7th Annual International IEEE EMBS Conference on Neural Engineering
Montpellier, France, 22 - 24 April, 2015

Validating Online Recursive Independent Component Analysis on EEG Data

Sheng-Hsiou Hsu, *Student Member, IEEE,* Tim Mullen, *Member, IEEE,*
Tzyy-Ping Jung, *Fellow, IEEE,* and Gert Cauwenberghs, *Fellow, IEEE*

***Abstract*— The needs for online Independent Component Analysis (ICA) algorithms arise in a range of fields such as continuous clinical assessment and brain-computer interface (BCI). Among the online ICA methods, online recursive ICA algorithm (ORICA) has attractive properties of fast convergence and low computational complexity. However, there hasn't been a systematic comparison between an online ICA method such as ORICA and other offline (batch-mode) ICA algorithms on real EEG data. This study compared ORICA with ten ICA algorithms in terms of their decomposition quality, validity of source characteristics, and computational complexity on the thirteen experimental 71-ch EEG datasets. Empirical results showed that ORICA achieved higher mutual information reduction (MIR) and extracted more near-dipolar sources than algorithms such as FastICA, JADE, and SOBI did while the performance of ORICA approached that of the best-performed Infomax-based algorithms. Furthermore, ORICA outperforms most of ICA methods in terms of the computational complexity. The properties of fast convergence and low computational complexity of ORICA enable the realization of real-time online ICA process, which has further applications such as real-time functional neuroimaging, artifact reduction, and adaptive BCI.**

I. INTRODUCTION

Independent Component Analysis (ICA) has wide applications in biomedical signal processing such as on electroencephalography (EEG) and electrocorticography (ECoG) data. ICA methods have been used to effectively separate artifacts such as eye blinks and muscle activity [1] and to study brain activities [2]. Applying ICA to EEG data is based on a reasonable assumption that multi-channel scalp EEG signals are mixtures of temporally independent non-Gaussian sources [3].

Many ICA algorithms have been developed to separate the sources from the channel signals [4], with the assumption of spatiotemporal stationarity of the data, such as widely used Infomax ICA [5] and FastICA [6]. For non-stationary EEG signals, ICA methods such as Adaptive Mixture ICA (AMICA) [7] can generate good component decomposition but are computationally expensive. However, the aforementioned ICA algorithms are all offline analyses. In many real-world applications, including real-time functional neuroimaging [8], artifact rejection, and adaptive brain-computer interface (BCI), online (sequential) source separation methods are needed. Desirable properties in these applications are fast convergence and low computational complexity.

Among the online ICA methods, online recursive ICA (ORICA) [9] has such attractive properties of fast convergence and low computational complexity because it employs a recursive-least-squares (RLS) type method derived from a fixed point solution to Infomax ICA. Subsequently, a computationally real-time and online ICA pipeline has been proposed [10] which incorporates the whitening step prior to the application of ORICA to improve the convergence speed and employs a block-update rule to accelerate the computational speed. It has been shown that the ORICA method can decompose 64-ch simulated EEG data and extract informative components from the 61-ch experimental EEG data. However, some questions remained unanswered: how consistent are the ORICA components across experimental EEG datasets and how is its performance compared to other ICA methods.

This study tested the ORICA algorithm on thirteen experimental 71-ch EEG datasets and compared the results to those obtained by ten popular batch-mode ICA or BSS algorithms, following the ICA comparison approach published by [3]. Both quantitative and qualitative performance of ORICA are given in terms of decomposition quality, validity of source characteristics, and computational complexity.

II. METHODS

With the standard ICA model $x = As$, complete decomposition is performed, i.e. estimate unmixing matrix W such that all sources are recovered by $y = Wx$, where x are scalp EEG observations, A is an unknown mixing matrix, and s and y are original and reconstructed source activation time courses respectively.

A. Online Recursive ICA (ORICA)

ORICA is an online ICA algorithm with a recursive update rule derived from a fixed-point solution to the natural-gradient Infomax ICA, proposed by [9]. The final learning rule is extended to a block-update form [10], with a block size L, to reduce the computational load:

$$W_{n+L} \approx \left(\prod_{l=n}^{n+L-1} \frac{1}{1-\lambda_l}\right) \cdot \left[I - \sum_{l=n}^{n+L-1} \frac{y_l \cdot f^T(y_l)}{\frac{1-\lambda_l}{\lambda_l} + f^T(y_l) \cdot y_l}\right] W_n \quad (1)$$

This work was in part by a gift by the Swartz Foundation (Old Field, NY), by the Army Research Laboratory under Cooperative Agreement Number W911NF-10-2-0022, by NIH grant 1R01MH084819-03 and NSF EFRI-M3C 1137279.

S.-H. Hsu (shh078@ucsd.edu) is with Dept. of Bioengineering (BIOE), Swartz Center for Computational Neuroscience (SCCN), and Institute for Neural Computation (INC) of University of California, San Diego (UCSD).

T. Mullen (tmullen@ucsd.edu) is with Dept. of Cognitive Science, SCCN and INC of UCSD.

T.-P. Jung (tpjung@ucsd.edu) is with BIOE, SCCN and INC of UCSD.

G. Cauwenberghs (gert@ucsd.edu) is with BIOE and INC of UCSD.

Followed the optimized parameters used in [10], we chose the block size $L = 8$, the heuristic time-varying forgetting factor $\lambda_n = 0.995/n^{0.6}$, and the nonlinear projection function $f(y) = -2\tanh(y)$ with the assumption that all EEG sources are supergaussian. The assumption is verified on the datasets and is generally hold for filtered, e.g. line-noise removed, EEG data.

A whitening step is added prior to ORICA to decorrelate the data and improve the ICA convergence. Pre-whitening provides a fair comparison between ORICA and other batch-mode ICA algorithms where data are also pre-whitened. ORICA with pre-whitening (denoted as *ORICA*) is highlighted in pink in Section IV). ORICA with two to three passes over the data (denoted as *ORICA-2 or ORICA-3*) are also tested to see the improvement of performance as more data points are available.

B. Other ICA or BSS Algorithms

Ten ICA or BSS algorithms from three groups plus principle component analysis (PCA) are used for comparisons. The first group is the natural gradient ICA algorithms, including AMICA [7], Infomax [5], and Extended Infomax [11], highlighted in yellow in Section IV. This group is known for its ability to extract dipolar sources from EEG data while it is computational expensive. The second group includes FastICA [6] that maximizes negentropy, JADE [12] that maximizes the fourth-order cumulants, and ThinICA (TICA) [13], highlighted in green. The third group belongs to the second-order time-delay approach [14] to the BSS problem including SOBI, icaMS, FOBI, and EVD, highlighted in blue. PCA is also tested to characterize the effect of pre-whitening, highlighted in purple. The detailed selection of parameters and references of the above algorithms are described and listed in [3].

C. Evaluation Methods

1) Mutual Information Reduction (MIR): ICA aims at maximizing the independence and mutual information reduction (MIR) provides an estimation of independence in terms of mutual information between sources, defined as

$$\begin{aligned} MIR = I(x) - I(y) = \log|\det W| + \\ [h(x_1) + \cdots + h(x_N)] - [h(y_1) + \cdots + h(y_N)] \end{aligned} \tag{2}$$

where I is mutual information and h is marginal entropy. Better ICA decomposition will have more independent sources and higher MIR.

2) Equivalent Dipole Modeling: It has been demonstrated that independent EEG sources are dipolar [3]. Hence to quantify the quality of the decomposed individual sources, the surface spatial distribution of each IC is fitted into a single equivalent dipole in a spherical four-layer head model using the DIPFIT plug-in in the EEGLAB toolbox [15]. The error of the dipole fitting is the residual variance (r.v.). Near-dipolar sources are defined as r.v. $< 5\%$. More near-dipolar components correlates with better ICA decomposition for EEG data.

III. Materials

A. EEG Data

The EEG datasets used in this study are available at http://sccn.ucsd.edu/wiki/BSSComparison, from thirteen subjects performing the visual working memory task described in [16]. Each dataset has EEG signals from 71 channels with sampling rate of 250Hz and data length of 296,000 to 315,000 samples (around 20 minutes of recording). The data have been high-pass filtered with cutoff frequency at 0.5Hz to remove trend, and the epochs with high-amplitude or high-frequency abnormalities have been removed [3].

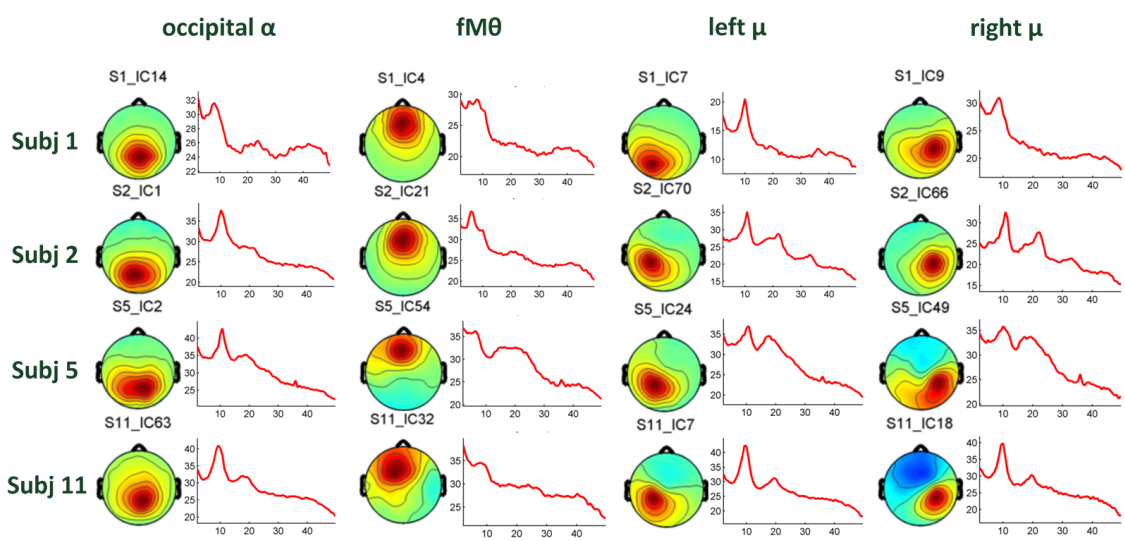

Fig. 1. Scalp maps and mean activity spectra of four ORICA component clusters (central occipital alpha, frontal midline theta, left mu and right mu) accounting for non-artifact brain sources from four subjects. Components in each cluster are selected by visual inspection according to their scalp maps and power spectra.

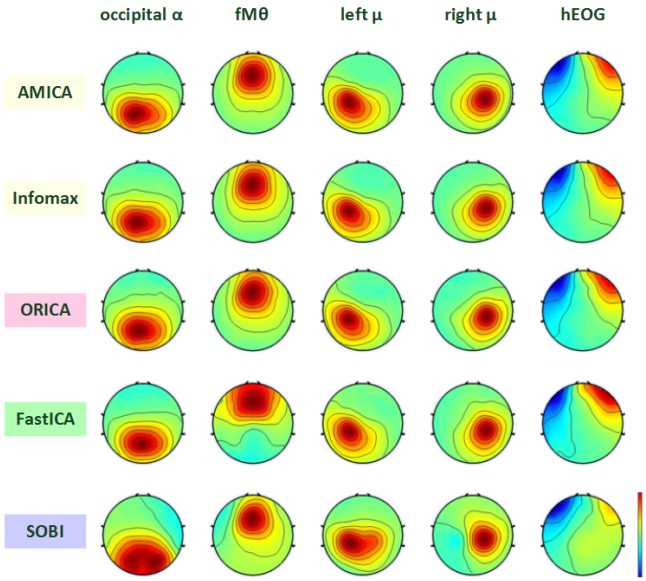

Fig. 2. Scalp maps of the five informative components - the occipital alpha, the frontal midline theta, the left mu, the right mu, and the horizontal eye movement (EOG) - decomposed by AMICA, Infomax, ORICA, FastICA, and SOBI from the dataset S2. The similar components are found by best-matching ICs with the results of ORICA.

B. Data Analysis

Data were loaded into MATLAB and analyzed by the open source EEGLAB toolbox [15]. We ran each algorithm on each dataset, recorded the runtime, solved for unmixing matrix, and evaluated the decomposition performance with the methods described in Sec. II-C. The MATLAB scripts for running the comparisons can be found at http://sccn.ucsd.edu/wiki/BSSComparison. The MATLAB functions of the algorithms in Sec. II-B can be downloaded from EEGLAB [15] (sccn.ucsd.edu/eeglab) and ICALAB [17] (www.bsp.brain.riken.jp/ICALAB).

IV. Results

The informative independent components (ICs) decomposed by ORICA from each dataset are manually selected based on their scalp maps and mean activity spectra and grouped into the following four non-artifact brain source clusters: central occipital alpha (8-12Hz), frontal midline theta (near 6Hz), left and right motor mu (near 10Hz with harmonics) clusters.

Each ORICA cluster contains similar ICs from 5-8 subjects, demonstrating the consistency of ORICA decomposition across the datasets. The results from the four selected subjects are shown in fig. 1. Due to variability across subjects, not every dataset returns the same informative ICs. However, the scalp maps and the source dynamics of the ORICA components are consistent with those from the previous study [3] using AMICA, which has shown to be the ICA algorithm with the best performance.

Fig. 2 gives the qualitative comparisons across algorithms by looking at the five informative ICs from ORICA on the datset S2 and their counterparts decomposed by AMICA, Infomax, FastICA, and SOBI algorithms from the same dataset. Similar components were found by best-matching the ICs, in terms of the correlation of the scalp maps, of ORICA and the other algorithms using the Hungarian method.

ORICA components strongly resemble the ICs from the best-performed ICA methods such as Infomax and AMICA. The FastICA returned the similar components but some ICs such as fMθ and hEOG are less homogeneous. Spatial distribution of the SOBI components are even worse. Results of the PCA components are not shown because PCA failed to decompose similar components. The components from AMICA, Infomax, and ORICA have lower r.v. values than those from FastICA, and SOBI.

Fig. 3 provides the quantitative comparisons of the five groups of the fourteen algorithms, assessed by mutual information reduction (MIR) and percentage of near-dipolar components with r.v. $< 5\%$ averaging across the thirteen datasets. Along the linear regressed curve, algorithms at top right have better decomposition performance and more near-dipolar sources. ORICA (pink) has worse performance than the standard Infomax-based algorithms (yellow), whereas it outperforms the other groups of algorithms such as FastICA (green), SOBI (blue), and PCA (purple).

It should be emphasized that even the ORICA with single-pass performs better than several batch-mode algorithms. With more passes through the dataset, the performance of ORICA-n improves and approaches the performance of the leading Infomax group, which has a few hundreds passes. This suggests ORICA has fast convergence characteristic.

Fig. 4 compares the computation time, i.e. complexity, and shows the number of near-dipolar sources averaged across the thirteen datasets for all algorithms. ORICA runs four times faster than Infomax does and is able to decompose 80% of Infomax's near-dipolar sources. Furthermore, ORICA outperforms FastICA (green) in both computation time (three times faster) and performance (13% more near-dipolar sources). SOBI and PCA run faster than ORICA but have poor decomposition quality. The result of AMICA is not shown because it requires parallel processing and runs for hours.

Algorithms that decompose more near-dipolar components require more computation time. In the performance-complexity relations, ORICA stands out with its low computational complexity while the performance remains at high level. It is worth noting that ORICA can on average decompose a 71-ch dataset with 300,000 samples in about 40 seconds on a laptop without the feature of parallel processing. The even more powerful feature of ORICA is that the computation time can be further reduced by increasing the block-update size [10].

V. Conclusions

This study provides both qualitative and quantitative evaluation of the performance of the ORICA method. Qualitatively, the results confirm that ORICA can extract the informative ICs (brain and artifact sources) from the 71-ch

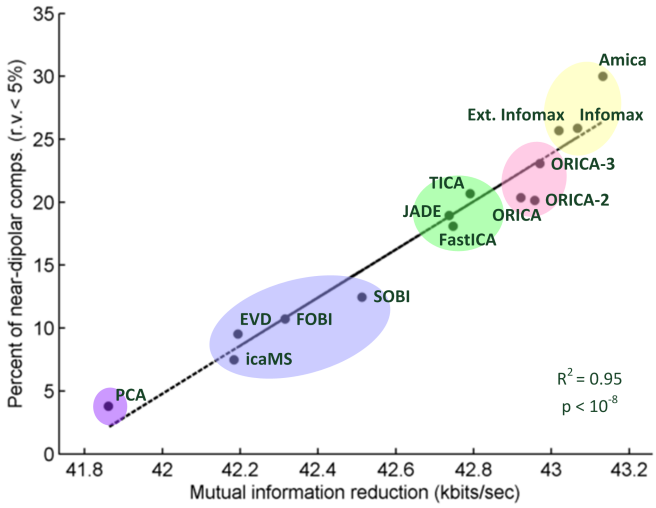

Fig. 3. Percentage of the components (out of 71 ICs) with near-dipolar scalp maps (residual variance $<= 5\%$) and mean mutual information reduction (MIR) averaged across thirteen datasets for the five groups of the fourteen algorithms described in the Methods Section. The color of each group is manually added. Results of ORICA with n-passes over the data (ORICA-n) are shown.

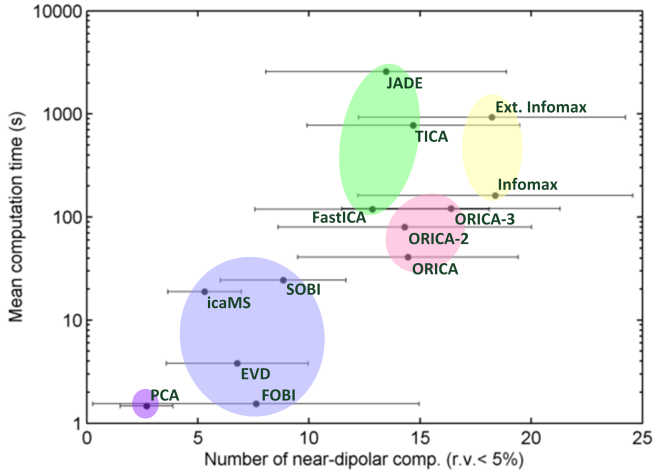

Fig. 4. The log-scale computation time and the number of near-dipolar scalp maps (r.v. $<= 5\%$) of the five groups of the fourteen algorithms are shown. The dots and bars show the mean and standard deviation of the number of near-dipolar sources averaged across the thirteen datasets. The standard deviation for the computation time is too small to plot. Algorithms at lower right are better. The datasets are processed in MATLAB 2012a on a dual core 2.5GHz Intel i5-3210M CPU with 8GB RAM.

experimental EEG data. The decomposed ICs are consistent across the datasets from different subjects and also agree with the previous study using AMICA. Quantitatively, in terms of decomposition performance (MIR) and validity of source characteristics (number of near-dipolar sources), ORICA's performance approaches the best-performed Infomax-based method as number of passes over data increases and is uniformly better than other algorithms. Furthermore, ORICA outperforms most of ICA methods in terms of computational complexity, except the algorithms with poor decomposition such as SOBI and PCA.

The results of this study indicate ORICA can achieve good decomposition with low computational complexity, which can be an alternative for ICA decomposition of high-density EEG data, compared to the standard ICA methods such as Infomax. These attractive properties of ORICA enable the realization of real-time online ICA process, which has further applications such as real-time functional neuroimaging, artifact rejection, and adaptive BCI.

ACKNOWLEDGMENT

We wish to thank Arnaud Delorme, Jason Palmer, Julie Onton, Robert Oostenveld and Scott Makeig for their datasets and MATLAB scripts for ICA comparisons (available at http://sccn.ucsd.edu/wiki/BSSComparison).

REFERENCES

[1] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. Mckeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.

[2] S. Makeig, M. Westerfield, T.-P. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T. Sejnowski, "Dynamic brain sources of visual evoked responses," *Science*, vol. 295, no. 5555, pp. 690–694, 2002.

[3] A. Delorme, J. Palmer, J. Onton, R. Oostenveld, and S. Makeig, "Independent eeg sources are dipolar," *PloS one*, vol. 7, no. 2, p. e30135, 2012.

[4] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons, 2004, vol. 46.

[5] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[6] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.

[7] J. A. Palmer, K. Kreutz-Delgado, B. D. Rao, and S. Makeig, "Modeling and estimation of dependent subspaces with non-radially symmetric and skewed densities," in *Independent Component Analysis and Signal Separation*. Springer, 2007, pp. 97–104.

[8] T. Mullen, C. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, G. Cauwenberghs, and T.-P. Jung, "Real-time modeling and 3d visualization of source dynamics and connectivity using wearable eeg." in *IEEE EMBS.*, vol. 2013, 2013, pp. 2184–2187.

[9] M. T. Akhtar, T.-P. Jung, S. Makeig, and G. Cauwenberghs, "Recursive independent component analysis for online blind source separation," in *Circuits and Systems (ISCAS)*. IEEE, 2012, pp. 2813–2816.

[10] S.-H. Hsu, T. Mullen, T.-P. Jung, and G. Cauwenberghs, "Online recursive independent component analysis for real-time source separation of high-density eeg," in *IEEE EMBS.*, 2014, pp. 3845–3848.

[11] T.-W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources," *Neural computation*, vol. 11, no. 2, pp. 417–441, 1999.

[12] J.-F. Cardoso and B. H. Laheld, "Equivariant adaptive source separation," *Signal Processing, IEEE Transactions on*, vol. 44, no. 12, pp. 3017–3030, 1996.

[13] S. Cruces and A. Cichocki, "Combining blind source extraction with joint approximate diagonalization: Thin algorithms for ica," in *Proceedings of 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003)*, 2003, pp. 463–468.

[14] A. Cichocki and S.-i. Amari, *Adaptive blind signal and image processing: learning algorithms and applications*. John Wiley & Sons, 2002, vol. 1.

[15] A. Delorme, T. Mullen, C. Kothe, Z. A. Acar, N. Bigdely-Shamlo, A. Vankov, and S. Makeig, "Eeglab, sift, nft, bcilab, and erica: new tools for advanced eeg processing," *Computational intelligence and neuroscience*, vol. 2011, p. 10, 2011.

[16] J. Onton, A. Delorme, and S. Makeig, "Frontal midline eeg dynamics during working memory," *Neuroimage*, vol. 27, no. 2, pp. 341–356, 2005.

[17] A. Cichocki, S.-i. Amari, K. Siwek, T. Tanaka, A. H. Phan, R. Zdunek, S. Cruces, P. Georgiev, Y. Washizawa, Z. Leonowicz *et al.*, "Icalab toolboxes," 2002.