33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

# Comparison of EEG Blind Source Separation Techniques to Improve the Classification of P300 Trials

Zach Cashero and Chuck Anderson

*Abstract*— **This paper provides a comparison of several blind source separation (BSS) techniques as they are applied to EEG signals. Specifically, this work focuses on the P300 speller paradigm and assesses the classification accuracies for the identification of P300 trials. Previous work has shown that BSS methods such as independent component analysis (ICA) are useful in extracting the P300 source information from the background noise, increasing the classification rates. ICA will be compared with two other BSS methods, maximum noise fraction (MNF) and principal component analysis (PCA). In addition to this, we will analyze the effect of adding temporal information to the original data, which allows these BSS algorithms to find more complex spatio-temporal patterns.**

## I. INTRODUCTION

The field of brain-computer interfaces (BCI) has emerged from the desire for new assistive technology, targeted at patients who are paralyzed and have lost all means of communication. The goal of a BCI is to establish a communication channel directly from the user's brain signals to the computer. We focus only on non-invasive recording methods using electroencephalography (EEG). Since the electrical signals must pass through the scalp, EEG data are inherently noisy, which presents many challenges for signal processing and classification.

The approaches for BCIs can be grouped into two categories. In the first category, the user switches between a small set of mental tasks, where each mental task is associated with a specific action. The most commonly used tasks are motor imagery, such as imagining a specific hand or foot moving. The second category, which this paper focuses on, uses evoked responses from external stimuli that are presented to the user. The user focuses their attention on a specific stimulus that elicits an event-related potential (ERP) in the EEG. The computer then detects this ERP to determine the user's desired action.

One of the more well studied ERPs is the P300 response that is elicited through the 'oddball' paradigm. Here, repeated stimuli are presented to the user that are either target or non-target, in which the targets must be displayed relatively infrequently. Each time the rare target stimulus is presented, a P300 response is produced in the user's EEG. Farwell and Donchin first incorporated this into a BCI called the P300 speller [1], which allows the user to type a single letter at a time. A grid of letters is displayed to the user with the rows and columns flashing in a random order eight times a second. The user focuses on the target letter that they wish to type, and each time it flashes, a P300 response is elicited. Therefore, the detection of the P300 response allows the system to locate the desired target letter.

Although the P300 speller has been studied extensively, a recent review of the field by Mak et al. [2] concludes that more work is still needed to optimize the speed, accuracy, and consistency before the P300 speller is practical to use with disabled patients. One of the main challenges of P300 classification is the low signal-to-noise ratio (SNR), which is usually overcome by averaging together many subsequent trials. However, this decreases the communication rate, and it is therefore desirable to find methods that can extract the true source information from the noise. The current goal of P300 research is to reliably detect P300s using fewer and fewer averaged trials, with the eventual goal of single-trial classification.

Although research from Xu et al. [3] and Li et al. [4] has shown that the application of ICA can improve the P300 classification accuracy, the literature lacks comparative studies of ICA to other preprocessing methods when it is applied to the P300 identification problem. Here, we will present an analysis comparing ICA, MNF, and PCA. Previous work has shown MNF to be effective at removing artifacts from EEG, but it has never been applied to P300 classification before.

The second contribution of this paper is to study the effects of time embedding by adding lags to the data, thereby allowing these methods to find linear spatio-temporal filters, rather than purely spatial filters. If we consider that the original data at a single time sample consists of purely spatial information, adding lags to this data will increase the dimensionality, effectively augmenting the original data with temporal information. These methods will be compared to the standard BSS methods that do not use time embedded data.

## II. RELATED WORK

Blind source separation (BSS) methods work under the assumption that the observed signals from a multi-channel recording are produced from a mixture of several distinct source signals. In the context of EEG recordings, many distinct brain sources are believed to contribute to the overall signal. Makeig et al. [5] find that ERP responses are comprised of several distinct and spatially independent brain processes. Since the electrodes are spatially distributed across the scalp, each electrode picks up a combination of these

Manuscript received April 15, 2011. This work was supported by NSF Grant No. 0841259 and NSF Grant No. 0208958.

Z. Cashero and C. Anderson are with the Department of Computer Science, Colorado State University, Fort Collins, CO 80523, USA zcashero@cs.colostate.edu, anderson@cs.colostate.edu

978-1-4244-4122-8/11/$26.00 ©2011 IEEE 

different brain processes, resulting in the observed signals. BSS methods attempt to transform the observed data into a set of original source signals.

BSS methods like ICA and PCA have been used for artifact removal by Jung et al. [6], which works well because artifacts are usually very strong sources relative to the ongoing brain activity. Maximum noise fraction (MNF) is a relatively new method introduced by Hundley and Kirby [7] that has also been shown to perform well when separating artifacts from the EEG [8], [9].

ICA has been used on P300 trials [3], [4] and shown to improve classification accuracies. Xu et al. [3] used *a priori* information about the spatial and temporal characteristics of the standard P300 response to select the most relevant components, whereas Li et al. [4] used an *a posteriori* template matching technique.

Since traditional ICA (referred to here as spatial ICA) must be used on multi-channel recordings, it is not possible to use with only a single channel. However, Davies and James [10] introduced the concept of single channel ICA that utilizes time embedded data. Time embedded data uses the data from the same channel lagged by one time sample to introduce a new dimension, essentially creating a new 'channel.' When only using a single channel with lagged data, ICA produces components which are a linear mixture of the lagged time samples, thereby creating a purely temporal filter. The authors applied single channel ICA to ictal EEG recordings [10] and P300 recordings [11] to find that it is able to successfully separate out the sources. Quantitative analysis showed that it is still inferior to spatial ICA, which is expected since more data is available from a multi-channel recording. James et al. further extend this idea to create spatio-temporal ICA [12] that uses time-embedded data from multi-channel EEG recordings to obtain a set of source components derived from spatio-temporal filters. These results show that spatio-temporal ICA is able to better isolate the ictal activity than traditional spatial ICA.

## III. METHODS

### A. Datasets

Dataset A is taken from the BCI Competition III, dataset II [13], which is a P300 dataset obtained using the speller paradigm as described by Farwell and Donchin [1]. Although there are two subjects, only the results from Subject A are shown here due to space limitations. The dataset contains 2550 target trials. It was recorded using a 64-electrode cap, although only 8 channels are used in these experiments (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8), which was shown to be sufficient by Krusienski et al. [14]. The data were sampled at 240 Hz and decimated by a factor of 2.

Dataset B was recorded at the Colorado State University Occupational Therapy lab using the Biosemi active electrode system. The paradigm for this study consisted of flashing letters where the letter in the middle of the screen changed each time it flashed. The subject was asked to count the occurrences of a specific target letter with a probability of occurrence of 0.25. A total of 540 target trials were collected. The data were recorded with a 32-electrode cap, but again only 8 channels were used (Fz, Cz, Pz, Oz, P3, P4, PO3, PO4). The data were sampled at 1024 Hz and decimated by a factor of 8.

Both datasets were bandpassed from 0.23 Hz to 30 Hz. The trials consisted of exactly one second of data after the stimulus onset. The single trials from each class were grouped together and 5 subsequent trials were averaged together to make up one trial in the training data. Both the training and testing sets were balanced, such that only a subset of the non-target trials were selected to match the number of target trials. For each experiment, the trials were randomly partitioned into training, validation, and test sets with a fraction of 0.3, 0.3, and 0.4, respectively, ensuring that the same random partition was used when comparing between methods.

### B. Blind Source Separation

One of the three methods, ICA, MNF, or PCA, was applied to the training data (with or without lags) to produce the source components. Data from both target and non-target trials in the training set were alternately concatenated together to produce the continuous time-series signal that BSS was applied to. With each of these methods, the number of source components is equal to the number of input dimensions. Using $n$ channels, this produces $n$ source components with spatial BSS and $nd$ source components if using spatio-temporal BSS with $d$ lagged dimensions. The most relevant components were then selected using the feature selection algorithm described below.

PCA is a commonly used method based on the singular value decomposition (SVD) that is many times used for preprocessing and dimensionality reduction. However, we use it here to extract source components from the data channels. ICA is a well known BSS concept that maximizes statistical independence, and we use the FastICA algorithm here [15]. We will not describe these two methods due to space limitations.

*1) Maximum Noise Fraction:* MNF is a BSS technique [7] that attempts to decompose the signal into source and noise components, based on the assumption that the observed signal $X$ is created by a combination of sources $S$ and noise $N$, as in $X = S + N$. The algorithm works to optimize the SNR as shown here, where $\alpha$ designates the eigenvector components:

$$SNR = \max_{\alpha \neq 0} \frac{\|S\alpha\|}{\|N\alpha\|} = \max_{\alpha \neq 0} \frac{\|X\alpha\|}{\|N\alpha\|} = \max_{\alpha \neq 0} \frac{\alpha^T X^T X \alpha}{\alpha^T N^T N \alpha}$$

This equation is only equivalent if the signal and noise are assumed to be orthogonal ($S^T N = 0$). Since the sample covariance $X^T X$ can be easily computed from the input data $X$, only the noise covariance $N^T N$ is unknown. The noise covariance is characterized by the covariance of the difference of the original signal with the same signal shifted by one time sample. This characterizes the noise as large fluctuations from one sample to the next, or essentially the higher frequency components.

### C. Time Embedded Data

As mentioned above, time embedded data is created by adding a new lagged dimension. This is the same data from one of the original channels, only shifted (or lagged) by one time sample. After the source components are found with the BSS method, it is possible to project these components back to the original data space as described in [10]. Since each source component contributes to the recorded signal, it is possible to reconstitute the original data by projecting all components back to the original data space and summing them together. Therefore, it is necessary to select only a subset of these components to use. For the feature selection step, each component is projected separately in order for the algorithm to determine the most relevant components.

### D. Feature Selection

The feature selection algorithm used here is based on the ANOVA statistical test. It was applied to all features using instances from two groups, target and non-target instances from the training set. The features are individual time samples from the projected time series of each component, and the ANOVA scores were averaged across all features associated with each component. The averaged ANOVA scores were then used to rank all of the components based on relevance. This way, the feature selection algorithm does not make any assumptions about the components and ranks them based on statistical differences of the associated features.

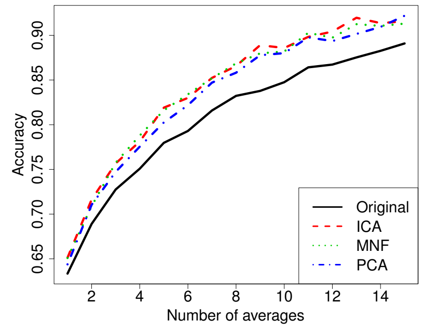
It is then necessary to determine the optimal number of top ranked components to use. The training set was used to train a support vector machine (SVM) with a Gaussian kernel, and a separate validation set was used to assess the accuracy. We used the top $n$ ranked components, as $n$ varied from 1 to a preset maximum of 20, to find the $n$ that resulted in the highest validation accuracy. Thus, the top $n$ ranked components were selected to be used for the final transform. If time embedded data is used, the final $n$ components are projected back and summed together in the original data space.

After finding the resulting linear transform, it was applied to the training data before training the final SVM model. In order to create a single sample for the SVM, we created a feature vector consisting of all time samples in the one second window from each of the eight channels. Once the SVM was trained, the same linear transform was applied to the unseen test data before it was classified with the SVM.

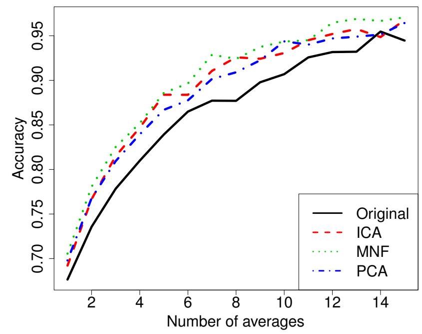
## IV. RESULTS

The results for each of the spatial BSS methods are compared with those of classifying on the original data in Figure 1. It should be noted that the ANOVA feature selection algorithm was still applied to the original data for channel selection. These results are shown as the number of averages in the test dataset increases from 1 to 15.

The results in Figure 2 show the effect of adding lagged dimensions to the data, increasing from 0 to 15. When the number of lags is equal to 0, the original data is used, resulting in the traditional spatial versions of the BSS




(a) Results on Dataset A




(b) Results on Dataset B

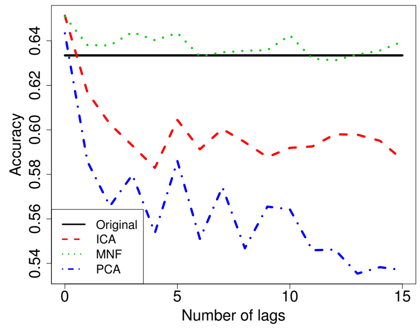
Fig. 1. This plot shows how the three BSS methods compare to classifying on the original untransformed data.

methods. The spatio-temporal versions of the BSS methods are used when the number of lags is greater than 0, and as the number of lags increases, the number of resulting source components also increases. Since lags are not added to the original data unless a BSS method is applied, the accuracy on the original data is unchanged as the number of lags varies and is shown by a flat line. Although the results here are only shown for single trial classification on the test set, the trends are almost identical as the number of averaged trials increases.

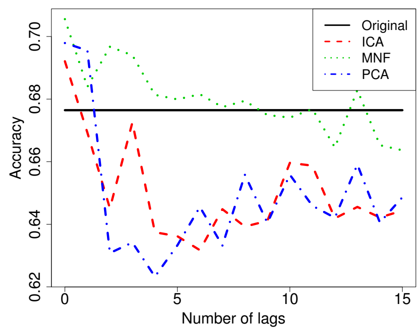
## V. DISCUSSION

First, it is important to point out that the BSS methods did improve the classification accuracy over using the original data. It is helpful to remember that the motivation for improving the accuracy is to reduce the number of averaged trials required. If we strive for a minimum accuracy of 0.85, we can look at where the curves intersect this point in Figure 1. In Dataset A, the application of one of the BSS methods would achieve this minimum accuracy with only 7 averaged trials, rather than 10 averaged trials if using the original data. Dataset B shows a reduction from 6 averaged trials down to 4. The results from Subject B in Dataset A (not shown) are similar, and the number of averaged trials can be reduced from 6 down to 4.

However, when comparing the three BSS methods, there is no distinguishable difference. This is interesting because ICA has become a well studied EEG signal processing technique in the BCI field, and is the most complex method of the




(a) Results on Dataset A




(b) Results on Dataset B

Fig. 2. This plot shows the effects of adding lagged dimensions to the original multi-channel data before applying one of the BSS methods. Here, the baseline performance of the original data is flat because it always uses 0 lags if no BSS algorithm is applied. The results here are shown for single trial classification.

three, utilizing higher-order statistics rather than second-order statistics. These results suggest that ICA does not gain any advantages over the relatively simple methods of PCA and MNF. The computation times of the BSS methods provide some insight into the relative complexities of the algorithms. PCA is the simplest method with the shortest computation time. The computation time for MNF is an order of magnitude larger than for PCA, and the computation time for ICA is two orders of magnitude larger than for PCA.

When looking at Figure 2, it is clear to see that adding lags does not improve the classification accuracy. It is also interesting to see that the performance of the spatio-temporal BSS methods are now clearly separated. The results from Subject B in Dataset A also show a similar separation of the three BSS methods. Although none of them improve with lags, the performance of MNF remains relatively stable, whereas ICA and PCA do worse. This suggests that the additional temporal information is irrelevant. We believe that the problem becomes too complex for any of these algorithms to handle the spatio-temporal information.

## VI. CONCLUSIONS AND FUTURE WORK

As shown by our results, for the problem of P300 classification, it does not appear that ICA shows any clear advantage over the simpler methods of PCA and MNF. The addition of temporal information with time embedded data reduces the classification accuracy. However, it is clear that all three spatial BSS methods improve the classification accuracy over using the original data. This allows the P300 speller to accurately classify target trials using fewer averages, thereby increasing the communication rate.

Future work consists of running experiments with more datasets and subjects. We are also currently looking at other feature selection algorithms. We are currently investigating the causes behind the differences between the three spatio-temporal versions of the BSS methods. Also, since the spatial BSS methods clearly show an improvement, we will look at how this is changed by expanding the number of electrodes beyond the current subset of eight.

## REFERENCES


[1] L. Farwell and E. Donchin, “Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials,” *Electroenceph. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, 1988.

[2] J. N. Mak, Y. Arbel, J. W. Minett, L. M. McCane, B. Yuksel, D. Ryan, D. Thompson, L. Bianchi, and D. Erdogmus, “Optimizing the P300-based brain-computer interface: current status, limitations and future directions,” *Journal of Neural Engineering*, vol. 8, no. 2, pp. 025003–+, 2011.

[3] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang, “BCI competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications,” *IEEE Transactions on Biomedical Engineering*, vol. 3, pp. 1067–1072, June 2004.

[4] K. Li, R. Sankar, Y. Arbel, and E. Donchin, “Single trial independent component analysis for P300 BCI system,” *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, pp. 4035–4038, 2009.

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

[6] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. McKeown, V. Iragui, and T. Sejnowski, “Removing electroencephalographic artifacts: comparison between ICA and PCA,” in *Neural Networks for Signal Processing VIII, 1998. Proceedings of the 1998 IEEE Signal Processing Society Workshop*, pp. 63–72, Aug. 1998.

[7] D. R. Hundley, M. J. Kirby, and M. Anderle, “Blind source separation using the maximum signal fraction approach,” *Signal Processing*, vol. 82, no. 10, pp. 1505–1508, 2002.

[8] J. Knight, “Signal fraction analysis and artifact removal in EEG,” Master’s thesis, Department of Computer Science, Colorado State University, Fort Collins, CO, 2003.

[9] L. Sun, J. Rieger, and H. Hinrichs, “Maximum noise fraction (MNF) transformation to remove ballistocardiographic artifacts in EEG signals recorded during fMRI scanning,” *NeuroImage*, vol. 46, no. 1, pp. 144–153, 2009.

[10] M. Davies and C. James, “Source separation using single channel ICA,” *Signal Processing*, vol. 87, no. 8, pp. 1819–1832, 2007. Independent Component Analysis and Blind Source Separation.

[11] C. James and S. Wang, “Performance analysis of a P300 BCI speller through single channel ICA,” *Advances in Medical, Signal and Information Processing, 2008. MEDSIP 2008*, pp. 1–4, July 2008.

[12] C. James, D. Abasolo, and D. Gupta, “Space-time ICA versus ensemble ICA for ictal EEG analysis with component differentiation via lempel-ziv complexity,” *Engineering in Medicine and Biology Society, 2007. EMBS 2007*, pp. 5473–5476, Aug. 2007.

[13] B. Blankertz, K.-R. Müller, D. Krusienski, G. Schalk, J. Wolpaw, A. Schlogl, G. Pfurtscheller, J. Millan, M. Schroder, and N. Birbaumer, “The BCI competition III: validating alternative approaches to actual BCI problems,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, pp. 153–159, June 2006.

[14] D. Krusienski, E. Sellers, D. McFarland, T. Vaughan, and J. Wolpaw, “Toward enhanced P300 speller performance,” *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 15–21, 2008.

[15] A. Hyvärinen and E. Oja, “A fast fixed-point algorithm for independent component analysis,” *Neural Comput.*, vol. 9, pp. 1483–1492, October 1997.