NeuroImage xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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

Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic resonance imaging (fMRI) using constraint independent component analysis (cICA)

Yves Leclercq [a,*], Evelyne Balteau [a], Thanh Dang-Vu [a,b], Manuel Schabus [a,c], André Luxen [a], Pierre Maquet [a,b], Christophe Phillips [a]

[a] Cyclotron Research Centre, University of Liege, Belgium
[b] Liege University Hospital, University of Liege, Belgium
[c] Division Physiological Psychology, University of Salzburg, Austria

ARTICLE INFO

Article history:
Received 5 June 2008
Revised 31 August 2008
Accepted 10 October 2008
Available online xxxx

ABSTRACT

Rejection of the pulse related artefact (PRA) from electroencephalographic (EEG) time series recorded simultaneously with fMRI data is difficult, particularly during NREM sleep because of the similarities between sleep slow waves and PRA, in both temporal and frequency domains and the need to work with non-averaged data. Here we introduce an algorithm based on constrained independent component analysis (cICA) for PRA removal. This method has several advantages: (1) automatic detection of the components corresponding to the PRA; (2) stability of the solution and (3) computational treatability. Using multichannel EEG recordings obtained in a 3 T MR scanner, with and without concomitant fMRI acquisition, we provide evidence for the sensitivity and specificity of the method in rejecting PRA in various sleep and waking conditions.

Introduction

In recent years, the combination of functional Magnetic Resonance Imaging (fMRI) and electroencephalography (EEG) has been increasingly used to characterize human brain function, with the objective of harvesting the high spatial resolution of the former and the exquisite temporal precision of the latter. Although also its use to characterize evoked response (Czisch et al., 2002; Liebenthal et al., 2003). The full benefit of combining EEG and fMRI is particularly obvious when one is interested in spontaneous brain states or events characterized by specific EEG oscillations or transients, the time course of which is hardly predictable. Accordingly, EEG–fMRI has proven particularly successful in characterizing the cerebral correlates of normal brain oscillations as waking brain rhythms (Goldman et al., 2000; Laufs et al., 2003; Moosmann et al., 2003; Niazy et al., 2004), or sleep transients (Schabus et al., 2007), as well as in localizing the hemodynamic correlates of epileptic activities (Seeck et al., 2001).

A major problem of EEG–fMRI consists of the contamination of EEG recordings by gradient switching (gradient artefacts) and pulse related artefacts (PRA). The former is related to the echoplanar sequence used in fMRI. The latter is thought to have multiple causes related to the interaction between the static magnetic field and the heart beat, such as pulsatile motion of recording electrodes, cardio-ballistic head movements and Hall effects in blood vessel (Allen et al., 1998; Goldman et al., 2000; Srivastava et al., 2005; Nakamura et al., 2006). Both artefacts are larger than the genuine EEG signal and have to be removed before EEG can be properly analyzed.

Because the gradient artefact is very reproducible across consecutive volume acquisitions, it is satisfactorily removed from raw data by subtracting an averaged artefact waveform (Average Artefact Subtraction (AAS)), followed by adaptive noise cancellation to reduce any residual artefact (Allen et al., 1998). In contrast, PRA is much harder to reject. Indeed, its morphology and its topography substantially vary from beat to beat. Because of this non stationarity, methods based on a simple AAS, although reasonably efficient, result in residual artefacts in the data because they assume that the PRA is a slowly changing signal that can be accurately characterized by a moving average. Various methods were derived from AAS modelled PRA using exponential weighted average (Goldman et al., 2000), median filter (Sijbers et al., 2000) or adaptive filters (Han et al., 2004; Wan et al., 2006) but eventually produced similar results (Grouiller et al., 2007).

To allow more flexibility in the characterization of PRA, other channel-wise model-explicit rejection algorithms were based on temporal Principal Component Analysis (PCA, Niazy et al., 2005) or a moving General Linear Model (mGLM, Vincent et al., 2007). Others used estimates of the motion artefact noise by piezoelectric motion

* Corresponding author. Fax: +32 4366 2946.
E-mail address: yves.leclercq@student.ulg.ac.be (Y. Leclercq).

1053-8119/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2008.10.017

sensors (Bonmassar et al., 2002; Masterton et al., 2007) and an adaptive filtering algorithm based on the Kalman filter (Bonmassar et al., 2002). These led to relatively efficient solutions but required extra equipment to be placed on and around the subject.

Another class of PRA removal techniques is based on blind source separation. Methods using Independent Component Analysis (ICA) decompose the EEG recordings into independent components (ICs) and suppress the PRA-related ones (Benar et al., 2003; Nakamura et al., 2006; Finelli et al., 2003; Han et al., 2004; Mantini et al., 2007). The problems of these methods are non-reproducibility of the results (Briselli et al., 2006; Grouiller et al., 2007) and the difficulty in objectively selecting the PRA-related ICs. To address the issue of ICA variability, it was suggested to apply the ICA algorithm iteratively and to average the results across multiple executions (MICA) but this procedure requires an overwhelming computational power (Briselli et al., 2006; Phillips et al., 2006) and is very time-consuming. To address the issue of IC selection, an automatic method based on the correlation between the IC and the ECG was developed (Srivastava et al., 2005) but this type of detection does not always perform well in detecting PRA-related ICs (Debener et al., 2007). Finally, it has also been argued that ICA could not be used to reject PRA for recordings obtained in a high-field scanner without additional pre-processing because ICA is a fixed spatial filter and the PRA was observed to have a variable within-cycle topography (Debener et al., 2008).

In the framework of sleep studies, the conditions for PRA rejection can be particularly difficult. During the deepest stages non rapid eye movement (NREM) sleep, also called slow-wave-sleep (SWS), EEG recordings are characterized by abundant high-amplitude slow waves. In that case, the difficulty in rejecting PRA from EEG recordings arises from the similarity between slow waves and PRA both in the temporal and frequency domains. Slow waves and PRA typically recur about every second and their amplitudes are in the same order of magnitude (typically several hundred μV peak to peak at 3 T). Similarly, the bulk of the spectral power of both sleep slow waves and PRA lies between 0.5 and 5 Hz. Consequently, the estimation of a model of PRA using classical methods (non-ICA based) is likely to be confounded by genuine neurally-generated slow waves and the ability of these methods to retrieve slow waves during NREM sleep is questionable.

Yet, slow waves of deep NREM sleep represent a fundamental brain activity associated with important functional properties. For instance, they have been used as a measure of sleep depth and their power density (between 0.75 and 4 Hz) is used as a reliable parameter quantifying the dissipation of the homeostatic sleep pressure (Dijk et al., 1997). Slow waves have also been implicated in offline memory processing either because they would allow for a local synaptic downscaling (Huber et al., 2004) or because they might promote the replay of neuronal firing sequence supposedly representing the learned material (Buzsaki, 1996).

In order to allow the study of SWS using EEG–fMRI, we developed a new efficient, robust, and computationally efficient PRA-rejection technique based on constrained ICA (cICA). This novel method is equally efficient in rejecting PRA for EEG data recorded during wakefulness and NREM sleep.

PRA rejection algorithm

In this section we will introduce our cICA-based PRA rejection algorithm as well as former ones which we will use for comparison. Since each of these techniques uses a QRS peak detection algorithm in ECG, we decided to use the same algorithm for each of them and describe this one in the next subsection. Note that we use the FMRIB plug-in for EEG, provided by the University of Oxford Centre for Functional MRI of the Brain (http://users.fmrib.ox.ac.uk/~rami/fmribplugin/index.html), for the practical implementation of the QRS peak detection and AAS and PCA-based methods for PRA rejection. We will then describe the available techniques used for comparison before introducing our new cICA algorithm.

QRS peak detection

The peak detection algorithm we use was developed by Niazy et al. (2005) and its performance has been evaluated in the same article. The ECG peak detection is a modification of the algorithm proposed by Christov (2004). In this algorithm, a complex lead is computed from several single ECG channels. Yet, in most EEG/fMRI application the ECG is measured by only one lead, so an alternative complex lead has to be found. This complex lead is computed using a *k*-Teager operator (Kim et al., 2004; Mukhopadhay and Ray, 1998) to the filtered (Bandpass filtered from 7 to 40 Hz) ECG channel and then setting all negative values to zero:

$$\mathrm{ECG}_{\mathrm{complex}}(t) = \max\left(E^2(t) - E(t-k)E(t+k), 0\right); \quad (1)$$

where $\mathrm{ECG}_{\mathrm{complex}}$ is the complex lead, t the time index, E the filtered ECG and k a frequency selection parameter (Kim et al., 2004). The peaks are then detected using the original adaptive threshold algorithm (Christov, 2004). This produces a binary vector **P** where the 1 in this vector indicates the peaks in the ECG channel.

False positive removal

The vector **P** is divided into segments of 20 s interval with an overlap of 5 s. For each segment, a peak is considered as false positive if the time between two peaks is smaller than $\mu_{RR} - 3\sigma_{RR}$ and **P** is corrected accordingly. μ_{RR} is the median of the R–R intervals of this segment and σ_{RR} is the standard deviation of the R–R intervals of this segment.

After **P** is corrected, the ECG is divided according to vector **P** and then the average is computed to form an average ECG template. This template is then taken as a reference and the peaks in the vector **P** are corrected to correspond to the maximum correlation between the ECG template and the original ECG.

False negative correction

As in the previous paragraph, the vector **P** is divided into 20 s segments with 5 s of overlap and σ_{RR} are then calculated for each segment. If an R–R interval exceeds 1.5 times μ_{RR} a peak is added in the **P** vector after a time μ_{RR} after the first peak of this interval. Once all false negatives were corrected, the peaks of vector **P** are adjusted to correspond to the maximum correlation between the ECG template computed using the new **P** and the original ECG as explained in the previous paragraph.

Note that the standard delay between an ECG peak and the occurrence of the PRA in EEG channel was observed to be 210 ms by Allen et al. (1998). The influence of the quality of ECG on QRS peak detection was investigated by Niazy et al. (2005).

Former PRA rejection algorithms

We have considered 3 reference methods (AAS, PCA and MICA) to assess the performance of our cICA-based correction method.

AAS

The PRA template in each EEG channel is computed by averaging the sections of EEG signals ± half the mean R–R interval centered on the ECG peaks delayed by the average time difference between the heart contraction and the occurrence of the PRA in the EEG. Before using an EEG section in the average computation, each section is tested for artefacts which would corrupt the PRA template. Sections of EEG with a mean amplitude 3 times higher than the minimum section are rejected. If fewer than two sections are artefact-free, the previous PRA template is used instead. Because high frequency phenomena are poorly rejected by an average subtraction, a first order polynomial is fitted to each section of EEG and subtracted before estimating the PRA template.

Once the PRA template is calculated, it is subtracted from the original EEG signal and this gives the corrected signal as a result.

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

PCA

After shifting the QRS peaks by 210 ms (to account for the standard delay between QRS complexes and artefact occurrence), a matrix **S** (dimension $p \times q$) is formed for each channel: each row is filled with one occurrence of the PRA. p is the number of artefact occurrences on this channel and q the number of points spanning each artefact interval.

Then PCA is performed on **S** such that the different variations of the artefact are captured by the principal components of **S** and ordered according to the variance explained by each component. The projection of **S** onto the principal component coefficients are referred to as PCs. The first 3 PCs are placed in matrix **B** (dimension $q \times 3$) and are considered as an optimal basis. Each segment of the original signal $\mathbf{Y}_j$ (which is equal to a row of **S**) can be expressed in term of **B** through a linear model:

$$\mathbf{Y}_j = \mathbf{B}\beta_j + \varepsilon_j \quad (2)$$

where β_j is a 3×1 parameter vector and ε_j is the error term for the segment j of the considered channel. The model is then fitted in a least squares sense to each row of **S** and subtracted to the corresponding artefact occurrence in the considered EEG channel. This process is repeated for each channel.

MICA

This section introduces ICA (for a more technical presentation see Bell and Sejnowski (1995); Hyvarinen (1999) and Hyvarinen (2000)) and MICA (cf. (Briselli et al., 2006; Phillips et al., 2006).

ICA stands for Independent Component Analysis and is used to decompose data (in our case, EEG recordings) as a linear combination of statistically independent sources. EEG data are represented by a matrix **X** of dimension $N_{ch} \times N_s$, where N_{ch} is the number of EEG channels and N_s the number of samples per channel. From a statistical point of view, EEG recordings are observations of $\mathbf{x} = (x_1, x_2, \ldots x_{N_{ch}})^T$, where each x_i is a random process. Each sample of the i^{th} row of the matrix **X** is thus considered as a realization of the x_i random process.

Let $\mathbf{x}=(x_1, x_2,\ldots, x_m)^T$ be the observations of an m-dimensional Random Process, and $\mathbf{s}=(s_1, s_2,\ldots, s_n)^T$ an n-dimensional random process so that the s_i are statistically independent from each other. Each s_i is called an Independent Component (IC) or a source. Each x_i is supposed to be a linear mixture of the s_i. Then the problem of ICA is to determine a constant weight (or demixing) matrix **W**, given **x** so that the following equation is verified:

$$\mathbf{s} = \mathbf{W}\mathbf{x} \quad (3)$$

If the mixing matrix is defined as $\mathbf{A}=\mathrm{W}^{-1}$ then:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (4)$$

Classically, ICA searches for as many statistically independent sources as there are for channels using an iterative scheme. Once all sources have been detected, artefact-related sources are identified and removed from EEG recordings by reconstructing EEG signals only from supposedly neural sources. In mathematical terms, the algorithm finds a demixing matrix **W** which is used to project the EEG from the signal space to the source space. The artefact sources are then replaced by a null vector, i.e. by sources generating no signal (void sources). Neural and void sources are projected back into the signal space, thereby reconstructing a clean EEG signal devoid of the artefactual contribution.

MICA consists in using the ICA algorithm iteratively t times. The main idea is to increase the reliability of the estimate of the weight vectors by selecting the ones that occur more often over all the executions. This technique requires a high number of executions of ICA to achieve a reliable estimate of the solution.

In our implementation, once the weight vectors for one execution have been computed, the ICs to be suppressed are identified by computing the correlation between the ICs and a ~~ECG~~ template as described in Srivastava et al. (2005). The cleaned data are then computed. This process is repeated for each execution of the ICA algorithm. After t versions of the cleaned data are available, their mean is computed.

cICA rejection algorithm

A mathematical account of constrained ICA (cICA) is presented in detail in Appendix 1. Independent Component Analysis assumes that EEG signals consist of a linear mixture of statistically independent (neural and artefact) sources. As the number of sources is generally assumed to be the same as the number of channels, its size is [N_{ch} channels × T time points]. x corresponds to recorded EEG signals [N_{ch} channels × T time points] and **W** a fixed scalar matrix, composed of N_{ch} row-vectors of dimension N_{ch} (square matrix of size N_{ch}).

Here we reasoned that, if EEG signals are assumed to be a linear mixture of two (neural and artefact) sets of sources, one could in principle identify one set of sources and deduce the second one in a more efficient way than through the cumbersome iterative search of all sources. The row vectors of **W** can be viewed as a non-orthonormal base of an N-dimensional space. During preprocessing, an orthonormal matrix is deducted from **W** (Appendix A.2). The identification of the artefact-related row-vectors in **W** would make it possible to find the complementary orthonormal set of vectors, i.e. the neural-related row-vectors of **W** using a Gram–Schmidt scheme. Because there are more neural than artefact sources (for a 62-electrode EEG recording system, 3 to 8 sources corresponding to PRA are typically identified) the computational load would be considerably reduced by computing only the latter.

In the present paper, constrained ICA was used to identify PRA sources. In contrast to classical ICA, cICA does not converge to the source closest to its random initialization but finds the closest source to a given constraint. This implies that the time course of the identified components closely matches the constraint. In our application, the constraint consists in a model of the PRA, as classically computed by the AAS method (Allen et al., 1998). For each channel an average PRA is computed across the sections of EEG signal centered on the last 10 electrocardiogram (ECG) QRS peaks. As stated above, the resulting model of the PRA is not perfect because the averaging does not account for the beat-to-beat variability in heart rate and in the shape of the PRA. However, it captures the essential features of PRA necessary to seed the cICA analysis.

The cICA algorithm provides N_{ch} estimates of the artefact sources. Since the number of artefact sources is largely inferior to the number of channels (64 in our case), some estimates will converge to the same artefact source. A clustering is thus performed through a classic K-means algorithm (Appendix A.3) to group corresponding estimates. The clustering algorithm gives also a measure of the variance expressed by the K clusters:

$$ev_K = 1 - \frac{\sum_{k=1}^{K} \frac{n_k}{N_{ch}} var_k}{Var} \quad (5)$$

where K is the number of clusters, var_k the variance in the kth cluster, Var the variance of the initial set of vectors, n_k the number of vectors in the kth cluster and N_{ch} the number of vectors in the initial set of vectors. Since each component was constrained by a template of the PRA, most of the variance (at least 70%) of the estimates should be expressed by a few clusters (typically less than 8). On the other hand if more clusters are necessary, it means that the estimates are more randomly or independently distributed, i.e. the data are unlikely to be contaminated by PRA.

Practically, our approach redirects the computational power used to find the neural sources in classic ICA algorithm to increase the robustness of the artefact source estimates. In other words, our algorithm acts in a similar way to multiple executions of an ICA algorithm (MICA), which has been shown to improve the robustness of the sources (Briselli et al., 2006; Phillips et al., 2006), but only for the artefact sources. Indeed, each

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

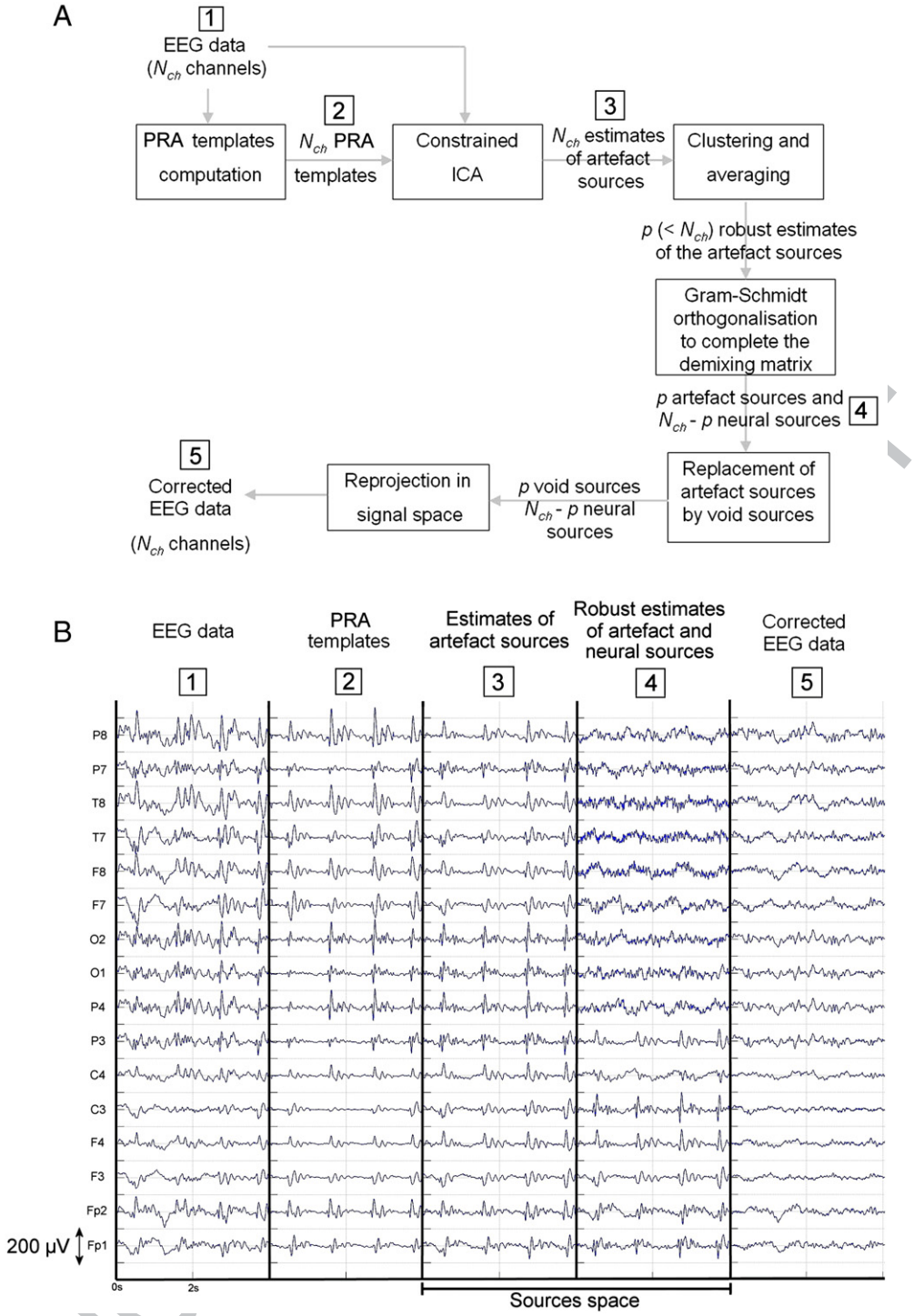

Fig. 1. (A) Block diagram of our cICA rejection algorithm, the numbers refers to the corresponding EEG time series in part B of the figure. (B) 4 s of signal are displayed in the different step of our algorithm. Note that the scale is not relevant in source space. B1 shows the original EEG recordings. B2 shows the constraint used to feed the cICA part of our algorithm. B3 shows the sources found using cICA. B4 shows the sources after clustering and Gram–Schmidt orthogonalization. Note that the last seven ICs are artefact-related. B5 shows the cleaned signal.

clustered artefact component is the average of various similar components. Crucially the problem of identifying the artefact related components (as met with classical ICA) is completely avoided here.

To summarize, our algorithm consists in (see Figs. 1A, B for a sample signal in each step):

- Computing a PRA template for each of the N_{ch} channels of the EEG signal;
- Computing the artefact related components based on these constraints using the cICA algorithm;
- Clustering the estimated artefact components and computing the average of each cluster to produce more robust estimates of the artefact sources;
- Finding the subspace of neural sources within the demixing matrix, i.e. the complementary space of the artefact sub-space, using a Gram–Schmidt procedure;

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

- Replacing artefact sources by void sources and reprojecting all sources in signal space.

This algorithm potentially leads to a more robust estimation than other ICAs of the PRA sources without increasing the computational cost. In addition, the identification of sources to suppress is automated, data driven and not observer-dependent.

Methods and materials

The section is organized in two parts. In the first one, EEG data recorded on 2 sleeping volunteers are processed. They were simply placed in the static field of the MRI scanner, but in the absence of fMRI acquisitions (in order to avoid any gradient artefact in the EEG signal) during NREM sleep. The second part uses EEG data acquired on 9 volunteers in various consciousness states (wakefulness, light and deep non rapid eye movement sleep), during simultaneous fMRI imaging. All volunteers were fully informed and gave their written consent. In both cases, EEG was recorded utilizing two MR-compatible 32-channel amplifiers (Brainamp MR plus, Brain Products GmbH, Gilching, Germany) and an MR-compatible EEG cap (Braincap MR, Falk Minow Services, Herrsching Breitbrunn, Germany) with 64 ring-type electrodes. EEG caps included 62 scalp electrodes as well as one electrooculogram (EOG) and one electrocardiogram (ECG) channel which were online referenced to FCz. Electrode-skin impedance was kept below 5 kΩ in addition to the 5 kΩ resistor built into the electrodes. EEG was digitized at 5000 Hz sampling rate with a 500 nV resolution. Data were analog-filtered by a bandlimiter low pass filter at 250 Hz (30 dB/Octave) and a high pass filter with 10 s time constant corresponding to a high pass frequency of 0.0159 Hz. Data were transferred outside the scanner room through fibre optic cables to a personal computer where the EEG system running Brain Recorder Software v1.03 (Brain Products GmbH, Gilching, Germany) was synchronized to the scanner clock. For analysis, EEG data were low-pass filtered (FIR filter, −36 dB at 70 Hz), and down-sampled to 250 Hz. The power spectra of the raw data were computed using fast Fourier transform on successive 4 s epochs, overlapping by 2 s and weighted by a Welch window.

EEG recordings without gradient artefact

Two healthy volunteers were completely sleep deprived during one night (corresponding to about 24 h of continuous wake). During the following morning, they were recorded in the scanner room (3 T MR scanner, Allegra, Siemens, Germany) in 2 consecutive conditions: (1) lying on the scanner couch in reverse position, i.e. the head outside the scanner[1] and (2) lying in standard position, i.e. installed head first in the scanner bore. Conditions (1) and (2) are therefore as close as possible for the volunteers and should not affect their sleep EEG. Nonetheless, the recordings in condition (1) are expected to be much less contaminated by PRA due to the lower magnetic field at electrodes level. In both cases, recordings were obtained without simultaneously scanning. For each session, recordings were carried on until subjects reached steady slow wave sleep, the deepest stage of non REM sleep. Analyses were run on two minute periods of these recordings. They consisted in cleaning the EEG time series with four different methods: AAS, PCA, MICA and cICA. To assess the quality of PRA rejection by these four methods, two quality indices were computed. Two indices of quality are then computed. Firstly, the Improvement of Normalized Power Spectrum introduced by Tong et al. (2001) measures the power attenuation in the corrected signal:

$$\mathrm{INPS}(\mathrm{Chan},N) = \frac{\sum_{j=1}^{N} P_j^{\mathrm{before}}}{\sum_{i=1}^{N} P_i^{\mathrm{after}}} \tag{6}$$

where P^{before} and P^{after} are mean power in a 1 Hz window centered on the jth harmonic of the heart frequency respectively before and after PRA removal for a particular channel Chan. The INPS thus represents the quality of power rejection in these windows, hypothesizing that most of the power in the latter are PRA-related.

For our computations, we chose $N=5$ to take into account the first four harmonics of the PRA, hypothesizing that most of the PRA-related power is taken into account. N is usually chosen between 4 and 6 in the literature (cf. Briselli et al., 2006; Srivastava et al., 2005; Nakamura et al., 2006). INPS is larger than 1 when power is around spectral modes characteristic of pulse artefact: the larger INPS, the larger the power attenuation.

Second, the Distortion Index introduced by Nakamura et al. (2006) measures the distortion introduced by the PRA rejection method:

$$\mathrm{DI}(f) = \frac{1}{N_{\mathrm{ch}}} \sum_{\mathrm{Chan}=1}^{N_{\mathrm{ch}}} \left(\frac{\int_{f-\Delta f}^{f+\Delta f} \sqrt{\phi_{\mathrm{cleaned,Chan}}(u)}\,du}{\int_{f-\Delta f}^{f+\Delta f} \sqrt{\phi_{\mathrm{w/oPRA,Chan}}(u)}\,du} \right)^2 \tag{7}$$

where $\phi_{\mathrm{w/o\ PRA,Chan}}(u)$ (resp. $\phi_{\mathrm{cleaned,Chan}}(u)$) is the value of the power spectrum frequency u at the electrode Chan ~~=1~~ for the dataset recorded in reverse position, with the head outside the scanner and with minimal PRA (respectively recorded with the head inside the scanner and corrected for the large PRA). N_{ch} is the number of EEG channels and the DI is computed for a frequency interval between $f-\Delta f$ and $f+\Delta f$. For our computation, we used $\Delta f = 0.25$ Hz, i.e. frequency windows of 0.5 Hz.

EEG recordings with gradient artefact

The aim of this experiment is to generalize the findings of the previous case to a larger set of data acquired during effective fMRI scanning, and to various states of consciousness characterized by typical oscillations other than slow waves, namely alpha rhythm during wakefulness and spindles during stage 2 sleep. EEG data used in this section were recorded from 9 healthy volunteers (4 females; age range: 18–25), who participated in an EEG/fMRI sleep study. Functional MRI time series were acquired continuously using a 3 T MR scanner (Allegra, Siemens, Germany). Multisliced, T2*-weighted fMRI images were obtained with a gradient echo-planar sequence using axial slice orientation and reduced slew rate for efficient gradient rejection (32 slices; voxel size: 3.4×3.4×3 mm^3; matrix size 64×64×32; TR=2460 ms; TE=40 ms; flip angle=90°; FOV=220×220 mm^2;). Some of these data have been previously published (Schabus et al., 2007) but have not been corrected by cICA. Gradient artefacts were removed in Vision Analyzer (Brain Products GmbH, Gilching, Germany) using a method adapted from Allen et al. (2000). Sleep staging followed standard criteria (Rechtschaffen and Kales, 1968), and identified periods of wakefulness, stage 2 and slow wave sleep. Analyses were conducted on 2-minute EEG epochs acquired in each of these three states and again assessed the PRA rejection by AAS, PCA, MICA and cICA. The INPS was computed as in the previous section. For 24 two-minute EEG epochs, an expert was blindly presented with the 4 differently corrected signals, knowing in which consciousness stage the EEG was recorded. His duty was then to sort the corrected recording by quality of the signal (1 being the worse correction and 4 the best one), taking into account the residue of PRA left in the signal and the suppression of stage-related activity.

Results

EEG recordings without gradient artefact

The PRA was obviously much larger when the head of the subject was placed within the MR scanner than outside the scanner (Fig. 2). The comparison of these recordings thus allowed for the estimation of the sensitivity and specificity of various artefact rejection methods.

[1] Since this is a head only scanner, the subject has his feet near the scanner bore and his head as far as possible from the magnet center (about 2m). The magnetic field at the head level has a theoretical value of about 20mT given the scanner's specifications.

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

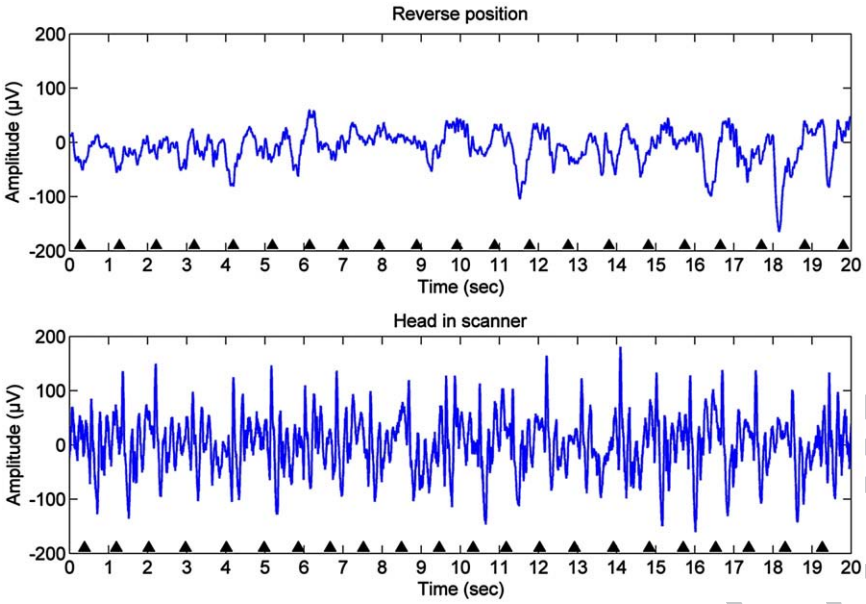

Fig. 2. EEG recording at electrode F8 with the head outside (top) and inside (bottom) the MR scanner, both without MR acquisition. The PRA is clearly identified in the latter condition. The black triangle indicates a QRS peak in the ECG channel.

We first characterized the power spectrum of PRA and its scalp topography (Fig. 3). We computed the power spectrum of data recorded inside and outside the scanner bore. The power spectrum of data recorded outside the scanner bore is characterized by a component around 1 Hz (0.5 Hz and 1 Hz in subjects 1 and 2, respectively), consistent with the prominent slow rhythm typical of deep NREM sleep (Figs. 3 A, B, red curves). A secondary mode is observed in subject 2 in the spindle frequency range (Fig. 3B, arrow). The power spectrum of data recorded within the scanner bore is dominated by several modes between 1 and 5 Hz as well as broadband modes around 8, 13 and 17 Hz (Figs. 3 A, B, blue curves). The scalp topography of PRA power, computed as the difference between the 2 conditions, is not homogeneous (Figs. 3 C, D). Most of the power difference is located in the right fronto–temporal (subject 1) and left fronto–temporo–occipital regions (subject 2). The low power observed around fronto–central electrodes is in part related to referencing to FCz but not entirely so. The origin of these variations has not been systematically evaluated and could also be linked to the orientation of the electrodes relative to the static field, or their position relative to the heart and blood vessels.

Fig. 4A shows the signal recorded in subject 1 for 20 s during slow wave sleep on electrode Fz before and after correction by different methods. This electrode was chosen because it suffers from a large PRA in this particular recording (see above). In the original recordings (top row), the PRA is so prominent that the sleep slow waves are barely identifiable.

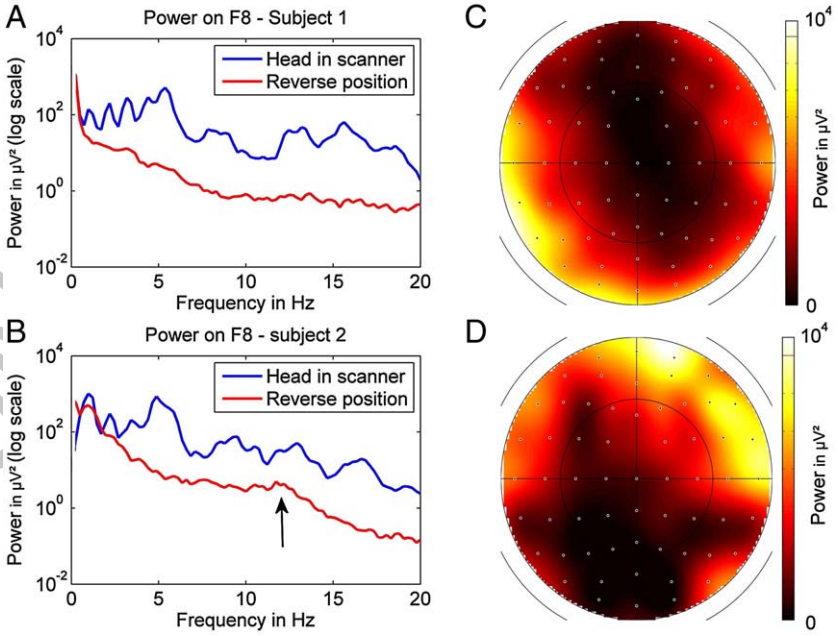

Fig. 3. Power spectrum at electrode F8 with the head inside or outside the MR scanner, both without image acquisition for 2 subjects. The topography of the PRA (computed as the power difference between the 2 conditions) is showed in C and D.

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

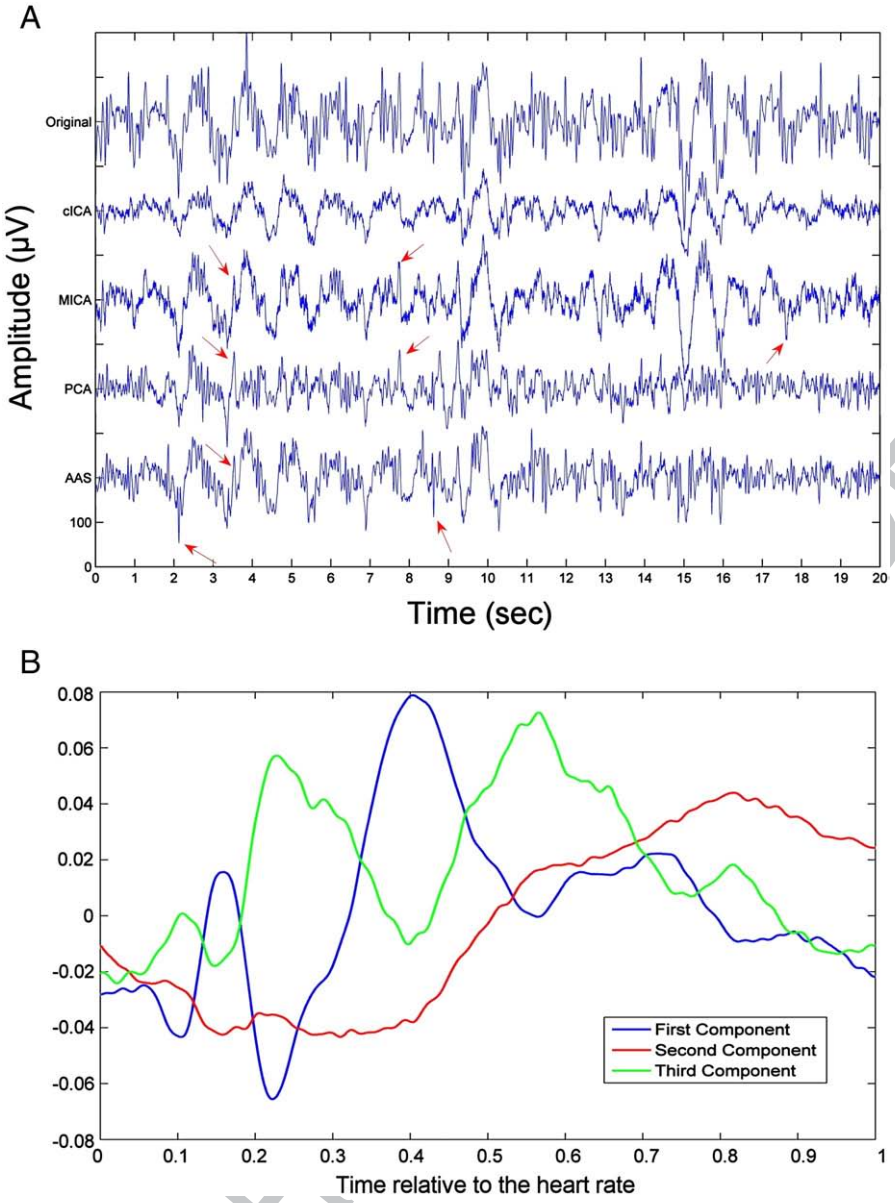

Fig. 4. (A) Example of original data, recorded on electrode Fz during slow wave sleep, and the corrected signal obtained by different method of PRA rejection. (B) First three components retrieved by the Principal Component Analysis method.

Slow waves are efficiently retrieved by all methods except PCA which suppresses most of them. With this technique (row 4), large-amplitude sleep waves are mistakenly included in the first principal components, and removed as PRA (Fig. 4B). On the other hand, PRA is still easily identifiable (red arrows) after artefact rejection by AAS (row 5) or MICA (row 3), but hardly so after cICA (row 2).

In the frequency domain, all correction methods globally decrease signal power (Fig. 5). As compared to other methods, cICA (blue line) is the method which removes the most power in all frequency bands and pleasingly provides the spectrum closest to the recording obtained outside the scanner bore (grey line). Constrained ICA removes marginally more power than AAS (red) and MICA (cyan) in the lowest frequency range although it retains the characteristic spectral mode of deep slow wave sleep around 0.8 Hz. In contrast, PCA considerably reduces the power density in the frequency range of sleep slow waves (green line and arrow).

To quantify these power changes after PRA correction and characterize their scalp topography, we computed the INPS over all channels for subject 1. Fig. 6 displays INPS topography over the entire scalp for the four methods.

More power is suppressed by PCA than by any other method as indicated by a large INPS value (8.49, mean over all EEG electrodes, Fig. 6B). However, as suggested earlier, the specificity of PCA in removing only PRA is questionable. A topographical structure can be observed for INPS computed with cICA, AAS or MICA (Figs. 6A, C, D, respectively), characterized by large INPS values in right frontal and left temporal areas, a scalp topography reminiscent of PRA power distribution (Fig. 3). In contrast, PCA suppresses power homogeneously over the scalp, suggesting a non specific rejection of both PRA and genuine cerebral oscillations. Characterized by a mean INPS over all channels of 5.93, cICA suppresses more power than the AAS and MICA (mean INPS of 1.92 and 3.96, respectively). However, cICA still preserves a regionally specific effect which again speaks for the selective removal of PRA.

To further compare the different PRA removal methods, the distortion index was computed from 0.5 Hz to 20 Hz, corresponding to the bandwidth of sleep-related activity, by steps of 1 Hz. An index equal to one

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

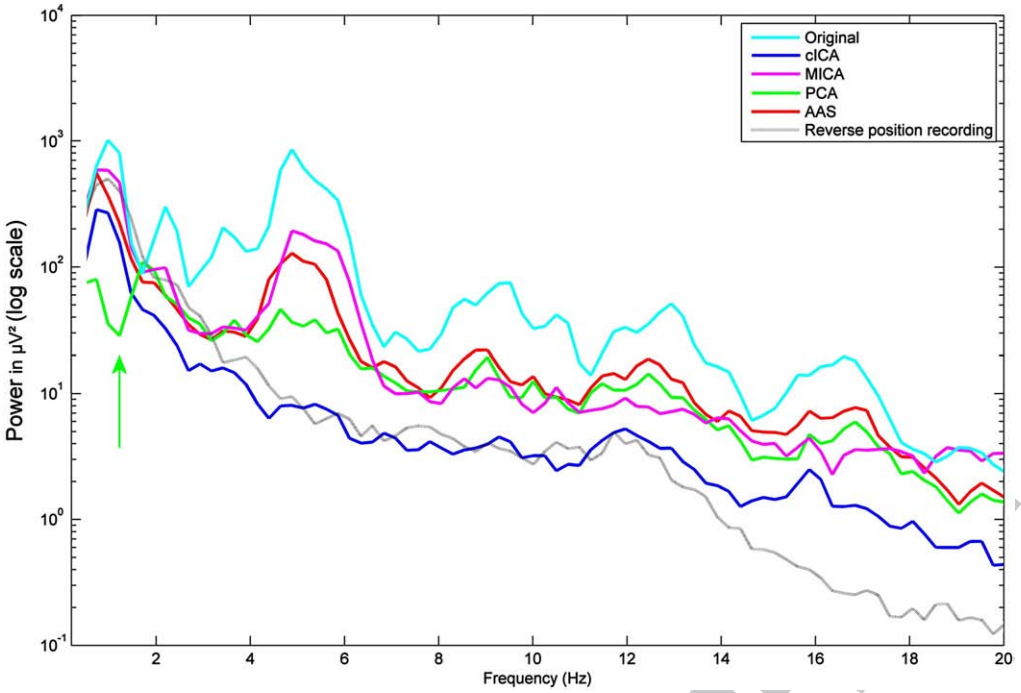

Fig. 5. Power spectrum of the EEG recorded at electrode F8 after correction by the different PRA rejection methods. The power spectrum of the original signal and signal recorded with the head outside the MR scanner are given as references.

would mean that there is the same amount of power in the reference signal without PRA as in the artefacted signal after cleaning. An ideal PRA suppression method would therefore have an index of 1 across all frequency bins. The results are shown in Fig. 7.

All correction methods suppress too much power ($I(f)<1$) at frequencies below 5 Hz and not enough power ($I(f)>1$) from 15 to 25 Hz. Constrained ICA clearly outperforms the other methods for frequencies from 5 Hz to 25 Hz. Note that the indices of the other methods have a peak around 5 Hz, which is successfully avoided by cICA (Fig. 7, inset). Below 5 Hz, the cICA algorithm removes more power than AAS or MICA (Fig. 7, inset). Although this is a potential problem for low frequency processes such as slow waves in deep sleep, we previously showed that cICA successfully preserves them albeit with a reduced amplitude compared with AAS and MICA (as shown on Figs. 4 and 5). In keeping with results presented above, PCA has an alarmingly low index (~~$I(f)@0.1$~~) around 1 Hz and strongly affects the slow waves.

EEG recordings with gradient artefact

Fig. 8 illustrates the effect of various PRA rejection methods on a typical individual recording during wakefulness. The upper panel shows EEG recorded at Oz during a 20 s period of wakefulness. Although PRA is

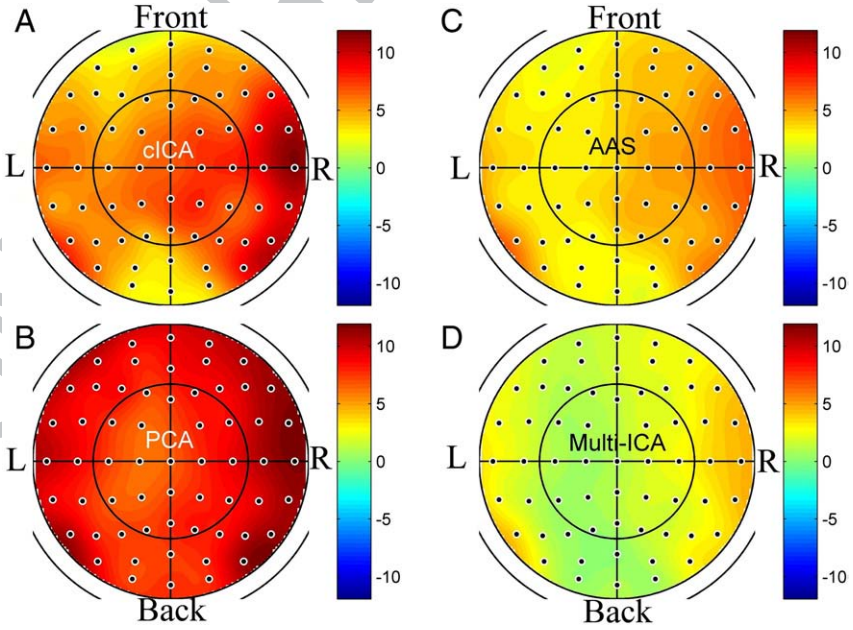

Fig. 6. INPS values interpolated (from the values calculated at each electrode) over a flat-map representation of the scalp for all the tested correction methods.

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

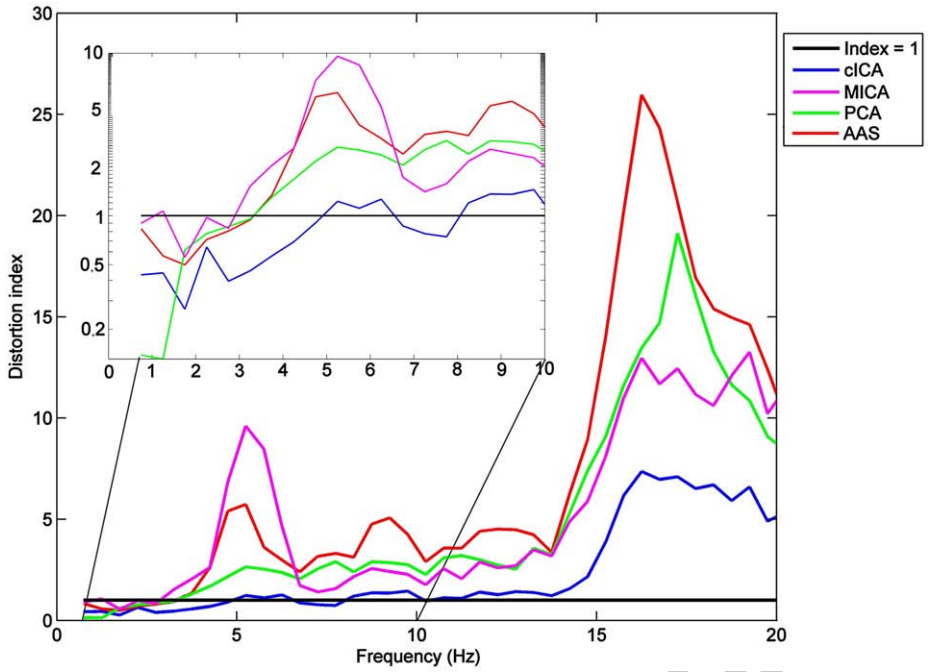

Fig. 7. Distortion index computed for frequencies from 0.5 to 20 Hz by steps of 1 Hz for all the tested methods. A value of 1 means no distortion.

considerably reduced by all techniques, it is still easily visually detected after rejection by AAS and MICA (arrows). PRA is hardly visible after correction by PCA and cICA which seems to reduce the artefact to a similar extent. The lower panel shows the corresponding power spectrum. The peak around 10 Hz, corresponding to the alpha rhythm, is most clearly identified after correction by cICA or MICA, and much less so after correction by AAS or PCA.

Likewise, during stage 2 sleep, sleep spindles are readily identified after correction by all methods (Fig. 9, top). From the power spectrum (Fig. 9, bottom), cICA rejects more power than all other correction methods for nearly all frequencies between 0 and 20 Hz. Nevertheless cICA clearly preserves power modes corresponding to typical activities of this state, i.e. the spindles and the slow waves. As already noticed above, PCA considerably reduces power between 0.5 and 2 Hz.

At the group level, we computed the power attenuation as the power ratio between corrected and original signals (Fig. 10), averaged across the 9 subjects. Although computed on an independent data set, this analysis confirms findings previously observed on the individual data of 2 subjects (cf. previous section). First, in all states of vigilance, cICA method rejects more power than all other methods (except PCA around 1 Hz), especially between 1 and 5 Hz, the frequency band mainly affected by PRA. Second, during stage 2 and slow wave sleep (respectively characterized by *K* complexes and slow waves), power around 1 Hz is marginally more attenuated by cICA than AAS or MICA

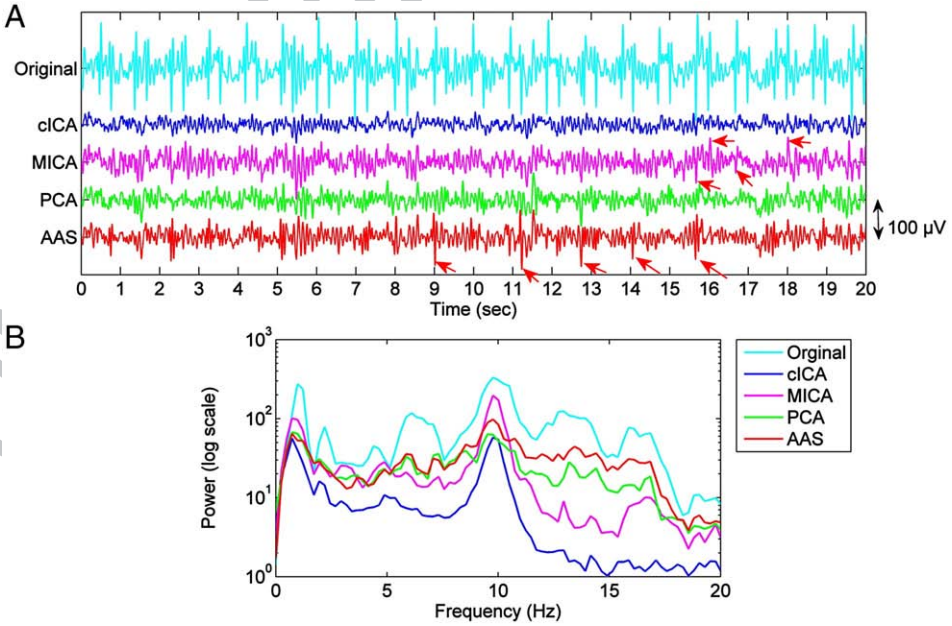

Fig. 8. Time course (panel A) and power spectrum (panel B) of EEG recording corrected by all tested PRA correction methods at electrode Oz during a wake state. The time course and the power of the original signal are given for reference.

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

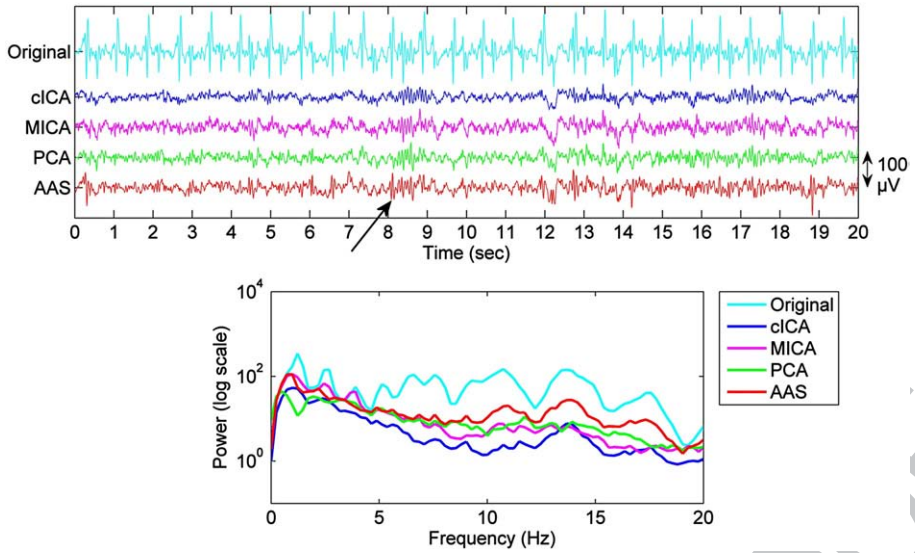

Fig. 9. Time course (~~panel A~~) and power spectrum (~~panel B~~) of EEG recording corrected by all tested PRA correction methods at electrode Oz during sleep stage 2 state. The time course and the power of the original signal are given for reference. The arrow indicates spindle activity. Q1

but much less than by PCA. As shown in Fig. 10, the cICA corrected signals are significantly different (Friedman test, $p<.05$) from the original ones between 0.5 and 4 Hz in every state and every tested electrode, i.e. over the frequency band traditionally associated with PRA. Moreover for the particular case of Oz in wake state, the correction is significant from 0.5 to 16 Hz. This is also the case for F8 in sleep stage 2 while the correction is significant for Cz from 0.5 to 12 Hz and from 9 to 13 Hz in sleep stage 2, suggesting that the frequency content of the artefact can be different across electrodes.

We computed the mean INPS for the 4 correction methods across all electrodes and subjects (Fig. 11). For all methods, INPS varies between states and progressively decreases from wakefulness to light and deep sleep stages. This result arises from the smaller proportion of total power related to PRA during sleep than wakefulness, due to the abundance of large amplitude slow waves during the former. During wakefulness, cICA INPS is larger than that of any other methods. During sleep states, cICA INPS is still larger than that of MICA and AAS. However, PCA INPS is now larger than cICA INPS because, as indicated before, PCA non selectively suppresses a large amount of power around 1 Hz. All these results are statistically significant (Wilcoxon signed-rank test; $p<0.01$).

Fig. 12 shows the mean score of the blind expert's classification for each correction methods. The cICA ranks significantly better than PCA (Wilcoxon signed-rank test; $p<0.0003$). However even if cICA has the

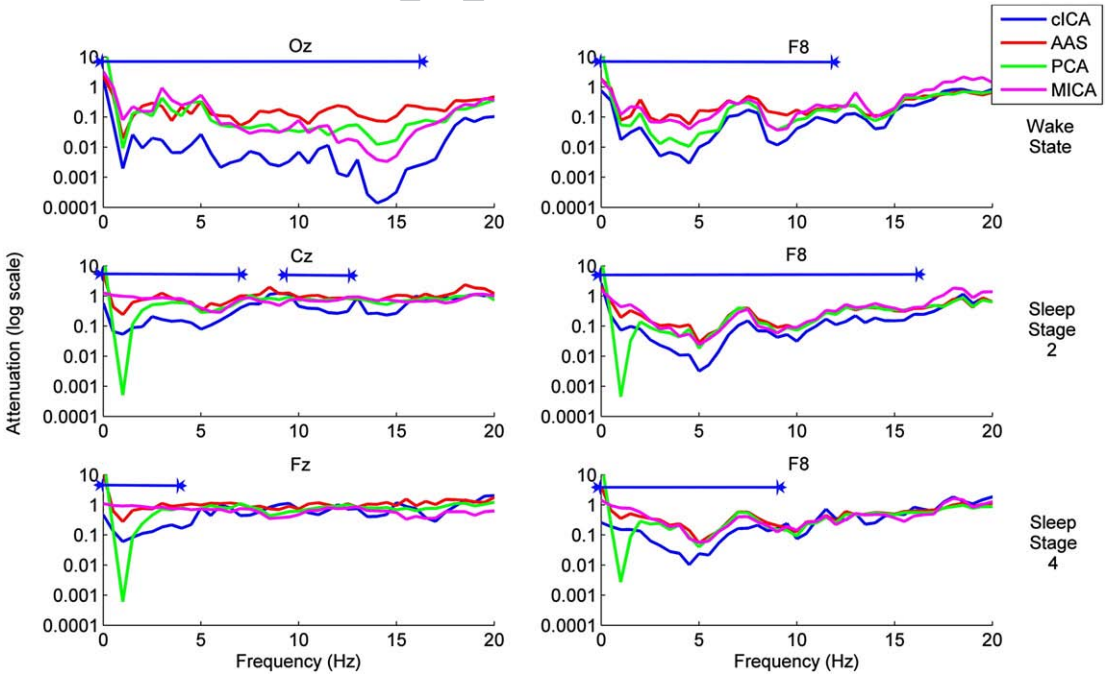

Fig. 10. Mean attenuation over 9 subjects for the different PRA correction methods and for 3 different wake/sleep stages. Two different electrodes are displayed per stage. The blue lines above the different plot shows over which frequency band is the cICA correction significant relative to no correction (Friedman test; $p<0.05$).

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

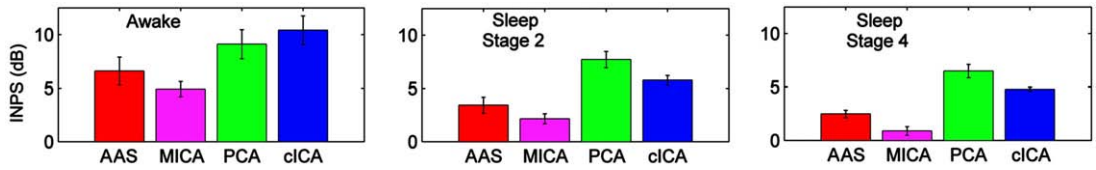

Fig. 11. Mean INPS over the 9 subjects and over all EEG electrodes for different sleep stages. The cICA mean INPS is significantly different from the other method mean INPS for all sleep stages (Wilcoxon signed-rank test; $p<0.01$).

best mean score, it is not significantly better than AAS (Wilcoxon signed-rank test; $p<0.1$) or MICA (Wilcoxon signed-rank test; $p<0.11$).

Discussion

On EEG recordings, deep non rapid eye movement sleep is characterized by large amplitude slow waves which typically lasts about 1 s. In the framework of EEG–fMRI studies on deep NREM sleep, removal of PRA from EEG recordings are problematic because the period and amplitude of sleep slow waves and PRA overlap. In this paper, we introduce constrained ICA as a novel method for PRA rejection from EEG recorded simultaneously with fMRI during slow wave sleep. We first evaluated cICA in datasets obtained during slow wave sleep in 2 subjects positioned consecutively in a large static magnetic field (head within the scanner) then in a negligible field (on the scanner table but head away from the bore). Only the former recording was detectably contaminated by PRA. Our evaluation indicates that cICA is both efficient and specific in rejecting PRA. The power spectrum obtained after correction by cICA was very close to the spectrum obtained away from the magnetic field, as confirmed by a favorable distortion index between 1 and 20 Hz. As indicated by the INPS, cICA remove more power than the AAS, PCA or MICA. Below 1 Hz, cICA retained the characteristic spectral mode of slow oscillation whereas PCA significantly reduced the power of slow waves. These findings were statistically confirmed on a larger dataset collected in 9 subjects not only during slow wave sleep but also during wakefulness and light non rapid eye movement sleep. Finally, the quality of PRA rejection by cICA also compared favorably with other techniques, when visually assessed by an expert.

The main advantages of cICA are that, although using the classical ICA machinery, (1) it considerably reduces the computational load of PRA rejection, relative to classical ICA; (2) it eschews the problem of component identification and (3) it provides an efficient and robust rejection of PRA, which by many aspects outperforms earlier classical methods. The only parameter, the threshold of the clustering step, is easily calculated given the number of channels. One disadvantage of cICA is that, as for ICA methods in general, it requires a large number of sensors (i.e., EEG channels). The cICA solution also relies on the accurate detection of the QRS peak detection which in turn depends on the quality of ECG recording and the detection algorithm used. We found that the solution proposed by Niazy et al. (2005) was very reliable for our data. Our simulation shows that the quality of PRA rejection substantially decreases when the number of channels is smaller than about 30. Another drawback of the method, also common to other ICA algorithms, is that the propagation of noise from a single bad channel (due to muscular artefact or electrode disconnection) to all other channels after PRA rejection. This problem is easily fixed by rejecting the bad channel before PRA removal.

Finally, this novel technique is equally efficient in removing PRA from continuous EEG time series recorded in other states of vigilance (wakefulness and light sleep). We also successfully tested our algorithm on an EEG recording obtained in an epileptic patient who fortuitously fell asleep during the EEG/fMRI session (see Supplemental material). However, the efficiency of this technique in other types of

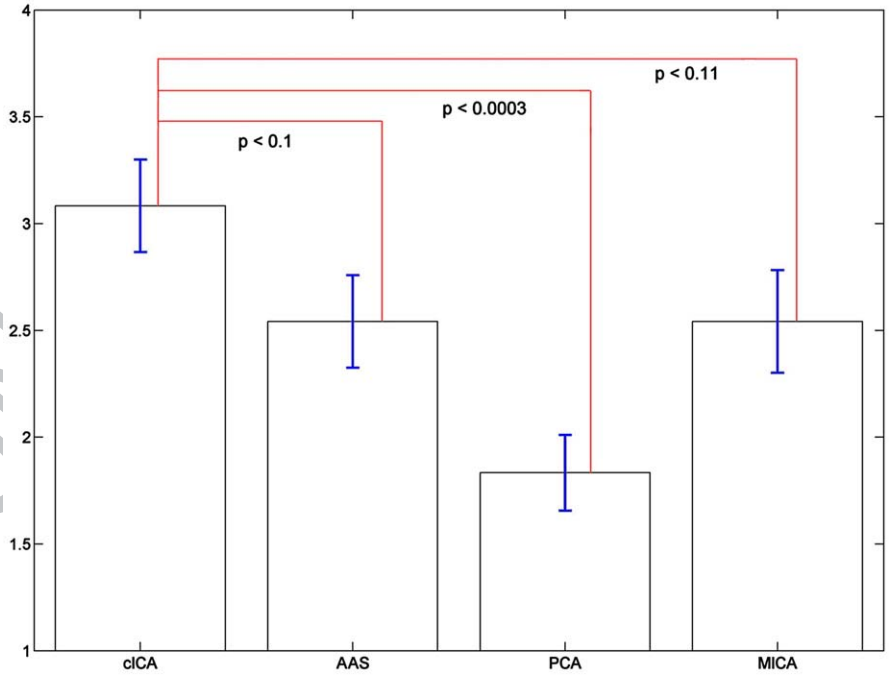

Fig. 12. Mean score obtained by the 4 correction methods with a blind expert's evaluation. The p-value, obtained by performing a Wilcoxon signed-rank test, indicates that cICA and PCA are significantly different but that cICA and AAS or MICA are not.

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

recording such as event-related potentials, for instance, has not been evaluated and should be subjected to further investigation.

Acknowledgments

This study was supported by the Belgian Fonds National de la Recherche Scientifique (F.R.S.-F.N.R.S.), the Fondation Médicale Reine Elisabeth, the Research Fund of ULg, and PAI/IAP Interuniversity Pole of Attraction. Y.L., E.B., T.D., P.M. and C.P. are supported by F.N.R.S. (Belgium).

Appendix A. Mathematical appendix

Introduction to cICA

The problem of Constrained Independent Component Analysis (cICA) consists in (1) maximizing the negentropy of the sources $\sum_{i=1}^{l} J(s_i)$, as in classical ICA algorithm but (2) subject to constraints $\mathbf{g}(\mathbf{W}) \leq 0$ and $\mathbf{h}(\mathbf{W}) = 0$, where

- l is the desired number of components to be extracted;
- $\mathbf{g}(\mathbf{W}) = (g_1(\mathbf{w}_1) \ldots g_l(\mathbf{w}_l))^T$ in which each $g_i(\mathbf{w}_i) = \varepsilon_i(s_i, r_i) - \xi_i$ for $i = 1, \ldots, l$;
- $\mathbf{h}(\mathbf{W}) = (h_1(\mathbf{w}_1) \ldots h_l(\mathbf{w}_l))^T$ with $h_i(\mathbf{w}_i) = E\{s_i^2\} - 1$ for $i = 1, \ldots, l$.

The first constraint forces the distance between each source estimates and its respective reference $\varepsilon_i(s_i, r_i)$ to stay under a specified distance ξ_i. The second constraint is just a normalization term.

To solve this problem, the Kuhn–Tucker conditions, i.e. the generalization of the Lagrange multiplier, are used. The augmented Lagrangian function $L_2(\mathbf{W}, \mu, \lambda)$ is given by:

$$L_2 = \sum_{i=1}^{l} \left(J(y_i) - \frac{\max^2\{0, \mu_i + \gamma_i g(\mathbf{w}_i)\} - \mu_i^2}{2\gamma_i} \right) - \boldsymbol{\lambda}^T h(\mathbf{W}) - \frac{1}{2}\gamma^T \|h(\mathbf{W})\|^2 \quad (8)$$

where $\mu = (\mu_1 \ldots \mu_l)^T$ and $\lambda = (\lambda_1 \ldots \lambda_l)^T$ are two sets of Lagrange multipliers for the inequality and equality constraints respectively, and $\gamma = (\gamma_1 \ldots \gamma_l)^T$ are the parameters to form the penalty terms to ensure that the optimization problem is held at the condition of local convexity assumption.

To find the decorrelation matrix, a Newton-like iteration process is used:

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \langle ss(\mathbf{W}) L'_w \rangle \mathbf{R}_{xx}^{-1} \quad (9)$$

where $\langle ss(\mathbf{W}) \rangle$ is a vector depending of the negentropy of each s_i and $\mathbf{R}_{xx}$ is the autocorrelation matrix of the original recordings $\mathbf{x}$.

$$\mu_{k+1} = \max\{0, \mu_k + \langle \gamma \rangle \mathbf{g}(\mathbf{W})\} \quad (10)$$

$$\lambda_{k+1} = \lambda_k + \langle \gamma \rangle h(\mathbf{W}) \quad (11)$$

The stopping criteria for the iterative process was simply that the relative difference between the norm of 2 successive estimates of $\mathbf{W}$ be smaller than 10^{-5}.

Decorrelation of the different estimates can be used to prevent the convergence of two estimates towards the same IC. In our case, the decorrelation is not used since different estimates are wanted to converge to the same independent components (to increase robustness of our estimates, see section cICA rejection algorithm).

Importantly, in classic cICA algorithms, the constraint on $\mathbf{g}$ is relaxed once, along the iterative process, the estimation of the independent component is in the neighborhood of the reference. This is similar to an ICA algorithm where the initialization is the reference. The sources the algorithm will find are the nearest ones to the reference since the Kuhn–Tucker conditions are a local maximum finder.

Yet, in our implementation the constraint is not relaxed and thus our algorithm does not find the statistically independent component but a tradeoff between the constraint and this component. The choice of the ξ_i is said to be critical in Lu and Rajapakse (2001). However since the constraint is not relaxed we choose $\xi_i = 0$ and whatever value we choose for ξ, the algorithm always tries to minimize the value of $\mathbf{g}$. Our algorithm is thus not an ICA algorithm, strictly speaking.

Preprocessing and mathematical assumptions

Assumptions

Since, the variance of the independent components cannot be recovered (see Hyvarinen, 2000), it is hypothesized that $E\{\mathbf{ss}^T\} = 1$ for simplicity. To ensure that it holds true, we have to fix the norm of each row-vector of the demixing matrix equal to one and to center and whiten the data.

Preprocessing

The data are preprocessed in the classical way: they are centered and whitened. After centering the x_i, all the s_i will have a zero mean since we find them through a linear combination of the x_i using the demixing matrix $\mathbf{W}$.

Whitening consists in transforming the data $\mathbf{x}$ linearly to obtain a new vector $\tilde{x}$ where the components of are uncorrelated and their variances equal one. A whitened vector $\tilde{x}$ has a covariance matrix equal to

$$E\left\{ \tilde{x}\tilde{x}^T \right\} = I \quad (12)$$

This whitening is commonly achieved using the eigenvalue decomposition of the covariance matrix $E\{\mathbf{xx}^T\} = \mathbf{EDE}^T$, where $\mathbf{E}$ is the orthogonal matrix of eigenvectors of $E\{\mathbf{xx}^T\}$ and $\mathbf{D}$ its eigenvalues, $\mathbf{D} = \mathrm{diag}(d_1, \ldots, d_n)$. Whitening is achieved through

$$\tilde{x} = \mathbf{ED}^{-1/2}\mathbf{E}^T x \quad (13)$$

which obviously leads to $E\left\{ \tilde{x}\tilde{x}^T \right\} = I$. In the framework of the ICA problem, whitening transforms the mixing matrix $\mathbf{A}$ into a new one: $\tilde{\mathbf{A}}$ which can be found using Eqs. (4) and (13)

$$\tilde{x} = \mathbf{ED}^{-1/2}\mathbf{E}^T \mathbf{A} s = \tilde{\mathbf{A}} s \quad (14)$$

The main advantage of the whitening step is that this new mixing matrix $\tilde{\mathbf{A}}$ is orthogonal:

$$E\left\{ \tilde{x}\tilde{x}^T \right\} = \tilde{\mathbf{A}} E\{ss^T\} \tilde{\mathbf{A}}^T = \tilde{\mathbf{A}}\tilde{\mathbf{A}}^T = \mathbf{I} \quad (15)$$

and if $\tilde{\mathbf{A}}$ is orthogonal, so is its inverse $\tilde{\mathbf{W}}$.

Sorting out the artefact sources using K-means clustering

We use a "*K*-means" (MacQueen, 1967) clustering method. The idea is to gather the sources which are similar together, in the same cluster. Once the sources have been sorted in clusters, the mean vectors of each cluster are calculated.

However instead of applying the clustering program to the observations of the sources directly (which have a length equal to the number of time bins in the time series), it is computationally more efficient and theoretically equivalent to apply the clustering algorithm to the row-vector of the demixing matrix (which has a length equal to the number of channels) provided by the cICA program.

K-means-based algorithms have the advantage to be quick and simple. However they have two drawbacks: they do not find the best clusters at each execution (due to local convergence and random

Please cite this article as: Leclercq, Y., et al., Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic..., NeuroImage (2008), doi:10.1016/j.neuroimage.2008.10.017

initialization) and the number K of clusters must be specified a priori. The quality of the clustering can be measured by the *expressed variance*, ev_K, which is defined by:

$$ev_K = 1 - \frac{\sum_{k=1}^{K} \frac{n_k}{N} \mathrm{var}_k}{\mathrm{Var}} \quad (16)$$

where K is the number, var_k the variance in the kth cluster, Var the variance of the initial set of vectors, n_k the number of vectors in the kth cluster and N the number of vectors in the initial set of vectors.

The idea is to initialize K at 2, to run the K-means algorithm a sufficient number of times and to retain only the best clustering (which has the maximum expressed variance ev_2). K is then incremented to 3 and the algorithm is run the same number of times and the maximum value of the expressed variance is stored: ev_3. The difference between ev_2 and ev_3 represents the gain of representing the set of vectors by 3 clusters instead of 2. If this gain is sufficiently high, we again increment k by one. We repeat these steps until the increment of expressed variance: $ev_{p+1} - ev_p$ is below an arbitrarily chosen threshold value and, the clustering of p clusters is kept. The algorithm returns the means of these p clusters which are p row-vectors projecting the EEG recordings onto source space resulting in robust estimation of the artefact sources.

We empirically fixed the stopping threshold at 3%. Too large a value of this threshold could lead to a poor description of the set of vector and different artefact-related estimates may be merged in the same cluster. Too small a value could lead to an over-description of the artefact-related estimates, reducing the dimension available to describe the neural related-activity. Note that for a fixed number N_{ch}, the ev_K is increasing with the number of cluster while the difference of consecutive ev ($ev_{p+1} - ev_p$) is decreasing. The rate of this decrease is dependent of the number of vectors in the set to cluster. The threshold used for the clustering stopping criteria should be empirically adjusted for the number of EEG channels, as a rule of thumb we found that with half the channels (32), we should use the double of the threshold (6%) to obtain similar results.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2008.10.017.

References

Allen, P.J., Polizzi, G., Krakow, K., Fish, D.R., Lemieux, L., 1998. Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. Neuroimage 8 (3), 229–239.

Allen, P.J., Josephs, O., Turner, R., 2000. A method for removing imaging artifact from continuous EEG recorded during functional MRI. Neuroimage 12 (2), 230–239.

Bell, A.J., Sejnowski, T.J., 1995. An information–maximization approach to blind separation and blind deconvolution. Neural Comput. 7 (6), 1129–1159.

Benar, C.G., Aghakhani, Y., Wang, Y., Izenberg, A., Al-Asmi, A., Dubeau, F., Gotman, J., 2003. Quality of EEG in simultaneous EEG–fMRI for epilepsy. Clin. Neurophysiol. 114 (3), 569–580.

Bonmassar, G., Purdon, P.L., Jääskeläinen, I.P., Chiappa, K., Solo, V., Brown, E.N., Belliveau, J.W., 2002. Motion and ballistocardiogram artifact removal for interleaved recording of EEG and EPs during MRI. Neuroimage 16 (4), 1127–1141.

Briselli, E., Garreffa, G., Bianchi, L., Bianciardi, M., Macaluso, E., Abbafati, M., Marciani, M.G., Maraviglia, B., 2006. An independent component analysis-based approach on ballistocardiogram artifact removing. Magn. Reson. Imaging 24 (4), 393–400.

Buzsaki, G., 1996. The hippocampo–neocortical dialogue. Cereb. Cortex 6, 81–92.

Christov, 2004. Real-time electrocardiogram QRS detection using combined adaptive threshold. Biomed. Eng. Online 3 (1), 28.

Czisch, M., Wetter, T.C., Kaufmann, C., Pollmacher, T., Holsboer, F., Auer, D.P., 2002. Altered processing of acoustic stimuli during sleep: reduced auditory activation and visual deactivation detected by a combined fMRI/EEG study. Neuroimage 16 (1), 251–258.

Debener, S., Strobel, A., Sorger, B., Peters, J., Kranczioch, C., Engel, A.K., Goebel, R., 2007. Improved quality of auditory event-related potentials recorded simultaneously with 3-T fMRI: removal of the ballistocardiogram artefact. Neuroimage 34 (2), 587–597.

Debener, S., Mullinger, K.J., Niazy, R.K., Bowtell, R.W., 2008. Properties of the ballistocardiogram artefact as revealed by EEG recordings at 1.5, 3 and 7 T static magnetic field strength. Int. J. Psychophysiol. 67 (3), 189–199.

Dijk, D.J., Shanahan, T.L., Duffy, J.F., Ronda, J.M., Czeisler, C.A., 1997. Variation of electroencephalographic activity during non-rapid eye movement and rapid eye movement sleep with phase of circadian melatonin rhythm in humans. J. Physiol. 505 (Pt 3), 851–858.

Finelli, L.A., Jung, T.P., Duann, J.R., Haist, F., Makeig, S., Sejnowski, T.J., 2003. Information theoric modeling and removal of artefacts from EEG recordings during fMR imaging. 9th Annu. Meeting Organization for Human Brain Mapping.

Goldman, R.I., Stern, J.M., Engel, J., Cohen, M.S., 2000. Acquiring simultaneous EEG and functional MRI. Clin. Neurophysiol. 111 (11), 1974–1980.

Grouiller, F., Vercueil, L., Krainik, A., Segebarth, C., Kahane, P., David, O., 2007. A comparative study of different artefact removal algorithms for EEG signals acquired during functional MRI. Neuroimage 38 (1), 124–137.

Han, J.Y., Lee, H.R., Lee, S.Y., Choi, H.Y., Kim, J.G., 2004. ICA-based spatial filters to reduce cardioballistic noises in EEG signals acquired inside the MRI magnet. Prc. Int. Mag. Reson. Med. 11, 1101.

Huber, R., Ghilardi, M.F., Massimini, M., Tononi, G., 2004. Local sleep and learning. Nature 430, 78–81.

Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Netw. 10 (3), 626–634.

Hyvarinen, A., 2000. Independent component analysis: algorithms and applications. IEEE Trans. Neural Netw. 13 (4–5), 411–430.

Kim, K.H., Yoon, H.W., Park, H.W., 2004. Improved ballistocardiac artifact removal from the electroencephalogram recorded in fMRI. Neurosci. Methods 135 (1–2), 193–203.

Laufs, H., Kleinschmidt, A., Beyerle, A., Eger, E., Salek-Haddadi, A., Preibisch, C., Krakow, K., 2003. EEG-correlated fMRI of human alpha activity. Neuroimage 19 (4), 1463–1476.

Liebenthal, E., Ellingson, M.L., Spanaki, M.V., Prieto, T.E., Ropella, K.M., Binder, J.R., 2003. Simultaneous ERP and fMRI of the auditory cortex in a passive oddball paradigm. Neuroimage 19 (4), 1395–1404.

Lu, W., Rajapakse, J.C., 2001. Constrained independent component analysis. Adv. Neural Inf. Process. Syst. 13, 570–576.

MacQueen, J.B., 1967. Some methods for classification and analysis of multivariate observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability.

Mantini, D., Perrucci, M.G., Cugini, S., Ferretti, A., Romani, G.L., Gratta, C., Del, 2007. Complete artifact removal for EEG recorded during continuous fMRI using independent component analysis. Neuroimage 34 (2), 598–607.

Masterton, R.A.J., Abbott, D.F., Fleming, S.W., Jackson, G.D., 2007. Measurement and reduction of motion and ballistocardiogram artefacts from simultaneous EEG and fMRI recordings. Neuroimage 37 (1), 202–211.

Moosmann, M., Ritter, P., Krastel, I., Brink, A., Thees, S., Blankenburg, F., Taskin, B., Obrig, H., Villringer, A., 2003. Correlates of alpha rhythm in functional magnetic resonance imaging and near infrared spectroscopy. Neuroimage 20 (1), 145–158.

Mukhopadhay, S., Ray, G.C., 1998. A new interpretation of non linear energy operator and its efficacy in spike detection. IEEE Trans. Biomed. Eng. 45 (2), 180–187.

Nakamura, W., Anami, K., Mori, T., Saitoh, O., Cichocki, A., Amari, S., 2006. Removal of ballistocardiogram artifacts from simultaneously recorded EEG and fMRI data using independent component analysis. IEEE Trans. Biomed. Eng. 53 (7), 1294–1308.

Niazy, R., Ianetti, G., Brady, M., Matthews, P., Smith, S., 2004. A multi-subject study of motor task modulation of EEG-alpha-rhythm/FMRI correlation. Tenth Int. Conf. on Functional Mapping of the Human Brain.

Niazy, R.K., Beckmann, C.F., Iannetti, G.D., Brady, J.M., Smith, S.M., 2005. Removal of FMRI environment artifacts from EEG data using optimal basis sets. Neuroimage 28 (3), 720–737.

Phillips, C., Thielen, G., Maquet, P., 2006. Iteratively applied ICA to reject ballisto-cardiogram artefact of EEG data acquired during sleep fMRI. Human Brain Mapping, Florence.

Rechtschaffen, A., Kales, A., 1968. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Washington DC.

Schabus, M., Dang-Vu, T.T., Albouy, G., Balteau, E., Boly, M., Carrier, J., Darsaud, A., Degueldre, C., Desseilles, M., Gais, S., Phillips, C., Rauchs, G., Schnakers, C., Sterpenich, V., Vandewalle, G., Luxen, A., Maquet, P., 2007. Hemodynamic cerebral correlates of sleep spindles during human non-rapid eye movement sleep. Proc. Natl. Acad. Sci. U. S. A. 104 (32), 13164–13169.

Seeck, M., Michel, C.M., Spinelli, L., Lazeyras, F., 2001. EEG mapping and functional MRI in presurgical epilepsy evaluation. Rev. Neurol. (Paris) 157 (8–9 Pt 1), 747–751.

Sijbers, J., Audekerke, J., Van, Verhoye, M., der Linden, A., Van, Dyck, D., Van, 2000. Reduction of ECG and gradient related artifacts in simultaneously recorded human EEG/MRI data. Magn. Reson. Imaging 18, 881–886.

Srivastava, G., Crottaz-Herbette, S., Lau, K.M., Glover, G.H., Menon, V., 2005. ICA-based procedures for removing ballistocardiogram artifacts from EEG data acquired in the MRI scanner. Neuroimage 24 (1), 50–60.

Tong, S., Bezerianos, A., Paul, J., Zhu, Y., Thakor, N., 2001. Removal of ECG interference from the EEG recordings in small animals using independent component analysis. J. Neurosci. Methods 108 (1), 11–17.

Vincent, J.L., Larson-Prior, L.J., Zempel, J.M., Snyder, A.Z., 2007. Moving GLM ballistocardiogram artifact reduction for EEG acquired simultaneously with fMRI. Clin. Neurophysiol. 118 (5), 981–998.

Wan, X., Iwata, K., Riera, J., Ozaki, T., Kitamura, M., Kawashima, R., 2006. Artifact reduction for EEG/fMRI recording: nonlinear reduction of ballistocardiogram artifacts. Clin. Neurophysiol. 117, 668–680.