# EEG artifact correction

Lisa Tostrams

July 2017

## 1 Problem description

One of the methods used in the CHDR to asses the effect of a drug is resting state EEG. For the duration of a few minutes, cortical activity is measured in subjects who close and open their eyes. From the power spectrum, some indications about the state of the central nervous system can be extracted. However, these measurements are often influenced by artifacts (e.g. eye and muscle movements). Parts of the EEG recording are rejected when an artifact has been (sometimes manually) detected.

Within CHDR there is a growing interest in acquiring Machine Learning techniques to analyze data. One possibly interesting application of such techniques is the automatic detection and correction of artifacts in EEG recordings.

The goal of this report is to provide an overview of methods to analyze data from functional brain imaging (EEG) in the presence of (ocular) artifacts, using Machine Learning and pattern recognition. The goal of these methods is to detect the presence and correct the effects of such artifacts. Besides providing an overview of methods, this report will recommend performance measurements and a modelling/experimental set up, and establish the amount of data needed to asses and implement such methods.

## 2 Electroencephalography

Electroencephalography (EEG) is a monitoring method to record electrical activity of the brain. The recording is non-invasive with electrodes placed on the scalp [8]. In measuring EEG, the interest typically lies in underlying neural potentials.

**TODO: add EEG normal patterns, show recording setup, explain what is measured exactly**

### Artifacts

Signals detected by the EEG recording that did not originate from the cortex are considered artifacts. The EEG signal is often contaminated with various physiological factors other than cerebral activity [11]. Ocular movements and eye blinks are some of the most common types of biological artifacts. The correction or rejection of artifacts is an important issue in EEG signal processing and an necessary component of most signal analysis.

#### Ocular artifacts

The eye forms a electric dipole, where the cornea is positive and the retina is negative. Eye movements change the electrical field around the eye producing an electrical signal known as the electro-oculogram (EOG). The electrical activity produced by eye movement is usually strong enough to be recorded along with the EEG [2]. As a result, EEG recordings are often significantly distorted. The eye-movement related voltages propagate to sites along the scalp with an inverse relation to the squared distance from the eyes. About 0.20 of the vertical eye-moment voltage reaches $F_z$, and this amount decreases posteriorly to about 0.05 at occipital sites.

Blinking is another cause for contamination of the EEG, in a more abrupt manner than eye movement and with a higher amplitude.

When the effect of ocular activity can be estimated, this signal can be removed from the EEG (EOG correction).

# 3 Performance Evaluation

### Compare to linear regression methods

# 4 Methods

## 4.1 EOG correction [2]

EOG correction generally refers to methods that assume the measured EEG is a linear combination of the true signal and ocular artifacts. The proportion of ocular activity present in each channel is calculated by regression. The correction is done by subtracting the regressed portion of the EOG reference wave forms from each EEG channel.

Advantage: robust

Disadvantage: need for EOG reference signal

## 4.2 Filtering

Advantages of filtering: no need for labeled training data, can be automated

Disadvantages of filtering: need a reliable model or measurement of ocular artifacts

### 4.2.1 Adaptive filtering [1] [4] [9]

An adaptive filter is a system with a linear filter that has a transfer function controlled by variable parameters and a means to adjust those parameters according to an optimization algorithm. The adaption of filter parameters is based on minimizing the mean squared error between the filter output and the desired signal. The most common adaptation algorithms are Recursive Least Square (RLS) and Least Mean Square (LMS). RLS offers a higher convergence speed but LMS offers computational simplicity. The filter is controlled by a set of coefficients or weights $w(n)$. LMS is based on gradient search according to the equation

$$w(n+1) = w(n) + \mu e(n)x(n)$$

where $w(n)$ is the weights vector at sample $n$, $x(n)$ is the input signal simple vector, $e(n)$ is the filters error and $\mu$ is the filters convergence factor.

To remove ocular artifacts with adaptive filtering, separately recorded vertical EOG and horizontal EOG signals are used as reference inputs.

Advantage: adaptive, minimal training, real time

Disadvantage: need for EOG reference signal

Data needed for training: none

### 4.2.2 Wiener filtering [5] [3] [9]

Wiener filtering produces a filter that minimizes the mean square error between the desired signal and its estimate. The filter is calculated from the spectra of both the signal $\phi_{signal}(\omega)$ and the noise $\phi_{noise}(\omega)$, which are in turn calculated from the averages of spectra of the measured signals or by obtaining pure eye blinking components with Independent Component Analysis. Therefore, Wiener filtering does not need a reference signal.

Advantage: no need for EOG reference

Disadvantage: need a reliable model of artifact and true signals
Data needed for training: representative true signal and artifact recordings

#### 4.2.3 Bayesian filtering [10] [11] [9]

Bayesian filtering refers to formulating the problem of estimating the state of a time-varying system that is indirectly observed in terms of Baying probability theory. Like the Kalman filter, the Bayesian filter produces an estimate of current state variables from a probability distribution. However, the observations are assumed to be independent and the estimate requires formulating a joint probability distribution over all previous measurements. Therefore, Bayesian filtering is limited by its computational complexity. For every observation $O_t$ and true state $S_t$ the sensor model is defined by $P(O_t|S_t)$. Using the Bayes rule, the distribution of probabilities for the current state $P(S_t|O_0, .., O_t)$ based on the observation history is computed. Bayesian filtering assumes the process is Markov, meaning that knowledge of the current state contains all relevant information about the true process. The Kalman and particle filters approximate the Bayesian filter.

Advantage: no need for EOG reference, only need a sensor model

Disadvantage: computationally intractable, need a prior distribution, need reliable sensor model

Data needed for training: a lot of representative data to calculate prior distributions

#### 4.2.4 Kalman filtering [12] [3] [7] [9]

The Kalman filter is a set of equations that provide a means to estimate the state of a process from a series of measurements in a way that minimizes the mean squared error. The algorithm for establishing the true state works in a two-step manner: in the prediction step, the filter produces an estimate of the current state variables and their uncertainties. In the update step, the measurement is used to update the estimates using a weighted average (the lower the uncertainty, the higher the weight). The Kalman filter needs a model for measurement noise and process noise variances to produce reliable predictions, but is only dependent on the current measurement and those models.

Advantage: no need for EOG reference (if measurement noise model is available)

Disadvantage: need reliable measurement and process noise models (if EOG reference is not available)

Data needed for training: data on measurement and process noise

#### 4.2.5 Particle filtering [9]

Particle filtering uses Monte Carlo sampling to calculate a state transition model of the process. N samples are selected from a probability density function, and are weighted according to the amplitude of the probability density function at their sample point. The weights of new particles are determined using an update equation that uses the sensor model, and the estimates state is the mean of the probability density function.

Advantage: no need for EOG reference, no need for process and measurement noise model

Disadvantage: sampling doesn't necessarily lead to a good state transition model, may not converge correctly

Data needed for training: representative data to calculate prior distributions

### 4.3 Blind Source Separation

Advantages: unsupervised learning

Disadvantages: a lot of assumptions needed

#### 4.3.1 Independent Component Analysis [11] [6]

Independent component analysis separates a multivariate signal into additive sub-components using only recorded information, by imposing statistical independence of the signal sources. This independence implies no spatial, temporal or time-frequency correlation between source signals. The ICA algorithms that exploit higher-order statistics (HOS-ICA) start by applying a orthogonal transformation (PCA), and then finding the linear transformation that results in the most independent estimated sources. ICA algorithms that apply second-order statistics are based on decorrelating the data in the time domain.

A drawback of ICA is the required intervention of a trained professional to manually identify the components related to artifacts. ICA can be combined with Bayesian classification to be automated, by computing the probability that an epoch represents EEG activity. Again, calculating the necessary probabilities for Bayesian learning is computationally intractable, but they can be estimated by using tree-augmented naive Bayesian networks that construct a maximum spanning tree based on feature dependencies. (Can be optimized further with Variable Elimination?)

Advantages: combines the best of ICA and Bayesian learning, automating a reliable and tested method

Disadvantage: violates the notion of unsupervised learning by introducing Bayesian learning

Data needed for training: around 200 recordings, according to [6]

#### 4.3.2 Principal Component Analysis

#### 4.3.3 Canonical Correlation Analysis

### 4.4

## References


[1] J.G. Avalos, J.C. Sanchez, and J. Velazquez. *Applications of Adaptive Filtering*. InTech, 2011.

[2] R.J. Croft and R.J. Barry. Removal of ocular artifact from EEG: a review. *Journal of Clinical Neurophysiology*, 30(3):5–19, 2000.

[3] J. Ferdous and S. Ali. A comparison of wiener and kalman filters for the artifact suppression from EEG signal. *International Journal of Science and Research*, 6(4):2029–2035, 2017.

[4] P. He, G. Wilson, and C. Russel. Removal of ocular artifact from electro-encephalogram by adaptive filtering. *Medical and Biological Engineering and Computing*, 42(3):407–412, 2004.

[5] S. Krieger, J. Timmer, S. Lis, and H.M. Olbrich. Some considerations of estimating event-related brain signals. *Journal of Neural Transmission*, 99(1-3):103–129, 1995.

[6] P. LeVan, E. Urrestarazy, and J. Gotman. A system for automatic artifact removal in ictal scalp eeg based on indepentet component analysis and Bayesian classification. *Clinical Neurophysiology*, (117):912–927, 2006.

[7] F. Morbidi, A. Garulli, D. Prattichizzo, C. Rizzo, and S. Rossi. Application of kalman filter to remove TMS-induced artifacts from EEG recordings. *Transactions on control systems technology*, 16(6):1360–1367, 2008.

[8] E. Niedermeyer and F.L. da Silva. *Electroencephalography: Basic Principles, Clinical Applications and related fields*. Lippincott Williams & Wilkins, 2004.

[9] K.T. Sweeney, T.E. Ward, and S.F. McLoone. Artifact removal in physiological signals – practices and possibilities. *Transactions on information technology in biomedicine*, 16(3):488–501, 2012.

[10] S. Särkkä. *Bayesian filtering and smoothing*. Cambridge University Press, 2013.

[11] J.A. Urigüen and B. Garcia-Zapirain. EEG artifact removal – state-of-the-art and guidelines. *Journal of Neural Engineering*, 12(3):23pp, 2015.

[12] G. Welch and G. Bishop. *An introduction to the Kalman Filter.* Department of Computer Science, University of North Carolina at Chapel Hill, 2006.