Constrained Blind Source Extraction of Readiness Potentials From EEG

Pouya Ahmadian, Saeid Sanei, *Senior Member, IEEE*, Luca Ascari, Lara González-Villanueva, and Maria Alessandra Umiltà

***Abstract*—One of the changes seen in electroencephalography (EEG) data preceding human voluntary movement is a cortical potential called readiness potential (RP). Detection of this potential can benefit researchers in clinical neurosciences for rehabilitation of malfunctioning brain and those working on brain–computer interfacing to develop a suitable mechanism to detect the intention of movement. Here, a constrained blind source extraction (CBSE) is attempted for detection of RP. A suitable constraint is defined and applied. The results are also compared with those of the traditional blind source separation in terms of true positive rate, false positive rate, and computation time. The results show that the CBSE approach in overall has superior performance.**

***Index Terms*—Bereitschaftspotential (BP), blind source separation (BSS), constraint blind signal extraction (CBSE), electroencephalography (EEG), readiness potential (RP), second-order blind identification (SOBI), template matching.**

I. Introduction

In THIS work, we have tried to determine whether a certain brain potential called readiness potential or RP (explained in a dedicated paragraph) exists in single trials of the electroencephalography (EEG). To do that we had to design an experiment setup including the necessary software and hardware to evoke the RP in EEG of the subjects. In this experiment, the subjects were asked to move their hands and press a button on the pad whenever they wanted; no specific cues were given and the subject had the complete freedom for the movement time. After data acquisition, first, the EEG data was segmented based on the start of hand movement. Then, after a minor preprocessing, two algorithms were applied to the single trials. Both algorithms are based on independent component analysis (ICA). However, in one of the methods, the RP is extracted by applying a proper constraint to the source extraction cost function. As it can be seen in the results section, both methods have shown to be resilient across different subjects and had good results on the trials which were expected to have RP in them (true positive rate).

On the other hand, we also prepared single trials in which the RPs were not present to evaluate the performance of algorithms against false detection. In this area, the constrained blind source extraction (CBSE) outperforms the traditional blind source separation (BSS) which are discussed later on.

The remainder of this paper is organized as follows. First, the basic concepts of the work are presented, which include a review of the RP and ICA problems and solutions. In Section II, the EEG data acquisition is introduced, including hardware and software setup and the simulation paradigm to obtain two different kinds of trials. In Section III, the overview of two algorithms used for automatic detection is presented and the necessary steps for automatic detection are illustrated. In Section IV, we explained the experiments carried out to evaluate the detection algorithms. This paper ends with the discussion about the results in Section V and the proposals for possible future works in Section VI.

A. Readiness Potential

Many of human everyday works and actions are considered voluntary. There are evidences from EEG data that show human brain is active even before the beginning of the voluntary movement. The first proof of this evidence was made by Kornhuber and Deecke in 1964 [1]. In their experiment, the subjects were asked to perform a repetitive movement (finger flexion) with their choice of speed while EEG and electromyogram (EMG) data were recorded simultaneously. Then, by averaging EEG segment before the EMG onset, which indicated the start of movement as a trigger, a potential preceding human voluntary movement was discovered and published with the name Bereitschaftspotential (BP) or RP, shown in Fig. 1. RP is a negative cortical potential seen in motor cortex which can be seen 1.5 to 1 s before the onset of a self-paced or voluntary movement in EEG data. Moreover, RP is an event-related potential (ERP), because its onset is time locked to an event such as movement [1]–[3]. RP is proven to be evoked not only when a movement is performed by the subject, but also when execution of an action by others is observed or even when the movement is imagined [4], [5].

Except our previous work [6], there is not much work done to detect RP in single trials. In [7] the authors implemented a user specific template matching structure as part of a method to detect movement planning. To build the template an RP waveform recorded via only one electrode placed between C3 and A1 in

Manuscript received May 23, 2012; revised August 05, 2012, September 18, 2012; accepted October 21, 2012. Date of publication November 15, 2012; date of current version July 02, 2013. The work of P. Ahmadian and L. González-Villanueva was supported by the European Commission (Marie Curie ITN MIBISOC, FP7 PEOPLE-ITN-2008, GA n. 238819).

P. Ahmadian is with Henesis, 43125, Parma, Italy, and also with the IbisLab, Department of Information Engineering, University of Parma, 43100 Parma, Italy (e-mail: pouya.ahmadian@henesis.eu).

S. Sanei is with the Faculty of Engineering and Physical Sciences, University of Surrey, GU2 7XH Guildford, U.K. (e-mail: s.sanei@surrey.ac.uk).

L. Ascari is with Henesis, 43125 Parma, Italy (e-mail: luca.ascari@henesis.eu).

L. González-Villanueva is with the Henesis, 43125, Parma, Italy, and also with IbisLab, Department of Information Engineering, University of Parma, 43100 Parma, Italy (e-mail: lara.villanueva@henesis.eu).

M. A. Umiltà is with the Department of Neuroscience, Section of Physiology, University of Parma, 43125 Parma, Italy (e-mail: mariaalessandra.umilta@unipr.it).

Digital Object Identifier 10.1109/TNSRE.2012.2227278

1534-4320/$31.00 © 2012 IEEE

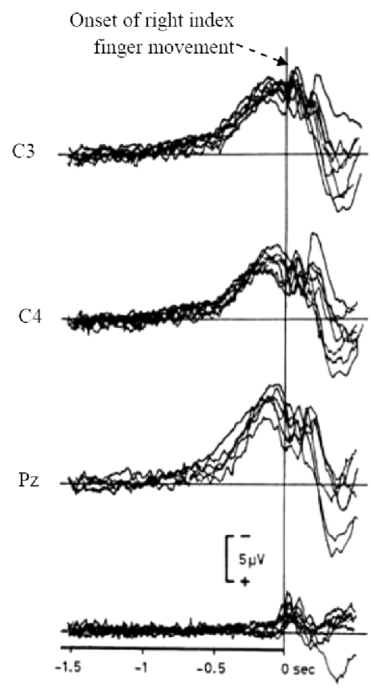


Fig. 1. RP preceding self-initiated movement (vertical line shows $t = 0$ s). C3, C4, and Pz indicate the place of electrodes in conventional 10–20 EEG electrode positions. The lowest plot shows the difference between RP in C3 and in C4, taken from [2].

the 10–20 International System was used. For each subject 12 movements during three or four sessions, each containing six right- and left-hand movements, were used to build the subject's template. A classifier was implemented to compare the template with EEG raw data and detect the movement in the test trails. This classifier was tested over five subjects and could predict movements with accuracy greater than 85% and a low false positive rate. However, the authors did not publish the results of RP detection in their experiment.

B. Independent Component Analysis

EEG signal is assumed to be formed by a number of independent components (ICs) or sources. This assumption justifies the use of ICA in EEG signal processing [8]. A review of different ICA algorithms used for EEG can be found in [9]. One major application of ICA is in two signal separation approaches called blind source separation (BSS) and blind signal extraction (BSE) which are frequently used in biomedical signal analysis and processing such as for EEG or magnetoencephalogram (MEG). Both techniques share the common word, blind, referring to the fact that the source signals are unobservable in the mixture signal, no prior information about them exists and the mixing process is unknown. This is in contrast to other techniques such as minimum-norm estimation (MNE) developed by Hämäläinen, [10], [11] that require a forward head model. However, some assumptions are made about the sources or mixing process such as independency of the sources to make the extraction procedures possible.

In BSS and BSE methods the problem is formulated as follows. Let us assume $\mathbf{x}(t)$ shows n continues time series from n EEG channels and thus, $x_i(t)$ shows EEG data recorded by ith sensor. Since human scalp performs like a conductor, the sources coming from the brain are summed together and any of $\mathbf{x}(t)$ can be assumed as a linear mixture of m $(n \geq m)$ unknown brain sources $\mathbf{s}(t)$ at discrete time t, mixed through unknown $n \times m$ mixing matrix $\mathbf{A}$. That is

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{v}(t) \tag{1}$$

where $\mathbf{v}(t)$ is an $n \times 1$ white noise vector independent of the source signals, $\mathbf{x}(t) = [x_1(t), \ldots, x_n(t)]^T$, $\mathbf{s}(t) = [s_1(t), \ldots, s_m(t)]^T$ and superscript T represents the transpose operator. Thus, assuming a noiseless environment, we can eliminate $\mathbf{v}(t)$ from the above formula.

Now that we have addressed the similarities between BSE and BSS methods, we focus on their differences on how they tackle the problem of extracting the sources in the next section.

1) Blind Source Separation Techniques: BSS approaches try to recover or estimate all the original sources, $\mathbf{s}_i(t)$, simultaneously from the mixture signals. Different algorithms have been developed for this approach and the challenge is to choose the suitable algorithm based on the application. After applying different algorithms such as JADE [12] and FastICA [13], and careful observation of their resulting components, we came to the conclusion that one of the components produced by second-order blind identification (SOBI) [8], [14] best match the features of RP and therefore, we used SOBI as the core for our automatic BSS-based method to detect the RP.

Each SOBI estimated component has a time course and a corresponding scalp projection which determines the effect of that specific component on all n electrodes. That is to show which electrodes are nearer to the component or where in the scalp the component is originated. Assuming a noiseless situation, the estimated sources, $\hat{\mathbf{S}}$, is given by

$$\hat{\mathbf{S}} = \mathbf{W}\mathbf{X} \tag{2}$$

where $\hat{\mathbf{S}}$ shows the time course of all the components. Note that $\mathbf{W}^{-1}$ entries are nonlinearly proportional to the inverse of the distance between the sources and the sensors. This is however subject to inherent scaling ambiguity of the BSS.

To calculate the matrix $\mathbf{W}$, SOBI considers the time coherence of signals. SOBI tries to find $\mathbf{W}$ by minimizing the squared sum of cross correlations between one component at time t and another component at time $t + \tau$. SOBI applies this calculation across multiple time delays and across all pairs of components. Since this cross correlation is sensitive to the temporal features of time based signal, temporal information of EEG data is enough for source separation. For more information on SOBI algorithm see [8] or [14].

SOBI is able to recover the components form the EEG data that are meaningful from physiological and neuroanatomical points of view. Validation of the components estimated by SOBI from 128 channel EEG data has been done by [15]. Joyce in [16] showed that SOBI is able to successfully extract the EEG artifact components and he used it as part of a procedure to automatically correct the ocular artifacts. SOBI is also used as part of a method for classification of ERPs awakened by a sequence of randomly mixed left, right, and bilateral median nerve stimulations [17]. Then, by feeding the SOBI extracted features to a

back-propagation neural network as the classifier, the accuracy of $83.3\% \pm 5.6\%$ among four subjects was measured.

2) Blind Signal Extraction Techniques: BSE approaches try to recover or estimate the original sources, $\mathbf{s}_i(t)$, one at a time from the mixture signals. Compared with BSS, BSE has the following advantages.

- By creating some constraints or criteria based on the features of the desired source signals, they can be extracted in a specified order.
- The approach is very flexible and can use different criteria in each step of the extraction, to separate the preferred sources in that step.
- It is possible to extract only the sources of interest and thus, it can save the computation time and resources.
- BSE algorithms are simpler than BSS algorithms and can be modified to be applied in a number of different situations.

In this work, we used sequential BSE with an optimization criterion based on the absolute value of normalized kurtosis. Kurtosis measures the deviation of the extracted source from normal Gaussian distribution and it is based on the fourth-order central moments of the signal [18]. A Gaussian signal has the kurtosis of zero, for positive values of kurtosis, signal is super-Gaussian or long tailed and for negative values, it is sub-Gaussian or short tailed. According to Central Limit Theorem, the distribution of the mixture signal is closer to Gaussian distribution than any of the original sources. Therefore, by maximizing the non-Gaussianity criterion such as absolute value of kurtosis, we can separate the source from the mixture signals. Since we are looking for one specific signal, we can formulate the output of the algorithm as

$$\hat{\mathbf{s}}_\mathbf{i} = \mathbf{y}_i = \mathbf{w}_\mathbf{i}^\mathbf{T}\mathbf{X} \tag{3}$$

where $\mathbf{y}_i$ is the $1 \times T$ ith output signal, $\mathbf{w}_i$ is the $1 \times n$ ith row of the unmixing matrix calculated by the algorithm, and $\mathbf{X}$ is the $n \times T$ mixture matrix consisting of the electrode signals. T denotes the number of time samples and n is the number of electrodes, as said before. Avoiding index i for simplicity, the goal is to minimize the following cost function:

$$J_m(\mathbf{w}) = -\frac{1}{4}|\mathrm{kurt}(\mathbf{y})| \tag{4}$$

where $\mathrm{kurt}(\mathbf{y})$ is the normalized kurtosis. Applying the standard gradient descent approach to minimize the cost function in (4), the following online learning rule is obtained [18]:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \eta(k)\varphi(y(k))\mathbf{x}(k) \tag{5}$$

where

$$\varphi(y(k)) = b\left(\frac{m_2(y(k))}{m_4(y)(k)}y(k)^3 - y(k)\right)\frac{m_4(y(k))}{m_2^4(y)(k)} \tag{6}$$

and the moments m_q can be iteratively estimated in an online application as

$$m_q(k) = (1-\beta)m_q(k-1) + \beta|y(k)|^q \tag{7}$$

where k is the iteration number and the sample number, m_q is the qth moment, and $\beta \in (0, 1]$ denotes the adjusting influence of the previous estimate of the moment and current estimate.

Now, in order to form the constraint to extract the RP first, we use prior knowledge about the shape and latency of RP in combination to the normalized kurtosis cost function. Thus, we created a reference signal of RP to calculate the actual unmixing vector, $\mathbf{w}_{\mathrm{opt}}$, which is used to minimize the Euclidean distance between the reference signal and the data with

$$J_c(\mathbf{w}_{\mathrm{opt}}) = \|\mathbf{y}_{\mathrm{ref}} - \mathbf{w}_{\mathrm{opt}}\mathbf{X}\|^2 \tag{8}$$

where $\|\cdot\|^2$ denotes the Euclidean distance. The solution to this is minimum norm [19]

$$\mathbf{w}_{\mathrm{opt}} = (\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}\mathbf{y}_{\mathrm{ref}}^T. \tag{9}$$

Therefore, we need to minimize the distance of the obtained $\mathbf{w}$ from (5) and $\mathbf{w}_{\mathrm{opt}}$ from (9). So we should minimize

$$d(\mathbf{w}) = \|\mathbf{w}_{\mathrm{opt}} - \mathbf{w}\|^2. \tag{10}$$

This constrained cost function replaces the original cost function (4) and the learning rule becomes

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \eta(k)\varphi(y(k))\mathbf{x}(k) + K(\mathbf{w}(k) - \mathbf{w}_{\mathrm{opt}}) \tag{11}$$

where K is the penalty parameter to make a balance between the two cost functions J_m and J_c and if chosen too high, it will overcome the effect of the main cost function while if chosen too low it will have a small effect. In Section IV, we describe how to choose K.

In [20], the authors designed a constrained ICA similar to the BSE algorithms to extract the components of brain activation in fMRI data. The cost function designed, is based on both the second-order and the higher-order statistics which uses the prior information about the desired components in the form of a reference signal. The authors in [21] proposed a constraint BSE algorithm which uses the cost function based on fourth-order moment (kurtosis for nonperiodic signals, or a new cost function for periodic ones) with a reference signal correlated with the desired source. The algorithm has been reported to be accurate in source extraction from simulated data and also in extraction of the cardiac artifacts from MEG data. In [19] a constrained BSE algorithm based on kurtosis is used to detect P300 components. They also included a reference signal based on the temporal shape of usual P300 subcomponent in their cost function and were able to correctly extract P3a and P3b signals from the EEG single trials. The CBSE-method developed here uses the same theory described in the Method section. Obviously, resembles the RP.

II. EEG Data Acquisition

We examined both the CBSE-based and the BSS-based methods on two kinds of datasets; the simulated EEG dataset and the real EEG data recorded from normal subjects. Here, we address how these datasets have been formed.

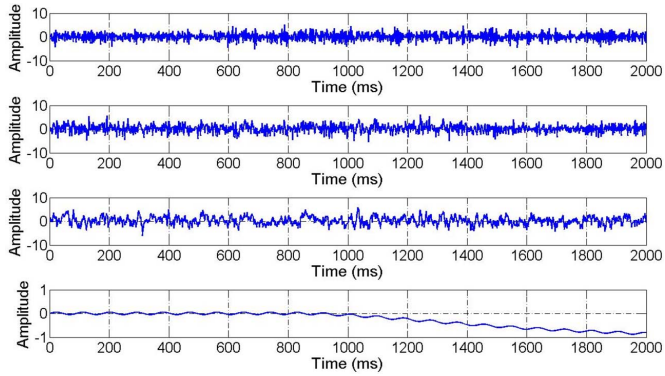


Fig. 2. The original sources used to create the simulated mixture EEG data. The bottom plot shows the simulated RP.

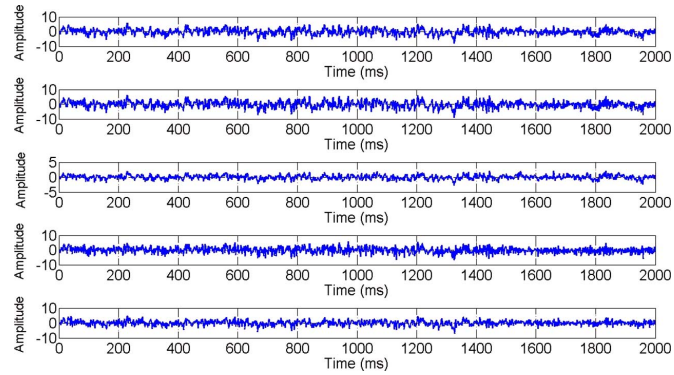


Fig. 3. The simulated EEG data obtained by mixing the sources of Fig. 2 using a random mixing matrix. Each plot simulates the data acquired by one EEG electrode.

A. Simulated EEG Data

EEG data was simulated using multivariate autoregressive (MVAR) modeling. In this approach a multichannel scheme is assumed. Thus, each signal sample is considered to be related to its previous sample and the previous samples of other channels. Here, we adopted the methods and parameters from [22] where the MVAR model uses two time varying parameters which are a step function and a positive triangle function

$$\begin{aligned} y_1(n) &= 0.5y_1(n-1) - 0.7y_1(n-2) \\ &\quad + c_{12}(n)y_2(n-1) + v_1(n) \\ y_2(n) &= 0.7y_2(n-1) - 0.5y_2(n-2) + 0.5y_1(n-1) \\ &\quad + c_{23}(n)y_3(n-1) + v_2(n) \\ y_3(n) &= 0.8y_3(n-1) + v_3(n) \end{aligned} \tag{12}$$

where

$$c_{12}(n) = \begin{cases} 0.5\frac{n}{L}, & n \leqslant \frac{L}{2} \\ 0.5\left(1 - \frac{n}{L}\right), & n > \frac{L}{2} \end{cases} \tag{13}$$

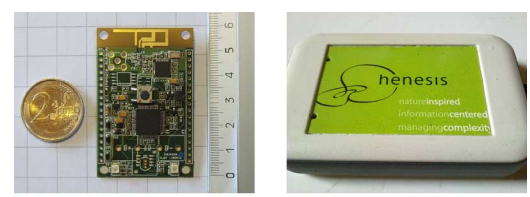
and

$$c_{23}(n) = \begin{cases} 0.4, & 0 \leqslant n \leqslant 0.7L \\ 0, & n > 0.7L \end{cases} \tag{14}$$

$c_{12}(n)$ is set to 0.2 and L is the number of samples which is set to 2000, same as EEG trials in real data. $vi(n)$ represents the white noise. For more information on the model and how to set the parameters, please see [22]. To this model we added one sinc signal which had the peak of RP before the movement. The sources are shown in Fig. 2. These sources, which are referred to as the original sources of the simulated data, were mixed using a randomly generated 5×4 mixing matrix, producing the five signals shown in Fig. 3, which were used in both algorithms for RP extraction. The results of both algorithms on this data are discussed in Section IV.

B. Real EEG Data

Three healthy subjects, one male and two females, volunteered in conducting the experiment. All subjects were right handed and aged between 26 and 30 years. None of the subjects had any history of neurological or psychological disorders.



(a) (b)

Fig. 4. HModule: External device design to mark the beginning of the hand movement on EEG data; (a) electronic board and (b) packaging.

1) Experiment Hardware and Software Setup: The acquisition framework used E-Prime and EGI's NetStation software to acquire EEG data from a 128-channel Sensor Net (Electrical Geodesic, Eugene, OR, USA) helmet. The E-Prime software was used for experimental paradigm and stimuli delivery. The main feature of this setup was the possibility of marking online the beginning of the hand movement. This could be done thanks to the use of an external device, called HModule, which is shown in Fig. 4. It is based on the Henesis WiModule [23], which has already been used for human motion acquisition in [24]. One of its main characteristics is that it contains a high performance three-axis accelerometer, which permits detection of the beginning of the movement. At the same time, it allows recording the movements by saving the accelerometer data. The accelerometer used has a precision of ± 1 mg with a range of ± 6 g. The sampling frequency was set to 160 Hz.

As it can be appreciated in Fig. 4(a), the HModule is small and wearable, which makes it ideal for being used in the experiments as it was attached to the moving hand of subjects. An algorithm based on thresholding and the exponential moving average (EMA) has been implemented. Using this technique in time series leads to smoothing out the short-term fluctuations and to highlight longer-term trends. For each sample of the accelerometer, the energy is calculated considering the data in its three axis. Afterwards, it is filtered, giving more weight to the latest data. This filtered value is the one compared with the set threshold. When a small hand movement is done (e.g., just when the subject starts to rise his/her hand), the given threshold is exceeded and the HModule triggers an event sending a signal to E-Prime to mark the EEG data. Choosing the correct threshold

is of the utmost important; as if it is too low, any minimal movement done by the subject would be caught, while if it is too high, the beginning of the movement would not be marked properly, but later on the data sequence. Several experiments were carried out by different subjects in order to choose the correct threshold for the given movement, "raising the hand and pressing a button on a pad."

The whole experiment had a duration between 10 and 11 min (the time variation is due to the fact that the subjects were free to move their hand whenever they wanted) and contained two parts each about 5 min. Preparing each subject and adjusting the sensor net on his/her head, took approximately between 25 and 30 min. Therefore, the whole experiment time did not take longer than 45 min per subject. The EEG data was recorded with the sampling frequency of 1000 Hz.

2) Experiment Paradigm Simulation: As mentioned above, the EEG data recording paradigm included two parts, each with different instructions.

- Part 1: Steady; the subject was asked to remain calm, breath normally with open eyes and try not to think of anything or move.
- Part 2: Hand movements; the subject was asked to move his/her hand whenever he/she wanted.

The instruction of what was asked in the experiment was given to subjects both orally and in writing at the beginning of the experiment. At the start of each part, a cue was shown on a monitor to inform the subject of what task they should perform. Other than these two cues, no other cues were given to subjects for two reasons. First, this was necessary to simulate the real world situation as accurately as possible. Second, giving a cue to the subjects might evoke the potentials which are uncalled for such as contingent negative variation (CNV), which is a slow negative wave that develops in the interval between a warning and a "go" stimulus and reflects preparation for *signaled* movements.

In part two, the EEG data to be analyzed started 2 s before the movements. Since no cue was given in the trials of this part, the HModule resting time was set to 5 s. This means the time duration between two sequential hand movements had to be at least 5 s. Otherwise, the HModule would not send any signal to E-Prime and the trial had to be repeated by subject. As the results, it was assured that 40 valid trials were obtained for each subject.

III. Methods

Algorithm 1 shows the overview of the BSS-based algorithm and Algorithm 2 shows the overview of the CBSE-based algorithm used for detecting the RP is single trials. All 128 channels acquired EEG data were used in our analysis. The EEG data was divided into two categories; premovement and steady trials. Pre-movement trials were obtained by cutting the EEG data based on HModule markers to 2 s before the beginning of the hand movements. Since we had 40 hand movements, we obtained 40 trials in this category. The steady trials were obtained by cutting the first 5 min of the steady data to 2 s nonoverlapping segments. Thus, we had 150 trials in the steady category.

Algorithm 1 BSS-based algorithm for RP detection

Step 1: Segment EEG data to single trials
Step 2: Create the $RP_{template}$ based on $t^{3-1}e^{-t/100}$
Step 3: For each single trial do:
  Step 3.1: Apply SOBI with $Source_{num} = m$
  Step 3.2: For $i = 1, ..., m$ Calculate
    $VCorr_i = Correlation(RP_{template}, IC_i)$
  Step 3.3: Find the $max_{corr} = max(VCorr_i)$
  If $max_{corr} > Threshold$ **then**
    RP exists in this trial
  end if

Algorithm 2 CBSE-based algorithm for RP detection

Step 1: Segment EEG data to single trials
Step 2: Create $\mathbf{y}_{ref}$
Step 3: For each single trial do:
  Step 3.1: Filter the data between 0.5 Hz and 70 Hz
  Step 3.2: Form $\mathbf{w}_{opt}$, based on $\mathbf{y}_{ref}$ with equation (9)
  Step 3.3: Initialize $\mathbf{w}$ randomly
  Step 3.4: Minimize the cost function in equation (10)
  if $d(\mathbf{w}) < Threshold$ **then**
    RP exists in this trial
  end if

Both categories, premovement and steady trials, were used in the evaluation of the algorithms.

A. BSS-Based Algorithm

In this algorithm (see Algorithm 1), after applying SOBI to each single trials of EEG data, we have a number of ICs which show the original sources that were active in the mixture signal. Unfortunately, the output of SOBI is not predictable. The sources extracted by SOBI are subjected to scale and sign change. Moreover, the orders of the extracted sources in different trials vary. To determine which one of the sources was similar to RP, we used a template matching technique. The shape of RP was simulated with the help of Gamma function $f(t; k, \theta) = (1)/(\Gamma(k))t^{k-1}e^{-t/\theta}$ with $k = 3$ and $\theta = 100$ and we added a small adjustments to that. Fig. 5 shows the final shape of the template used in BSS-based algorithm. Then, we calculated the correlation between each ICs in the output of SOBI with this template. When the value of correlation exceeded a $Threshold$ we assumed that RP is present in that trial. For more information please refer to [6].

B. CBSE-Based Algorithm

In this algorithm (see Algorithm 2), first, we filtered the EEG data in the range from 0.1 to 70 Hz. The filtering is done so the CBSE algorithm can be in line with [19] and can be eliminated. After that, we created the RP reference signal, $\mathbf{y}_{\mathrm{ref}}$. Both Gamma wave (see Fig. 5) and an average over 40 trials containing RP (see Fig. 6) have been used as reference signal. Then, we calculated the $\mathbf{w}_{\mathrm{opt}}$, with (9). The actual unmixing vector, $\mathbf{w}$, was initialized with random values. In each iteration the new $\mathbf{w}$ was computed using (11) and after each iteration of the algorithm it was compared with $\mathbf{w}_{\mathrm{opt}}$. If after some iteration,

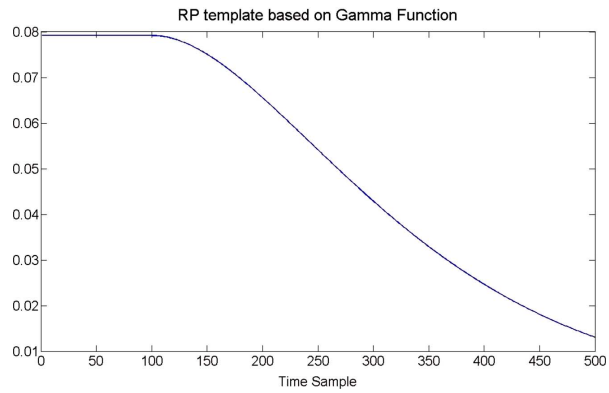


Fig. 5. The final shape of the template based on Gamma function.

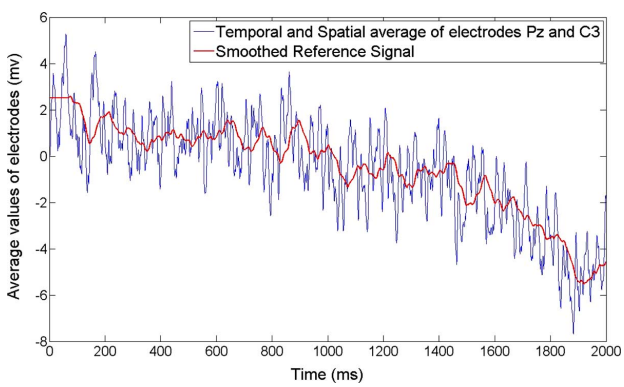


Fig. 6. Shape of signals used for creating the reference signal.

the Euclidean distance between $\mathbf{w}_{\text{opt}}$ and $\mathbf{w}$ were less than the $Threshold$ for a trial, we would assume that an RP existed in that trials. Otherwise it would be assumed that there was no RP in that trial.

IV. Results

Both methods were applied to two datasets, the the simulated data and real data. In simulated data, the goal was to extract the RP source as close as to the actual source as possible. In real data, both categories of trials were used to evaluate the detection rate of the algorithms. The trials where RP was expected to be found, meaning pre-movement trials, were used for true positive rate evaluation and the trials which RP was not expected to be found, meaning steady trials, were used to obtain the false positive rate of the algorithms.

The K parameter in CBSE algorithm has been set by trial and error. If the K parameter in the CBSE is set to zero, the algorithm will be unconstrained and the output is not RP. To set this parameter, we started from a small value for K (10^{-5}) and observed the extracted RP. At this small value, the results are not satisfactory. We increased K slowly until we reached a practical value (10^{-4}) where the extracted RP is good. The β parameter in (7) was set to 0.5. The learning rate, $\eta(k)$ in (11), was set to 10^{-3} at the beginning of the algorithm and was reduced by 1% in every iteration. These values were the same for simulated and real EEG data and were not changed across the subjects or different categories of trials.

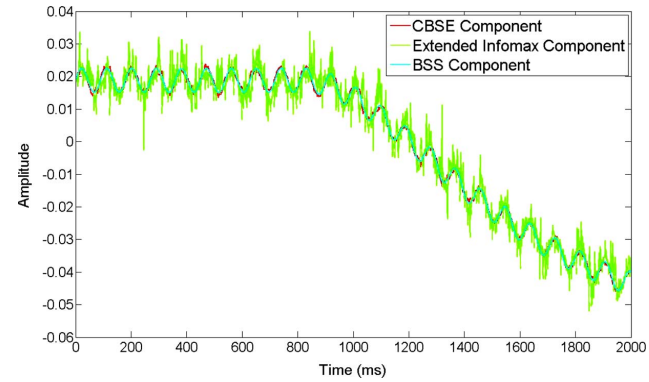


Fig. 7. The extracted RP in simulated data: the result of SOBI and CBSE method overlap each other while the result of Extended Infomax is noisy.

TABLE I
MSE and SNR of Three Blind Algorithms for Simulated Data

	Extended Infomax	SOBI	CBSE
MSE	-42.6112	-45.6082	-45.6062
SNR	6.7727	10.2771	10.2751

A. Simulated Data

For the first experiment, we used the simulated EEG data shown in Fig. 3. The goal was to extract the source with RP peak simulated, shown in Fig. 2, bottom plot. We applied the SOBI and Extended Infomax algorithm described in [25] as well as our constrained BSE-algorithm. Extended Infomax uses a learning algorithm for a feed-forward neural network that blindly separates the linear mixtures of independent sources with different distributions using information maximization theory. The results of the extracted sources are showed in Fig. 7. Comparing the three extracted signals with the original source, it is visually comprehensible that the CBSE-based method has a good performance.

Moreover, the performances of three algorithms on the simulated data were measured with respect to mean square error (MSE) and signal-to-noise ratio (SNR)

$$\text{MSE} = 10\log\frac{1}{N}\sum_{i=1}^{N}(s(i)-\hat{s}(i))^2 \tag{15}$$

$$\text{SNR} = 10\log\left(\frac{\sum_{i=1}^{N}(s(i))^2}{\sum_{i=1}^{N}(s(i)-\hat{s}(i))^2}\right) \tag{16}$$

where N is the number of time samples, $s(i)$ shows the original source, and $\hat{s}(i)$ is the extracted source. In order to calculate the error for the three algorithms, the original source was used as the reference. Since SOBI output components are subjected to scale change, the extracted sources and the original source were normalized before calculation of the errors. The results have been shown in Table I, as it can be seen in this table, SOBI and CBSE have similar results while Extended Infomax results on this data is noisy (lower SNR and higher MSE).

Here, we assumed that the number of independent sources are less than or equal to number of EEG channels. If otherwise, none of the components driven from SOBI or Infomax algorithm will have any similarities to the simulated RP while it is still possible to extract RP with our CBSE-based algorithm.

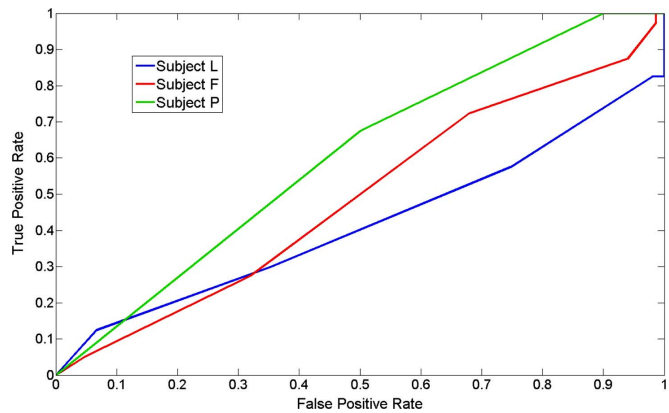


Fig. 8. ROC curves of BSS method for three subjects. As it can be seen the change of $Threshold$ affects the true positive and false positive rate similarly.

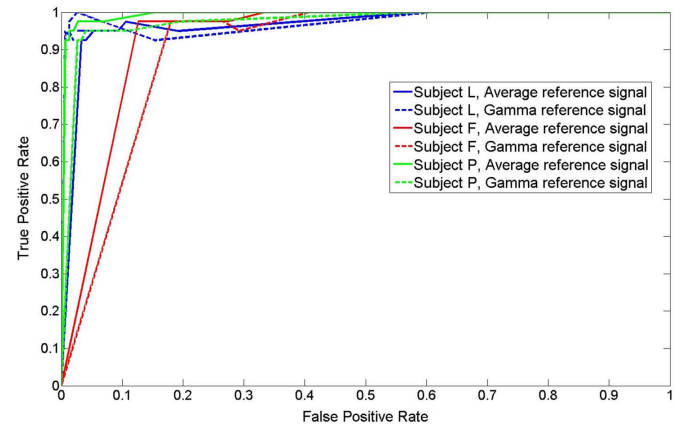


Fig. 9. ROC curves of BSE method for three subject using two different $\mathbf{y}_{\mathrm{ref}}$. The average signal and the general Gamma function. As it can be seen the change of $Threshold$ affects true positive and false positive rate independently.

B. Real Data

For the second experiment, we used real data recorded from human brain with the features described in EEG data acquisition. Here, we used receiver operating characteristic (ROC) curves to explain the performance of the binary classifier at different values of the $Threshold$ for both methods.

In order for BSS-based algorithms to be applicable, the number of sources active in the brain should be less than the number of electrodes. Thus, we used the very robust and well-known algorithm developed in [26] to obtain the number of independent sources from EEG data. The number of independent sources obtained by this algorithm in a randomly chosen 10 s segment of data was not more than 52. Since we used all 128 electrode channels and the EEG segments were 2 s, it is possible to apply both algorithms to the data.

Fig. 8 shows the results of BSS-based algorithm detection rate and Fig. 9 shows the results of CBSE-algorithm detection rate. For the BSS-based method, since the value of correlation ranges from 0 to 1, the $Threshold$ was changed between 0.1 and 0.9, increased 0.1 in each step and the values of false positive rate and true positive rate were calculated accordingly. As it can be seen in Fig. 8, the change of $Threshold$ affects both rate similarly. This means that both false positive rate and true positive rate increase or decrease simultaneously by the change of $Threshold$. The curves shows the poor performance of this method for all three subjects.

For the CBSE method, the $Threshold$ was changed between 10^{-8} and 1. In every step the value of threshold was multiplied by 10. The ROC curves for the three subjects are shown in Fig. 9; for each subject the results for two different $\mathbf{y}_{\mathrm{ref}}$, one based on average and another based on Gamma function is plotted. As it can be seen, in this method the true positive rate and false positive rate are independent from each other. The performance of binary classifier was acceptable in the range of $10^{-7} < Threshold < 10^{-4}$. However, lower values of $Threshold$ indicated more number of iterations in the optimization process and thus require more computation time. A practicable value for $Threshold$ can be 10^{-5}, which produces good results with acceptable computation time (0.26 s for each single trial on average).

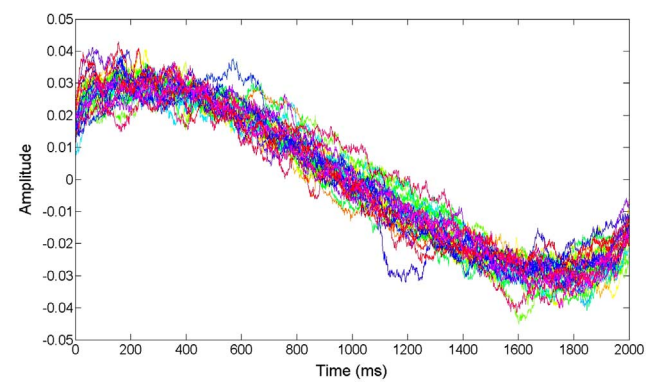


Fig. 10. Twenty extracted RPs in 20 single trials via BSS-based algorithm.

As it can be seen in Figs. 8 and 9, both algorithms can produce good results in the case of correct classification. However in the case of false detection, meaning detection of RP in the trials with no RP, CBSE-based algorithm had a much better performance. This is due to the fact that in the case of BSS-based algorithm, first, all of the sources are extracted, which are usually plenty, and some of them pass the criteria to be miscounted as RP.

The time of extraction in both cases also differ significantly. Average extraction time of RP in one trial via the CBSE-based algorithm took 0.26 s ($Threshold = 10^{-5}$) while for BSS-based algorithm it was 51.90 s. Again, this is due to simultaneous extraction of all sources by applying SOBI.

Figs. 10 and 11 show 20 extracted RPs in 20 single trials from premovement category via BSS-based algorithm and CBSE-based algorithm, respectively.

V. Discussion

The shape of the RP signal extracted by both algorithms is different. This is because in CBSE-based algorithm we designed the reference signal, $\mathbf{y}_{\mathrm{ref}}$, as the trade-off between having maximum true positive rate and minimum false positive rate. Thus, we set all the values of $\mathbf{y}_{\mathrm{ref}}$ except the last 400 to zero. Setting less values to zero gives us a more similar shape to Fig. 1 or Fig. 10, but also increases the false positive rate.

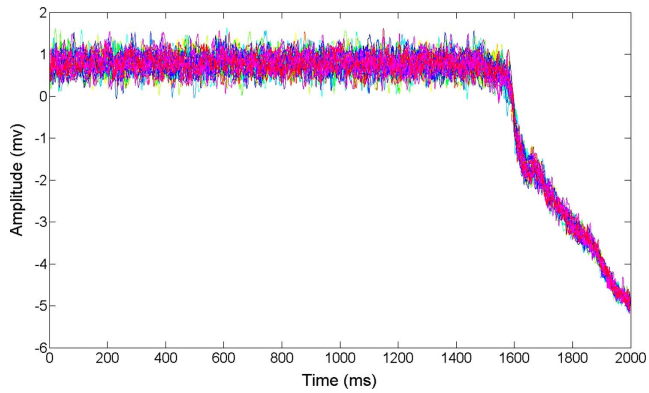


Fig. 11. Twenty extracted RPs in 20 single trials via CBSE-based algorithm.

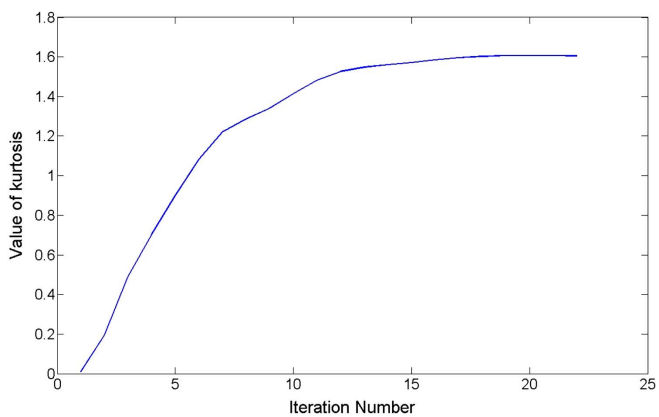


Fig. 12. Convergence of CBSE-algorithm in one single trial: the value of Kurtosis is increasing in every iteration with respect to the new learning rule.

In the application, we had for the ICA in both implemented algorithms, longer segments of EEG data could not be used. This is because ICA is based on the assumption that EEG signals are short-term stationary [8], [27]. Generally speaking, the sources within different parts of the brain are not always active. They may even move and go deep or be propagated to the surface. So, it may not be possible to localize them in a long window. On the other hand, the RP is a locally restricted source [28] or a focal source. As mentioned before the RP is a local event which is initiated approximately 1.5 s before the movement onset. The RP may vanish or be replaced by other brain events. Therefore, choosing long windows for the EEG signals may work for strong and somehow distinct events like eye-blinks or repetitive ERPs but not for a short duration potential like RP. Moreover, the number of active sources in a short time can be considered fixed. However, in a long segment the number of sources may change and the change in the number of sources within a time segment is not considered by ICA methods. Thus, it is better to use short windows to detect the RP. Consequently, we had to divide the EEG data to short segments such as 2-s segments and then, apply the algorithms to the result.

The changes for the values of Kurtosis of the extracted signal in every iteration are shown in Fig. 12 ($Threshold = 10^{-5}$). This figure shows the convergence of the CBSE-based algorithm after only 23 iterations, when the method meets its stopping criterion, when the value of Kurtosis is at its maximum.

VI. Conclusion and Future Work

In this work, we showed the superiority of CBSE-based algorithm in detection of brain potential compared with BSS-based algorithm. To sum up, in BSS-based algorithm, first we extract all the possible sources in the EEG single trial and then look for the one that is in our interest. CBSE-based algorithm performs differently by adding some constraint to the cost function, we extract only the sources of interest and ignore the other active potentials or noise from the very beginning.

The SOBI algorithm gives the corresponding scalp projections of the extracted sources as well as their temporal time course of the signal. Some of the extracted sources which are mistaken with RP shape are transmitted from other regions of the brain than the motor cortex. This information can be used in the BSS-based algorithm to reduce the false detection rate. However, this modification would increase the computation time.

By detecting RP in various trials we can correctly infer that the intention of the subject is rising and anticipate an upcoming movement or movement imagination and be prepared for it. Finding RP can be the first step in predicting the subject's next behavior. By predicting the human movement, it is possible to evaluate the next motor task before it happens and, if necessary, stop it to prevent any problem. This feature could have a useful application on platforms where human error has fatal consequences, such as during driving a car or plane. Moreover, detecting the upcoming movement has applications in rehabilitation systems that want to wake other devices, such as, monitoring the body movement of the patient.

In addition, automatic detection of RP in single trials can benefit neuroscientist to better understand the human perception since this is the first stage of brain activation against an intention. A robotic arm controlled by the brain can be nicely initialized for a smooth take off if this component is correctly detected. Finally, for people with physical challenges, this will be the main clue of having an intention to move.

Acknowledgment

The first author would like to thank Dr. M. Ahmadian for English proofreading and also, the reviewers of the article for their constructive comments.

References

[1] H. Shibasaki and M. Hallett, "What is the bereitschaftspotential?," *Clin. Neurophysiol.*, vol. 117, pp. 2341–2356, 2006.

[2] L. Deecke and H. H. Kornhuber, "Human freedom, reasoned will, and the brain: The bereitschaftspotential story," in *The Bereitschaftspotential: Movement-Related Cortical Potentials*, M. Jahanshahi and M. Hallett, Eds. New York: Kluver Academic/Plenum, 2003, pp. 283–320.

[3] M. Jahanshahi and M. Hallett, "The bereitschaftspotential: What does it measure and where does it come from," in *The Bereitschaftspotential: Movement-Related Cortical Potentials*, M. Jahanshahi and M. Hallett, Eds. New York: Kluver Academic/Plenum, 2003, pp. 1–17.

[4] J. A. Pineda, B. Z. Allison, and A. Vankov, "The effects of selfmovement, observation, and imagination on mu rhythms and readiness potentials (RP's): Toward a brain-computer interface (BCI)," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 219–222, Jun. 2000.

[5] J. M. Kilner, C. Vargas, S. Duval, and S.-J. Blakemore, "Motor activation prior to observation of predicted movement," *Nature Neurosci.*, vol. 7, pp. 1299–1301, 2004.

[6] P. Ahmadian, S. Sanei, L. Mussi, L. Ascari, and M. A. Umiltà, "Automatic detection of readiness potential," presented at the 9th IASTED Int. Conf. Biomed. Eng., Innsbruck, Austria, 2012.

[7] C. Haw, D. Lowne, and S. Roberts, "User specific template matching for event detection using single channel," presented at the Graz 2006 Conf. Brain Comput. Interface, Graz, Austria, 2006.

[8] S. Sanei and J. A. Chambers, *EEG Signal Processing*. New York: Wiley, 2007.

[9] J. Ontona, M. Westerfieldb, and J. T. S. Makeig, "Imaging human EEG dynamics using independent component analysis," *Neurosci. Biobehav. Rev.*, vol. 30, pp. 808–822, 2006.

[10] C. M. Michela, M. M. Murraya, G. Lantza, S. Gonzaleza, L. Spinellib, and R. G. de Peraltaa, "EEG source imaging," *Clin. Neurophysiol.*, vol. 115, no. 10, pp. 2195–2222, 2004.

[11] F.-H. Lin, T. Witzel, S. P. Ahlfors, S. M. Stufflebeam, J. W. Belliveau, and M. S. Hämäläinen, "Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates," *Neuroimage*, vol. 31, pp. 160–171, 2006.

[12] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEEE Trans. Signal Process.*, vol. 140, no. 6, pp. 362–370, Dec. 1993.

[13] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol. 9, no. 7, pp. 1483–1492, 1996.

[14] A. Belouchrani, K. Abed-Meraim, J.-F. C. Cardoso, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, Feb. 1997.

[15] A. C. Tang, M. T. Sutherland, and C. J. McKinney, "Validation of SOBI components from high density EEG," *NeuroImage*, vol. 25, pp. 539–553, 2004.

[16] C. A. Joyce, I. F. Gorodnitsky, and M. Kutas, "Automatic removal of eye movement and blink artifacts from EEG data using blind component separation," *Psychophysiology*, vol. 41, pp. 2317–2321, 2004.

[17] Y. Wang, M. T. Sutherland, L. L. Sanfratello, and A. C. Tang, "Singletrial classification of ERPs using second order blind identification (SOBI)," in *Proc. 3rd Int. Conf. Mach. Learn. Cybern.*, Shanghai, China, 2004, vol. 7, pp. 4246–4251.

[18] A. Cichocki and S. I. Amari, *Adaptive Blind Signal and Image Processing, Learning Algorithms and Applications*. New York: Wiley, 2002.

[19] L. Spyrou and S. Sanei, "A robust constrained method for extraction of p300 subcomponents," in *Proc. IEEE ICASSP, 2006*, 2006, vol. 2.

[20] W. Lu and J. C. Rajapakse, "Approach and applications of constrained ICA," *IEEE Trans. Neural Netw.*, vol. 16, no. 1, pp. 203–212, Jan. 2005.

[21] A. K. Barros, R. Vigrio, V. Jousmki, and N. Ohnishi, "Extraction of event-related signals from multichannel bioelectrical measurements," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 5, pp. 583–588, May 2000.

[22] A. Omidvarnia, M. Mesbah, J. M. OToole, P. Colditz, and B. Boashash, "Analysis of the time-varying cortical neural connectivity in the newborn EEG: A time-frequency approach," in *Proc. 7th Int. Workshop Syst., Signal Process. Appl. (WoSSPA 2011)*, 2011, pp. 179–182.

[23] Henesis WiModule Apr. 26, 2012 [Online]. Available: http://www.henesis.eu/prodwimodule-eng.htm

[24] L. González-Villanueva, L. Chiesi, and L. Mussi, "Wireless human motion acquisition system for rehabilitation assessment," in *Proc. 25th IEEE Int. Symp. Comput.-Based Med. Syst.*, Jun. 2012, pp. 1–4.

[25] T.-W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources," *Neural Comput.*, vol. 11, pp. 417–441, 1999.

[26] X. Bai and B. He, "Estimation of number of independent brain electric sources from the scalp EEGs," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 10, pp. 1883–1892, Oct. 2006.

[27] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Adv. Neural Inf. Process. Syst.*, vol. 8, pp. 145–151, 1996.

[28] I. Rektor, "Scalp-recorded bereitschaftspotential is the result of the activity of cortical and subcortical generators a hypothesis," *Clin. Neurophysiol.*, vol. 113, pp. 1998–2005, 2002.

Pouya Ahmadian received the M.S. degree in artificial intelligence from Amirkabir University of Technology, Tehran, Iran. She is currently working toward the Ph.D. degree in information technologies at the University of Parma, Parma, Italy.

Her research interests include biomedical signal processing, medical image processing, bio-inspired and soft computing algorithms and computer vision.

Saeid Sanei (SM'05) received the Ph.D. degree in biomedical signal processing from Imperial College London, London, U.K.

He is a pioneer in signal processing and machine learning with major application to EEG. He has published two monograms and over 250 papers in the field. Currently, he is a Reader within Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, U.K.

Luca Ascari received the MScEE degree, in 2000, and the Ph.D. degree in bioengineering, in 2005.

He is co-founder and head of research in HENESIS. His scientific interests, in the field of artificial perception, and focused on the use of neuromorphic computation in control and processing systems with massive sensorial input. He co-authored eight industrial patents and more than 30 scientific papers on international peer-reviewed journals and conferences.

Lara González-Villanueva received the M.S. degree in telecommunication engineering, in 2009, and the M.S. degree in electronic engineering, in 2013, both with honors from the University of Las Palmas de Gran Canaria, Spain. She is currently working toward the Ph.D. degree in information technologies at the University of Parma, Italy.

Her research interests include wearable sensing systems for rehabilitation therapies, bio-inspired algorithms, and medical image processing.

Maria Alessandra Umiltà received the Ph.D. degree in neuroscience from the University of Parma, Parma, Italy, 2000.

Since 2003, she has been Senior Researcher in the Department of Neuroscience, Section of Physiology, University of Parma, Parma, Italy. Her main research interests include the studies of the cognitive properties of the cortical motor system and of the neural basis of social cognition.