Journal of Neuroscience Methods 262 (2016) 14–20

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Basic Neuroscience

Sparse time artifact removal

Alain de Cheveigné [a,b,c,*]

[a] Laboratoire des Systèmes Perceptifs, UMR 8248, CNRS, France
[b] Département d'Etudes Cognitives, Ecole Normale Supérieure, France
[c] UCL Ear Institute, United Kingdom

HIGHLIGHTS

- The STAR algorithm addresses channel-specific noise that is sparse in time.
- It removes electrode or sensor noise and certain forms of myogenic artifact.
- In contrast to other techniques, few data are lost and the dimensionality of the data is preserved.
- The STAR algorithm complements component analysis techniques such as ICA.

ARTICLE INFO

Article history:
Received 25 October 2015
Received in revised form
23 November 2015
Accepted 2 January 2016
Available online 8 January 2016

Keywords:
EEG
MEG
LFP
ECoG
Artifact
Myogenic
ICA
Sensor noise

ABSTRACT

Background: Muscle artifacts and electrode noise are an obstacle to interpretation of EEG and other electrophysiological signals. They are often channel-specific and do not fully benefit from component analysis techniques such as ICA, and their presence reduces the dimensionality needed by those techniques. Their high-frequency content may mask or masquerade as gamma band cortical activity.
New method: The sparse time artifact removal (STAR) algorithm removes artifacts that are sparse in space and time. The time axis is partitioned into an artifact-free and an artifact-contaminated part, and the correlation structure of the data is estimated from the covariance matrix of the artifact-free part. Artifacts are then corrected by projection of each channel onto the subspace spanned by the other channels.
Results: The method is evaluated with both simulated and real data, and found to be highly effective in removing or attenuating typical channel-specific artifacts.
Comparison with existing methods: In contrast to the widespread practice of trial removal or channel removal or interpolation, very few data are lost. In contrast to ICA or other linear techniques, processing is local in time and affects only the artifact part, so most of the data are identical to the unprocessed data and the full dimensionality of the data is preserved.
Conclusions: STAR complements other linear component analysis techniques, and can enhance their ability to discover weak sources of interest by increasing the number of effective noise-free channels.

1. Introduction

Among the many sources of noise and artifact that plague studies involving human and animal electrophysiology, some affect only one channel at a time. This paper addresses such *channel-specific* artifacts, leaving aside other types that impinge on several channels such as eyeblink, heartbeat or background neural activity. Signals recorded by multichannel recording techniques such as electroencephalography (EEG), magnetoencephalography (MEG), electrocorticography (ECoG), invasive electrode arrays or optical techniques are related to underlying sources by a linear mixing process:

$$x_j(t) = \sum_i s_i(t) u_{ij}, \tag{1}$$

where t is time, $s_i(t)$ are the source signals, and u_{ij} are mixing coefficients. We call "channel-specific" a source s_i for which the u_{ij} are zero for all channels j except one.

Channel-specific noise includes electrode contact noise, pulsation noise, and certain forms of muscle artifact in EEG, as well as sensor noise in MEG or other techniques. Electrode-tissue contact artifacts are usually temporally sparse, occurring as isolated events or bursts of events that affect one channel at a time.

* Correspondence to: Audition, DEC, ENS, 29 rue d'Ulm, 75230 Paris, France. Tel.: +33 144322672/+44 7912504027.
E-mail address: Alain.de.Cheveigne@ens.fr

http://dx.doi.org/10.1016/j.jneumeth.2016.01.005
0165-0270/

Muscle artifacts may also be channel-specific if they are produced by motor units proximal to a single electrode (Yilmaz et al., 2014), although activity from deeper muscles may affect multiple electrodes. Such artifacts often have a spectrum that extends to high frequency, but they may also include low-frequency components that overlap with low-frequency cortical activity (Goncharova et al., 2003), so that spectral filtering is not sufficient to eliminate them. Electromyogenic noise is particularly troublesome as it can be confused with high-frequency cortical activity (Yuval-Greenberg et al., 2008; Whitham et al., 2007; Muthukumaraswamy, 2013; Yilmaz et al., 2014), particularly as muscle activity may correlate with cognitive state (Whitham et al., 2008; Yuval-Greenberg et al., 2008; McMenamin et al., 2011). Cephalic skin potential (Corby et al., 1974) may also covary with cognitive state, and contact noise may correlate with behavior.

A standard approach to dealing with spatio-temporally sparse artifacts is to discard either the offending channel, or the offending time interval or trial (Junghöfer et al., 2000). This entails loss of data, particularly if artifacts affect multiple channels and/or are widely distributed in time, and it also complicates analysis and interpretation stages, that need to be made tolerant to the missing data.

Another approach is to apply multichannel linear analysis techniques such as independent component analysis (ICA), beamforming, or joint diagonalization (JD) (de Cheveigné and Parra, 2014) to isolate noise components. Component signals $y_k(t)$ obtained by these methods are related to observations as:

$$y_k(t) = \sum_j x_j(t) w_{jk}, \tag{2}$$

where t is time and the w_{jk} are weights. The J observation channels span a space that contains all such linear combinations, and the *dimensionality* of the data is the number of dimensions of this space (J or less). Components belong to this space. Different methods (PCA, ICA, beamforming, etc.) differ in how they find the appropriate weights to apply to the data. ICA in particular has been proposed to remove artifacts including myogenic (Delorme et al., 2007; Ma et al., 2012; Crespo-Garcia et al., 2008).

The appeal of these linear techniques is that a noise source x_i can potentially be perfectly canceled: a component y_k is insensitive to source i as long as $\sum_j u_{ij} w_{jk} = 0$, where u_{ij} are the mixing coefficients and w_{jk} the component weights (Eqs. (1) and (2)). With high-dimensional data (lots of channels) there is considerable flexibility in satisfying this constraint, and the strength of analysis algorithms such as ICA lies in their ability to find such sets of weights. However, if a noise source is specific to a single channel j, it can only be cancelled by setting w_{jk} to zero for every component k, effectively discarding that channel. In this situation, component analysis offers little over the time-honored practice of discarding noisy channels.

Component analysis itself is vulnerable to channel-specific noise because it relies on the dimensionality of the data (determined by the number of channels) to resolve the various sources. If channels are discarded due to artifacts, analysis may be impaired, whereas if they are *not* discarded (due to lack of knowledge or the need to conserve enough channels), the artifact is injected into the extracted components via Eq. (2). The presence of artifacts may also interfere with the ability of the algorithm to find the optimal w_{jk}. For example an algorithm such as CSP (Koles et al., 1990), that searches for components that differ in power between two intervals, may lock on to an artifact that is present in one interval but not the other. Finally, the artifacts may interfere with the ability to estimate the topography associated with cortical activity, possibly compromising source modeling. Channel-specific noise thus limits the ability of linear methods to improve the signal to noise ratio (SNR) of weak brain activity.

These considerations lead us to focus on channel-specific noise, leaving other techniques such as ICA, JD, or beamforming to deal with noise sources that impinge on multiple electrodes. This is important scientifically, to obtain a more accurate picture of brain activity, and also for applications such as brain-computer interfaces (BCI), prediction of epileptic seizures, wearable brain-monitoring devices, and so on. The method described here is effective, fully automatic, and amenable to an online implementation for applications that involve realtime monitoring.

2. Methods

2.1. Signal model and assumptions

Each observation $x_j(t)$ is the sum of signals from multiple sources i within the brain or the environment (Eq. (1)). We make several restrictive assumptions. Each noise source $n_i(t)$ affects only one particular channel (*assumption 1*). Noise activity is temporally sparse so that artifacts on different channels do not temporally overlap (*assumption 2*), and for a significant proportion of time the data are artifact-free (*assumption 3*). Finally, we assume that in the absence of artifacts the data are *linearly dependent* such that each channel belongs to the subspace spanned by the other channels. In other words for each channel j there exist $a_{jj'}$ such that $x_j(t) = \sum_{j' \neq j} a_{jj'} x_{j'}(t)$ (*assumption 4*). This is plausible for neurogenic activity due to source-to-sensor mixing. Of course, many kinds of noise do *not* meet these assumptions; the focus here is on those that do. In real data, these assumptions will be met only as an *approximation*, for example because of non-stationarity of the brain and noise processes underlying the data. The quality of the outcome depends on the quality of the approximation.

2.2. The STAR algorithm

The algorithm proceeds in two phases. The first phase detects the presence of channel-specific artifacts, the second phase corrects them.

Phase 1. The covariance matrix of the data is estimated, and from it is calculated the matrix **A** that projects each channel on the subspace spanned by the other channels. The projection $\bar{x}_j(t)$ of channel j is the weighted sum of the channels $j' \neq j$ that best approximates $x_j(t)$. In the absence of an artifact we should have $\bar{x}_j(t) - x_j(t) = 0$ as a result of the linear dependence assumption, so a significant deviation indicates the presence of an artifact. Values of $\bar{x}_j(t) - x_j(t)$ are fit by a zero-mean Gaussian distribution, and values eccentric from this distribution (relative to a predefined threshold θ) are flagged as artifactual. This is repeated for all channels, and the union of eccentric time samples is labeled as artifact-contaminated. The covariance matrix is initially estimated from the entire data, and subsequently reestimated on the part of data labeled as artifact-free. A few iterations of this process lead to a stable partition of the time axis between artifact-free and artifact-contaminated parts.

Phase 2. The artifact-contaminated part is further divided according to which channel is most degraded at each time sample. For this purpose, a second eccentricity measure is calculated for each channel as the ratio of instantaneous power to power averaged over the artifact-free portion. The channel with the highest score at a given time sample "owns" that sample, and its data are replaced with its projection on the subspace spanned by the other channels, using projection coefficients calculated from the artifact-free part. Data replacement occurs only at the part corresponding to the artifact: at other times the data are left intact, so most of the data remain untouched by the processing.