Coupled Basis Learning and Regularized Reconstruction for BCG Artifact Removal in Simultaneous EEG-fMRI Studies

Hongjing Xia, Dan Ruan, Mark S. Cohen

Bioengineering Department, University of California, Los Angeles

Abstract

The ballistocardiogram (BCG) is a major artifact in electroencephalographic (EEG) data acquired inside a magnetic resonance imaging (MRI) scanner, and is several times larger in magnitude than the actual EEG signals. Removing the BCG artifacts remains an unresolved challenge, especially in studies of continuous EEG recordings. In this work, we propose a Direct Recording-Joint Incoherent Basis (DRJIB) method to decompose the observed noisy EEG measurements into BCG and underlying EEG components. We compare its preliminary performance quantitatively with that of the benchmark Optimal Basis Set (OBS) method. Without assuming orthogonality or independence of the BCG and EEG subspaces, as in conventional methods, our approach learns the bases faithfully from BCG-only and EEG-only signals acquired from our new experimental setup. Specifically, to promote subspace separability, a paired set of low-dimensional and semi-orthogonal (BCG, EEG) basis representations is obtained by minimizing a cost function consisting of group sparsity penalties for automatic dimension selection and an energy term for encouraging incoherence. Reconstruction is subsequently obtained by fitting the contaminated data to a generative model using the learned bases subject to regularization. In the challenging non-event-related EEG studies, our DRJIB method outperforms the OBS method by nearly 12-fold in separating and preserving the continuous BCG and EEG signals.

Index Terms—Ballistocardiogram, simultaneous EEG-fMRI, artifact removal, basis learning, group sparsity, incoherence

1. Introduction

Electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) provide complementary yet functionally linked information about underlying brain activity. Simultaneous acquisition of data from both modalities has been investigated intensively recently and has proven its value in numerous applications [1-6]. While artifacts introduced to the MRI data are relatively easy to manage [7, 8], artifacts appearing in the EEG data recorded inside the MR scanner still present daunting obstacles. Unlike event-related experiments, where artifacts may be damped by averaging around randomly distributed events, the prominent magnetically-induced artifacts generating the ballistocardiogram (BCG) remain extremely difficult to remove in continuous experiments, such as studies of ongoing spontaneous brain rhythms. It has been suggested [9] that BCG artifacts may be related to cardiac pulsation and breathing, exhibiting high temporal irregularity and several times larger in magnitude than normal brain EEG signals.

The most widely used means of suppressing the BCG artifacts is the Optimal Basis Sets method (OBS) [10]. It applies principal component analysis (PCA) to the contaminated EEG data and removes the first few (usually 3) principal components (PCs) as contributions from BCG artifacts. Therefore OBS can be interpreted to assume (1) complete orthogonality between the BCG and EEG subspaces, and (2) separation of these two spaces with pure energy concentration, as indicated by partitioning the PCs sequentially. Earlier developed adaptive template approaches [11] can also be interpreted as weighted PCA. A commonality among all existing methods is to extract the BCG and EEG signal components directly from contaminated data, with no, or minimal, characterization of the true subspaces of the BCG and EEG components.

By contrast, we address the BCG removal challenge from the perspective of signal separation. Our DRJIB method consists of two phases: (1) a basis-learning phase where the bases for BCG and EEG signals are optimized jointly to represent the BCG/EEG-only signals, and to be as independent as possible to promise better separability with an incoherence term; (2) a reconstruction phase where the BCG and EEG components are estimated from the contaminated EEG data using the learned bases and regularization on the structures of the coefficients.

We developed a new experimental setup introduced in Sec. 2.2, to generate BCG/EEG-only prior data for the purpose of basis learning. The jointly learned bases were obtained with an optimization framework with an objective function to encode low-dimensional subspace modeling and incoherence regularization, reported in Sec. 2.3. Reconstruction with coefficient regularization is reported in Sec. 2.4.
2. METHODS

2.1. Generative model for contaminated data
It is reasonable to assume that the BCG $X_{bcg}$ and brain EEG $X_{eeg}$ signals are generated from independent sources. The collected data $Y$ can be modeled as their superposition subject to noise contamination, according to

$$ Y = X_{bcg} + X_{eeg} + noise = B_{bcg}C_{bcg} + B_{eeg}C_{eeg} + noise, \quad (1) $$

where $B$ and $C$ are the basis and coefficient matrices.

2.2. Experimental setup and simulation

We acquire EEG (256-channel, Electrical Geodesics Inc. (EGI) GES300MR) recordings at 250 Hz from both inside and outside the scanner (Siemens Trio MRI). EEG-only (no MR-induced artifacts) recordings from outside the scanner are acquired prior to placing the subject inside the scanner and used as prior data for EEG characterization. During acquisition within the scanner, we isolate a subset of the EEG electrodes from the scalp by inserting a thin plastic insulating layer in between to record only the MR-induced artifacts (no brain EEG). In the absence of a running MRI sequence, there are no other MR-induced artifacts, as caused by MR gradient or RF, and BCG is the only MR artifacts that contribute to the signals collected from those insulated EEG electrodes. As such, we obtain EEG-only $X_{eeg}$ and BCG-only $X_{bcg}$ signals from outside and inside the scanner as prior data without cross-contamination. As the electrical activity of the brain of the same subject is reasonably consistent regardless of the scanning condition, it is safe to assume the characteristics of EEG brain signals are consistent inside/outside scanner. Furthermore, the BCG-only recordings from the isolated electrodes are obtained simultaneously as the other non-blocked channels, and we expect them to have consistent temporal signature with the underlying BCG component at an adjacent channel, assuming no change of major blood vessels structure in between.

To evaluate different approaches quantitatively, we simulate contaminated data by (1) acquiring additional EEG-only signals, $X_{e}$ from channel A outside the scanner as the EEG signals are assumed to be responsibly consistent from inside and outside the scanner, (2) acquiring insulated BCG-only signal, $X_{b}$ from channel A inside the scanner, and (3) synthesizing the contaminated data according to the generative model, $Y=X_{b}+X_{e}$. To reflect the fact that no “ground-truth” BCG/EEG is available for contaminated recordings, we use the BCG-only data, $X_{b,prior}$, acquired inside the scanner from channel B, and the EEG-only data $X_{e,prior}$ from channel A outside scanner at different times as prior data for our basis learning step. The learned bases are then used in the reconstruction stage. Figure 2 illustrates this process.

![Fig. 2. BCG-only data (blue) from channel B and EEG-only (blue) data from channel A are used as prior BCG $X_{b,prior}$ and prior EEG $X_{e,prior}$ data. Simulated data $Y$ (red) is synthesized from summing BCG $X_{b}$ (red) and EEG data $X_{e}$ (red) from channel A.](image)

2.3. Basis learning

We aim to generate a pair of bases: $B_{b,prior}$ for BCG-only and $B_{e,prior}$ for EEG-only signals, respectively, such that (1) they properly represent prior data $X_{b,prior}$ and $X_{e,prior}$, (2) they span subspaces that are orthogonal to each other as possible, and yet allow overlap when strongly supported by data, and (3) their intrinsic dimensions are low. To this end, we consider the following optimization problem:

$$ \begin{align*}
\text{minimize} & \quad \frac{1}{2} \left\| B_{b,prior}^T B_{b,prior} \right\|_F^2 + \lambda_1 \left\| C_{b,prior} \right\|_1 + \lambda_2 \left\| C_{e,prior} \right\|_1 \\
\text{subject to} & \quad \left\| X_{b,prior} - C_{b,prior} B_{b,prior} \right\|_{l_2} \leq \sigma_1 \\
& \quad \left\| X_{e,prior} - C_{e,prior} B_{e,prior} \right\|_{l_2} \leq \sigma_2 
\end{align*} \quad (2) $$

with the $\ell_2$-norm defined as $\left\| C \right\|_{l_2} = \sum_{i=1}^{m} \left\| C[i,:] \right\|_2$, where $i \in \{1, 2, ..., m\}$ is an index set corresponding to the $i^{th}$ group (row), and $m$ is the row number of $C$.

In (2), the data fidelity constraints ensure proper representation of the prior data. The $\ell_2$-norm encourages group sparsity structure on the coefficients of the $C_{b,prior}$ and $C_{e,prior}$, leading to the concentration of significant values to a few rows of the coefficient matrices – this structure effectively “nullifies” the contribution of the basis columns.

Fig. 1. (a) Normal portion of an EEG net that collects contaminated data with both brain EEG and MR-induced BCG and (b) insulated portion of the EEG net that collects BCG-only signals.

After applying data drift removal with a high pass filter (cutoff at 0.1Hz), we extract alpha band (8-12Hz) from EEG-only recordings with a FIR filter. All data are divided into column vectors according to detected heartbeats and then concatenated to form data matrices (the same procedure as in the OBS method [10]).
corresponding to the insignificant rows of the coefficient matrices, and results in low-dimensional representations of the spanned subspaces. Motivated by incoherent dictionary learning in Ramiriez, et al. [12], we penalize the coherence between the two bases embodied by \( \| B_{b\text{ prior}}^T B_{e\text{ prior}} \|_F \) to encourage orthogonality of the spanned subspaces for the purpose of signal separation.

Based on the \( \ell_2,1 \)-norm promoted sparse entries of \( C_{b\text{ prior}} \) and \( C_{e\text{ prior}} \), we perform an explicit column reduction to \( B_{b\text{ prior}} \) and \( B_{e\text{ prior}} \) according to the row structures in the corresponding coefficient matrices, and denote the column-reduced bases as \( B_{b\text{ prior\_sub}} \) and \( B_{e\text{ prior\_sub}} \).

2.4. Reconstruction

We use an \( \ell_2 \) penalty on estimating BCG coefficients to encourage the reconstructed BCG to have similar temporal behavior to the simultaneously acquired BCG reference data \( X_{b\text{ prior}} \) (c.f., Figure 1), to best capture the strong temporal variation present in BCG artifacts, as illustrated in Figure 3. With the learned basis vectors \( (B_{b\text{ prior\_sub}} \) and \( B_{e\text{ prior\_sub}} \), we reconstruct the signal components by seeking \( C_{b} \) and \( C_{e} \) to

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} \| B_{b\text{ prior\_sub}} C_{b} + B_{e\text{ prior\_sub}} C_{e} - X \|_F^2 \\
\text{subject to} & \quad \| C_{b\text{ prior}} - C_{b} \|_F \leq \sigma, \\
\end{align*}
\]

where \( \sigma \) is a parameter controlling the resemblance of the BCG coefficients. Subsequently, we recover the BCG and EEG components of the contaminated data \( Y \) with

\[
\begin{align*}
\hat{X}_b = B_{b\text{ prior\_sub}} C_b, \quad \hat{X}_e = B_{e\text{ prior\_sub}} C_e.
\end{align*}
\]

We evaluate the reconstruction performance by comparing the reconstructed BCG and EEG signal with the ground-truth data used to generate the simulation, using relative error \( RE \) when \( RE = \| \hat{X} - X_{\text{true}} \|_F / \| X_{\text{true}} \|_F \).

![Fig. 3. BCG power from channel A in time-frequency domain reveals the non-stationary characteristics of the BCG artifacts](image)

3. RESULTS AND DISCUSSION

3.1. Assessment of validity of OBS assumptions

The experiment of Sec. 2.2 provides observations of BCG-only and EEG-only signals, and offers an opportunity to examine the validity of the assumptions of the widely used OBS method and its variants. To this end, we assess the degree of orthogonality presented in the BCG and EEG subspaces. We apply PCA to the BCG-only and EEG-only data respectively and calculate the orthogonality index,

\[
\text{index}_{\text{orthogonality}} = \frac{\| B_e^TB_e \|_F}{\| B_e^TB_b \|_F},
\]

between the principal component matrices. We evaluate this index for different choices of included PCs.

Figure 4 shows that the orthogonal index starts from 0.128 rather than 0, suggesting that the BCG and EEG subspaces are not orthogonal even if represented with a single basis vector.

We further compare the feasibility of the bases from the DRJIB and the OBS method by calculating the residual energy by projecting ground-truth BCG/EEG-only data onto the basis vectors. The bases from the OBS method leave as much as 64.19% residual energy in representing the BCG-only data, and 48.12% for the EEG-only signals when the first 3 PCs from the contaminated data are considered as the BCG bases and the remaining PCs are used as the EEG bases. These large residuals confirm the model mismatch introduced from deriving BCG/EEG bases from contaminated data with improper assumptions of subspace relationships. By contrast, our jointly learned bases leave only 2.579x10^{-5} % and 1.373x10^{-13} % residual energies for the BCG and EEG signals, demonstrating the significant benefit of deriving bases from pure signals.

![Fig. 4. Orthogonality indices vs. number of PCs included.](image)

3.3. Reconstruction results

Figure 5 shows the reconstruction results from the DRJIB and the OBS methods. In the DRJIB reconstruction, we used 73 (out of 73) of the learned BCG basis and 23 (out of 73) of the learned EEG basis. The results of the OBS method use typical default parameters (3 PCs for BCG and the remaining 70 PCs for EEG reconstruction). These reconstructions are conducted without further residual removal techniques.

Without further parameter tuning, our results already have produced approximately 12-fold improvement in decreasing the relative errors compared to that of the OBS method. Furthermore, Figure 5 clearly illustrates that the DRJIB method successfully recovers the qualitative behavior of the EEG signal while the result from the conventional OBS method is hard to interpret. Unlike the OBS method that suffers from model mismatch based on the orthogonal subspace assumption, the DRJIB learns the bases faithfully from proper prior data and performs...
reconstruction on a different data set than the priors, showing no or minimal overfitting from model mismatch.

Such significant improvement may be attributed to the relaxation of OBS assumptions of the direct orthogonality and sequential diminishing energy assumption, the coupling (or rather decoupling) by learning the bases jointly, the increased robustness with the lower-dimensional characterization of each signal space and, lastly the integration of prior knowledge via regularization in the reconstruction.

**Fig. 5.** Comparison of the results from the DRJIB and the OBS. The top and bottom figures display BCG and EEG signals, respectively. The first, middle and bottom rows in each figure show, respectively, ground-truth signals used in simulation, reconstructed signals from the DRJIB and reconstructed signals from the OBS. The DRJIB yields only 5.699% and 11.75% relative errors for the reconstructed BCG and EEG signals, while the OBS generates 67.33% and 149.2% relative errors for BCG and EEG.

**Acknowledgement:** NIH NIBIB, AACR, TRDRP. Thanks to Wotao Yin for insightful discussions on the incoherence design.

**4. REFERENCES**


