



BLIND SEPARATION OF BLINKS FROM EEG DATA: EVALUATION OF INFOMAX AND FASTICA ALGORITHMS

R. Frank¹ G. Frishkoff² C. Davey^{2,3} K. Glass¹ D. Tucker³
¹ Neuroinformatics Center ² Department of Psychology ³ Electrical Geodesics, Inc.

INTRODUCTION

ICA (Independent Components Analysis) is a mathematical procedure that can be used to identify and remove blinks (and other *artifacts*) from EEG data.

ICA decomposes the data into statistically independent temporal streams (the *independent components*) using methods of *blind separation*. The independent components are then matched with a *blink template* — a spatial vector that is defined a priori to characterize the topography of a typical blink. Activations that closely match the template are then zeroed out, or removed from the data. This process is referred to as *data cleaning*.

ERP DATA ACQUISITION & PREPROCESSING

EEG Acquisition:

- 256 scalp sites; vertex recording reference (Geodesic Sensor Net).
- .01 Hz to 100 Hz analogue filter; 250 samples/sec.

EEG Preprocessing:

- All trials with artifacts detected & eliminated.
- Digital 30 Hz bandpass filter applied offline.
- Data subsampled to 34 channels & ~50,000 samples

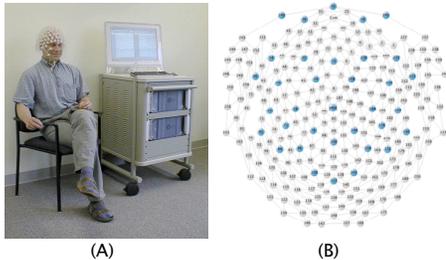


Figure 1. (A) Net Station dense-EEG system; (B) Layout for 256-channel sensor array (nose and eyes at top)

ANATOMY OF A BLINK

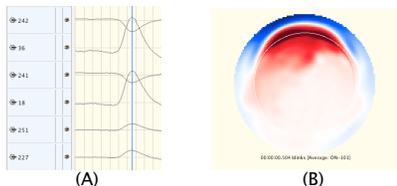


Figure 2. (A) Timecourse of a blink (1sec); (B) Topography of an average blink (red = positive; blue = negative)

SYNTHESIZED DATA

Creation of Blink Template

- ~15-20 blinks manually marked in the raw EEG.
- Data segmented into 1sec epochs, timelocked to peak of blink.
- Blink segments averaged to create a blink template.

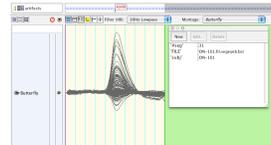


Figure 3. Butterfly plot for channels overlaid (~1sec)

Creation of Synthesized Data

- A: "clean" data (32ch, ~50k time samples)
- B: "blink" data (created from template)
- C: The derived "blink" data were added to the clean data to create a synthesized dataset, consisting of 34 channels x 50,000 time samples

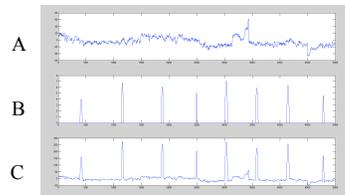


Figure 4. Input to ICA: Synthesized data, consisting of cleaned EEG plus artificial "blinks" created from blink template (Fig. 2).

ICA ALGORITHMS

The *Infomax* and *FastICA* algorithms provide two different approaches to the same problem. They are mathematically equivalent, but use different procedures for obtaining linearly independent sources. The goal of the present comparison was to verify that the two algorithms give equivalent results when applied to EEG data to remove blinks.

Infomax Procedure

Infomax uses an information-theoretic approach, which minimizes mutual information. This is equivalent to separating the data channels into linearly independent components.

FastICA Procedure

FastICA achieves linear independence through a change in the basis set. This procedure finds the basis for which the coordinates of the observations have a probability distribution that is maximally non-Gaussian (using 4th order moments).

ICA RESULTS

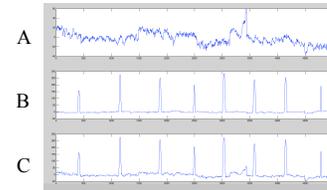


Figure 5. Results from Infomax. (A) "Cleaned" data. (B) Blink activations extracted using Infomax. (C) Synthesized data.

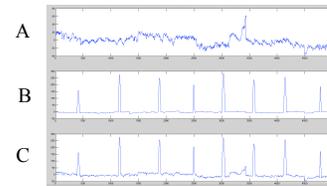


Figure 6. Results from FastICA. (A) "Cleaned" data. (B) Blink activations extracted using FastICA. (C) Synthesized data.

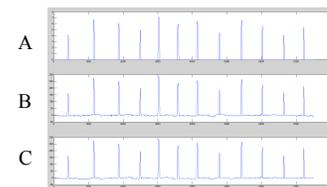


Figure 7. (A) Original "blink-free" data. (B) Blink activations extracted using Infomax. (C) Blink activations extracted using FastICA.

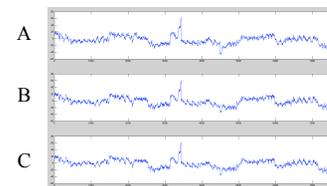


Figure 8. Comparison of algorithms. (A) Original blink-free data. (B) Infomax cleaned data. (C) FastICA cleaned data.

CONCLUSIONS

- ❑ The FastICA algorithm is mathematically "simpler" than the Infomax procedure.
- ❑ The Infomax and FastICA algorithms have been shown to be theoretically equivalent.
- ❑ The present tests verify that the two ICA algorithms give qualitatively and quantitatively similar results.
 - ICA succeeded in removing blinks that had been added to the original "blink-free" data & successfully regenerated the blink template
 - ICA preserved the original features of the data

FUTURE DIRECTIONS

Future experiments with ICA will involve multiple simulations, designed to test the effects of various computational and data-specific variables:

- ❑ Alternative methods for generating blink templates
 - Optimizing creation of averaged blink template
 - Using Principal Components Analysis (PCA)
- ❑ Use of prior spatial and temporal constraints
 - SOBI (Second-Order Blind Independence)
 - Seeding FastICA with the blink template
- ❑ Comparing results using a range of blink thresholds
 - Determining number of ICA activations related to blink activity (there may be more than one)
- ❑ Examining efficacy of ICA for removal of components with different frequency characteristics
 - Use of sub-Gaussian distributions for removal of slow-wave activity due to blink recovery
 - Use of super-Gaussian distributions for removal of fast components, such as EMG (muscle artifact)

References

- James, C. J. and O. J. Gibson (2003). "Temporally Constrained ICA: An Application to Artifact Rejection in Electromagnetic Brain Signal Analysis." *IEEE Transactions on Biomedical Engineering* **50**(9): 1108-1116.
- Bell, A. J. and T. J. Sejnowski (1995). "An information-maximization approach to blind separation and blind deconvolution." *Neural Computation* **7**(6): 1004-1034.
- Hyvarinen, A. (1999). "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis." *IEEE Transactions on Neural Networks* **10**(3): 636-634.