



**HAL**  
open science

## A robust method to filter various types of artifacts on long duration EEG recordings

Samuel Boudet, Laurent Peyrodie, Philippe Gallois, Christian Vasseur

► **To cite this version:**

Samuel Boudet, Laurent Peyrodie, Philippe Gallois, Christian Vasseur. A robust method to filter various types of artifacts on long duration EEG recordings. ICBBE 2008, May 2008, shanghai, China. pp.2357 - 2360, 10.1109/ICBBE.2008.922 . hal-00838829

**HAL Id: hal-00838829**

**<https://hal.science/hal-00838829v1>**

Submitted on 27 Jun 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A robust method to filter various types of artifacts on long duration EEG recordings

Samuel Boudet<sup>\*</sup>, Laurent Peyrodie<sup>\*</sup>, Philippe Gallois<sup>‡</sup> and Christian Vasseur<sup>§</sup>

<sup>\*</sup> HEI-ERASM (Hautes Etudes d'Ingénieur) and LAGIS, Lille, 59000, France

Email: samuel.boudet@gmail.com and laurent.peyrodie@hei.fr

<sup>‡</sup> St Vincent Hospital, Lille, France

<sup>§</sup>LAGIS UMR-CNRS 8146, University of Lille1, Villeneuve d'Ascq, 59655, France

**Abstract**—The EEG is a system to measure electrical brain activity using multiple electrodes placed on the scalp. Unfortunately, the signal can be easily contaminated by a lot of noises called artifacts. These can be generated by various actions such as eye blinks, eye movements, muscle activities or small electrode movements. This paper presents a global artifact removal method corresponding to an evolution of the AFOP method (Adaptive Filtering by Optimal Projection) in order to improve its stability. This evolution automatically filters ocular, muscular and heart beat artifacts. The results are validated on long duration EEG recordings containing pathological activity. An expert analysis shows that the cerebral signal is well conserved while a lot of artifacts are removed.

## I. INTRODUCTION

Artifacts represent a major problem on Electroencephalography analysis. Artifacts are defined by signal alterations caused by non-cerebral activity and can have several origins. There are the ocular artifacts that appear when the patient blinks or moves his eyes, muscular artifacts that appear when the patient contracts his jaw or his forehead, and electrode artifacts (including heart beat artifact) generated by electrode mechanical movement.

Artifact amplitude is often much higher than EEG amplitude, thus making the analysis difficult and making the automatic processing impossible. Our team has been working on this subject for years [1] in order to improve automatic treatment of EEG for anticipated detection of epilepsy seizures.

This paper deals with the filtering of long duration EEG recordings (3 to 6 hours) for clinical application in order to ease the neurophysiologic interpretation. The long duration examinations are often more artifacted than standard examination (20 min) because the patient are not asked to relax. The signal over noise ratio can then be very low and all studies show that it is impossible to completely filter high amplitude artifacts without removing EEG signal [2].

This paper presents a global method enabling to filter a large amount of artifacts while preserving the major part of EEG activity and particularly pathological activity. This method is an improvement of AFOP method (Adaptive Filtering by Optimal Projection) that the authors have published in [1], which improves its stability. The AFOP method filters ocular and muscular artifacts but is not well designed to filter electrode artifacts. This article integer a way to filter heart beat artifacts described in sec. II, using the idea of Y. Wang [3].

The medical context is described in the first part, in order to understand what has to be filtered or not. Then, a brief overview of spatial filtering is introduced by presenting background on ICA and the principle of AFOP and its improvement is presented on a third part. Finally a comparison between manual ICA and this method is made to prove that this method can improve filtering in several cases.

## II. MEDICAL CONTEXT

The authors have been working on a long duration examination recording with a 19 electrodes 10/20 system. Two electrodes are added on each wrist to measure Electrocardiographic activity. This examination lasts from 3 to 6 hours. During this time, the patient is exposed to flashing lights at various frequencies and also carries out hyperventilation, which consists in breathing quickly and deeply. Sometimes the patient is also asked to sleep and may enter in the firsts sleeping phases. These tests can reveal pathological trouble or cause epilepsy seizures.

During the standard examination, some elements must be kept due to the fact that they correspond to cerebral activity. It can be noticed especially:

- **Paroxysms** (epilepsy indicators). They are graphical elements that can be seen under various forms. The most common one is called spike wave. The paroxysms belongs to a very large frequency band (1 – 30 Hz).
- **Reactivity to eye closure**. This is characterized by an alpha rhythm apparition (8 – 13 Hz) mainly located in the occipital region.
- **Slow waves** (< 4 Hz). These waves can be seen as pathological for the adult. However, this activity is normal for people under 25 years old during hyperpnoea.

The main purpose of the filtering process is to keep these signals as well as possible and to erase the ones with artifactual origins. The artifact types can be broken down as follows:

- **Ocular artifacts**. Two kinds of ocular artifacts can be observed: eye blinks and eye movements. Eye blinks are represented by a low frequency signal (< 4 Hz) with high amplitude. It is a symmetrical activity mainly located on the front electrodes (FP1, FP2) with a low propagation. Eye movements are also represented by a low frequency signal (< 4 Hz) but have a higher propagation, especially

on the temporal electrode. It is caused by the fact that eyes represent dipoles and their movements lead to an alteration of the electrical field. It is characterized by a dissymmetry between the two hemispheres.

- **Muscle artifacts.** Muscular activity creates high frequency signals ( $> 13$  Hz). There are a lot of muscles all over the head, but the muscles in the forehead and jaw are strongly marked. Jaw muscles are powerful and can produce an important signal on temporal area. Forehead muscles are less powerful but are closer to electrode and produce a signal on frontal area.
- **Mechanical and electrode movements.** It may be caused by moving the wire or by a bad connection between the electrode and the skin. In this case the signal is located on an only electrode. Mechanical artifact may also occur when the patient contracts his muscle. A movement of the facial skin or of head part can then be generated, which will result to an artifact signal on several electrodes. And finally, mechanical artifacts can be caused by heart beats if an electrode is located on a vein. This will create an activity synchronized with the heart but have a more sinusoidal shape. Mechanical and electrode artifacts are mainly low frequency signals ( $< 4$  Hz) due to the fact that they directly correspond to the mechanical movement.

### III. STATE OF ART

For an awake adult, the main part of measured signal is on the frequency band  $\alpha$  (8 - 13 Hz) and the artifacts are either on higher frequencies ( $> 13$  Hz) for muscle artifacts, or on lower frequencies ( $< 8$  Hz) for ocular and mechanical artifacts. However, some EEG signals (particularly pathological element like slow waves and spikes) can occur on the artifactual frequency bands and make a frequency filtering impossible. This is why a lot of work is realized on spatial filters which use information of signal repartition on various channels. Most of the methods use Independent Component Analysis (ICA) to realize this filter [4].

#### A. Spatial filter by ICA

ICA supposes that the various signal channels are a linear mixing of sources [5]:

$$\mathbf{V} = \mathbf{M}\mathbf{S} \quad (1)$$

$\mathbf{V}$  is the signal matrix where lines represent channels and columns time samples and  $\mathbf{S}$  is the signal matrix of sources. The matrix  $\mathbf{M}(m, n)$  is called mixing matrix. If number of sources ( $n$ ) is inferior or equal to number of channels ( $m$ ),  $\mathbf{M}$  will be invertible. The pseudo-inverse matrix  $\mathbf{W} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T$  is called separating matrix.

ICA aims to estimate this separating matrix  $\mathbf{W}$  so that sources ( $\mathbf{S} = \mathbf{W}\mathbf{V}$ ) would be independent. There exists many methods of ICA and each of them uses a different measure of independence called contrast function [5]. For example, T. Jung et al. uses Infomax method which maximizes negentropy of sources [4]. Once sources are defined, they are identified as artifactual or not and the artifactual ones are canceled:  $\mathbf{S}' = \mathbf{D}\mathbf{S}$ , where  $\mathbf{D}$  is a diagonal matrix with 0 on

artifactual components and 1 on the others. The signals are then reconstructed by inverse transformation.

$$\mathbf{V}' = \mathbf{M}\mathbf{D}\mathbf{S} \quad (2)$$

$\mathbf{V}'$  represents filtered signals. A filtering matrix is built using:

$$\mathbf{F} = \mathbf{M}\mathbf{D}\mathbf{W} \quad (3)$$

So, a special filter consists in building a filtering matrix which is a projection matrix (ie.  $\mathbf{F}\mathbf{F} = \mathbf{F}$ ). The signal can then be directly filter by applying this matrix:

$$\mathbf{V}' = \mathbf{F}\mathbf{V} \quad (4)$$

Projection matrices have interesting properties: they are diagonalizable and the eigenvalues are all 1 or 0. They can then be defined by two eigenspaces  $E_1$  and  $E_2$  of respective dimension  $n_1$  and  $n_2$  ( $n_1 + n_2 = n$ ). The subspace,  $E_1$  corresponding to cerebral sources is defined by  $\forall x \in E_1, \mathbf{F}x = x$  and the subspace  $E_2$ , corresponding to artifactual sources is defined by  $\forall x \in E_2, \mathbf{F}x = 0$ .

It is important to notice that the  $n_1$  first line vectors of  $\mathbf{W}$  matrix (eq. 3) represents a base of the subspace  $E_1$  and the  $n_1$  first column vectors of  $\mathbf{M}$  are a base of the orthogonal subspace of  $E_2$ . In the same way, the  $n_2$  last line vectors of  $\mathbf{W}$  matrix represents a base of the subspace  $E_2$  and the  $n_2$  last column vectors of  $\mathbf{M}$  are a base of the orthogonal subspace of  $E_1$ . The mixing of artifactual sources is then equivalent to the separation of cerebral sources and reciprocally.

#### B. ICA automation

One of the major problems of ICA is the necessity to manually identify each component as artifactual or not. Many papers are concentrated on this subject but most of them are concentrated only on ocular artifacts. There is only the method of P. Le Van et al. [2] that treats all types of artifacts. It consists in extracting some features of each source, characterizing location, frequencies, amplitude or other properties. Then a Bayesian classifier is used to decide from these features if a source is artifactual or cerebral.

### IV. AFOP METHOD

The method proposed in this article is not based on ICA but on Common Spatial Pattern (CSP) theory. It consists in constructing a filtering matrix using a training step to learn distribution of artifactual activity. On this purpose, two EEG periods are compared. The first one contains resting EEG and the second one contains EEG contaminated with artifacts. The artifactual sources are then defined by the sources with the most important variance increase between rest instant and artifact instant. On the other way, the EEG sources are the sources with approximate constant variance.

#### A. Protocol

This method has to learn the most common artifacts and requires a protocol realized on each beginning of the recording. During two minutes, the patients have to carry out four artifact types several times. This period is then compared to a rest period (without artifacts) for the learning step. The four artifact types are: eye blinks, eye movements, jaw clenching and forehead movements.

## B. Standard AFOP

The AFOP method [1] is decomposed in two steps. The first one consists in determining a subspace  $E_1$  by learning the separation of cerebral sources with Common Spatial Pattern (CSP) method. The second step consists in determining the  $E_2$  subspace with a least squared method on a sliding window. This  $E_2$  subspace corresponds to the best distribution of the cerebral sources defined by  $E_1$ .

For the first step, the CSP method enables to find sources whom variance increases the less between two learning periods corresponding to matrix  $\mathbf{V}_1$  and  $\mathbf{V}_2$ . The two covariance matrices are then calculated by  $\mathbf{C}_1 = \mathbf{V}_1 \mathbf{V}_1^T$  and  $\mathbf{C}_2 = \mathbf{V}_2 \mathbf{V}_2^T$ . The subspace  $E_1$  correspond to sources increasing the less and it is defined by the eigenvectors  $([\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots] = \mathbf{W}^T)$  corresponding to the smallest eigenvalues of:

$$\mathbf{C}_1^{-1} \mathbf{C}_2 \quad (5)$$

The eigenvalues correspond to the variance ratios between two periods. It is possible to select only eigenvectors whom eigenvalues are inferior to a threshold (empirically fixed to 2.5) in order to determine automatically the dimensions  $n_1$  and  $n_2$  of each subspaces.

Once the subspace  $E_1$  defined, the orthogonal subspace of  $E_2$  is determined by finding the best distribution of cerebral sources by a least squared method on the period to filter. A sliding window of 20 seconds (of signal matrix  $\mathbf{V}_t$ ) is used to cover the entire recordings. The best distribution is given by:

$$\mathbf{M} = \mathbf{C}_t \mathbf{W}^T (\mathbf{W} \mathbf{C}_t \mathbf{W}^T)^{-1} \quad (6)$$

Where  $\mathbf{C}_t = \mathbf{V}_t \mathbf{V}_t^T$  is the covariance matrix. It can be noticed that if the period  $\mathbf{V}_t$  correspond to one of the period  $\mathbf{V}_1$  or  $\mathbf{V}_2$ , the subspace  $E_2$  is given by the  $n_2$  last eigenvectors of eq. 5.

For the artifact filtering application, ones can suppose that both cerebral and artifactual sources are static due to the fact that the brain and the muscles are always on the same place. However, it is possible that some sources would not be activated during the resting period of the learning step. For example the sources of paroxysms may not appear on this period. By the way, there is only the artifactual distribution space that will be considered as constant. The sec. III-A shows that considering artifact distribution as constant is equivalent to consider cerebral separation as constant.

## C. Heart beat artifacts

The AFOP method does not learn artifacts which are constant during the entire recording. This is the case of heart beat artifacts due to the fact that they are as important during the rest period as during artifact period. It is then possible to use the method of J. Wang [3] by determining times of each heart beat by using Electrocardiogram (ECG). The mean of signals taken one second after each beat is computed in order to minimize the amplitude of sources which are not synchronized with ECG. A CSP is then carried out between the entire period to filter and the mean signals. Sources with a variance ratio closed to 1 are artifactual sources and are added to artifact repartition subspace (orthogonal of  $E_1$ ).

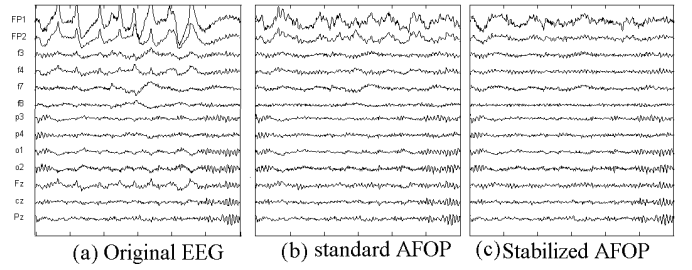


Fig. 1. Illustration of stability improvement

## D. Stabilized AFOP

In the case where the artifact distribution is slightly different to the one of the learning period, some instability may occur. This can be explained by the fact that a small artifact appears on cerebral sources and the AFOP method tries to rebuild the original signal by incorrectly amplifying those sources. The fig. 1.b illustrates the result of instability problem. A meaningless signal appears on channel FP1 and FP2 because the method tries to rebuild the artifact.

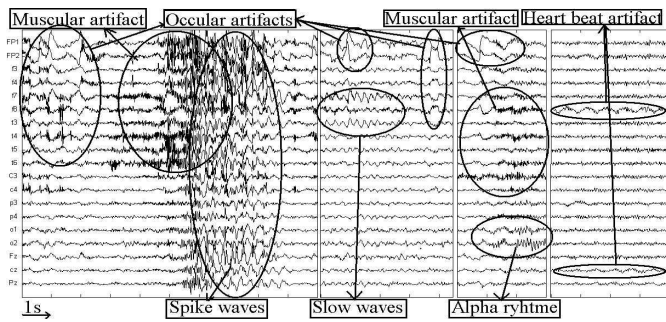
This instability can be translated by a small angle between the two subspaces  $E_1$  and  $E_2$ . It is in general due to the fact that the cerebral sources are neglecting compared to the artifactual ones. There are then a lot of liberty degrees to rebuild the artifacts. These liberty degrees can then be deleted by doing a Principal Component Analysis (PCA) on the current period. Then, a reduced space based on the first components of PCA (those of bigger eigenvalues) is considered and the subspace  $E_1$  is projected on this base. The normal process is then carried out on this reduced space.

The number of components to keep with the PCA decreases from  $n$  until the angle between the two subspaces would be superior to a threshold (fixed empirically to  $7^\circ$ ) or until too much signal would be deleted with the PCA. The fig. 1.c illustrates the result of this method.

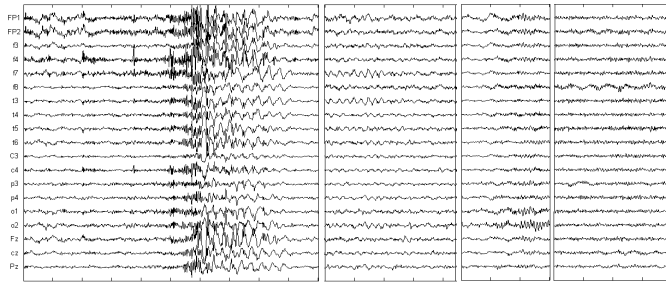
## E. Band frequency decomposition

The described method enables to spatially characterize the artifacts. However, artifacts can be defined as well by their frequencies. This is the reason why each of these steps can be carried out on frequency band decomposition. The signals in each frequency band are treated independently and when all frequency bands are spatially filtered, the final EEG signal will be their sum. The selected frequency bands are the followings:

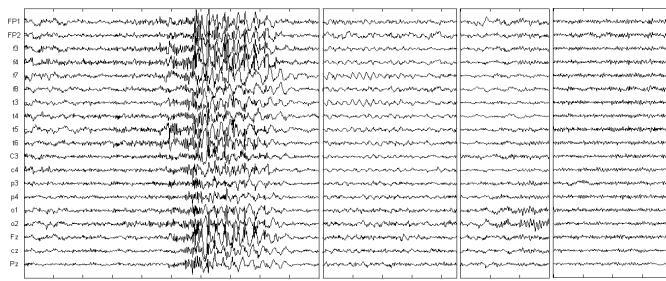
- 1) The band (1-8 Hz) corresponding to neurologic band  $\Delta$  (1-4 Hz) and  $\theta$  (4-8 Hz). These signals are on low quantity on normal EEG but their presences often reveal abnormalities. All the eye and electrode artifacts are contained on this band.
- 2) The band (8-13 Hz) corresponding to neurologic band  $\alpha$ . Most of "Normal" EEG signal is in this band, generally poor in artifacts.
- 3) The band  $\beta$  ( $> 13$  Hz). There are very few EEG  $\beta$  signals that can be measured at the brain surface. This band contains mainly muscular artifacts but during paroxysms the spikes are on this band. Taking account



(a) Artifacts original signal



(b) Filtered signal with ICA



(c) Filtered signal with stabilized AFOP with frequency band decomposition

Fig. 2. Examples of filtering results with ICA and stabilized AFOP

of its width, a more precise segmentation is carried out for the followed frequency bands (13-20 Hz), (20-30 Hz) and (> 30) Hz).

For the band (1-8 Hz), only ocular artifacts and heart beats artifacts are considered so the matrix  $V_2$  is composed only of ocular artifacts. On the other bands, only muscular artifacts are considered and the stabilization is less important than in low frequency.

## V. RESULTS

Six recordings of epileptic adult patients have been treated with this method and examined by an expert. Two of these recordings contain spike wave paroxysms, and three contain slow waves activities. The fig. 2(c) illustrates some examples of result that can be obtained with Stabilized AFOP and the fig. 2(b) shows results that can be obtained by manual ICA. The ICA filtering has been made on 20 second pages EEG with JADE algorithm. The automatic method using ICA gives worst result than manual ICA since these methods can only add classification error of components.

This figure shows that both methods filter most of artifacts and remove only a small quantity of EEG signals. It can

be noticed that muscular artifacts are often less filtered with ICA method and they are about equivalent concerning ocular artifacts. This can be explained by the fact that the frequency band decomposition treats each frequency independently. In general case, ICA trends to badly filter short duration artifacts because there is not enough time to learn them. By the other way, AFOP cannot filter electrode artifacts because the spatial distribution is always different. These examples show that both AFOP and manual ICA reduce the amplitude of pathological elements of about 25%. Given that the AFOP method continually applies the filter, this reducing is present even if there is no artifact contrary to ICA but the results of AFOP are a little better if there is a mixing of EEG and artifacts because ICA may not manage to correctly separate the sources. On all realized tests, the  $\alpha$  rhythm is never reduced whereas ICA can reduce it in presence of many artifacts. The last part of this figure shows that AFOP method perfectly filter heart beat artifacts whereas it is not always the case with ICA.

This method has been programmed in Matlab using EEGLab toolbox [6]. The computation of AFOP is instantaneous and there is only the filtering step for the frequency band decomposition that takes about 10 minutes on a Pentium 4 3GHz, for a 3 hours recording.

## VI. CONCLUSION

The AFOP method filters automatically a large variety of artifacts. However, some instability can occur when artifacts are too important compared to the EEG signal or when the training is not precise enough. This paper has presented a method increasing this stability. Using the frequency band decomposition, the results are often better than those realizable with ICA. There is only mechanical movement artifacts that this method is badly defined to filter due to the fact these artifacts are not static. The study on long duration recording shows that it is possible to filter a large amount of artifacts while keeping most of the EEG signal.

## ACKNOWLEDGMENT

The authors gratefully acknowledge the medical personnel of Hospital St Vincent and St Philbert of Lille. They are also obliged for the financial support of FRM.

## REFERENCES

- [1] S. Boudet, L. Peyrodie, P. Gallois, and C. Vasseur, "Filtering by Optimal Projection and application to automatic artifact removal from EEG," *Signal Process.*, vol. 87, no. 8, pp. 1978–1992, 2007.
- [2] P. Le Van, E. Urrestarazu, and J. Gotman, "A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification," *Clinical Neurophysiology*, vol. 117, no. 4, pp. 912–927, 2006.
- [3] Y. Wang, "Reduction of cardiac artifacts in magnetoencephalogram," *Proc. of the 12th International Conference on Biomagnetism*, 2000.
- [4] T. Jung, S. Makeig, C. Humphries, T. Lee, M. McKeown, V. Iragui, and T. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163–178, 2000.
- [5] P. Common, "Independent component analysis, a new concept," *Sig. Pro.*, vol. 36, pp. 287–314, 1994.
- [6] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.