Eeg Artifacts Removal By Ica

Sagar A. More, Jagadish B. Jadhav, Shailaja A. Patil, Vijaya Ahire, Kiran H. Sonawane

Abstract— A common problem encountered in data analysis and signal processing is finding an appropriate representation of multivariate data. Most popular linear transformations are principal component analysis or projection pursuit. Comparatively newly developed nonlinear method is that the Independent Component Analysis (ICA), during which the components of the specified representation have minimal stochastically dependence. In this paper, we are throwing light on the appliance of ICA to electroencephalographic data. Eye movements, eye blinks, cardiac signals, muscle noise, and line noise present serious problems for electroencephalographic interpretation and analysis when rejecting contaminated EEG segments leads to an unacceptable data loss. Use of principal component analysis (PCA) has been proposed to get rid of eye artifacts from multi-channel EEG. However, PCA cannot completely separate eye artifacts from brain signals, especially once they have comparable amplitudes. Here, we are reviewing a replacement and usually applicable method for removing a good sort of artifacts from EEG records supported blind source separation (BSS) by independent component analysis (ICA).

Index Terms— Independent component analysis, Principle Component Analysis, Blind source separation, Gaussian, Artifacts.

1 INTRODUCTION

The Electroencephalogram (EEG) is that the recording of the electrical potential of the brain, employing either invasive (internal) or non-invasive (external) detection methods. The externally recorded EEG is low amplitude (order of µV), highly attenuated and contaminated with artifacts e.g. ocular artifacts (EOG) or skeletal muscle movements, line noise [1]. Numbers of techniques are applied, on the removal of artifacts like thresholding, regression, wavelet thresholding [2, 3 and 4] as a way of distinguishing between the EEG and therefore the artifact signals. An alternative and more accurate route that doesn't end in any loss of valuable data is that the application of Blind Source Separation (BSS), and more commonly Independent Component Analysis (ICA) [5, 6, and 7]. we will discuss how the artifacts are far away from EEG data using higher statistical approach. A replacement higher order statistical tool Independent Component Analysis separates a group of knowledge into its statistical independent components. These components can then be studied and those identified as artifacts, can be removed.

2 BLIND SIGNAL SEPARATION

Trying to separate a group of mixed signals without knowing anything about the amount of original signals or how they're mixed together is named Blind Signal Separation or Blind Source Separation.

Let \( X = \) recorded signal, \( S = \) original source signal, \( H = \) mixing of artifacts (noise).

In matrix form:

\[
X = HS
\]

The entire point of Blind Signal Separation is that we all know neither \( H \) nor \( S \) which makes it tons more complex. A method to unravel this problem would be to use the upper statistical tool Independent Component Analysis (ICA).

2.1 INDEPENDENT COMPONENT ANALYSIS

ICA are often defined because the method of decomposing a group of multivariate data into its underlying statistically independent components. The statistical “latent variables” [10] model are often defined as,

\[
x_i = a_{i1}s_1 + a_{i2}s_2 + ... + a_{in}s_n
\]

For all \( i = 1, …, n \)

Where \( a_{ij} \), \( j = 1… n \). By definition, the sources \( s_i \) is statistically independent.

If \( S = [s_1, s_2, s_3, …, s_n]^T \) represents the original multivariate data that is transformed through some transformation matrix \( H \) producing \( X \) such that:

\[
X = HIS
\]

Then ICA tries to identify an unmixing matrix \( W \) such that:

\[
W \approx H^{-1}
\]

So that the resulting matrix \( Y \) is:

\[
Y = WX = W(HS) = S' \approx S
\]

(Since, \( W \approx H^{-1} \))

The only thing ICA demand is that the original signals \( s_1, s_2, …, s_n \) are to be satisfied following assumptions.

- The independent components are statistically independent and the mixing is linear.
- There is no quite one Gaussian signal among the latent variables and therefore the latent variables have cumulative density function not much different from a logistic sigmoid.
- The ICA model works ideally when \( n = m \) [5].
- The mixing matrix is of full column rank, which suggests that the rows of the blending matrix are linearly independent. If the blending matrix isn't of full rank then the mixed signals are going to be linear multiples of 1 another.
- The propagation delay of the mixing medium is negligible.

3 ILLUSTRATION OF ICA WITH PROBABILITY DENSITY FUNCTIONS

Let us consider a source matrix consisting of two statistically independent and uniform random variables, \( S = [s_1, s_2]^T \) (Figure3.1). It is clear that \( s_1 \) and \( s_2 \) are statistically independent. Now let’s mix these independent components with any real valued mixing matrix \( H \), Figure 3.2 shows the
resulting density function, we see that mixing the independent components somewhat skews the density function.

This is analogous to our matrix operation:
\[ Y = WX \]

For our work with EEG data the logistic sigmoid function \[ g(x) = \frac{1}{1 + e^{-u}} \]

Where \( w = \) slope of \( y \) (also called the weight)
\( w_0 = \) bias weight to align the high density parts of the input with \( y \).
The pdf of the output \( f_y(y) \) can be written as a function of the pdf of the input \( f_x(x) \) as:

\[ H(y) = \int_{-\infty}^{\infty} f_y(y) \ln f_y(y) dy \]

We now would like to maximize \( H(y) \) of equation 4.5 for statistical independence. The change in slope, \( \Delta w \), necessary for maximum change in entropy is then:

\[ \frac{\partial H}{\partial w} = \int_{-\infty}^{\infty} \left( \frac{f_y(y)}{f_x(x)} \right) \left( f_y(y) \right) \frac{\partial f_y(y)}{\partial w} dy \]

**Figure 10:** K-means Clustered PWC

**Figure 11:** Fuzzy distance of a person walking at angle 36\(^\circ\) from PWCP

**Figure 12:** Normalized Fuzzy Membership Vector
5 CONCLUSION

In this work we extracted 1D signals from the movement of hands, shoulders and legs of all the gait sequences with all covariate conditions of multi view gait database. Then Partial Wavelet Coherent Poses of those sequences having similar coherence extracted by K-means clustering. The fuzzy distance between partial wavelet coherence of every sequence and clustered partial-wavelet coherent poses preserves discriminant information of the walking subject. We got 97.5% mean identification rate. Table 1 shows comparison of proposed method with earlier methods. Even though we got encouraging results the self-occlusion couldn't identified and removed. This might flow from to the binary nature of image. In future we'll attempt to resolve the matter of self-occlusion.

7 REFERENCES


