexity of the model and it is computed according to

$$\beta(k) = \frac{1}{2}\ln(N)\left[Mk - k^2/2 + k/2 + 1\right] \tag{6}$$



(a) Channel T4 (b) Channel T3

Figure 2. Power Spectral Density (dB/Hz) versus Frequency (Hz) of 2 channels of segment 2: (a) original, extracted artifact and corrected signals; (b) original



(a) Channel FP1 (b) Channel F7

Figure 3. Power Spectral Density (dB/Hz) versus Frequency (Hz) of 2 channels of segment 2: original, extracted artifact and corrected signals

4 Results

4.1 Data Description

The EEG signals were chosen from a database of epileptic patients collected during long-term EEG monitoring sessions. The EEG signals were recorded using 19 electrodes placed according to the $10-20$ system (common ground reference at Fz). The signals are filtered and digitalized (sampling rate-$128Hz$), stored as European Data Format (EDF) using the EEG Galileo recording system. Monopolar (common Cz reference) brain signals were visualized using EEGLAB [2] and processed using local SSA methodology. We will present results using three data segments (with N=1280 samples) of a patient with paroxysmal epileptiform activity. This patient was with a partial complex seizure from the right temporal focus. The three segments correspond to EEG signals preceding the epileptic seizure onset and are corrupted by artifacts: the first segment starts 28 minutes before seizure onset, the second 24 minutes before and the last starts at seizure onset.

	Number of Clusters		
EEG Channels	**1st Seg**	**2nd Seg**	**3rd Seg**
Fp2-Cz	6	10	8
F4-Cz	3	2	6
C4-Cz	3	-	7
P4-Cz	3	-	8
O2-Cz	7	-	-
F8-Cz	6	6	3
T4-Cz	5	6	-
T6-Cz	3	-	10
Fp1-Cz	6	5	7
F3-Cz	6	7	7
P3-Cz	3	-	-
O1-Cz	3	-	-
F7-Cz	3	8	4
T3-Cz	6	-	-
T5-Cz	6	-	-

Table 1. Number of clusters in each processed channel ("-"indicates not processed).

4.2 Parameters of the algorithm

The algorithm was implemented in MATLAB according to the steps described in section 2.2, considering an embedding dimension $M = 41$ and starting with $q_{max} = 10$. Each channel is processed separately and table (1) presents the number of clusters in each processed signal for the three processed segments. The output of the algorithm, in each channel, forms an estimate of the extracted artifact $\hat{x}[n]$ which is subtracted from the original to obtain the corrected EEG signal.

Segment 1 All channels are processed one after the other by the algorithm (see figure 4), and we can see that $\hat{x}[n]$ (fig. 4- (b)) exhibits the high-amplitude components of the original signals. In most of the channels, an instantaneous frequency analysis (spectrogram) of $\hat{x}[n]$ shows that the frequency contents is mainly in the low frequency range ($< 10Hz$) and also around $50Hz$. But there are channels where $\hat{x}[n]$ also contains contributions in other frequency ranges. The corrected EEG (see fig. 4-(c)) mainly possesses the high frequency ($> 10Hz$) contents of the original signal. However, in $T4$ and $T6$ bursts of theta ($3-7Hz$) waves can be seen (around 00:53s). The figure 1 shows the power spectral density of two temporal channels where distinct frequency characteristics are visible. In $T4$, the power spectral density of the extracted artifact drops from $10dB$ (at $10Hz$) to $-30dB$, while the corrected EEG maintains an amplitude higher than $10dB$ in the whole frequency range. In $T3$, the power spectral density of the corrected EEG maintains that characteristics but the PSD of the extracted artifact did not drop so strongly as in the corresponding PSD of $T4$.

Segment 2 This segment shows typical frontal (eye blinking) artifacts. Hence only the frontal channels and temporal channel $T4$ were processed. Fig 5 shows 4 sec-

onds of this analysis: the extracted signal is only related with an EOG artifact and $50Hz$ line noise (fig. 5-(b)) and the corrected EEG only shows the lower amplitude components of the signal. It is possible to compare the $T4$ channel (fig.5-(c)) is similar with the segment preceding the seizure onset (fig.6-(a) or (c)) and in such case indicating the epileptogenic focus. The power spectral analysis of two channels are shown in figure 2: one channel ($T4$) was processed but the other ($T3$) was not. It is obvious that at high-frequencies ($> 10Hz$) both channels present the same trend. Furthermore the high frequency contents of the extracted artifact in $T4$ has an amplitude $30dB$ lower than in the corrected EEG.

Segment 3 This segment precedes the seizure onset and shows paroxysmal epileptiform activity in the temporal right regions. The frontal channels show ocular artifacts and $C4$, $P4$ and $T6$ show an electrode artifact (a drift in baseline) of low frequency and high amplitude. It can be seen in fig. 6(b) that all those artifacts as well as the $50Hz$ line noise could be extracted. Furthermore, in the corrected EEG (fig. 6-(c)) in channels $Fp1$ and $Fp2$ also spike waves can be seen that were masked by the high amplitude artifacts in the original signal. In $T4$ the spike bursts are even more clearly visible in the corrected EEG. The power spectral density of two channels (see fig 3) shows that for frequencies higher than $7Hz$ the corrected EEG has an amplitude similar to the original signals. The extracted artifact signal is similar to the original before $7Hz$ and at $50Hz$.

5 Conclusion

In this work we presented a method to remove high-amplitude and low frequency artifacts from EEG recordings. The method is applied separately to each channel and our results show that it has good performance when high-amplitude eye movements or low frequency electrode artifact are present. We process the raw data (without any digital filtering processing) and can extract the $50Hz$ line noise in conjunction with low frequency high amplitude components. It has to be pointed out that this evaluation is mainly based on visual comparison between the original, the extracted artifact and the corrected EEG as well as the respective spectrograms (not presented here). The real EEG (without any artifact) is not defined and cannot be practically measured. Therefore no standard exists to quantitatively evaluate the performance of the algorithm. With the method proposed we got the possibility of having the signals separated in two components: the artifacts and the corrected EEG. Furthermore, the user can choose to process a subset of channels keeping others unprocessed which also allows a comparison of the outcomes of the algorithm with non-processed channels . In summary, we present a method that can help the visual inspection of the EEG recordings which might be useful in some critical segment analysis like the onset of the ictal seizure.

References

[1] Christopher M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, 1995.

[2] A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, 134:9–21, 2004.

[3] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky. *Analysis of Time Series Structure: SSA and Related Techniques*. Chapman & HALL/CRC, 2001.

[4] P. He, G. Wilson, and C. Russel. Removal of ocular artifacts from electroencephalogram by adpative fi ltering. *Medical & Biological Engineering & Computing*, 42:407–412, 2004.

[5] C. J. James and D. Lowe. Extracting multisource brain activity from a single electromagnetic channel. *Artificial Intelligence in Medicine*, 28:89–104, 2003.

[6] C. A. Joyce, I. F. Gorodniysky, and M. Kutas. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. *Psychophysiology*, 41:313–325, 2004.

[7] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. Mckeown, V. Iragui, and T. j. Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37:163–178, 2000.

[8] T.-P. Jung, S. Makeig, M. Westerfi eld, J. Townsend, E. Courchesne, and T. J. Sejnowski. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, 111:1745–1758, 2000.

[9] A. P. Liavas and P. A. Regalia. On the behavior of information theoretic criteria for model order selection. *IEEE Transactions on Signal Processing*, 49(8), 2001.

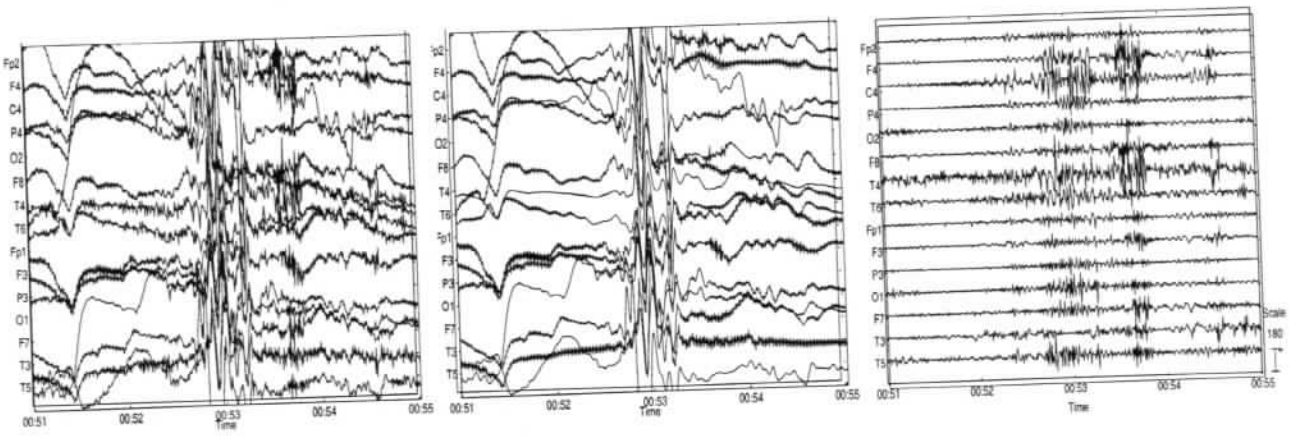
[10] A. Teixeira, A. M. Tomé, E.W.Lang, P. Gruber, and A. M. d. Silva. On the use of clustering and local singular spectrum analysis to remove ocular artifacts from electroencephalograms. In *IJCNN2005*, pages 2514–2519, Montréal, Canada, 2005. IEEE.

[11] E. Urrestarazu, J. Iriarte, M. Alegre, M. Valencia, C. Viteri, and J. Artieda. Independent component analysis removing artifacts in ictal recordings. *Epilepsia*, 45(9):1071–1078, 2004.

[12] R. N. Vigário. Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalography and Clinical Neurophysiology*, 103:395–404, 1997.

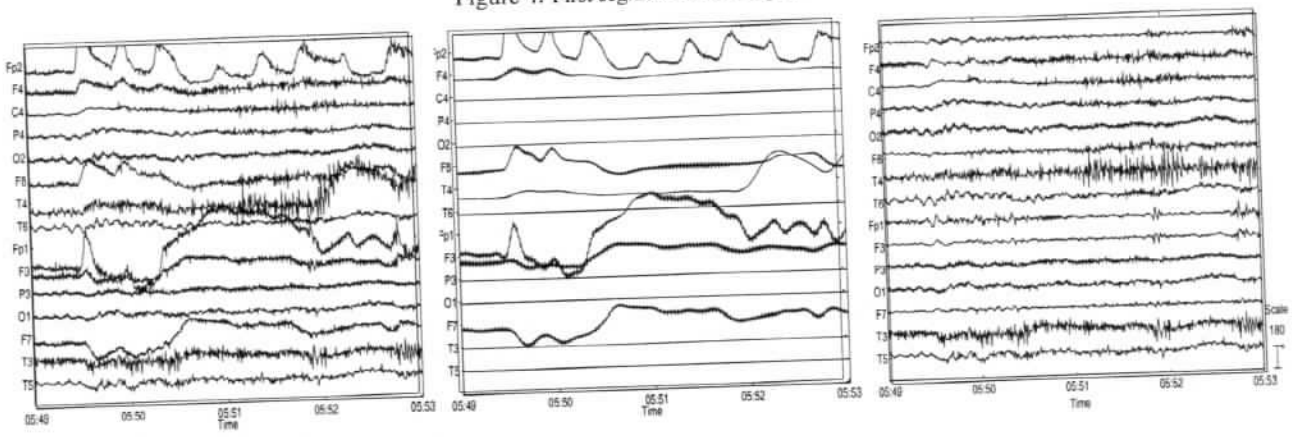
[13] G. L. Wallstrom, R. E. Kass, A. Miller, J. F. Cohn, and N. A. Fox. Automatic correction of ocular artifacts in the EEG: a comparison of regression and component-based methods. *International Journal of Psychophysiology*, 53:105–119, 2004.

[14] W. Zhou and J. Gotman. Removing eye-movement artifacts from the EEG during the intracarotid amobarbital procedure. *Epilepsia*, 46(3):409–414, 2005.



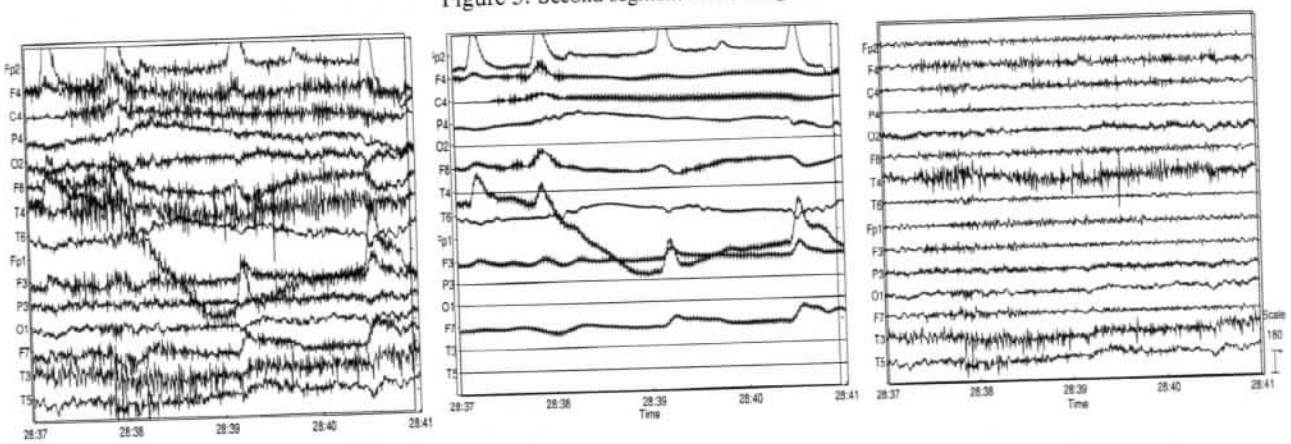
(a) Original Signal (b) Extracted Artifacts (c) Corrected Signal

Figure 4. First segment of EEG signal



(a) Original Signal (b) Extracted Artifacts (c) Corrected Signal

Figure 5. Second segment of EEG signal



(a) Original Signal (b) Extracted Artifacts (c) Corrected Signal

Figure 6. Third segment of EEG signal