Eye Blink Artifact Removal from Cognitive EEG Data using ICA (Independent Component Analysis)

Sagar S. Motdhare[1], Dr. Garima Mathur[2]

[1]Research Scholar, EEE Department, Poornima University, Jaipur, India
[2]Professor, EEE Department, Poornima University, Jaipur, India
[1]sagar_motdhare@yahoo.com, [2]drg.mathur@poornima.edu.in

Abstract

Eye blinking artifacts make electroencephalographic (EEG) elucidation as well as examination more difficult. In this paper, eye blinking artifacts are removed from cognitive EEG records by means of ICA. The blinks of eyes approximately at all times coincide with the event-related potentials (ERP) due to the experiment's design, generating ICA difficulties. To build a new data set, we combined spontaneous blink data with sole tryout ERP data. According to our findings, applying ICA to the combined data set results in a more logical separation, making EEG interpretation and analysis easier.

Keywords: - EEG, Eye-blink Artifact, ICA.

INTRODUCTION:

Electroencephalographic recordings are frequently contaminated by movements of eyes, eye blinks, muscular action as well as line noise. These aberrations can cover a broad area of the scalp and be many commands in size greater than the signals of concern. When various aspects of the signals (amplitude, latency, etc.) must be investigated, as in event-related potentials (ERP) analysis, this presents a substantial issue. As a result, it's preferable to get rid of the artifacts as totally as promising while preserving the fundamental brain impulses.

The use of a voltage threshold to reject EEG epochs with artifacts is a prevalent method. However, if some individuals' blinks and muscle activities occur too repeatedly, the quantity of data that is disregarded may become unsatisfactory [1]. A number of methods for removing ocular artifacts have been developed; all of them are based on time or frequency domain regression [3]-[4]. One of the most important drawbacks in all regression approaches is that the baseline pathways, which are usually EOG, include together eye in addition to brain activity, resulting in significant alteration of critical brain signals. [1][5]-[6].

For reducing ocular artifacts in EEG data, PCA has been projected as a decomposition technique [7]. The primary restriction of PCA is that the data component projections must be orthogonal, which is not always the case with artifacts along with brain signals. Source modeling, PCA, and artifact averaging [8] surpass the individual approaches outlined above, but it necessitates prior incident associated activity of the brain modeling and a considerable quantity of data on calibration. [1].

ICA has lately emerged as a viable solution to the problem of BSS. Makeig and Vigario et al. [9]-[10] independently demonstrated that ICA can distinguish activity in the brain from muscles and blink artefacts using EEG and MEG data. Jung et al. employed the bigger Infomax ICA approach to remove ocular flashes, muscular noise, heart impulses and line noise from unstructured EEG data. In the domain of EEG data processing, artefact reduction depending on ICA has now emerged as a feasible option. [1]

INDEPENDENT COMPONENT ANALYSIS (ICA)

ICA is a mathematical means for quantifying a linear transformation of data with as many independent outputs as achievable. It is exceptionally effective method for BSS in situations where the statistically independent sources are available. The following is the vital instantaneous linear model of ICA:

$$\mathbf{x} = \mathbf{As}$$

Where $\mathbf{x} = (x_1, x_2, ...x_m)^T$ is the pragmatic m-directional vector; $\mathbf{s} = (s_1, s_2, ...s_m)^T$ is an unidentified random vector with ICs in lieu of the sources; $\mathbf{A}$ is an unidentified m x m constant un-mixing matrix.

The explanation of fundamental un-mixing matrix can be uttered as:

$$\mathbf{y} = \mathbf{Wx}$$

Where $\mathbf{W}$ is m x m de-mixing matrix.

There have been a slew of methods developed to compute the de-mixing matrix W, the most popular of which are Informax, FastICA, JADE, and SOBI [11]–[14]. The expanded Infomax [14], which is an expansion of the Infomax method, is briefly described here. The expanded Infomax, which is based on the idea of information maximisation, can discriminate between resources with sub Gaussian with super Gaussian populations. In the expanded Infomax, a learning rule switches between the two types of distributions:

$$\Delta \mathbf{W} \, \alpha \, [\mathbf{I} - \mathbf{K} \tanh(\mathbf{y})\mathbf{y^T} - \mathbf{yy^T}]\mathbf{W}$$

For $k_i = 1$, Super-Gaussian

For $k_i = -1$, Sub-Gaussian

The diagonal matrix $\mathbf{K's}$ entries are obtained as follows:

$$K_i = \text{sign} \{E[\text{sech}^2 (u_i)] \, E[u_i^2] - E[\tanh(u_i) \, u_i]\}$$

In specified cases, separation is done by maximizing system diversity, which is identical to the maximum probability technique. Minimizing reciprocal information between data is well known as increasing the probability of independent and identically distributed data. As a result, the enlarged Infomax approach can be used to distinguish independent sources. [1]

APPLYING ICA TO EEG AND ARTIFACT REMOVAL

The application of ICA to the aforementioned scenario necessitates the availability of substantially different sources, their linear immediate mixing at the sensors, and the mixing mechanism' stationarity. At EEG frequencies, they have been argued to be plausible assumptions [16]. Multiple independent components (IC) typically exhibit artifact-like features when ICA is useful to EEG data. Examining the component's time course is one technique to separate artifacts from brain processes. Eye blinks, for example, frequently feature transient big mono-polar peaks. Another crucial consideration is the component's relative projection strengths at each of the scalp sensors, which are shown by the W^{-1} column. These coefficients determine the scalp topography of each component, presenting proof for the function's neurological origin [6]. Blinks of the eyes, for example, should be directed to the far frontal parts of the head. When an independent component is determined as being

an artifact, the artifact can be removed by merely extrapolating the components backside to every scalp sensor and removing it from the unusual EEG data. [1]

THE COGNITIVE EEG DATA AND SOME PRACTICAL ISSUES

To capture the EEG of one healthy person, 128-channel Geometric Sensor Networks were placed at 128 scalp electrodes. For analysis, all 128 channels were transmitted to Cz and digitally processed at 100Hz. All channels were subjected to a bandpass filter with a frequency range of 0.1 to 60 Hz before being transformed to average reference. We chose 33 out of 129 channels (including Cz) for ultimate analysis using the standardized International 10-20 Electrode Positioning System. The volunteers were subjected to an auditory startle stimuli (a 50 ms, 95 dB burst of white noise with an instantaneous increasing time) while seeing pictures with varying valence [17]. The single-trial ERP data set we used had 122 trials, each lasting 1000 milliseconds and starting 250 milliseconds before the startle event. [1]

Due to the manner the experimentation was set up, the participant would blink once every time the startle stimulus was administered. As a result, the eye blink becomes "event-related" and time-locked to the stimulus's start. This means that the stimulus synchronises blinks and ERP (particularly, the N100 wave), and that both have mono-polar forms with peaks around the same time. This is problematic for ICA since the fact outlined above undermines the premise of independence. In reality, ICA can be used to extract single components that account for co-occurring occurrences when numerous topographically visible sources almost invariably co-occur in the facts. [18].

Previous studies utilising ICA to remove ocular eye movements from ERP that is time locked to the same stimuli have been restricted, if at all. To address this issue, we combined the single-trial ERP data with 159 trials of spontaneous blinks (without any stimulation) and sent the full data set to ICA. In the next part, we'll utilise ICA to analyze the performance of artefact eradication for the integrated collection and single-trial ERP set of data.

RESULTS

In this part, the results of ICA on the mixed set of data and single-trial ERP testing dataset are presented. In this example, we're mainly concerned with decreasing eye blink artefacts because they have a major impact on ERP analysis. The EEGLAB software suite [18] was used to generate the ICA, which is a useful tool for evaluating solitary trial Data sets.

Blinking's capacity to successfully project to far frontal regions of the skull and produce enormous brief monopolar potentials in time has been extensively documented. We carefully studied the brain surface features and frequencies of all 33 ICs after the ICA decomposition and manually selected the component that best mimicked the blink artifact's features.

Figure 1 and figure 2 illustrate the IC1 parameters derived from a solitary trial ERP data set and a collective data set (a). Blink artefacts were found in both cases. In all repeated runs, the enlarged Infomax approach consistently recovered these 2 independent components as IC1 for the two data sets. Figure 1 shows an ERP picture. In addition to the original 122-trial ERP data, there are 159 trials of EEG data that simulate spontaneous blinks is depicted in 2(a). The fact that IC1 shows blink artifacts has been proven.

Surprisingly, when ICA was applied to the merged set of data, the ERP was detected as a single IC, as seen in Fig. 2. (b). The startle ERP projects most powerfully to the deepest area of the scalp, according to the IC4 scalp topography, and the two peaks in the waveform correspond to the N100 and P300 waves, respectively.

However, we must exercise caution while working with IC4 because other ICs may extract some ERP, albeit insignificantly.
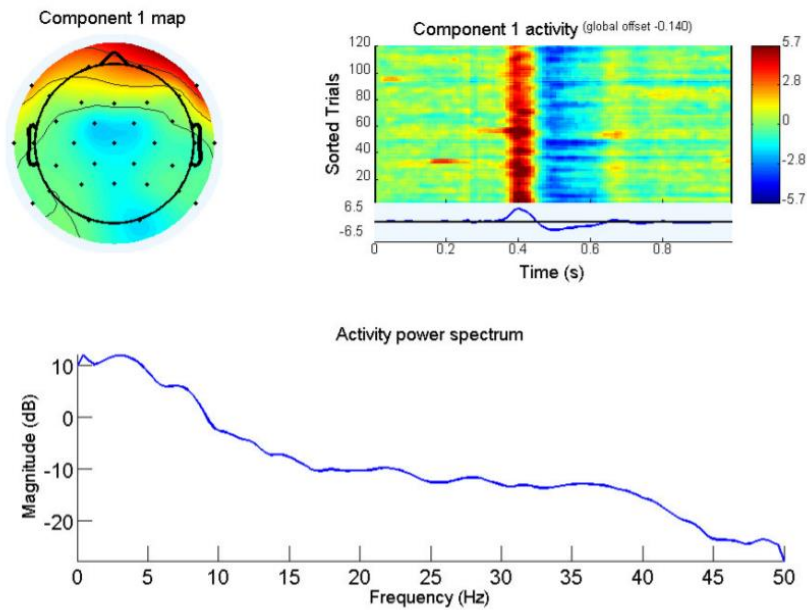

Fig. 1: Independent component 1 derived from ERP data from a single trial. Topography of the scalp at the upper left. ERP picture and average waveform, upper right. The average power spectrum is at the bottom.
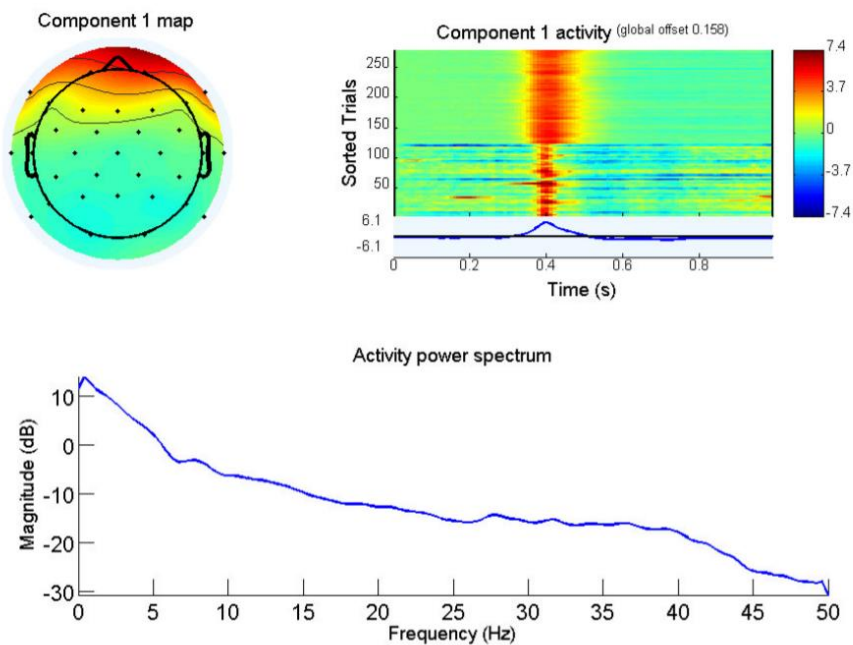

Fig. 2(a): Independent Components obtained from combined dataset IC1
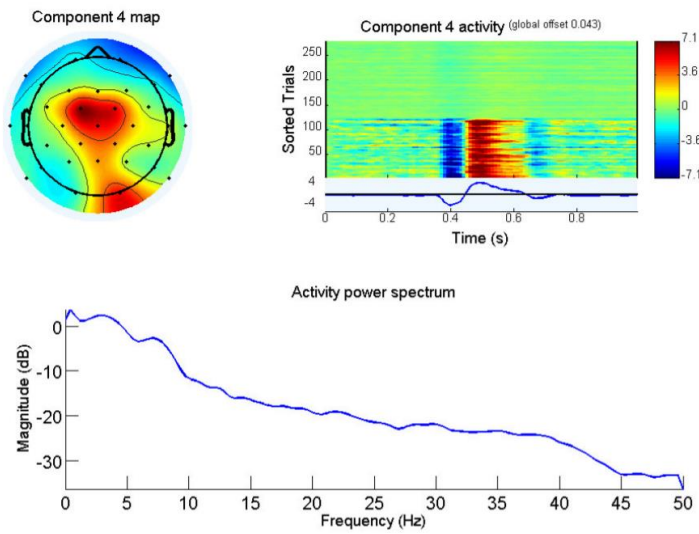
Fig. 2(a): Independent Components obtained from combined dataset IC4

The two data sets are then compared to see how well ICA's blink removal works. Figs. Fp1, Fz, and Cz, Pz, respectively, are located at Fp1, Fz, and Cz, Pz. The average ERP before and after blink removal are depicted in Figures 3 and 4. Blink removal with only single-trial ERP data failed right away since the majority of ERP was also eliminated, in addition to eye blinks. Given that the IC1 derived from solitary ERP data shows blinks and ERP co-activation, this is feasible . This demonstrates that ICA is incapable of distinguishing between sources that virtually invariably co-occur in the data. When employing the merged data set, ICA, on the other hand, produced more reasonable findings. Significant ERP presentations at Cz, Pz, and even Fp1, in which the N100 and P300 waves were collected, were revealed by eye blinks (make a note of the negative ERP protrusion at Fp1). While the blinks look to be gone at Fz, understanding it is still difficult. It's possible that the baseline for this channel was set incorrectly.
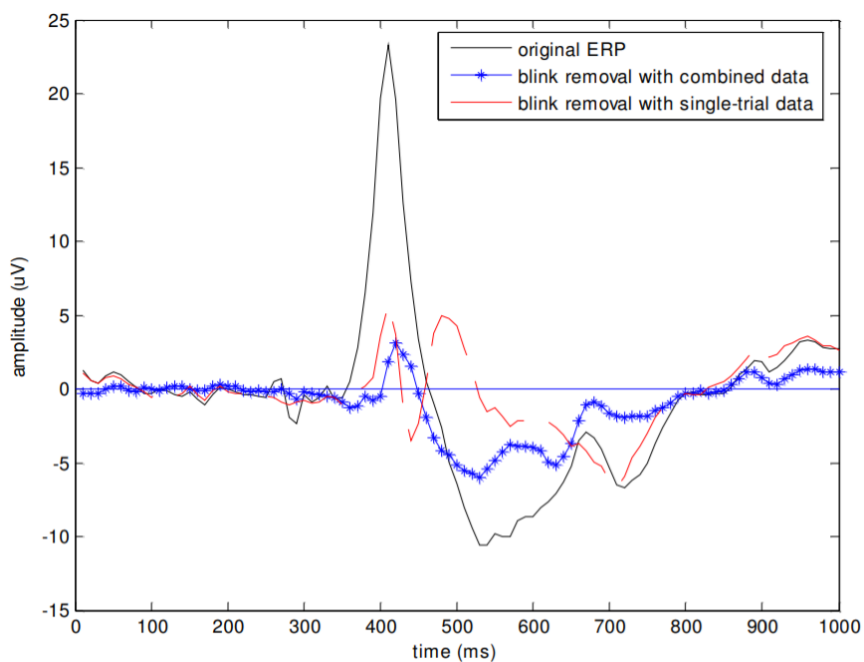

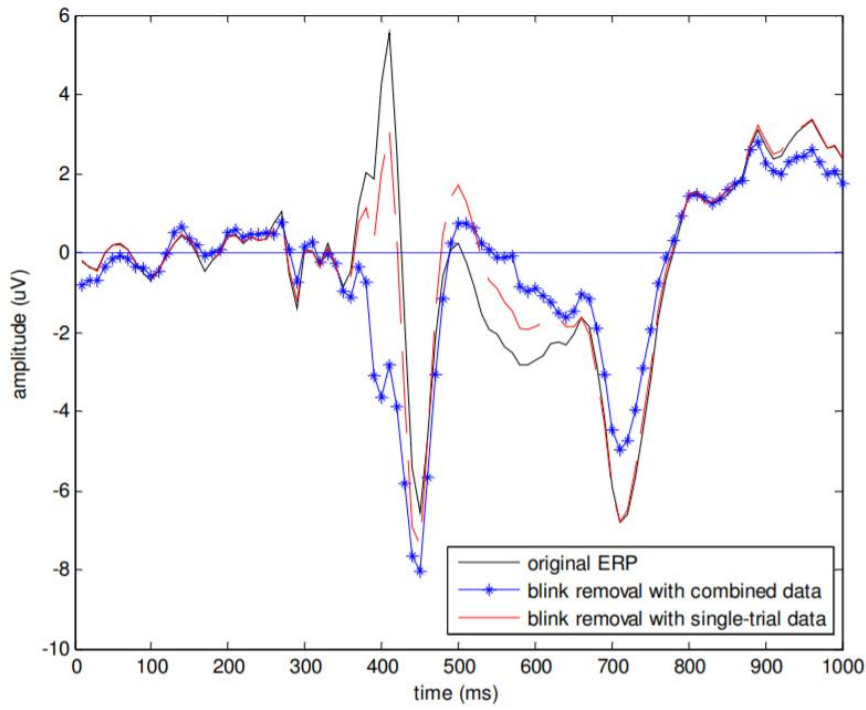**Fig. 3(a): Average ERP before and after blink removal at Fp1**

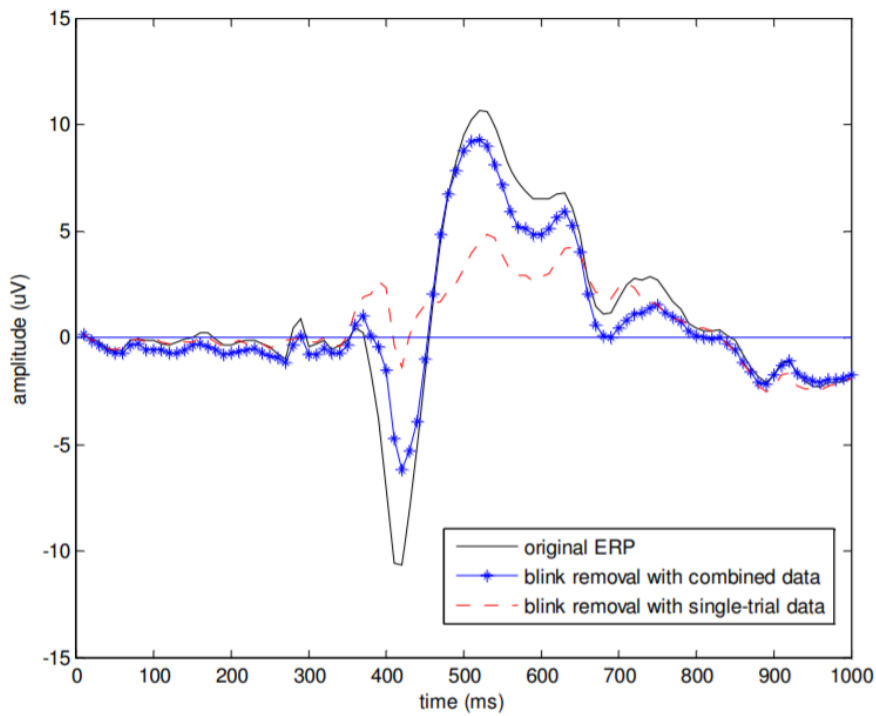**Fig. 3(b): Average ERP before and after blink removal at Fz**

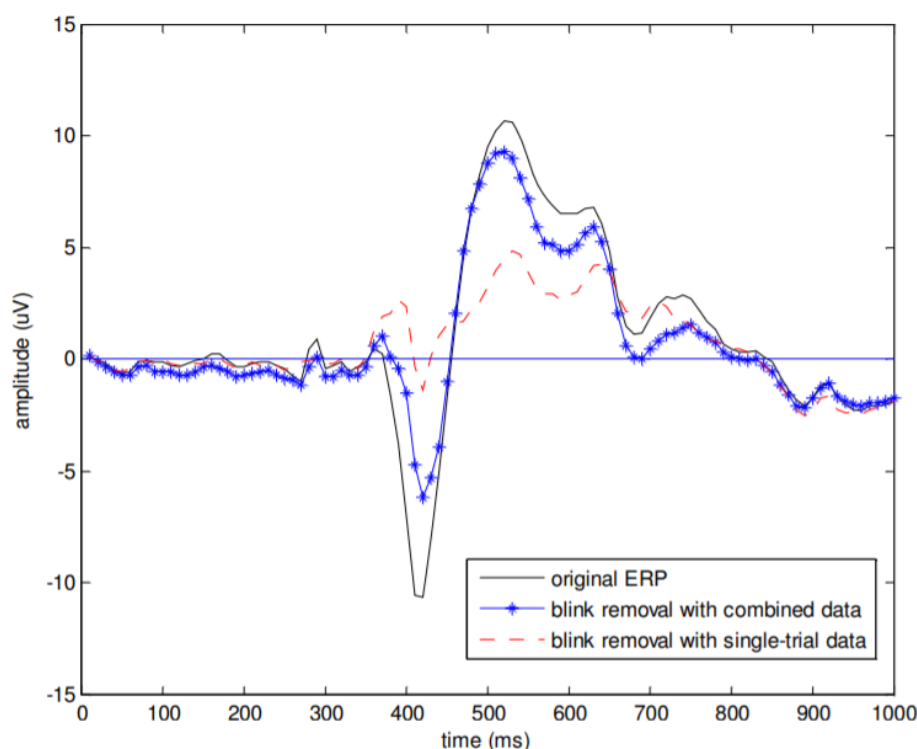**Fig. 4(a): Average ERP before and after blink removal at Cz**

Fig. 4(b): Average ERP before and after blink removal at Pz

DISCUSSION

Our early labors to remove the eye-blink artifacts from ERP are shown in the above research. On the continuous EEG data, we may have utilized ICA. Pruning the EEG data, on the other hand, would take a long time, and ICA could provide difficult-to-understand conclusions. When particular assessment ERP data is paired with blink records, ICA only receives the data necessary to distinguish the two sources. We believe that when blinks and ERP are synchronised with the same input, ICA becomes "ill-conditioned," making it harder to distinguish between the two sources. Each row of W indicates a vector in the space encompassed by all of the image pixels if we view the de-mixing matrices W as a group of spatial filters. When two sources covariate with time, the two rows of W relating to them will point in local thresholding space orientations that are so close together so that W estimation becomes inadequately, and ICA will eventually incorporate various spatial filters and identify the different sets as a single IC.

However, when we feed the impulsive blink data to ICA computations, the updated info, which was not available in the solitary ERP data, draws ICA's attention, making the separation problem 'better-conditioned.' Despite their temporal proximity, the constituent mappings of IC1 and IC4 generated from the merged set of data (see Figure 2) reveal that their scalp features are substantially distinct. ICA looks for discrete geographic expansions of blinks and ERP to identify the two sources.

It's worth mentioning that the strategy to solve the separation difficulty mentioned above is still based on the premise of independence. For single-trial ERP data, the separation is not guaranteed to be ideal because the blinks as well as ERP become fundamentally coupled due to the same stimulus. Working in the context of conditional independence is one viable option. The stimulus is a shared parent node in graphical models for both blinks and ERP because there are causal relationships seen between stimulus and the various responses.

As a result, when the identical stimulus is delivered, the blink and the ERP turn into provisionally independent. Bach [20] proposed a modern TCA model that could be effective in this situation. Another alternative is to tackle the problem using temporally constrained ICA, which takes into account the blinks' distinct temporal structure.

REFERENCES

[1] Li, Ruijiang, and Jose C. Principe. "Blinking Artifact Removal in Cognitive EEG Data Using ICA." 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, 2006, https://doi.org/10.1109/iembs.2006.260605.

[2] J.G. Small, Sensory evoked responses of autistic children. Infantile Autism 224-39, 1971.

[3] S.A. Hillyard, R. Galambos, Eye-movement artifact in the CNV. Electroencephalographic Clinical Neurophysiology 28(2):173-82, 1970.

[4] J.L. Whitton, F. Lue, H. Moldofsky, A spectral method for removing eye-movement artifacts from the EEG. Electroencephalographic Clinical Neurophysiology;44:735-41, 1978.

[5] P. Berg, M. Scherg, A multiple source approach to the correction of eye artifacts. Electroencephalographic Clinical Neurophysiology; 90:229-41, 1994.

[6] 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;37:163-78, 2000.

[7] P. Berg, M. Scherg, Dipole models of eye activity and its application to the removal of eye artifacts from the EEG ad MEG. Clin Physiol Meas;12(Suppl A):49-54, 1991.

[8] P. Berg, M. Scherg, A multiple source approach to the correction of eye artifacts. Electroenceph clin Neurophysiol;90(3):229-41, 1994.

[9] S. Makeig, T-P Jung, A.J. Bell, D. Ghahremani, T.J. Sejnowski, Blind separation of event-related brain responses into independent components. Proc Natl Acad Sci USA;94:10979-84, 1997.

[10] R.N. Vigario, Extraction of ocular artifacts from EEG using independent component analysis. Electroenceph clin Neurophysiol;103:395-404, 1997.

[11] A. Bell, T. Sejnowski, An Information Approach to Blind Separation and Blind Deconvolution, Neural Compuationt., vol. 7: 1129-59, 1995.

[12] A. Hyvärinen, E. Oja, A Fast Fixed-Point Algorithm for Independent Component Analysis, Neural Computation, vol. 9: 1483–92, 1997.

[13] J.-F Cardoso, High-Order Contrasts for Independent Component Analysis. Neural Computation, vol. 11, 157–92, 1999.

[14] A. Belouchrani, K. Abed-Meraim, J-F Cardoso, E. Moulines, A Blind Source Separation Technique Based on Second-Order Statistics, IEEE Trans. on Sig. Proc., vol.45, 434–44, 1997.

[15] T-W Lee, M. Girolami, T. Sejnowski, ICA Using an Extended Infomax Algorithm for Mixed Sub- and Supergaussian Sources. Neural Computation., vol. 11, 417–41, 1999.

[16] R. Vigario, J. Sarela, V. Jousmaki, M. Hamalainen, E. Oja. Independent component approach to the analysis of EEG and MEG recordings. IEEE Trans Biomed Eng. 47(5):589-93, 2000.

[17] B. Cuthbert, H. Schupp, M. Bradley, M. Mcmanis, P. Lang. Probing affective pictures: attended startle and tone probes, Psychophysiology; 35, 344-7, 1998

[18] T-P Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, T.J. Sejnowski. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clinical Neurophysiology;111:1745-58, 2000.

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

[20] F. Bach, M. Jordan, Beyond independent components: trees and clusters, Journal of Machine Learning Research 4: 1205-33, 2003