GRADIENT ARTIFACT REMOVAL IN CONCURRENTLY ACQUIRED EEG DATA USING INDEPENDENT VECTOR ANALYSIS

Partha Pratim Acharjee¹, Ronald Phlypo¹, Lei Wu^{2,3} Vince D. Calhoun^{2,3} and Tülay Adalı¹

¹ University of Maryland, Baltimore County, Dept. of CSEE, Baltimore, MD 21250
² Mind Research Network, Albuquerque, New Mexico, USA
³ Dept. of ECE, University of New Mexico, Albuquerque, NM 87131, USA

ABSTRACT

We consider the problem of removing gradient artifact from electroencephalogram (EEG) signal, registered during a functional magnetic resonance imaging (fMRI) acquisition, by calculating and utilizing the statistical properties of the artifacts. We propose a new approach to EEG data organization for extracting artifactual components using independent vector analysis. This new approach estimates the gradient artifact signal as a single component thus alleviating the need of using advanced order selection algorithm before back reconstruction of EEG data. Experimental results are compared with average artifact subtraction method on real EEG data collected concurrently with fMRI data.

Index Terms— Gradient artifact, Independent vector analysis, AAS, EEG

1. INTRODUCTION

Concurrently collected electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI) data enables us to examine the relationship between electrical and hemodynamic signals [1, 2]. Typically, these are independently recorded but research has shown that recording the two modalities simultaneously has proven more effective in localizing the generators of EEG events [3]. However, concurrent recordings bring additional inevitable artifacts relative to those seen when the techniques are used separately. Specifically, concurrent recording introduces gradient artifact (GA) and radio frequency (RF) artifacts which significantly degrades the signal to noise ratio (SNR) of EEG signal. The rapid switching of magnetic gradient for fMRI recording results in the presence of a very large GA in the EEG data. This artifact is introduced during each of the fMRI slice acquisitions.

As the slice acquisitions are periodic (one period for each fMRI timepoint), this provides the motivation to create GA templates by using the onset of each GA cycle as the time locking event and averaging over the different cycles. This

general template mapping procedure is known as average artifact subtraction (AAS) [4]. As an extension, adaptive filtering [5] and frequency domain methods [6] have also been developed. However, the amplitude and periodicity of the GAs vary from epoch to epoch. This variability of GA leads to the idea of using multiple periods in a single analysis while weighting each period equally in terms of importance. Our independent vector analysis (IVA) based method is designed considering the variability of the GA from epoch to epoch. We organize data in a form that allows for IVA to effectively separate artifacts from source components. We show that composing each dataset of samples from a single lead, which are indexed by epoch and time, we are able to extract the gradient artifact as a single component. This means that, given the lead, the gradient artifact is a scaled constant function. By using IVA, we exploit the dependency of the gradient artifact functions across leads, while guaranteeing minimal dependence between the GA and the signals of physiological origin with other artifacts. The combination of IVA with our proposed data organization extracts the GA as a single component and alleviates the need of any advanced order selection algorithms to decide the number of artifact components. This is important as improper order selection may lead to have residual GA or accidental removal of EEG sources.

2. METHOD

2.1. EEG Acquisition and Preprocessing

A 32-channel BrainAmp MR-compatible system was used for EEG recordings using the BrainCap electrode cap . Ring-type sintered nonmagnetic Ag/AgCl electrodes were placed on the scalp according to the international 10/20 system. Data were collected for nine minutes at 5 kHz sampling rate in 32 channels; band-pass filtering from 0.016 to 250 Hz was applied. To avoid temporal jitter the EEG amplifier and fMRI were synchronized using an in-house device. This dataset is also used in [7, 8].

The preprocessing was performed in Matlab with the software package EEGLAB [9]. Some visualization and analytic plug-ins of EEGLAB are also used [10, 11]. Continuous EEG

This work was supported by the National Science Foundation grants NSF-CCF 1117056 and NSF-III 1017718.



Fig. 1. Illustration of epoching on continuous EEG data to construct dataset for IVA algorithms and IVA decomposition to obtain estimated components for gradient artifact removal. Formation of SCVs from estimates are also shown. ICs: Independent Component, SCVs: Source Component Vector

data was segmented in standard 2 s window synchronized with the fMRI scanning repeat time (TR) to achieve sufficient frequency resolution in EEG frequency bands of interest and each window will be referred to as an epoch from here onward. Between adjacent epochs 24 ms overlap was allowed. After epoching, 256 epochs are extracted from one recording per channel and each epoch has 10240 data points. From epoched EEG data, channel baseline means are removed from each channel. Then the EEG data was down-sampled to 1000 Hz which results 2048 time points.

2.2. Independent Vector Analysis

Joint blind source separation (JBSS) is used in fields such as medical imaging, telecommunications, hyper-spectral data, speech and audio signal processing. The goal of JBSS is to separate mixed sources under certain assumptions on the sources and/or the mixing process from multiple observations. JBSS algorithms are useful in estimating sources using information from multiple observations [12, 13]. IVA was introduced as a natural extension of independent component analysis (ICA) to concurrently extract independent components from multiple datasets [14, 15]. In IVA, corresponding components extracted from different dataset are maximally dependent while components within a dataset are independent from each other. Thus, component independence within a dataset and corresponding component dependence across datasets are maximized simultaneously, which cannot be achieved by separate ICA of each dataset [16]. Furthermore, IVA has been extended to enable decomposition of complex-valued data [17].

We propose a new IVA-based method and data structure for analyzing EEG data in order to achieve JBSS and extract spatial components from multiple channels concurrently. We now formulate the IVA problem for the multichannel EEG analysis. Suppose, N number of epochs from kth channel are observed. Here, K is the number of channels yielding K different datasets. Suppose, N number of epochs from kth channel form a dataset which are linear mixtures of Nunknown sources,

$$\mathbf{X}^{[k]} = \mathbf{A}^{[k]} \mathbf{S}^{[k]}, 1 \le k \le K$$

where $\mathbf{S}^{[k]} = [\mathbf{s}_1^{[k]}, \mathbf{s}_2^{[k]}, \dots, \mathbf{s}_N^{[k]}]^T$ is the zero mean underlying source vector that are linearly mixed by $N \times N$ mixing matrix **A**. The source vector consists N number of independent sources with T number of datapoints in each. Therefore, dimension of source vectors are $N \times T$. To associate components across all channels, the *n*th source component vector (SCV) is formed by taking *n*th component vectors from each dataset, i.e., $\mathbf{S}_n = [\mathbf{s}_n^{[1]}, \mathbf{s}_n^{[2]}, \dots, \mathbf{s}_n^{[k]}]^T$, $n = 1, 2, \dots, N$. The goal of IVA is to find the $N \times N$ demixing $\mathbf{W}^{[k]}$ for all datasets and the corresponding estimation of observations, $\mathbf{Y}^{[k]} = \mathbf{W}^{[k]} \mathbf{X}^{[k]}, 1 \leq k \leq K$, such that the *n*th estimates, $\mathbf{Y}_n = [\mathbf{y}_n^{[1]}, \mathbf{y}_n^{[2]}, \dots, \mathbf{y}_n^{[k]}], n = 1, 2, \dots, N$, are maximally dependent and estimates from same dataset are maximally independent from each other, where $\mathbf{y}_n^{[k]} = (\mathbf{w}_n^{[k]})^T \mathbf{X}^{[k]}$ and $(\mathbf{w}_n^{[k]})^T$ is the *n*th row of $\mathbf{W}^{[k]}$. IVA decomposition can be achieved by minimizing the mutual information among source component vectors (SCV),

$$\begin{split} J_{\text{IVA}} &= \sum_{n=1}^{N} \mathcal{H}(\mathbf{y}_n) - \mathcal{H}(\mathbf{y}_1, ..., \mathbf{y}_N) \\ &= \sum_{n=1}^{N} \left(\sum_{k=1}^{K} \mathcal{H}(y_n^{[k]}) - \mathcal{I}(\mathbf{y}_n) \right) \\ &- \sum_{k=1}^{K} \log |\text{det} \mathbf{W}^{[k]}| - \mathbf{C} \end{split}$$

where $\mathcal{H}(.)$ is entropy and $\mathcal{I}(\mathbf{y}_n)$ is the mutual information within the *n*th SCV, and *C* is a constant term that depends only on $\mathbf{X}^{[k]}, k = 1, ..., K$. By minimizing this IVA cost function, the dependence among components within each SCV, $\mathcal{I}(\mathbf{y}_n), 1 \le n \le N$, are maximized concurrently.

The IVA algorithm, IVA-L [12] utilizes a multivariate Laplace prior for distribution of SCVs. Thus, the components within each SCV are assumed to be second-order uncorrelated and IVA-L exploits only higher-order statistics. On the contrary, IVA-G exploits only second-order statistical information by multivariate Gaussian assumption. To take advantages of both algorithms, we initialize IVA-L with a solution from IVA-G to achieve IVA-GL decomposition. IVA-GL can yield more robust joint blind source separation than using either IVA-L or IVA-G alone [15].

2.3. JBSS Architecture and Motivation

In this paper, performance of the proposed method is shown for EEG concurrently recorded with fMRI. EOG and ECG channels are ignored thus 30 channels, 2048 time points and 20 epochs are considered for each IVA run. The data organization we introduce for IVA is able to estimate the GA signal as a single component which can be easily removed by back reconstruction of the EEG signal by eliminating the component from the back reconstruction. Configuration of JBSS input is shown in Fig. 1. From continuous EEG data, N epochs are extracted with 24 ms overlap between adjacent epochs and each epoch duration is 2 s. All epochs from one channels are reshaped to form a dataset $\mathbf{X}^{[k]}$ and K channels provides K datasets for the IVA algorithm. From each dataset, IVA provides N estimated sources each of which has T samples. Sources from the same dataset are maximally independent and shown in different color. Permutation ambiguity of sources across datasets is resolved by IVA through full use of dependence across the datasets. Corresponding sources from different datasets are highly dependent and shown in same color in Fig 1. Corresponding sources are gathered to a set of component vectors, to form SCV such that N SCVs can be formed from N sources which are shown in the right most box of the figure. According to our dataset formation, k = 1, ..., 30 and n = 1, ..., 20.

An initial thought may be using epochs as the multiple datasets for IVA but using channels as the multiple datasets is useful for a number of reasons. First, sources are estimated for each channel by viewing the epoch data. Each channel captures brain activity from a specific part of the brain. Sources affecting each channel are spatially separated and are not expected to have a large effect on all channels. Channelwise source estimation can capture more local sources from a particular brain region than epoch-wise source estimation. Second, in this setup, SCVs are formed by corresponding sources from different channels. As IVA tries to maximize dependence of sources within SCVs, it effectively extracts the



Fig. 2. (a) The first component has the highest MI within SCV and almost zero MI with other SCVs. (b) First SCV has the highest MI measure with all channels at all epochs where other SCVs has zero MI measure.

GA as a component that is highly present over all channels. GA is not spatially separable from the other sources (large spatial overlap, spatially smeared out), hence the classic approach of GA removal would not work.

2.4. Normalized mutual information

After estimating source component vector for all channels, dependence between SCVs are calculated by mutual information (MI). Calculated mutual information is normalized as [18]

$$\lambda_{n,n'}^{[k,k']} = \sqrt{1-\exp(-2\mathcal{I}(y_n^k,y_{n'}^{k'}))},$$

where (n, n') = 1, ..., N; (k, k') = 1, ..., K and $\mathcal{I}(y_n^k, y_{n'}^{k'})$ is the mutual information between *n*th estimated component from *k*th channel and *n*'th estimated component from *k*'th channel. Normalized MI, λ , is in the range [0,1) where zero means completely independent.

3. EXPERIMENTAL RESULTS

3.1. Normalized MI of SCVs

After estimating components from the EEG data, normalized mutual information within and across SCVs are calculated and used to form a normalized mutual information matrix such that diagonal 30×30 sub-blocks are the measure within SCVs and off-diagonal blocks are the measure across SCVs. IVA tries to maximize dependency within SCVs and minimize dependency across SCVs. So, diagonal 30×30 sub-blocks should have higher mutual information than off-diagonal sub-blocks. From Fig. 2(a), it is clear that first component has highest normalized mutual information within its SCV than other components and the first SCV has almost zero normalized MI with other SCVs. Thus, corresponding component of first SCV is highly dependent and the first component is almost independent from other components for all epochs. The highest mutual information within the first



Fig. 3. Frequency spectrum of EEG signal before and after artifact removal only using IVA-GL algorithm. Magnitude peaks at artifact band is suppressed significantly.

SCV indicates that the first component of all dataset are triggered by the same source. Almost zero MI measure between first SCVs with other SCVs indicates that the first estimated source is different than any other source extracted from any channel. We also show MI measures between estimated components and raw EEG data for all channels and epochs in Fig. 2(b). It is clear that the whole SCV of first component has the highest MI measure with raw EEG signals from all channels whereas other components are almost independent. As GA is present in all channels and all epochs of the raw EEG data, highest dependence implies that the first component is very similar to the gradient artifact. Independence of raw EEG data with SCVs from other components shows that other components are less likely to be GA component.

3.2. Performance of the Proposed IVA-Based Approach

Frequency spectrum of the reconstructed signal using the IVA-based method and the raw EEG signal can be compared to observe the energy at the artifact band. The average frequency spectra for all channels are shown in Fig. 3. It is clear that the IVA-based method removes significant artifact energy from the artifact frequency band. In the low frequency band, the frequency spectrum is almost unchanged. Thus, it is observed that the IVA-GL removes primarily the artifact portion and keeps the original EEG signal unchanged. From frequency spectrum of raw EEG at Fig. 3, we see that 14 Hz, 28 Hz, 42 Hz and 56 Hz are the artifact band center frequencies.

The AAS method is a template matching and subtraction procedure based GA removal method. For performance comparison of AAS and IVA based approach, ten separate experiments, using 20 epochs for each, are performed. Five consecutive epochs are used to estimate the template for AAS method. After removing GA, a spectral density at artifact



Fig. 4. Remaining percentage of energy at artifact frequency band for multiple runs using AAS and IVA based approach. On each box, the central mark is the median, the whiskers extend to the most extreme data points not considered outliers.

band is calculated to evaluate the performance. The percentage of energy at artifact frequency band for all runs using both method are shown as box plot in Fig. 4. On average, after our method, percentage of remaining artifact energy is less than the AAS method for all runs except last one. The distribution of the remaining artifact energy from both the AAS and IVA based methods is symmetric but not identical. To compare the results in terms of statistical significance, Wilcoxon signed rank test is performed on the data with false discovery rate (FDR) correction for multiple tests using 5% significance level. Supporting the null hypothesis of the test indicates that both distributions come from same median. Calculated p values are corrected for multiple comparison via the FDR correction method with 5% significance level which are used as α for individual test. It is observed that 2nd, 4th and 6th support null hypothesis and others reject the null hypothesis. Thus, with 5% statistical significance, IVA performs better in six runs, not performing worse in three runs and in one run AAS performs better.

4. DISCUSSION

In this paper, we propose an effective approach based on JBSS to remove the GA from EEG data. The GA is captured in a single component that alleviates the need to use order selection algorithms to find the true number of artifact components. The proposed IVA-based approach eliminates artifacts without affecting the low frequency EEG signal in comparison with the AAS method. With variability in GA, extracted artifact using IVA approach is shown to perform better. In future work, a constrained IVA can be developed to use GA template as reference for sources, and may provide even better estimation of artifact component since it uses some prior knowledge about the GA.

5. REFERENCES

- Maria J Rosa, James Kilner, Felix Blankenburg, O Josephs, and W Penny, "Estimating the transfer function from neuronal activity to BOLD using simultaneous EEG-fMRI," *NeuroImage*, vol. 49, no. 2, pp. 1496– 1509, 2010.
- [2] René J Huster, Stefan Debener, Tom Eichele, and Christoph S Herrmann, "Methods for simultaneous EEG-fMRI: an introductory review," *The Journal of Neuroscience*, vol. 32, no. 18, pp. 6053–6060, 2012.
- [3] Stefan Debener, Markus Ullsperger, Markus Siegel, and Andreas K Engel, "Single-trial EEG-fMRI reveals the dynamics of cognitive function," *Trends in Cognitive Sciences*, vol. 10, no. 12, pp. 558–563, 2006.
- [4] Philip J Allen, Giovanni Polizzi, Karsten Krakow, David R Fish, and Louis Lemieux, "Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction," *NeuroImage*, vol. 8, no. 3, pp. 229–239, 1998.
- [5] Jan Sijbers, Johan Van Audekerke, Marleen Verhoye, A Van der Linden, and Dirk Van Dyck, "Reduction of ECG and gradient related artifacts in simultaneously recorded human EEG/MRI data," *Magnetic Resonance Imaging*, vol. 18, no. 7, pp. 881–886, 2000.
- [6] A Hoffmann, L Jäger, KJ Werhahn, M Jaschke, S Noachtar, and M Reiser, "Electroencephalography during functional echo-planar imaging: detection of epileptic spikes using post-processing methods," *Magnetic Resonance in Medicine*, vol. 44, no. 5, pp. 791– 798, 2000.
- [7] Lei Wu, Tom Eichele, and Vince D Calhoun, "Reactivity of hemodynamic responses and functional connectivity to different states of alpha synchrony: a concurrent EEG-fMRI study," *NeuroImage*, vol. 52, no. 4, pp. 1252–1260, 2010.
- [8] David A Bridwell, Lei Wu, Tom Eichele, and Vince D Calhoun, "The spatiospectral characterization of brain networks: Fusing concurrent EEG spectra and fMRI maps," *NeuroImage*, 2012.
- [9] Arnaud Delorme and Scott Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.
- [10] Scott Makeig, Stefan Debener, Julie Onton, and Arnaud Delorme, "Mining event-related brain dynamics," *Trends in Cognitive Sciences*, vol. 8, no. 5, pp. 204–210, 2004.

- [11] Tom Eichele, Vince D Calhoun, and Stefan Debener, "Mining EEG-fMRI using independent component analysis," *International Journal of Psychophysiology*, vol. 73, no. 1, pp. 53–61, 2009.
- [12] Jong-Hwan Lee, Te-Won Lee, Ferenc A Jolesz, and Seung-Schik Yoo, "Independent vector analysis (IVA): multivariate approach for fMRI group study," *NeuroImage*, vol. 40, no. 1, pp. 86–109, 2008.
- [13] Yi-Ou Li, Tülay Adalı, Wei Wang, and Vince D Calhoun, "Joint blind source separation by multiset canonical correlation analysis," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3918–3929, 2009.
- [14] Taesu Kim, Torbjørn Eltoft, and Te-Won Lee, "Independent vector analysis: An extension of ICA to multivariate components," in *Independent Component Analy*sis and Blind Signal Separation, pp. 165–172. Springer, 2006.
- [15] Matthew Anderson, Tülay Adalı, and Xi-Lin Li, "Joint blind source separation with multivariate Gaussian model: Algorithms and performance analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1672–1683, 2012.
- [16] Matthew Anderson, Xi-Lin Li, and Tülay Adalı, "Nonorthogonal independent vector analysis using multivariate Gaussian model," in *Latent Variable Analysis* and Signal Separation, pp. 354–361. Springer, 2010.
- [17] Matthew Anderson, Xi-Lin Li, and Tülay Adalı, "Complex-valued independent vector analysis: Application to multivariate Gaussian model," *Signal Processing*, vol. 92, no. 8, pp. 1821–1831, 2012.
- [18] Andreia Dionisio, Rui Menezes, and Diana A Mendes, "Mutual information: a measure of dependency for nonlinear time series," *Physica A: Statistical Mechanics and its Applications*, vol. 344, no. 1, pp. 326–329, 2004.