Shared-control brain-computer interface for a two dimensional reaching task using EEG error-related potentials

Iñaki Iturrate,   Luis Montesano,   Javier Minguez

Abstract— One of the main problems of EEG-based brain computer interfaces (BCIs) is their low information rate, thus for complex tasks the user needs large amounts of time to solve the task. In an attempt to reduce this time and improve the application robustness, recent works have explored shared-control strategies where the device does not only execute the decoded commands, but it is also involved in executing the task. This work proposes a shared-control BCI using error potentials for a 2D reaching task with discrete actions and states. The proposed system has several interesting properties: the system is scalable without increasing the complexity of the user's mental task; the interaction is natural for the user, as the mental task is to monitor the device performance to promote its task learning (in this context the reaching task); and the system has the potential to be combined with additional brain signals to recover or learn from interaction errors. Online control experiments were performed with four subjects, showing that it was possible to reach a goal location from any starting point within a 5x5 grid in around 23 actions (about 19 seconds of EEG signal), both with fixed goals and goals freely chosen by the users.

I. INTRODUCTION

Over the last years, brain-computer interfaces (BCIs) have made significant progress in controlling devices using both invasive and non-invasive techniques [1], [2]. Non-invasive approaches are mainly dominated by EEG-based BCIs that have demonstrated the control of 2D virtual cursors [3], mobile robots [4], [5] and robotic wheelchairs [6], [7]. For EEG-based systems, the BCI maps the electrical signals measured at the scalp to control commands.

The major design decision of such a system is to select the mental task the user will carry out to generate the EEG signals. Successful control has been achieved using self-generated activity such as the motor imagery of body limbs [8] or the performance of cognitive mental tasks [9]; and with event-generated activity, such as the P300 event-related potential (ERP) [7] and the steady-state-visually-evoked potential (SSVEP) [10].

In all cases, the EEG-based BCIs are affected by a low information rate, mainly due to the poor signal-to-noise ratio of the EEG, and to the fact that the performance of the system varies with time due to EEG non-stationarities [11]. Furthermore, they require large amounts of time to perform a task [7] and usually a high and/or constant concentration during operation [9]. To alleviate these limitations, recent systems started to explore shared-control strategies where

the device does not only execute the decoded commands, but is also involved in executing the task [1] (e.g. taking into account the environment while reaching a target or avoiding an obstacle [4], [7]).

The use of shared-control schemes opens the door to a new use of error-related potentials as an alternative or complementary signal for a BCI system. Error-related potentials are ERPs elicited by an unexpected outcome (i.e. after making a mistake or observing one) [12]–[15]. Due to its nature, this type of potentials are particularly well suited as supervision or feedback signals during the execution of a task. For instance, they have been already used as reward signals to teach a device within a reinforcement learning framework [14], [15].

This work presents a shared-control BCI of a 2D reaching task of a cursor over a discrete grid of possible targets, using error-related potentials as supervision signals. Under this control scheme, the user has to evaluate whether the cursor is correctly moving towards the goal or is making wrong movements. Based on this signal, the device estimates which is the desired target while moving towards it based on the optimal motion policies for each of the possible goals. The results show that all users were able to reach predefined and self-selected goal locations in around 23 actions (about 19 seconds of EEG signal). The use of error potentials has some interesting properties. First, the user's mental task is a supervision of a device operation, making the system very scalable, as long as the user can assess the device actions as correct or wrong. Second, the system is natural, since the user is assessing the actions performed by a device to guide it to a desired position. Third, within a shared-control strategy it can be combined with other brain signals to correct decoding failures or recover from wrong or ambiguous decisions of the device.

II. METHODS

A. Data Recording

Electroencephalographic (EEG) and electrooculographic (EOG) activity were recorded using a gTec system (3 synchronized gUSBamp amplifiers). For the EEG, 32 electrodes were recorded, distributed according to an extended 10/20 international system (FP1, FP2, F7, F8, F3, F4, T7, T8, C3, C4, P7, P8, P3, P4, O1, O2, AF3, AF4, FC5, FC6, FC1, FC2, CP5, CP6, CP1, CP2, Fz, FCz, Cz, CPz, Pz and Oz), with the ground on FPz and the reference on the left earlobe; for the EOG, 6 monopolar electrodes were recorded (placed above and below each eye, and from the outer canthi of the left and right eyes [16]), with the ground on FPz and the
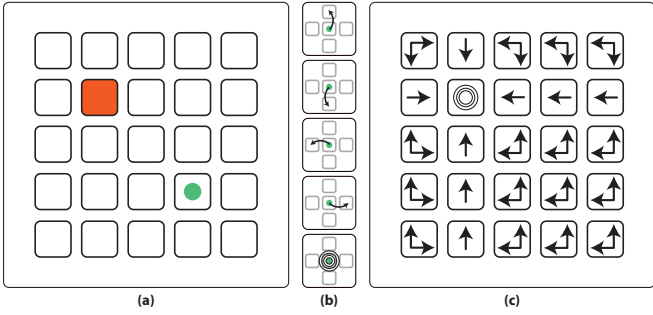
Fig. 1. (a) Experimental protocol designed. The protocol showed a 5x5 grid with a virtual cursor (green circle) and a goal location (shadowed in red). The up-left-most position is the [1 1] (row and column). (b) The cursor could perform five different actions (from top to bottom, move one position up, down, left or right, or performing a goal-reached action). (c) Correct actions (i.e. optimal policy) from each state for the goal exemplified on (a).

reference on the left mastoid. The EEG and EOG signals were digitized with a sampling frequency of 256 Hz, power-line notch filtered, and band-pass filtered at $[0.5, 10]$ Hz. The EEG was also common-average-reference (CAR) filtered. Additionally, the horizontal, vertical, and radial EOG were computed as in [16] to remove the EOG from the EEG using a regression algorithm [17]. The data acquisition and online processing was developed under a self-made BCI platform.

B. Experimental Protocol

Four subjects (mean age 26 ± 2 years) participated in the experiments. The participants were seated on a comfortable chair one meter away of a computer screen that displayed all the information related to the experiment. The experimental protocol is shown in Fig. 1. The protocol consisted of a 5x5 squares grid, a virtual cursor (green circle), and a goal location shadowed in red. The cursor performed five different instantaneous actions: move one position left, right, up or down; and a goal-reached action, represented as concentric blue circumferences (see Fig. 1b). The time between two actions (inter-action interval) was random within the range $[3, 3.5]$ s. The role of the subjects was to assess the cursor actions as correct when the cursor performed (i) a movement towards the goal position, or (ii) a goal-reached action over the goal position; and as incorrect otherwise (see Fig. 1c). The users were instructed not to move their eyes during the cursor actions, and to restrain blinks only to the resting periods.

The experiment consisted of two phases: the training phase used to calibrate a classifier able to distinguish between correct and error user's assessments; and the control phase, where the user controlled the cursor to a goal position. During the control phase, two different groups of goal locations were tested: (i) the first group was fixed for all the subjects, and consisted of five predefined goals and initial cursor positions (see table I); (ii) for the second group, each user was asked to freely choose five different initial cursor positions and goals to reach. During this group of goals, the goal position was not shadowed in red, since it was the user who chose it.

	Run 1	Run 2	Run 3	Run 4	Run 5
Initial position	[1 1]	[2 3]	[1 2]	[1 4]	[3 3]
Goal position	[5 5]	[3 1]	[3 3]	[4 1]	[3 3]

C. Calibration of error potentials

For the calibration of the error potential detection, a training phase was first executed to acquire sufficient examples of potentials while the user assessed the device actions. During this phase, the virtual cursor performed random actions, with a 20% of probability of performing an erroneous action. This phase lasted for 30 minutes, acquiring around 80 correct and 320 erroneous trials.

Once the training data was acquired, features were extracted from eight fronto-central channels (Fz, FC1, FCz, FC2, C1, Cz, C2, and CPz) within a time window of $[200, 800]$ ms downsampled to 64 Hz, forming a vector of 312 features. The feature vectors of all trials were normalized and decorrelated using PCA, retaining those that explained 95% of the variance. Finally, a regularized linear discriminant (LDA) [18] classifier was trained using the retained features. The classifier output has the form $y(\mathbf{x}) = \mathbf{w}'\mathbf{x} + b$, where $y(\mathbf{x}) \geq 0$ is classified as a correct assessment (class 1), and $y(\mathbf{x}) < 0$ as an error assessment (class 0). This output $y(\mathbf{x})$ was transformed into the probability that a trial $\mathbf{x}$ belonged to the correct class, $p(c = 1|\mathbf{x}) = \frac{1}{1+e^{-y(\mathbf{x})}}$ [19].

D. Shared-control for a reaching task

This section describes the proposed shared-control strategy with error potentials. Although these potentials provide feedback about the device actions, the amount of information conveyed by them is limited. In particular, the decoder of Section II-C does not contain information about direction or magnitude. Furthermore, it has a non-negligible number of misdetections. Therefore, the proposed shared control uses memory to accumulate evidence about possible goals while executing a trajectory. The proposed approach consists of two phases. The first one computes offline optimal trajectories (i.e. policies) for each potential target, while the second one ranks them during execution using error potentials elicited for wrong actions.

1) Offline policies computation: The offline phase was used for the computation of the optimal motion policies (see Fig. 1c) for all possible goal locations. Let $\pi_i^*(s)$ be the optimal policy for reaching the goal position i from the state s, whose output is the best action to compute in s (see Fig. 1c for a representation of the optimal policy when the goal location was [2 2]). The optimal policies can be represented with their associated Q-values $Q_i^*(s,a)$, which represents the value of executing the action a in state s when the goal location is i. These Q-values were converted into probabilities, following a soft-max normalization [20]:

$$\hat{Q}_i^*(s,a) = \frac{e^{Q_i^*(s,a)/\tau}}{\sum_b e^{Q_i^*(s,b)/\tau}}, \tag{1}$$
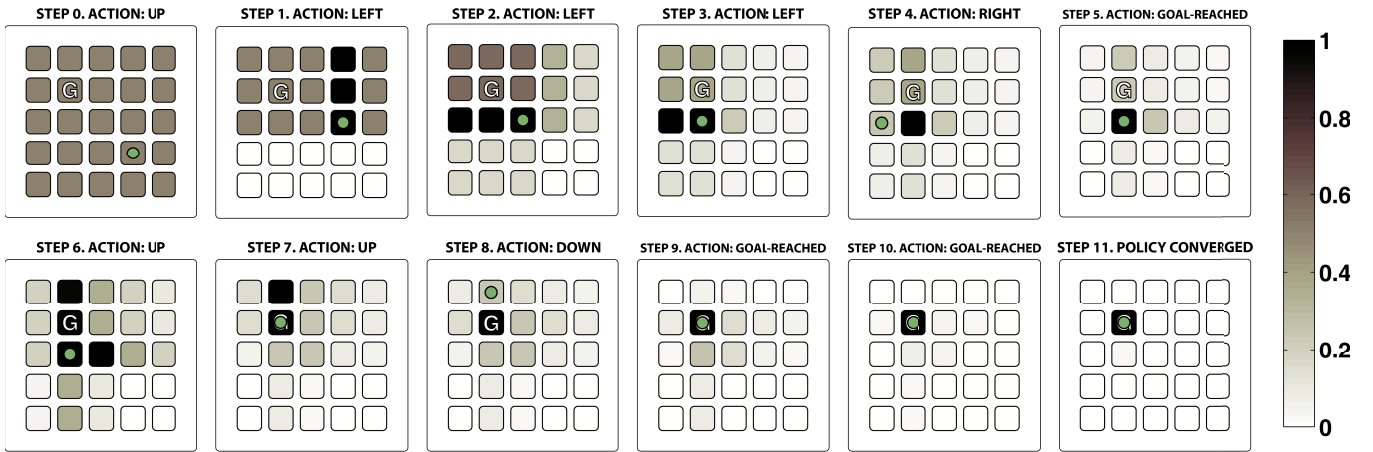
Fig. 2. Device control simulation for the goal shown in Fig. 1a ([2 2], marked with letter G). At each step t, the device was on a state s_t and executed an action a_t (shown over each individual plot). Each state i is colour-encoded as the probability of the goal being at that state i, $p(\pi_i^*|a_{t...1}, s_{t...1})$. The classifier was ideal, with its output $p(c = 1|x) = 0.8$ for correct actions, and $p(c = 1|x) = 0.2$ for incorrect actions. Initially, all the probabilities were equiprobable. While more steps were executed several policies were discarded, whereas others increased their probability. After 11 steps, the policy convergence criterion was reached, and the device learned the goal position.
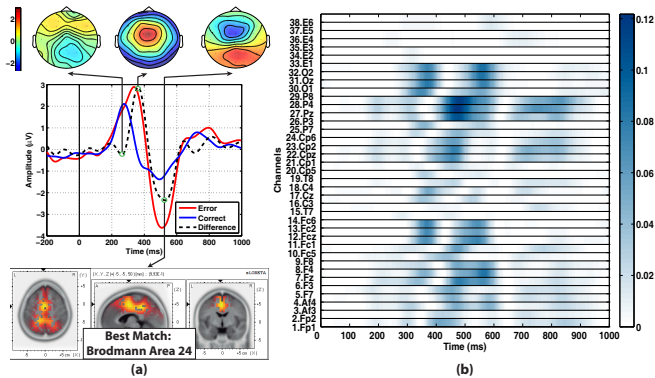

Fig. 3. (a) Action-locked averages from channel FCz (error, correct and difference in red, blue and black respectively), together with the color-encoded topographic interpolation of the three most prominent peaks of the difference average. Time 0 ms indicates when the cursor executed an action. The source localization for the most prominent negative peak of the difference average is also shown. (b) Color encoded r^2 discriminability test for each time point (x-axis) and channel (y-axis), including the 6 monopolar EOG channels recorded (denoted as E1 to E6).

where τ is denoted the temperature (fixed to $\tau = 0.3$ for the current experiment). This parameter served as a degree of reliability of the observed information (classifier output). For the designed protocol, the Q-values for all the possible goal locations can be computed offline prior to the control phase using the Q-learning reinforcement learning algorithm [20].

2) Online control: The online cursor control was performed by updating the probabilities of having the goal location on a state i $p(\pi_i^*)$, which where initially equiprobable (see Figure 2, step 0). At each time step t, the device performed an action a_t from state s_t. This action was assessed by the user, and the EEG classified following the procedure described in subsection II-C, obtaining $p(c = 1|x, a_t, s_t)$. This probability was used to update $p(\pi_i^*)$ for each goal location i (see Figure 2, steps 1 to 10). Finally, the execution

finished when a probability $p(\pi_i^*)$ reached a convergence criterion, $p_c = 0.9$ (see Figure 2, step 11). At time t, the probability of each policy i being the correct one given an action a_t executed in a state s_t is:

$$p(\pi_i^*|(a, s, \mathbf{x})_{1...t}) \propto p(a_t|\pi_i^*, (s, \mathbf{x})_t) \cdot p(\pi_i^*|(a, s, \mathbf{x})_{1...t-1}), \quad (2)$$

where the likelihood is computed as:

$$p(a_t|\pi_i^*, (s, \mathbf{x})_t) = p(c = 1|\mathbf{x}) \cdot \hat{Q}_i^*(s_t, a_t) + p(c = 0|\mathbf{x}) \cdot (1 - \hat{Q}_i^*(s_t, a_t)), \quad (3)$$

Notice that the first term of the likelihood represents how we should increase the policy π_i^* if the user's assessment was correct, while the second term penalized the policy π_i^* weighted by the probability of having and incorrect user's assessment. Fig. 2 shows an example of the device control when reaching the goal shown in Fig. 1a.

III. RESULTS

A. Electrophysiology analysis

Fig. 3a shows the error, correct and difference grand averaged potentials (error minus correct averages) in channel FCz, averaged for all the subjects. The difference grand average was characterized by three components: a negative deflection at around 250 ms, a positive deflection at around 400 ms, and a another negative component at around 500 ms. The topographic interpolations of the two broader peaks of the difference average showed that the signals were generated mostly in fronto-central channels. These results were in agreement with previous studies using error potentials [13]. The sLoreta source localization [21] was applied to the most prominent negativity of the difference average, confirming that the signals were generated in the anterior cingulate cortex (ACC, Brodmann area 24), which is the area thought to be the main generator of error-processing brain activity [13]. Fig. 3b shows the r^2 discriminability test of the signals obtained. The test revealed two main

	s1	s2	s3	s4	mean $\pm$ std
♯ **Targets reached (out of 5)**	5	5	5	5	5 ± 0
♯ **Actions**	16 ± 2	43 ± 9	23 ± 12	17 ± 5	25 ± 13
Net time (s)	12.96 ± 1.91	34.56 ± 7.10	18.40 ± 9.73	13.60 ± 4.38	19.88 ± 10.75
Total time (s)	52.65 ± 7.76	140.40 ± 28.83	74.75 ± 39.54	55.25 ± 17.80	80.76 ± 73.68
Accuracy correct / error (%)	82.46 / 79.17	60.98 / 78.49	82.43 / 63.41	71.19 / 76.92	74.27 ± 10.32 / 74.50 ± 7.45

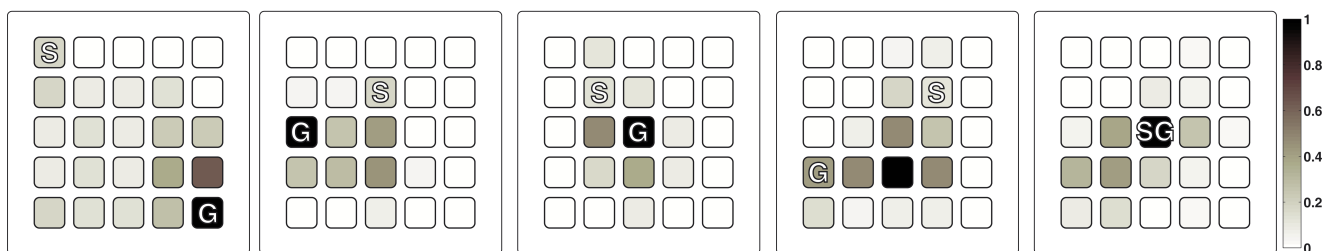|	s1	s2	s3	s4	mean $\pm$ std
♯ **Targets reached (out of 5)**	5	5	5	5	5 ± 0
♯ **Actions**	23 ± 9	21 ± 7	16 ± 6	23 ± 11	21 ± 8
Net time (s)	18.08 ± 7.37	16.48 ± 5.84	12.96 ± 5.10	18.08 ± 8.44	16.40 ± 6.61
Total time (s)	73.45 ± 29.93	66.95 ± 23.73	52.65 ± 20.73	73.45 ± 34.29	66.63 ± 26.85
Accuracy correct / error (%)	75.00 / 72.97	75.00 / 83.87	83.93 / 84.00	70.37 / 76.19	76.08 ± 5.67 / 79.26 ± 5.56

Fig. 4. States visited by all the subjects, for each of the five runs executed with the fixed goals (from left to right, runs 1 to 5). Darker colors indicate more visited states. The range was normalized from 0 to 1 according to the most visited state for each run. The initial and goal positions are marked with an S and a G respectively.

zones of discriminability between correct and error signals, at around 350 and 500 ms, agreeing with the two broader peaks appearing in the difference average. Furthermore, no discriminable information was found on the EOG channels (E1 to E6 in Fig. 3b), thus indicating that no EOG activity was contaminating the recorded potentials, and that the EOG was not involved in the device control.

B. Control phase analysis

For each group of goals (fixed and freely-chosen), five metrics were evaluated: (i) Number of goals reached; (ii) number of actions needed to reach the goal, (iii) EEG seconds needed to reach the goal (net time); (iv) total time needed to reach the goal; and (v) classifier accuracy, measured as the percentage of detection of correct and erroneous signals.

Table II shows the results for each subject (averaged for goals). The first result was that the device was able to reach all the target location, independently of their relative location. The users needed 25 ± 13 actions to reach the goal. With the proposed inter-action interval (around 3.25 s), the total time needed to reach the goals was of 80.76 ± 73.68 s. Nonetheless, the net time (i.e., the seconds of EEG signal used for decoding) was of 19.88 ± 10.75 s. Note that the difference between the net and total times was the seconds

belonging to inter-action intervals, which could be easily removed. The mean classifier accuracy across subjects was of 74.27 ± 8.94 and 74.50 ± 6.45 for correct and erroneous actions respectively. As expected, there was a significant negative correlation between the classifier mean accuracy and the time needed to reach the task ($r = -0.47, p = 0.038$). Notice that a random walk strategy that selected random actions at each step would require around 150 actions for a 5x5 grid size. Fig. 4 shows the states visited for the five runs performed. As can be seen, not all the states needed to be visited to reach the goal. For instance, during run 3, mostly all the central states were visited, whereas the peripheral states were not. This is an interesting property of the proposed control scheme, since it allows for a better scalability of the system, due to the fact that each step of the trajectory provides information for multiple goals simultaneously. As a result, the percentage of visited states would decrease for larger state spaces.

Table III shows the results for each subject. As with the fixed goals, all the targets were correctly reached by the device. The number of actions needed to reach the goal was 21 ± 8, with a total time of 66.63 ± 26.85 s, and a net time of 16.40 ± 6.61 s, and thus slightly better than the results with fixed goals. However, these differences were not significant (paired t-test, $t_{19} = 1.05, p = 0.31$). Regarding

the classifier accuracy, the detection rate was better in this case: 76.08±5.67 and 79.26±5.56 for correct and erroneous actions. The error accuracy was significantly better than the one obtained for the fixed goals, (paired t-test, $t_{19} = -2.27, p = 0.04$), but the correct accuracy was not (paired t-test, $t_{19} = -1.12, p = 0.28$). The fact of this slight increase of accuracy could be due to a protocol habituation, since the users always executed the free goals after the fixed ones. Again, there was a significant negative correlation between the mean accuracy and the time needed to reach the task ($r = -0.79, p = 3 \cdot 10^{-5}$).

IV. CONCLUSIONS AND FUTURE WORK

In this work, we have presented a shared-control BCI for a 2D virtual cursor based on error potentials. Most of the intelligence of the system was moved to the device side by comparing optimal trajectories with the executed one based on the detected error potentials (actions perceived as wrong by the user). The proposed approach compensates the low information transfer rate of error potentials by exploiting the structure of trajectories. The experimental results showed that, for a 5x5 grid, the system reached any goal after only 23 actions on average (less than a minute in our protocol).

The proposed shared-control BCI might scale to more complex scenarios because: i) it is not necessary to explore every single trajectory or potential goal; and ii) the user only has to monitor the device actions and evaluate if they are right or wrong. The authors are currently applying this BCI approach to real devices (e.g. mobile robots or manipulators) with larger and continuous state-action spaces.

REFERENCES

[1] JdR Millán, R Rupp, GR Müller-Putz, R Murray-Smith, C Giugliemma, M Tangermann, C Vidaurre, F Cincotti, A Kübler, R Leeb, C Neuper, KR Müller, and D Mattia, "Combining brain–computer interfaces and assistive technologies: state-of-the-art and challenges," *Front Neurosci*, vol. 4, 2010.

[2] MA Lebedev and MAL Nicolelis, "Brain-machine interfaces: past, present and future," *Trends Neurosci*, vol. 29, no. 9, pp. 536–546, 2006.

[3] N Birbaumer et al., "The thought translation device (TTD) for completely paralyzed patients," *IEEE Trans Rehabil Eng*, vol. 8, no. 2, pp. 190–193, June 2000.

[4] L Tonin, T Carlson, R Leeb, and JdR Millán, "Brain-controlled telepresence robot by motor-disabled people," in *Conf Proc IEEE Eng Med Biol Soc (EMBC), Boston, USA*, 2011.

[5] C Escolano, JM Antelis, and J Minguez, "A telepresence mobile robot controlled with a noninvasive brain-computer interface," *IEEE Trans Syst Man Cybern B Cybern*, vol. 42, no. 3, pp. 793 –804, june 2012.

[6] F Galán, M Nuttin, E Lew, PW Ferrez, G Vanacker, J Philips, and JdR Millán, "A brain-actuated wheelchair: Asynchronous and non-invasive brain-computer interfaces for continuous control of robots," *Clinical Neurophysiology*, vol. 119, no. 9, pp. 2159–2169, 2008.

[7] I Iturrate, J Antelis, A Kübler, and J Minguez, "Non-invasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation," *IEEE Transactions on Robotics*, vol. 25, no. 3, pp. 614–627, 2009.

[8] JR Wolpaw and DJ McFarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans," in *Proc Natl Acad Sci USA*, 2004, vol. 101, pp. 17849–17854.

[9] JdR Millán, F Renkens, J Mouriño, , and W Gerstner, "Noninvasive brain-actuated control of a mobile robot by human EEG," *IEEE Trans Biomed Eng*, vol. 51, no. 6, June 2004.

[10] GR Müller-Putz and G Pfurtscheller, "Control of an electrical prosthesis with an SSVEP-based BCI," *IEEE Trans Biomed Eng*, vol. 55, no. 1, pp. 361–364, 2008.

[11] C Vidaurre, M Kawanabe, P von Bünau, B Blankertz, and KR Müller, "Toward unsupervised adaptation of LDA for brain-computer interfaces," *IEEE Trans Biomed Eng*, vol. 58, no. 3, pp. 587 –597, 2011.

[12] M Falkenstein, J Hoormann, S Christ, and J Hohnsbein, "ERP components on reaction errors and their functional significance: A tutorial," *Biological Psychology*, vol. 51, pp. 87–107, 2000.

[13] PW Ferrez and JdR Millán, "Error-related EEG potentials generated during simulated Brain-Computer interaction," *IEEE Trans Biomed Eng*, vol. 55, no. 3, pp. 923–929, March 2008.

[14] I Iturrate, L Montesano, and J Minguez, "Robot reinforcement learning using EEG-based reward signals," in *IEEE Int Conf Robot Autom (ICRA)*, 2010, pp. 4822–4829.

[15] R Chavarriaga and JdR Millán, "Learning from EEG error-related potentials in noninvasive brain-computer interfaces," *IEEE Trans. Neural Syst. and Rehab.Eng.*, vol. 18, no. 4, pp. 381–388, 2010.

[16] RJ Croft and RJ Barry, "EOG correction of blinks with saccade coefficients: a test and revision of the aligned-artefact average solution," *Clinical neurophysiology*, vol. 111, no. 3, pp. 444–51, Mar. 2000.

[17] A Schlögl et al., "A fully automated correction method of EOG artifacts in EEG recordings," *Clinical neurophysiology*, vol. 118, no. 1, pp. 98–104, Jan. 2007.

[18] B Blankertz, S Lemm, M Treder, S Haufe, and KR Müller, "Single-Trial Analysis and Classification of ERP Components-a Tutorial," *NeuroImage*, 2010.

[19] CM Bishop, *Pattern recognition and machine learning*, Springer, 2011.

[20] RS Sutton and AG Barto, *Reinforcement learning: An introduction*, MIT Press, 1998.

[21] RD Pascual-Marqui, "Standardized low resolution brain electromagnetic tomography (sLORETA): Technical details," *Methods Find Exp Clin Pharmacol*, pp. 5–12, 2002.

[22] I. Iturrate, R. Chavarriaga, L. Montesano, J. Minguez, and J. del R Millan, "Latency correction of error potentials between different experiments reduces calibration time for single-trial classification," in *Int Conf of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2012.

[23] I Iturrate, L Montesano, and J Minguez, "Task-dependent signal variations in EEG error-related potentials for brain-computer interfaces," *Journal of Neural Engineering*, vol. 10, no. 2, pp. 026024, 2013.