Elimination of eye artifacts using the Independent Component Analysis (ICA) procedure

Abstract

This report addresses the issue of eye artifacts in EEG signals and proposes a solution based on Independent Component Analysis (ICA). The project employs ICA to decompose EEG signals into independent components, aiming to identify and eliminate components associated with eye artifacts, particularly blinks. Through this procedure, the resulting EEG signals are expected to be more refined, enhancing the precision of subsequent analyses focused on neural activity. The effectiveness of the proposed method is demonstrated on empirical EEG data, showcasing its potential contribution to improving the quality of EEG signal processing in the presence of eye-related artifacts.

1 Introduction

In the realm of electroencephalography (EEG) signal processing, the presence of eye artifacts poses a significant challenge to accurate analysis. These artifacts, arising primarily from eye blinks, can distort the EEG signals, potentially leading to misinterpretation of underlying neural activity. This study focuses on mitigating such artifacts through the application of Independent Component Analysis (ICA), a powerful technique capable of isolating distinct sources contributing to the observed EEG signals. By leveraging ICA, we aim to separate eye artifacts from the genuine neural signals, providing a clearer and more accurate representation of cerebral activity.

2 Methods

2.1 Data

The dataset utilized in this study comprises EEG recordings from three subjects, specifically records from subjects S001, S006, and S053. These recordings were obtained from the EEG Motor Movement/Imagery Dataset (EEGMMIDB) [1], a publicly available dataset designed for motor movement and imagery studies.

2.2 ICA

Independent Component Analysis (ICA) [3] is a computational technique employed in the analysis of multivariate data to uncover underlying independent components. In the context of EEG signal processing, ICA is widely used to separate mixed signals into their constituent components, revealing distinct neural sources contributing to the recorded EEG signals. The fundamental idea behind ICA is to find a linear transformation that maximizes the statistical independence of the components.

Given an observed signal matrix $\mathbf{X}$ of dimensions $N \times T$, where N represents the number of channels and T denotes the number of time samples, ICA seeks to factorize $\mathbf{X}$ into the product of two matrices $\mathbf{A}$ and $\mathbf{S}$:

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S} \tag{1}$$

Here, $\mathbf{A}$ is the mixing matrix representing the spatial patterns or weights of the sources, and $\mathbf{S}$ is the source matrix containing the independent components. The objective is to estimate $\mathbf{A}$ and $\mathbf{S}$ such that $\mathbf{S}$ has statistically independent columns.

Independent Component Analysis (ICA) can be calculated with various algorithms, including Infomax ICA, JADE, Extended Infomax ICA, Radical ICA, SOBI, and CuBICA. Each algorithm employs distinct techniques to achieve blind source separation and enhance independence among components.

FastICA Algorithm

The FastICA algorithm [2] is a popular implementation of ICA known for its efficiency. It maximizes non-Gaussianity, aiming to find sources that are as independent as possible. The core steps of the FastICA algorithm involve the following:

1. **Whitening:** Preprocess the data by whitening the observed signals, transforming them into uncorrelated variables with unit variance.
2. **Orthogonalization:** Apply orthogonalization to further decorrelate the signals and enhance non-Gaussianity.
3. **Nonlinearity:** Use a nonlinear function (contrast function) to maximize non-Gaussianity. Common contrast functions include negentropy and kurtosis.
4. **Iterative Update:** Iteratively update the weight matrix to maximize the non-Gaussianity of the estimated sources.

Mathematically, the FastICA algorithm aims to maximize the negentropy of the sources, defined as:

$$J(\mathbf{y}) = H(\mathbf{y}_{\text{gaussian}}) - H(\mathbf{y}) \tag{2}$$

where $\mathbf{y}_{\text{gaussian}}$ is a Gaussian random variable with the same covariance matrix as $\mathbf{y}$, and $H(\cdot)$ is the entropy function.

The iterative nature of FastICA ensures the convergence of the algorithm, resulting in an estimate of the mixing matrix $\mathbf{A}$ and the source matrix $\mathbf{S}$.

In the analysis, the FastICA algorithm was used to separate EEG signals into independent components, facilitating the identification and removal of eye artifacts from the recorded data.

2.3 Artifact Removal

After applying the FastICA algorithm to decompose the observed EEG signals into independent components, the next step is to identify and eliminate components (1) associated with artifacts, such as those arising from eye blinks. In the implemented approach, we leveraged user interaction to visually inspect the components and selectively exclude those corresponding to eye artifacts.

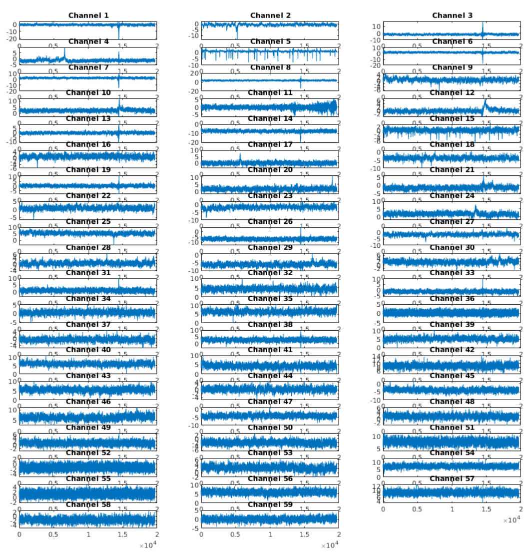
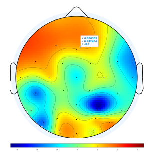


Figure 1: Visual exploration of individual channels of the signal decomposed by FastICA

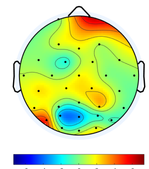
Visualization of $\mathbf{W_1}$ Matrix

The matrix $\mathbf{W_1}$ obtained from the FastICA algorithm represents the unmixing matrix, which contains the weights assigned to each channel for each independent component. Visualizing $\mathbf{W_1}$ provides insights into the spatial patterns of the components, allowing to identify those associated with specific brain regions, including the frontal lobe near the eye sockets.

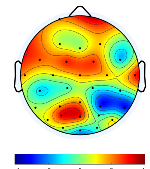
To visualize $\mathbf{W_1}$, topographic plots were employed that display the distribution of weights across the scalp. Each subplot corresponds to a different independent component, and the topographic plot depicts the spatial distribution of weights for that component. Components with prominent activity around the frontal lobe region are indicative of potential eye blink artifacts (2).



(a) Eye Blink 1



(b) Eye Blink 2



(c) Eye Blink 3

Figure 2: Identified Eye Blinks

User-Driven Artifact Exclusion

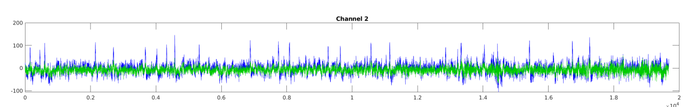
To enhance the flexibility of our artifact removal process, we incorporated user interaction for component exclusion. The user is prompted to visually inspect the topographic plots and select components that exhibit frontal lobe activity consistent with eye artifacts.

Once the user identifies components for exclusion, the corresponding indices are used to construct a modified unmixing matrix $\mathbf{W}_{\text{ap}}$ and a new set of independent components $\mathbf{Y}$. This refined signal is then reconstructed, providing a cleaned EEG signal with the targeted removal of eye blink artifacts.

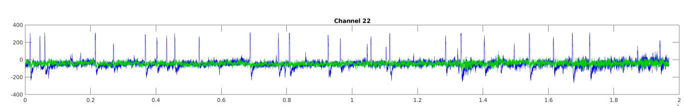
3 Results

The results of our artifact removal process are visualized through a series of plots. First, we present the original EEG signals alongside the decomposed independent components. The spatial patterns of the components are shown using topographic plots based on the $\mathbf{W_1}$ matrix. Components identified for exclusion are marked, facilitating transparency in the artifact removal process.

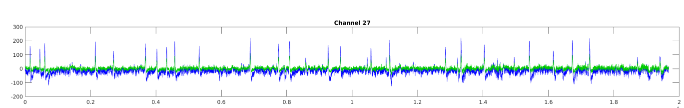
Next, we display the cleaned EEG signals obtained after excluding the selected components. Figure (3) illustrates three sample corrected signals.



(a)



(b)



(c)

Figure 3: Cleaned EEG Signals after Artifact Removal

4 Discussion

4.1 Current Limitations and Challenges

While the Independent Component Analysis (ICA) procedure coupled with user-driven artifact exclusion has proven effective in mitigating eye blink artifacts from EEG signals, there are inherent limitations and challenges. The success of artifact removal heavily relies on the user's ability to visually identify and exclude relevant components. This manual process can be time-consuming, subjective, and may vary across users.

Additionally, the ICA method assumes statistical independence between components, which may not always hold in real-world scenarios. Complex and overlapping neural sources might lead to challenges in separating genuine neural activity from artifacts solely based on statistical independence.

4.2 Enhancing Artifact Removal with Machine Learning

To overcome current limitations and refine artifact removal, integrating machine learning techniques presents a promising avenue. These methods can offer automated solutions for identifying and eliminating artifacts, reducing the need for manual intervention.

Supervised Learning for Artifact Classification

Supervised learning models, trained on labeled datasets containing both artifact-free and contaminated EEG signals, can learn to distinguish between the two. Input features, such as spectral characteristics and temporal patterns, enable classifiers (e.g., SVM or deep learning) to automatically identify and remove artifact-containing segments.

Unsupervised Approaches for Anomaly Detection

Unsupervised methods, like clustering or anomaly detection, identify abnormal patterns in EEG data. Algorithms such as k-means or density-based methods group EEG segments based on characteristics, flagging outlying clusters as potential artifacts.

4.3 Conclusion and Future Directions

While this work demonstrates progress in EEG artifact removal with ICA, integrating machine learning holds significant potential for improvement. Future research should explore and compare various machine learning methods, balancing automation, accuracy, and complexity. Developing user-friendly tools for seamless integration of machine learning-based artifact removal into EEG analysis pipelines could enhance reliability and efficiency in EEG data preprocessing.

References

[1] A. L. Goldberger et al. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals". In: *Circulation* 101.23 (2000). [Circulation Electronic Pages; http://circ.ahajournals.org/cgi/content/full/101/23/e215]; 2000 (June 13), e215–e220.

[2] A. Hyvärinen. "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis". In: *IEEE Transactions on Neural Networks* 10.3 (1999), pp. 626–634. DOI: 10.1109/72.761722.

[3] A. Hyvärinen and E. Oja. "Independent Component Analysis: Algorithms and Applications". In: *Neural Networks* 13.4-5 (2000), pp. 411–430. DOI: 10.1016/S0893-6080(00)00026-5.