# An Application of Gaussian Processes on Ocular Artifact Removal from EEG*

Saman Noorzadeh[1,2], Bertrand Rivet[1], Pierre-Yves Guméry[2]

*Abstract*— **Consequences of eye movements are one of the main interferences that distort the brain EEG recordings. In this paper, a multi-modal approach is used to estimate the ocular artifacts in the EEG: both vertical and horizontal eye movement signals recorded by an eye tracker are used as a reference to denoise the EEG. A Gaussian process, i.e. a second order statistics method, is assumed to model the link between the eye tracker signals and the EEG signals. The proposed method is thus a non-linear extension of the well-known adaptive filtering and can be applied with a single EEG signal contrary to independent component analysis (ICA) which is extensively used. The results show the applicability and the efficiency of this model on the ocular artifact removal.**

## I. INTRODUCTION

The non-invasive electroencephalogram (EEG), which measures the electrical brain activities through surface electrodes, can be distorted by several sources of noise that make it difficult to interpret [1]. Eye movements are among these interferences, and there is a rapidly growing literature on the removal of such artifacts. A common way to avoid the ocular artifacts is to restrict eye movements during the data recording; however, this method can limit the experiments for which the EEG is recorded [2]. It should also be mentioned that a band-pass filter is not capable of removing this artifact completely since it may remove also some information in the brain activity, or may lead to a step-like change at the eyeball rotation moment, due to the impulse response of the filter. Among other studies that have tackled this problem, concerning signal processing methods, one can name regression method [3] or blind source separation methods like Principal Component Analysis (PCA) [1] and Independent Component Analysis (ICA) [4], [5]. ICA, which is more popular in the literature, can decompose a noisy EEG into statistically independent components [6], [7], [8], [9]; however, its limitations would make it impractical because, first, it needs a large number of sensors and, second, some studies have assumed that the data recorded via certain electrodes around eyes contain only the ocular activities (Electrooculography). Since this assumption is not completely true (EEG signal is also present in the mentioned channels), it may lead to the subtraction of a portion of the brain activity in the final estimation of EEG. Another method which has been used in this regard is the adaptive filter [10], [11] which removes the artifacts but can also attenuate some parts of the wanted brain activities. Moreover, the EEG signals and the reference signals are assumed to be correlated by a linear system, which is not necessarily a correct assumption.

This paper intends to investigate the application of Gaussian processes (GPs) [12], [13] on modeling the ocular artifacts to remove them from the EEG signal. The GP is a flexible, non-parametric model and a practical tool to work with. One of these practical characteristics is that it can be defined only by the first and the second order moments of the process [12]. Another property of GP is that it can be used in a Bayesian setting where the GP is a prior on the functions and can provide a probability measure over the function space. Therefore, GP modeling provides the possibility of flexible models. Afterwards, the prediction can be done in a straightforward way within a Bayesian framework. Another interesting feature of GP model is its ability to solve nonlinear estimation [14] meaning that the flexible prior function would handle the nonlinearity, and the nonlinear estimation can be treated within a Bayesian framework.

Considering the mentioned characteristics, in [15], [16] the Gaussian process has been used to extract the fetal Electrocardiogram (ECG) from the abdominal observations by modeling the maternal and fetal ECGs using GPs. With the same reasoning, the GP model can also handle the problem of identification and removal of ocular artifacts. In this study, we use a two-channel approach in which the eye movement data act as a reliable input of the Gaussian Process. Using this reference channel, the ocular signal existing in the EEG observation is modeled as a GP. Having the GP model, this signal can then be estimated and removed from the observation channel and, as a result, the brain activity signal without ocular distortions is obtained.

The rest of the paper is organized as follows: in section II the Gaussian process model is introduced, and section III shows the application of this model on removing ocular artifacts from EEG. In section IV the results of this approach are shown on real data, and are compared with the results obtained from adaptive filters and with ICA, and finally in section V the conclusions and perspectives are presented.

## II. GAUSSIAN PROCESS MODEL

Gaussian process can describe a signal only using its second order statistics. Considering $s(t)$ as a real process,

*This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) and ERC AdG-2012-320864- CHESS.

[1]The authors are with the GIPSA-lab, CNRS UMR 5216, Joseph Fourier University, Grenoble, France `firstname.lastname@gipsa-lab.grenoble-inp.fr`

[2] The authors are with the TIMC-lab, UMR-CNRS 5525, Joseph Fourier University, Grenoble, France `firstname.lastname@imag.fr`

978-1-4244-9270-1/15/$31.00 ©2015 IEEE

it can be defined by its mean and covariance functions [12]:

$$m(t) = \mathbb{E}[s(t)],$$
$$k(t,t') = \mathbb{E}[\big(s(t)-m(t)\big)\big(s(t')-m(t')\big)].$$

If we assume the mean function to be zero for all the inputs ($m(t)=0$), it is only the covariance function which has to describe the data characteristics. This assumption is logical for the EEG signal, and also makes the notations simpler. To give an illustration of the model, let us consider the square exponential covariance function which is a popular kernel in describing physiological processes. This covariance function is defined by equation (1), for the real process $s(t)$:

$$k(t,t') = exp\big(-\frac{||\vec{u}(t)-\vec{u}(t')||_2^2}{2l^2}\big), \tag{1}$$

where,$||.||_2^2$ denotes the $L_2$ norm, and $l$ is the length-scale which defines the smoothness of the function. This covariance function is defined as the function of the input space, $\vec{u}(t)$. The input space can be either the time points where $\vec{u}(t)=t$, or any other inputs which can be mapped by a function $f$ to the output $s(t)$ : as $s(t)=f(\vec{u}(t))$. Having defined a covariance and mean function, the set of real valued functions, $s(t)\in\mathbb{R}$, can then be described as a Gaussian Process:

$$s(t) = \mathcal{GP}(0,k(t,t')). \tag{2}$$

Turning now to the problem of separation of sources, consider the signal $s(t)$ in which we are interested to extract from a noisy observed signal called $x(t)$:

$$x(t) = s(t) + n(t), \tag{3}$$

where $n(t)$ is considered as the white noise defined here as a GP: $\mathcal{GP}(0,\sigma_n^2\mathbf{I})$, with $\sigma_n^2$ as the power of the noise and $\mathbf{I}$ as the identity matrix. Having the prior defined by the GP for $s(t)$ in equation (1), the posterior distribution on functions $s(t)$ can be computed considering the set $\{x(t_i),\vec{u}(t_i)\}_{1\leq i\leq T}$ of $T$ pairs of the observation $x(t_i)$ and the related input $\vec{u}(t_i)$. Consequently, the prediction of $s(t^*)$ at any time $t^*$ would be the maximum of the posterior function defined as:

$$\hat{s}(t^*) = \underset{s(t^*)}{\mathrm{Argmax}}\, p\big(s(t^*)|t^*,\{x(t_i),\vec{u}(t_i)\}_{1\leq i\leq T}\big). \tag{4}$$

It is worth noting that the time $t^*$, defined for the estimated signal, can be different from that of the observations. The final estimation of the $s(t^*)$ signal can be expressed as the mean of the posterior function as:

$$\hat{s}(t^*) = \mathbf{k}(t^*)(K+\sigma_n^2\mathbf{I})^{-1}\mathbf{x}, \tag{5}$$

where $K\in\mathbb{R}^{T\times T}$ is the covariance matrix calculated from equation (1) for every pair of the observation input points, i.e. $K_{i,j}=k(t_i,t_j)$, $\mathbf{k}(t^*)=[k(t^*,t_1),\cdots,k(t^*,t_T)]$ is the covariance vector and $\mathbf{x}=[x(t_1),\cdots,x(t_T)]^\dagger$, with $\cdot^\dagger$ the transpose operator. This model is used to handle the ocular artifacts in the next section.

## III. REMOVING OCULAR ARTIFACTS

In this section, the possibility of the extraction of ocular artifacts from the EEG signal is investigated. For this purpose, besides EEG observation which we intend to denoise, we have also considered another data as the reference for the ocular artifacts. Here, the reference is considered to be the eye movements data. This data consists of two time series indicating the vertical and the horizontal eye movements, which are indicated as $e_x(t)$ and $e_y(t)$ respectively. As previously mentioned in section II, we can model the ocular artifact existing in the EEG if we can have a GP model with the covariance function that fits this data. Let us define the observed EEG, $x(t)$, and the ocular artifact, $s(t)$, as in equation (3). Here the $n(t)$ signal expresses the rest of the EEG separated from the ocular artifact, and $s(t)$ and $n(t)$ are assumed to be decorrelated. The $s(t)$ signal is then expressed to be non-linearly correlated with both the $e_x(t)$ and $e_y(t)$ as:

$$s(t) = f(\vec{e}_N(t)) + \epsilon(t), \tag{6}$$

where $f(\cdot)$ is the non-linear relationship between the eye position and the ocular artifact in the EEG recordings. $\vec{e}_N(t)$ is the input space related to the eye position defined as the concatenation of $e_x(t)$ and $e_y(t)$ for a window of length $N$, and $\epsilon(t)$ is an additive noise to model the remaining error. The non-linear mapping $f(\cdot)$ is defined as a GP

$$f(\vec{e}_N(t)) = GP(0,k(t,t')), \tag{7}$$

with the covariance function $k(\cdot,\cdot)$ defined as:

$$k(t,t') = \sigma^2 exp\Big(-\frac{1}{2}\big(\vec{e}_N(t)-\vec{e}_N(t')\big)^\dagger \Sigma^{-1}\big(\vec{e}_N(t)-\vec{e}_N(t')\big)\Big). \tag{8}$$

This covariance function is different from equation (1) in the sense that here we have two references which can have a correlation with each other and their relation is expressed in the covariance matrix $\Sigma$. This matrix is a block matrix of four diagonal matrices whose diagonal values define the smoothness behavior of the GP functions according to the relation of the references. $\sigma$ also defines the power of the GP. The hyper-parameters of the model ($\sigma$, $N$, and the diagonal values of the blocks of $\Sigma$) can be estimated according to the maximum likelihood framework or using a prior on parameters to maximize the posterior [15]. If the hyper-parameters are well defined, we have a covariance function and, consequently, we can obtain a GP which can describe the ocular activity. This activity can then be estimated as $\hat{s}(t^*)$ by the mean of the posterior function (as defined in equation (5)). Subsequently, this estimation can be subtracted from the EEG observed channel to have an estimation of the denoised EEG, $\hat{n}(t^*)$:

$$\hat{n}(t^*) = x(t^*) - \hat{s}(t^*) \tag{9}$$

It should be noted that the GP does not assume a linear relationship between the input and the output of the GP, meaning that the dependency between the two modalities,

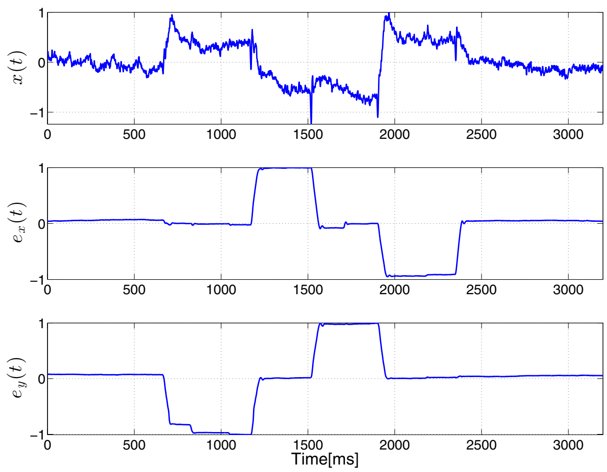


Fig. 1. EEG and eye movement signals recorded simultaneously. The figure on top shows the brain activities distorted by ocular artifacts captured with an EEG electrode. The next two figures show the horizontal and vertical eye movements captured with the eye tracker respectively.

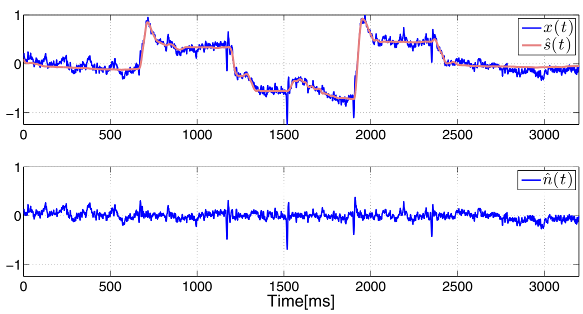


Fig. 2. Estimation of ocular artifacts and EEG using the proposed method. The figure on top shows the noisy EEG, $x(t)$, and the estimation of ocular artifacts, $\hat{s}(t)$. The next figure shows the estimated remaining brain activity, $\hat{n}(t) = x(t) - \hat{s}(t)$.

EEG and eye movements, can also be a non-linear relation [14].

## IV. RESULTS

The extraction of the ocular artifact using GP is tested on the data collected from a real subject. The data is taken from [17]. The participants have followed a given pattern on the screen with the eyes, and meanwhile EEG and eye-tracking signals were jointly recorded. Fig. 1 shows an EEG signal channel, referred to as $x(t)$, and the horizontal and vertical eye movements, referred to as $e_x$ and $e_y$ in the previous section. All the results which are shown are normalized in amplitude.

Using the proposed method described in Section III, the recorded EEG signal is separated to the brain activity and ocular artifacts. The reference eye movements which are used, are the ones in Fig. 1. This result is shown in Fig. 2. As one can see, the estimated brain signal $\hat{n}(t)$ does not contain some signals synchronized with the eye movements.

This result can be compared with the results obtained from other methods. For instance, the eye tracking signals of

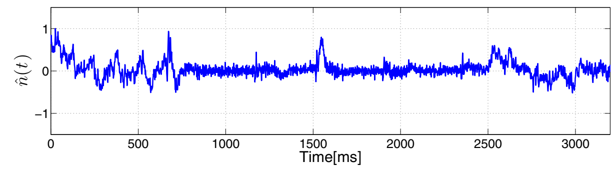


Fig. 3. Removal of ocular artifacts from EEG using adaptive filter. $\hat{n}(t)$ shows the estimated EEG.

Fig. 1 are given sequentially to two FIR adaptive filters as the inputs. 150 coefficients are considered for each filter and the Least Mean Square algorithm is used for the estimation of these coefficients. From Fig 3, it is clear that the estimation of the brain activity is not as good as the results of the GP since not only the final estimation of brain activity still contains some ocular artifacts, but also the middle part of the estimated signal has been attenuated. This is because of the fact that this portion is considered to be noisy and adaptive filter has also removed some part of the brain activity considering it as the ocular noise. GP outperforms the adaptive filter since a linear filter is not sufficient to extract the two modalities, however the GP gives a non-linear estimation.

ICA can also be compared to the previous results. We have used the FastICA algorithm [18]. Two experiments are designed: using only three mixtures $e_x(t)$, $e_y(t)$ and a single EEG sensor (Figure 4) and using seven mixtures $e_x(t)$, $e_y(t)$ and five additional EEG sensors (Figure 5). In the first case (Figure 4), two estimated independent components (ICs) are clearly related to the eye movements and the third one is then the estimation of the brain activity. However, one can see that the "eye" components do not only contain ocular artifact but also some brain activity (i.e. some small variations around the step-like curves) and that some ocular artifacts are remaining into the "brain" component. In the second case (Figure 5) some extra EEG sensors are added so that a total of five EEG signals are used in addition to the two eye-tracker signals. In this case two ICs are identified as related to the ocular artifacts while the five remaining ICs are associated with brain activity. However, even if the "brain" ICs are better estimated using more EEG sensors, the "eye" ICs still contain some brain activities. Consequently, removing these components to reconstruct the brain signal in the EEG channels will result in a loss of information in the desired estimated EEG signal.

## V. CONCLUSIONS AND PERSPECTIVES

A Gaussian process model is presented and adapted to handle the problem of ocular artifacts removal from the brain activities. A single EEG channel is used in a multi-modal approach using eye tracking signals as reference. This approach is a two-channel approach in which two modalities of data are used: EEG, recorded non-invasively through the electrode on the scalp, and the horizontal and vertical eye movements are captured with an eye tracker. The ocular activity is modeled using a GP with a zero mean function and the two eye movements are concatenated

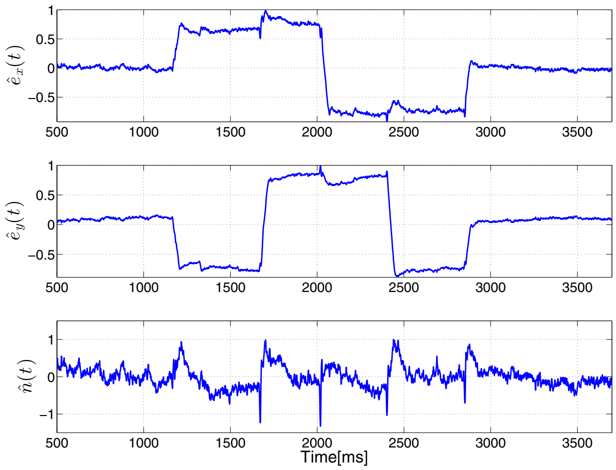


Fig. 4. Estimation of ocular artifacts and EEG using ICA with 3 mixtures: $x(t), e_x(t), e_y(t)$. The three estimated independent components are plotted: the first two ones are related to the eye movements and the third one is the estimation of the brain activity.

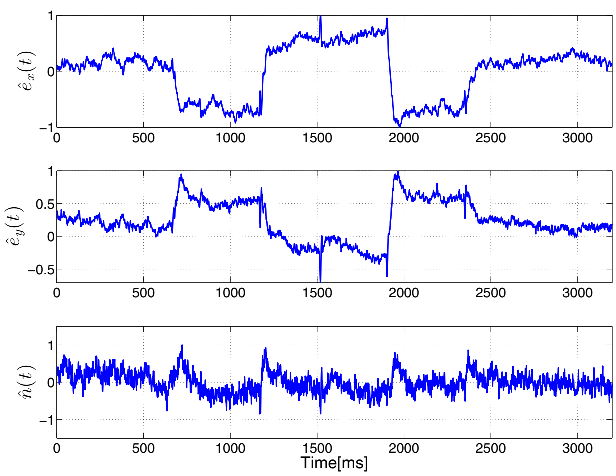


Fig. 5. Estimation of ocular artifacts and EEG using ICA with seven mixtures: $e_x(t)$, $e_y(t)$ and five EEG channels. The two figures on top show the ICs identified as "eye" components. The next figure is one of the ICs associated with the brain activity.

for the computation of the covariance function. The eye movement is then used as the input of the GP, since it is correlated with the ocular artifact existing in the brain activity. Having considered a proper model for the ocular artifact, and the noisy observation, this artifact is estimated using the mean of the posterior probability density function and is finally subtracted from the observed signal. The result of this subtraction is the extraction of the EEG which is denoised from the ocular activity. The GP method uses a non-linear relationship between the input and the output of the GP. The experiment shows that it can work with only a single EEG observation channel which makes it a practical method and it outperforms more classical adaptive (linear) filters or ICA when few sensors are used.

Future work will involve using other kinds of input to the GP, since the reference signals that are used in this work can be replaced by other kinds of signals as long as they are correlated with the ocular artifacts. This can include the EOG (Electrooculography) signals which are captured by sensors placed around the eyes and contain most of the ocular activities.

## References


[1] T. Jung, C. Humphries, T. Lee, S. Makeig, M. McKeown, V. Iragui, and T. Sejnowski, "Removing electroencephalographic artifacts: comparison between ica and pca," in *Neural Networks for Signal Processing VIII, 1998. Proceedings of the 1998 IEEE Signal Processing Society Workshop*. IEEE, 1998, pp. 63–72.

[2] O. Dimigen, W. Sommer, A. Hohlfeld, A. Jacobs, and R. Kliegl, "Coregistration of eye movements and EEG in natural reading: analyses and review." *Journal of Experimental Psychology: General*, vol. 140, no. 4, p. 552, 2011.

[3] H. V. Semlitsch, P. Anderer, P. Schuster, and O. Presslich, "A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP," *Psychophysiology*, vol. 23, no. 6, pp. 695–703, 1986.

[4] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[5] R. Croft and R. Barry, "Removal of ocular artifact from the EEG: a review," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 30, no. 1, pp. 5–19, 2000.

[6] C. Joyce, I. Gorodnitsky, and M. Kutas, "Automatic removal of eye movement and blink artifacts from EEG data using blind component separation," *Psychophysiology*, vol. 41, no. 2, pp. 313–325, 2004.

[7] R. Vigário, "Extraction of ocular artefacts from EEG using independent component analysis," *Electroencephalography and clinical neurophysiology*, vol. 103, no. 3, pp. 395–404, 1997.

[8] T. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. Sejnowski, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clinical Neurophysiology*, vol. 111, no. 10, pp. 1745–1758, 2000.

[9] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings," *Clinical neurophysiology*, vol. 118, no. 1, pp. 98–104, 2007.

[10] B. Widrow, J. R. Glover Jr, J. McCool, J. Kaunitz, C. Williams, R. Hearn, J. Zeidler, E. Dong Jr, and R. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, 1975.

[11] P. He, G. Wilson, and C. Russell, "Removal of ocular artifacts from electro-encephalogram by adaptive filtering," *Medical and biological engineering and computing*, vol. 42, no. 3, pp. 407–412, 2004.

[12] C. Rasmussen, "Gaussian processes for machine learning," 2006.

[13] M. Seeger, "Gaussian processes for machine learning," *International Journal of Neural Systems*, vol. 14, no. 02, pp. 69–106, 2004.

[14] F. Pérez-Cruz, S. Van Vaerenbergh, J. Murillo-Fuentes, M. Lázaro-Gredilla, and I. Santamaria, "Gaussian processes for nonlinear signal processing: An overview of recent advances," *Signal Processing Magazine, IEEE*, vol. 30, no. 4, pp. 40–50, 2013.

[15] S. Noorzadeh, M. Niknazar, B. Rivet, J. Fontecave-Jallon, P. Guméry, and C. Jutten, "Modeling quasi-periodic signals by a non-parametric model: Application on fetal ECG extraction," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 2014, pp. 1889–1892.

[16] S. Noorzadeh, B. Rivet, and P. Gumery, "A multi-modal approach using a non-parametric model to extract fetal ECG," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 2015.

[17] E. Kristensen, A. Guerin-Dugue, and B. Rivet, "Comparison between Adjar and Xdawn algorithms to estimate eye-fixation related potentials distorted by overlapping," in *International IEEE EMBS Neural Engineering Conference*. IEEE, 2015.

[18] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.