Removal of Ballistocardiogram Artifact from Electroencephalograms Exploiting Heart Rate Variability

A. Kosma, K. Nazarpour, and S. Sanei

Centre of Digital Signal Processing, Cardiff University, Wales, UK

ABSTRACT

An approach for eliminating the cardiac pulsation (ballistocardiogram, BCG) effect from electroencephalograms (EEG) is proposed. The artifact results from functional magnetic resonance imaging (fMRI) scanner during simultaneous EEG-fMRI recordings. The BCG artifact exhibits similar periodic physiological structure. The proposed blind source extraction (BSE) algorithm employs second order statistics to exploit the heart rate variability and focuses on the contaminated EEG segments. This method is compared with some previously developed, well-known algorithms such as FastICA, SOBI and JADE, and a qualitative comparison is given. The methods are examined for both synthetic and real data.

1. INTRODUCTION

Simultaneous EEG and fMRI exploit high temporal resolution of EEG as well as high spatial resolution of fMRI for the study of both functional and cognitive abnormalities of the brain. However, using this technique, EEG is affected by different fMRI scanner artifacts including BCG artifact, which is mainly caused by cardiac pulsation. BCG waveforms are so prominent that they greatly obscure the EEG signals. Mitigating this effect has been under research recently. An early attempt has been made by [1] based on average artifact subtraction (AAS), which creates a template of the artifact that is then subtracted from the EEG signal. However AAS relies on the assumption that BCG waveforms are similar in all scans. This is usually not the case and the pattern is quasi-periodic following the corresponding electrocardiogram (ECG) trend, having time varying period. Adaptive filtering has also been used by [2] and [3]. In the case of adaptive filtering having a good quality reference channel is essential for effective noise cancellation. Also very high temporal resolution of the signals is needed which means high sampling rates, as well as a long period of recordings. Recently, independent component analysis (ICA) has been used for removal of the BCG artifact from EEG signals [4]-[8]. PCA has also been used for BCG artifact removal. In a paper by [9] principal component analysis (PCA) is combined with AAS. A method based on temporal PCA estimates an optimal basis set that contains templates of the artifact's temporal structure, such as shape and amplitude of BCG peaks, at different time intervals to compensate with its variability over time. The most appropriate template is then chosen on a peak-to-peak basis and is subtracted from the EEGs. Compared to blind source separation (BSS) methods based on ICA, a blind source extraction (BSE) approach is suggested here which is capable of extracting only the Source of Interest (SoI) by

exploiting the heart rate variability. The BSE algorithm used here allows the use of both some prior information of the temporal structure and the mixing process of the desired sources. BCG artifact exhibits a degree of periodicity while the distance between successive BCG peaks slightly changes from beat to beat.

2. PROPOSED BSE ALGORITHM

The simplest BSE model assumes the existence of N unknown spatially uncorrelated sources, $s_1(t), s_2(t), \dots, s_m(t)$ and the observations of sensor signals or mixtures $x_1(t), x_2(t), \dots, x_n(t)$. Here, for simplicity, the number of sources is considered to be the same as the number of observations i.e. $N=n=m$. The mixtures are assumed to be linear and instantaneous i.e. $x_i(t) = \sum_{j=1}^n a_{ij}s_j(t)$ for each $i=1, \dots, n$.

The mixing process is therefore:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (1)$$

where $\mathbf{A}$ is an $N \times N$ non-singular mixing matrix. The BSE algorithm estimates the column vectors $\mathbf{w}^T$ such that:

$$y(t) = \mathbf{w}^T \mathbf{x}(t) \quad (2)$$

where $y(t)$ is a scaled estimate of the SoI. Since $\mathbf{x}(t)$ in (2) is a linear combination of the column vectors of $\mathbf{A}$, the i th source can be extracted by projecting $\mathbf{x}(t)$ onto direction $\mathbf{a}_i$ along the space span $\{\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_N\}$, that is a space orthogonal to all columns of $\mathbf{A}$ except $\mathbf{a}_i$ [10]. By defining vector $\mathbf{q} \perp \{\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_N\}$ and adopting oblique projector notation [11]:

$$y(t)\mathbf{a}_i = \mathbf{E}_{\mathbf{a}_i | \mathbf{q}_i^\perp} \mathbf{x}(t) \quad (3)$$

where $\mathbf{q}_i^\perp$ is the space spanned by $\{\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_N\}$ orthogonal to $\mathbf{q}$ and $\mathbf{E}_{\mathbf{a}_i | \mathbf{q}_i^\perp}$ is the oblique projector onto direction $\mathbf{a}_i$ along $\mathbf{q}_i^\perp$. By setting the directions $\mathbf{a}_i \equiv \mathbf{t}$ and $\mathbf{q}_i^\perp \equiv \mathbf{q}^\perp$ Eqn. (3) becomes:

$$y(t)\mathbf{t} = \mathbf{E}_{\mathbf{t} | \mathbf{q}^\perp} \mathbf{x}(t) \quad (4)$$

After substituting $\mathbf{E}_{\mathbf{t} | \mathbf{q}^\perp} = (\mathbf{t}\mathbf{q}^T / \mathbf{q}^T \mathbf{t})$ [11] in (4) and omitting the immaterial scalar $(1/\mathbf{q}^T \mathbf{t})$, the resulting equation is given by:

$$y(t) = \mathbf{q}^T \mathbf{x}(t) \quad (5)$$

At this stage the purpose of the algorithm is to iteratively estimate the vectors $\mathbf{t}$ and $\mathbf{q}$ in such a way that the extracted source is as much uncorrelated as possible with the rest of the components that are in line with the space span $\{\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_N\}$. Assuming that the sources are spatially uncorrelated the autocorrelation matrix $\mathbf{R}_\tau$ can be defined by [11]:

$$\mathbf{R}_\tau = E[\mathbf{x}(t)\mathbf{x}^T(t-\tau)] = \sum_{i=1}^N r_i(\tau) \mathbf{a}_i \mathbf{a}_i^T \quad (6)$$

where $r_i(\tau) = E[s_i(t)s_i(t-\tau)]$ and $\tau = 0, 1, \dots, K$ is the autocorrelation function of the i th source at time lag τ . A source can be perfectly extracted if $\mathbf{R}_\tau \mathbf{q}$ is orthogonal to the space spanned by $\{\mathbf{a}_1, \dots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \dots, \mathbf{a}_N\}$ and consequently in line with direction $\mathbf{t}$ or $\mathbf{a}_i$. Vectors $\mathbf{t}$ and $\mathbf{q}$ can be estimated by minimizing the cost function given by [11]:

$$J = (\mathbf{t}, \mathbf{q}, \mathbf{d}) = \sum_{\tau=0}^K \|\mathbf{R}_\tau \mathbf{q} - d_\tau \mathbf{t}\|^2 \quad (7)$$

where $\mathbf{d} = [d_0, d_1, \dots, d_K]$ is a vector of unknown scalars that defines the length of vector $\mathbf{t}$ and $\|\cdot\|$ denoted the Euclidean norm. In order to find the optimum $\mathbf{q}$ and extract one source as in Eqn. (2) the cost function in Eqn. (7) should be minimized with respect to $\mathbf{q}$. The estimation error $\mathbf{e}_\tau$ is defined as the difference between vectors $\mathbf{R}_\tau \mathbf{q}$ and $d_\tau \mathbf{t}$ as given by:

$$\mathbf{e}_\tau = \mathbf{R}_\tau \mathbf{q} - d_\tau \mathbf{t} \quad (8)$$

The cost function in (7) can then be rewritten as $J = \sum_{\tau=0}^K \mathbf{e}_\tau^T \mathbf{e}_\tau$. By differentiating it with respect to $\mathbf{q}$ and noting that J attains its minimum when all its elements are simultaneously zero, the optimal $\mathbf{q}$ is given by:

$$\frac{\partial J}{\partial \mathbf{q}} = 2 \sum_{\tau=0}^K \mathbf{R}_\tau \mathbf{e}_\tau = 0 \quad \text{for } \tau = 0, 1, \dots, K \quad (9)$$

By employing sequential AJD the cost function J can be minimized by alternatively adjusting its parameters $\mathbf{q}$, $\mathbf{t}$ and $\mathbf{d}$. This process can be completed in three steps as follows [11]:

- First, both $\mathbf{t}$ and $\mathbf{d}$ remain constant while $\mathbf{q}$ is iteratively adjusted to get the optimal vector $\mathbf{q}$ as can be seen in equation (10) below

$$\mathbf{q} \leftarrow \mathbf{H} \left(\sum_{\tau=0}^K d_\tau \mathbf{R}_\tau \right) \mathbf{t} \quad (10)$$

where $\mathbf{H} = \left[\sum_{\tau=0}^K \mathbf{R}_\tau^2 \right]^{-1}$ and $a \leftarrow b$ denotes replacing a by b .

- Second, both $\mathbf{q}$ and $\mathbf{t}$ remain constant and $\mathbf{d}$ is adjusted. In order to avoid the immediate global point of the cost function of equation (7) that is for $\mathbf{t} = \mathbf{q} = \mathbf{0}$, $\mathbf{d} = \mathbf{0}$, $\mathbf{t}$ and $\mathbf{d}$ are constrained as $\|\mathbf{t}\| = 1$ and $\|\mathbf{d}\| = 1$.

For an optimal $\mathbf{d}$ and considering the constraint $\|\mathbf{d}\| = 1$ the following Lagrange function can be minimized:

$$J_{\lambda_d} = J + \lambda_d \left(\sum_{\tau=0}^K d_\tau^2 - 1 \right) \quad (11)$$

where λ_d is the Lagrange multiplier. Differentiating with respect to $\mathbf{d}$, the optimal $\mathbf{d}$ can be estimated when:

$$\frac{\partial J_{\lambda_d}}{\partial d_\tau} = -2r_\tau^T t + 2(\lambda_d + 1)d_\tau = 0, 0 \leq \tau \leq K \quad (12)$$

By freezing both $\mathbf{q}$ and $\mathbf{t}$, $\mathbf{d}$ is calculated iteratively as

$$\mathbf{d} \leftarrow \mathbf{d} / \|\mathbf{d}\| \quad (13)$$

where $\mathbf{d} = [\mathbf{r}_0^T t, \mathbf{r}_1^T t, \dots, \mathbf{r}_K^T t]^T$ and $\mathbf{r}_\tau = \mathbf{R}_\tau \mathbf{n}$.

- Third, $\mathbf{q}$ and $\mathbf{d}$ are kept constant and $\mathbf{t}$ is adjusted. Taking into account the constraint $\|\mathbf{t}\| = 1$ the cost function to be minimized becomes the expression:

$$J_\lambda = J + \lambda_t (\mathbf{t}^T \mathbf{t} - 1) \quad (14)$$

Again, differentiating with respect to $\mathbf{t}$ and

$$\frac{\partial J_\lambda}{\partial \mathbf{t}} = -2 \sum_{\tau=0}^K d_\tau \mathbf{r}_\tau + 2(\lambda_t + 1)\mathbf{t} = 0$$

and normalizing as $\mathbf{t} = \mathbf{v} / \|\mathbf{v}\|$, where $\mathbf{v} = \sum_{\tau=0}^K d_\tau \mathbf{r}_\tau$, the optimal $\mathbf{t}$ is achieved. Following the above process the optimal vectors $\mathbf{q}$, $\mathbf{d}$ and $\mathbf{t}$ can be alternatively estimated. The process can be repeated until convergence and through equation (2) one source can be extracted. After the extraction of one source a deflation process follows to remove this source from the rest of the mixtures such that re-estimation of the same source is avoided. In the same way one by one, part or all the sources can be estimated. The deflation process updates the autocorrelation matrices by removing the extracted source as [12]:

$$\mathbf{R}_\tau^{\text{new}} = \mathbf{B}^T \mathbf{R}_\tau \mathbf{B} \quad (0 \leq \tau \leq K) \quad (15)$$

where $\mathbf{B} = \mathbf{I} - \frac{\mathbf{R}_\tau \mathbf{w}_i \mathbf{w}_i^T}{\sigma_{y_i}^2}$ and $\mathbf{I}$ is the identity matrix.

The method used here incorporates the time varying period of BCG in equation (7) after detecting the time instants where it occurs as [13][14]:

$$J = (\mathbf{t}, \mathbf{q}, \mathbf{d}) = \sum_{p=1}^K \|\mathbf{R}_{p\tau_i} \mathbf{q} - d_\tau \mathbf{t}\|^2 \quad (16)$$

where p_{τ_i} denotes the time varying period. In this case the set autocorrelation matrices are calculated for time lags τ_i .

Peak identification technique— A simple threshold peak identification technique [15] is firstly applied on EEG signals. It tracks the time varying period τ_i of BCG peaks by detecting their locations and incorporate it into BSE algorithm. Six of the EEG channels with clear BCG artifacts are chosen and averaged to obtain one average complex channel. Low pass filtering is applied to the complex channel to smooth the signal and emphasize the

peaks. Also, all the samples with negative amplitudes are set equal to zero to reduce the redundant peaks. A sample is identified as a peak if it has the maximal value but also if it is preceded to the left by a sample value lower than the threshold. A peak correction technique follows in order to correct any falsely detected peaks. Peaks that are detected within the time interval between two consecutive BCG peaks are identified as false and are deleted. As shown by Figure 1 the technique yields satisfactory results as all the peaks have been correctly identified.

Figure 1: Complex channel (blue), complex channel after preprocessing (black), detected peaks (red).

3. EXPERIMENTAL RESULTS

The results of the proposed BSE algorithm as well as those of FastICA, SOBI and JADE algorithms are presented here for both synthetic and real EEG signals. The synthetic signal consisted of a sine wave, a part of an EEG signal, random noise, a saw-tooth and a pulse train waveform, as shown in Figure 2, and the real EEG signal consisted of 62 channels containing only the BCG artifact. Some chosen channels are shown in Figure 5.

Application of BSE algorithm on synthetic signal— In the first stage the sources were linearly mixed by a random matrix, as shown in Figure 3. The results on the synthetic signal, after applying all the algorithms, in terms of the Euclidian distance between the original and estimated sources are shown in Table 1.

Figure 2: Synthetic source signals.

Since the sources were linearly combined and uncorrelated, all the algorithms performed well. As can be seen in Figure 4 the estimated sources by BSE have minor difference with the original sources in Figure 2. In the second stage the performance of the algorithms is tested

again by adding nonlinearity to the mixing process, according to which the mixing matrix is altered slightly every certain number of samples. As expected, the performance of all the algorithms degraded. The difference between the original and the estimated sources after adding nonlinearity in the mixing process are shown in Table 2.

Table 1: Comparison between the original and estimated sources in terms of Euclidean norm using different algorithms.

norm	FastICA	JADE	SOBI	BSE
Sinusoid	0.8417	4.2784	0.1395	0.2441
EEG	1.5242	1.3304	0.6352	0.5638
Noise	2.7725	1.1314	2.4612	2.3819
Sawtooth	2.1483	7.3676	1.1838	1.8352
Pulse train	0.2022	0.3355	0.1232	0.1153

Figure 3: Mixtures of the synthetic signals.

Figure 4: Estimated sources using the BSE algorithm after linear mixing.

Table 2: Difference between the original and estimated sources after adding nonlinearity.

norm	FastICA	JADE	SOBI	BSE
Sinusoid	5.6214	5.862	2.8327	2.6419
EEG	2.2015	2.5237	2.4327	1.8165
Noise	5.6634	21.829	0.9581	5.7092
Sawtooth	1.3532	7.5995	8.2314	2.1203
Pulse train	5.1817	2.0063	2.0846	3.1986

Application of BSE algorithm on real EEG signal— The above four algorithms were also tested on real EEG signal. The subjective results for a number of chosen channels are shown in Figure 6.

Figure 5: EEG signals contaminated by BCG artifact

Figure 6: Cleaned EEG signal using the BSE algorithm.

The BSE algorithm using the time varying period of the BCG extracted one source shown in Figure 11. It is obvious that the estimated source follows the pattern of the BCG artifact and has similar physiological characteristics and temporal structure. The estimated artifact was subtracted from the mixtures by the deflation process (15).

Figure 7: Extracted BCG artifact, using the BSE algorithm.

4. CONCLUDING REMARKS

The BCG artifact has been mitigated by the proposed BSE algorithm. The remaining BCG might be due to the fact that the BCG artifact tends to exhibit opposite polarity in the left and right hemisphere of the head. The algorithm tries to estimate the BCG artifact as the source of interest and thus sometimes the estimated source might not represent some BCG components in some of the channels at the same time. The difference between the BSE algorithm and other BSS algorithms is that it can trace the time varying period of the artifact and therefore results in a more accurate artifact identification and subtraction. A peak detection technique was employed to identify the BCG peaks and estimate the lags.

Acknowledgement—The authors wish to acknowledge Dr Andrew Bagshaw for providing the data.

REFERENCES

- [1] P. J. Allen, G. Polizzi, K. Krakow, D. R. Fish, and L. Lemieux, "Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction," *Neuroimage* 8 (3), pp. 229–239, 1998.
- [2] G. Bonmassar, P. Purdon, I. Jaaskelainen, K. Chiappa, V. Solo, E. Brown and J. Belliveau, "Motion and ballistocardiogram artifact removal for interleaved recording of EEG and EPs during MRI," *Neuroimage*, 16, pp. 1127–41, 2002
- [3] H. Myung, Y. Soo, and B. Young, "Ballistocardiogram artifact removal from EEG signals using adaptive filtering of EOG signals," *Physiol. Meas.* 27, pp. 1227–1240, 2006.
- [4] G. Srivastava, S. Crottaz-Herbette, K. M. Lau, and G.H. Menon, "ICA-based procedures for removing ballistocardiogram artifacts from EEG data acquired in the MRI scanner," *Neuroimage* 24 (1), pp. 50 - 60, 2005.
- [5] W. Nakamura, K. Anami, T. Mori, O. Saitoh, A. Cichocki, and S. Amari, "Removal of ballistocardiogram artifacts from simultaneously recorded EEG and fMRI data using independent component analysis," *IEEE Trans. Biomedical Engineering* 53 (7), pp. 1294–1308, 2006.
- [6] D. Mantini, M. G. Perrucci, Cugini, S., Ferretti, A., Romani, G.L., Del Gratta, C., 2007, "Complete artifact removal for EEG recorded during continuous fMRI using independent component analysis," *Neuroimage* 34 (2), 598 - 607.
- [7] E. Briselli, G. Garreffa, L. Bianchi, M. Bianciardi, E. Macaluso, M. Abbafati, M. Grazia Marciani, and B. Maraviglia, "An independent component analysis-based approach on ballistocardiogram artifact removing," *Magnetic Resonance Imaging* 24 (4), pp. 393 - 400, 2006.
- [8] C. G. Bénar, Y. Aghakhani, Y. Wang, A. Izenberg, A. Al-Asmi, F. Dubeau, and J. Gotman, "Quality of EEG in simultaneous EEG–fMRI for epilepsy," *Clin. Neurophysiol.* 114, pp. 569–580, 2003.
- [9] R. Niazy, C. Beckmann, G. Iannetti, J. Brady, and S. Smith, "Removal of FMRI environment artifacts from EEG data using optimal basis sets," *Neuroimage* 28 (3), 720–737, 2005.
- [10] T. Tsalaile S. M. Naqvi, K. Nazarpour, S. Sanei and J. Chambers, "Blind Source Extraction of heart sound signals from lung sound recordings exploiting periodicity of the heart sound," *Proc. EUSIPCO*, 2008.
- [11] X. Li and X. Zhang, "Sequential blind extraction adopting second-order statistics," *IEEE Signal Process. Lett.*, vol. 14, pp. 58–60, 2007.
- [12] A. Cichocki and S. Amari, *Adaptive blind signal and image processing: Learning algorithms and applications*, Wiley, 2002.
- [13] T. Tsalaile, R. Sameni, S. Sanei, C. Jutten and J. Chambers, "Sequential Blind Source Extraction For Quasi-Periodic Signals With Time-Varying Period" To appear in *IEEE Trans. Biomed. Eng.*, available in IEEE Xplore.
- [14] K. Nazarpour, H. Mohseni, C. Hesse, J. Chambers, and S. Sanei, "A novel semiblind signal extraction approach for the removal of eye-blink artifact from EEGs," *EURASIP J. Adv. Sig. Proc.*, vol. 2008. doi: 10.1155/2008/857459.
- [15] E. Billauer, *Peakdet: Peak detection using MATLAB*. [WWW] <http://www.billauer.co.il/peakdet.html> [Accessed 11 August 2008].