A Framework for Evaluating ICA Methods of Artifact Removal from Multichannel EEG

Kevin A. Glass[1], Gwen A. Frishkoff[2], Robert M. Frank[1], Colin Davey[3], Joseph Dien[4], Allen D. Malony[1], and Don M. Tucker[1,2],

[1] NeuroInformatics Center, 5219 University of Oregon, Eugene, OR 97403
{kglass,rmf,malony}@cs.uoregon.edu

[2] Department of Psychology, University of Oregon, Eugene, OR 97403

[3] Electrical Geodesics Inc.,1600 Millrace Dr. Suite 307, Eugene, OR 97403

[4] Department of Psychology, University of Kansas, Lawrence, KS 66045

Abstract. We present a method for evaluating ICA separation of artifacts from EEG (electroencephalographic) data. Two algorithms, Infomax and FastICA, were applied to "synthetic data," created by superimposing simulated blinks on a blink-free EEG. To examine sensitivity to different data characteristics, multiple datasets were constructed by varying properties of the simulated blinks. ICA was used to decompose the data, and each source was cross-correlated with a blink template. Different thresholds for correlation were used to assess stability of the algorithms. When a match between the blink-template and a component was obtained, the contribution of the source was subtracted from the EEG. Since the original data were known a priori to be blink-free, it was possible to compute the correlation between these "baseline" data and the results of different decompositions. By averaging the filtered data, time-locked to the simulated blinks, we illustrate effects of different outcomes for EEG waveform and topographic analysis.

1 Introduction

Accurate assessment of signal decomposition methods such as ICA should account for multiple parameters that affect the decomposition, including characteristics of the input data (properties of the signal and noise activity) and properties of different ICA algorithms and implementations (e.g., contrast functions, tolerance levels). The theoretical underpinning of ICA and its various algorithms have been extensively discussed in the literature [1,2,3], and experiments have been designed to demonstrate the effectiveness of the procedure (for example, see [4]). However, there are few empirical studies measuring the effectiveness of ICA algorithms, and even fewer discussing these measures in the context of specific applications. One reason for the lack of empirical studies is the lack of empirical measures of effectiveness [5].

To this end, the present paper describes a new method for evaluation of ICA decompositions and applies this method to the problem of artifact extraction from multichannel EEG (electroencephalographic) data. The goal of this application was to compare the efficacy of two ICA algorithms, FastICA [6] and Infomax [3], in removing blinks from EEG signals. However, the procedure can be generalized to other

C.G. Puntonet and A. Prieto (Eds.): ICA 2004, LNCS 3195, pp. 1033–1040, 2004.

problems and algorithms. Our technique, described below, is similar to Harmeling, et al. [5] and Zibulevski and Zeevi [7] except that our approach uses realistic data, thus giving the user a familiar basis for qualitative comparisons.

The results of our tests demonstrate the quantitative and qualitative utility of measures in evaluating ICA decomposition. With this method, it is possible to characterize the sensitivity of different ICA methods to multiple variables and perhaps, in future applications, to determine the appropriateness of different ICA methods for particular data analysis goals. Further, in addition to quantitative measures, we evaluated the effects of different ICA results on EEG waveforms and topographies. This allowed us to visualize the results and to examine the practical implications of different statistical outcomes.

2 Methods

EEG Acquisition and Preprocessing. EEG data were acquired from 256 scalp electrodes EEG net (Electrical Geodesics, Inc) referenced to Cz in a language task described elsewhere. Data contaminated by blinks were manually marked and removed, providing a blink-free EEG ("baseline") for evaluating the success of the blink removal. The EEG was downsampled to 34 channels, making it feasible to examine spatial and temporal properties of all 34 extracted components.

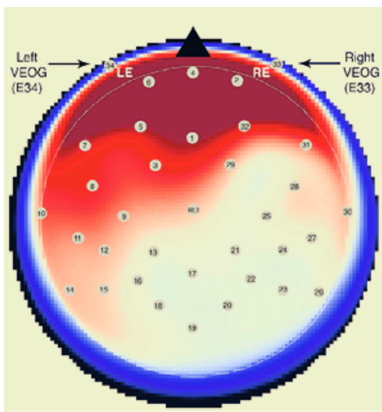

Fig. 1. Blink topography. Red, positive. Blue, negative. LE = left eye. RE = right eye.

Creating the Blink Template. Thirty-two segments of data with representative blinks were segmented from the continuous data. The segments were aligned to the peak of each blink and averaged to derive a blink template (Fig. 1).

Construction of Synthesized Datasets. To construct the synthesized data, the raw EEG data were inspected for ocular artifacts, and all trials contaminated with blink activity were removed from the recording, resulting in a "blink-free" EEG, to which a stream of blinks with known spatial and temporal characteristics was added (Fig 2).

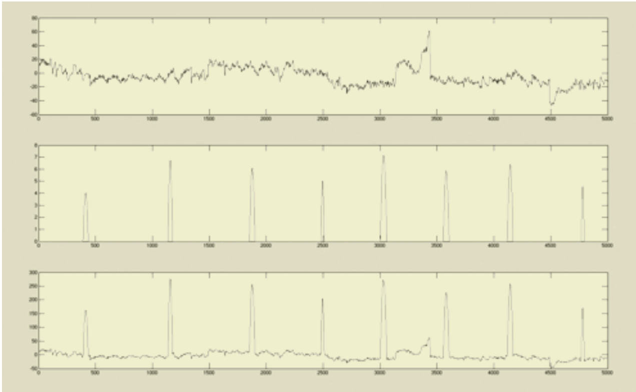

Fig. 2. Construction of "synthetic" data. Top panel, original data (~10 sec). Center panel, simulated blinks (Dataset #7). Bottom panel, original data plus simulated blinks.

To assess the robustness of the two algorithms and their sensitivity to data parameters, seven such datasets were constructed. The datasets differed with respect to blink amplitude, blink duration, and inter-blink interval. Datasets 1-5 contained blink activations of constant duration with inter-blink spacing of 400 milliseconds and 5000 milliseconds, respectively. Intensity of the blink activations ranged from 25% (Set #1) to 400% (Set #5) the intensity of the largest non-blink activity. Datasets 6 and 7 contained blinks of variable duration, spacing and intensity (Table 1).

Table 1. Test data set characteristics.

Data Set	Blink Strength	Inter-blink Spacing (ms)	Blink Duration (ms)
1	25%	5000	400
2	50%	5000	400
3	100%	5000	400
4	200%	5000	400
5	400%	5000	400
6	50%-200%	635-2500	312-5000
7	255-400%	312-5000	25-400

ICA Algorithms and Blink Removal Procedures. Both ICA algorithms were implemented in Matlab. The Infomax code [8] is an enhanced version of the Infomax algorithm of Bell and Sejnowski [2]; the FastICA code [9] uses a fixed-point algorithm. To remove blinks, we used a modified version of the ICABlinkToolbox [10,11].

The FastICA decomposition was performed using two contrast functions, the cubic (default) contrast function and a hyperbolic tangent (tanh) function. In the initial tests, the tanh function outperformed the cubic function. Therefore, in subsequent analyses, we used the tanh contrast function only. The Infomax decomposition used the developer's default settings. The projections of the components onto the EEG detector array ("spatial correlates" for short) were correlated with the blink template. Then contribution of the highest correlated component was removed from the dataset and the cleaned and original datasets were compared to measure the quality of the ICA algorithm's decomposition.

Metrics. The covariance between corresponding channels of the ICA-filtered EEG data and the original EEG data was computed for each dataset. To provide qualitative metrics for comparison of the different algorithms, we averaged the original and ICA-filtered data, time-locking the averages to the peak of the simulated blinks. The resulting averages should therefore accentuate residual blink activity after data cleaning. This procedure provides a visual reference for the significance of the correlation values.

3 Results

The overall (grand average) correlation between the original and cleaned data, for both ICA algorithms was 0.95 or better for FastICA and 0.969 for Infomax. When broken down for the separate electrodes, the lowest correlations occurred for channels 2, 4, and 6: depending on the particular dataset, and the threshold for blink identification, correlations at these channels ranged from about 0.55 to about 0.70. This is not unexpected, since these channels are located just above the eyes (Fig. 1).

A more detailed comparison of the results for FastICA and Infomax revealed several important differences. The most salient difference is that Infomax decompositions varied little across the datasets, whereas the FastICA decompositions showed considerable variation (Fig. 3). This suggests that changes in the properties of the blink data may affect factor extraction, allocation of variance across the factors, or both. As mentioned previously, FastICA implemented with the default (cubic) contrast function fared considerably worse than the implementation with the tanh contrast function. Therefore, subsequent analyses focused on the comparison of Infomax and FastICA using the tanh contrast function. Figure 4 demonstrates that the periorbital channels show the worst correlations. In addition, the largest differences between Infomax and FastICA are observed over these same channels, where blink activity is most pronounced.

Infomax was similarly robust to changes in tolerance (threshold for correlation with blink template), whereas FastICA on average showed worse accuracy at lower tolerances (data not shown here). In general, Infomax was more stable and more robust to changes in properties of the data and ICA implementation.

Further inspection of the ICA decompositions revealed that where FastICA was less successful, more than one spatial projector correlated strongly with the blink template, a strong correlation being any correlation above the experimentally determined threshold of 0.90. For example, as illustrated in Figure 5 above, FastICA-1, one of the least successful decompositions performed for this report, contained 6 projectors that matched the blink template > 0.90 as compared to InfoMax and FastICA-1, which each contained only one.

To illustrate the effects of successful and less successful ICA decompositions, we examined the ICA-cleaned data for different FastICA and Infomax runs (Dataset 5) after removing the source that was perfectly correlated with the blink template. Because FastICA gave more variable results across runs, we selected one example of a successful FastICA run (run 2) and one example of a less successful run (run 1). Although "the same" source was removed from the data in each case, the effects were very different, reflecting misallocation of variance when additional sources showed a close (but less than perfect) match to the blink template, as illustrated in Figure 6.

The failure of FastICA (run 1) that is evident in the averaged waveforms is also visible in the topographic distribution of the filtered data (Fig. 7). Note the resemblance of the topography for FastICA (run 1) to the blink template (Fig. 1). This outcome appears to reflect misallocation of variance to additional components in the decomposition [5].

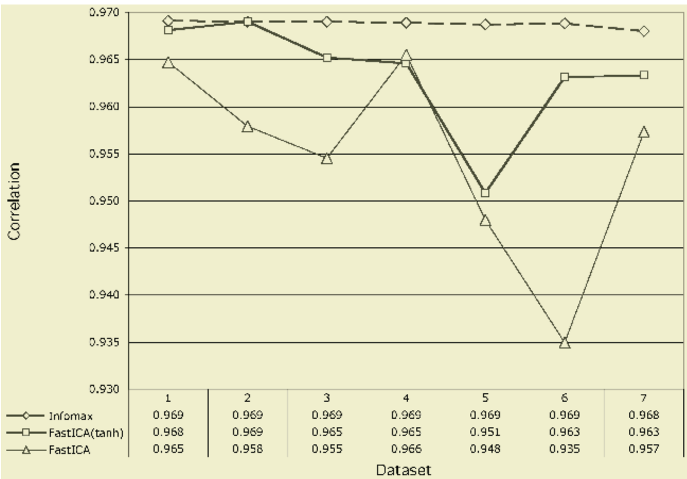

Fig. 3. Graph of correlations between original and ICA-filtered data across the seven datasets. Thin line, FastICA with cubic constrast function. Thick line, FastICA with tanh contrast function. Dotted line, Infomax.

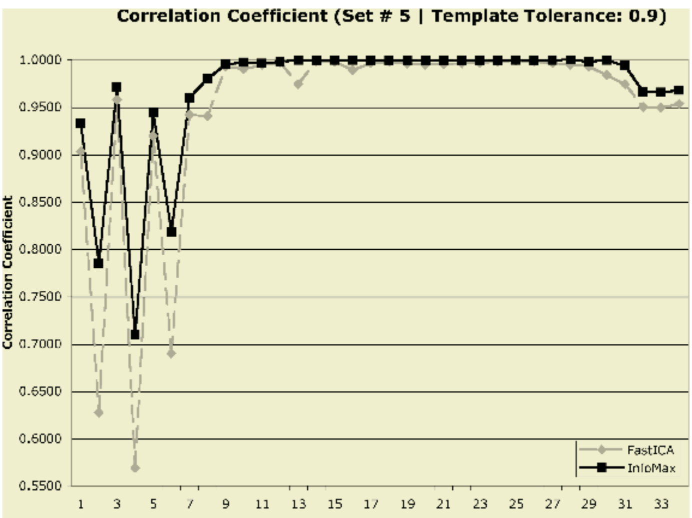

Fig. 4. Correlation between original & ICA-filtered data across the 34 electrodes.

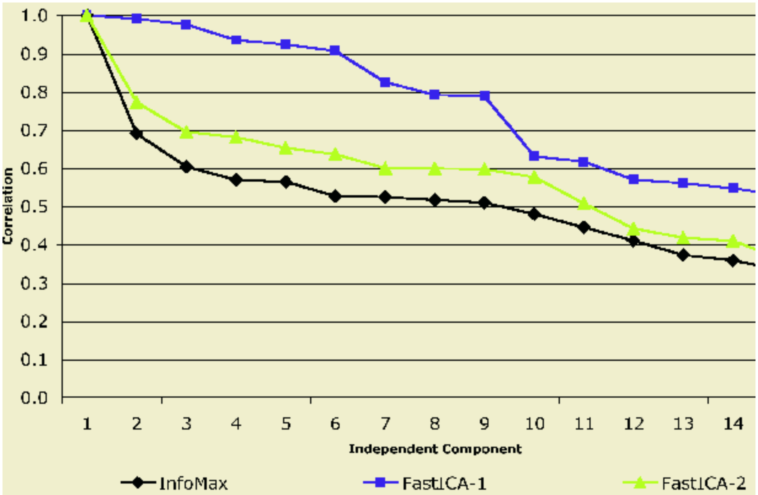

Fig. 5. Correlation between the spatial projectors of the independent component activations and the synthetic blink activitity template. The figure shows the 14 components with the strongest correlations.

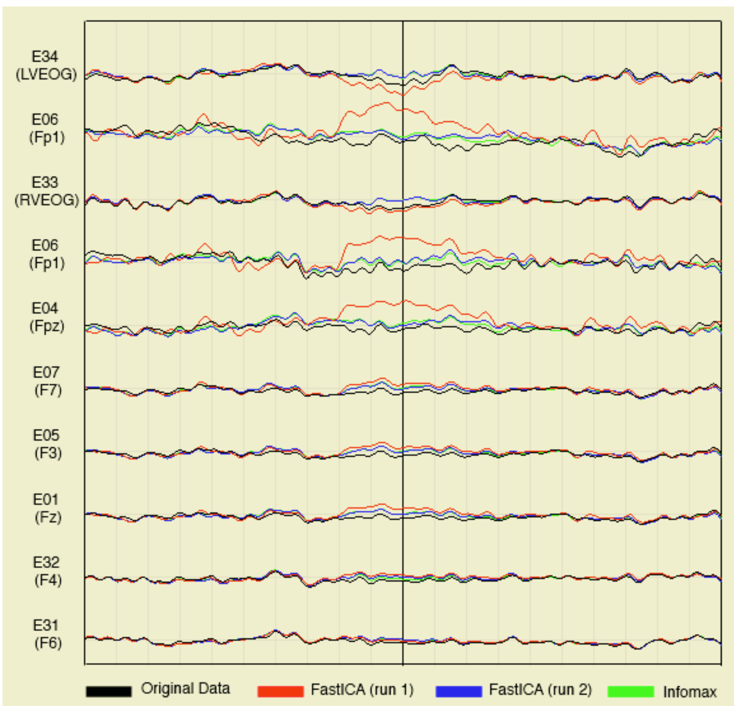

Fig. 6. EEG waveforms, averaged to the peak of the blink activity. Note residual blinks in run 1 for FastICA, where more than one source was strongly correlated with the blink template, and the source activations revealed misallocation of variance (cf. Fig 5).

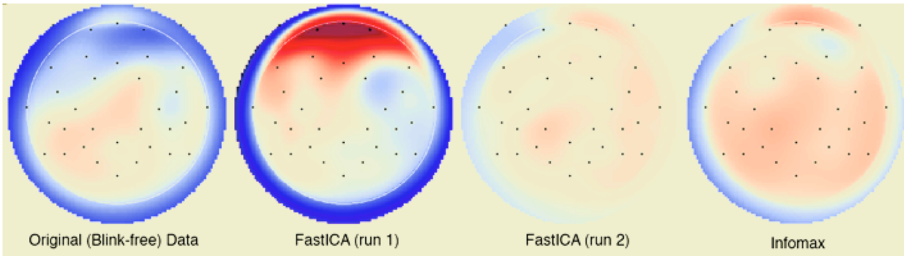

Fig. 7. Topography of blink-averaged data, centered at peak of blink activity. Red, positive voltage. Blue, negative voltage. FastICA run1 is the less successful decomposition. Note the remaining blink activity at this time point.

4 Discussion

In this report we have demonstrated a new method for evaluation of ICA for removal of blink activity from multichannel EEG. The grand average correlation suggest that Infomax and FastICA were highly accurate in their ability to separate out the simulated blinks from the EEG. In every ICA run, exactly one of the extracted components showed a perfect correlation with the blink topography used to construct the simulated blinks. On the other hand, the activations corresponding to this source differed across runs and across ICA algorithms and implementations. In every case, the source activations were less than perfectly correlated with the time series for the simulated blinks. Infomax showed the closest correspondence, while FastICA was more variable, showing excellent correspondence on some runs, and misallocation of variance on other runs. Future studies will examine causes of misallocation of variance, extend this method to account for other data parameters, and compare results for Infomax and FastICA with other ICA algorithms and implementations.

References

1. Cardoso, J.-F. and P. Comon (1996). "Independent Component Analysis, A Survey of Some Algebraic Methods." Proc. ISCAS'96, vol.2, pp. 93-96.
2. Hyvarinen, A. and E. Oja. (1999). "Independent Component Analysis: A Tutorial." Downloaded from http://www.cs.helsinki.fi/u/ahyvarin/papers/index.html, March 12, 2004.
3. Bell, A. J. and T. J. Sejnowski (1995). An Information Maximization Approach to Blind Separation and Blind Deconvolution. San Diego, Institute for Neural Computation, UCSD, San Diego CA: 1-38.
4. Aichner, R., H. Buchner, S. Araki, S. Makino (2003). On-line Time-domain Blind Source Separation of Nonstationary Convolved Signals." 4rd Intl. Conf. on Indep. Comp. Ana. and Blind Sig. Sep., Nara, Japan.
5. Harmeling, S., F. Mieinecke, et al. (2003). "Analysing ICA Components by Injecting Noise." 4rd Intl. Conf. on Indep. Comp. Ana. and Blind Sig. Sep., Nara, Japan.
6. Hyvarinen, A. (1999). "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis." IEEE Trans. on Neural Nets. 10(3): 626-634.
7. Zibulevski, M. and Y. Y. Zeevi (2001). Source Extraction Using Sparse Representation. 3rd Intl. Conf. on Indep. Comp. Ana. and Blind Sig. Sep., San Diego, CA.

8. EEGLab download site. http://www.sccn.ucsd.edu/eeglab/downloadtoolbox.html. March 12, 2004.
9. FastICA download site. http://www.cis.hut.fi/projects/ica/fastica/. March 12, 2004.
10. ICAToolBox download site. http://people.ku.edu/~jdien/. March 12, 2004.
11. Dien, J. (1998). Issues in the application of the average reference: Review, critiques, and recommendations. Behav. Res. Methods, Instruments, and Computers, 30(1), 34-43.