Adaptive assistance for brain-computer interfaces by online prediction of command reliability

Sareh Saeedi, Ricardo Chavarriaga, Robert Leeb and José del R. Millán *

Center for Neuroprosthetics, École Polytechnique Fédérale de Lausanne (EPFL), Geneva, Switzerland

Abstract. One of the challenges of using brain-computer interfaces (BCIs) over extended periods of time is the variation of the users' performance from one experimental day to another. The goal of the current study is to propose a performance estimator for an electroencephalography-based motor imagery BCI by assessing the reliability of a command (i.e., predicting a 'short' or 'long' command delivery time, CDT). Using a short time window (< 1.5 s, shorter than the delivery time) of the mental task execution and a linear discriminant analysis classifier, we could reliably differentiate between short and long CDT (Area under the sensitivity-specificity curve, AUC ≈ 0.8) for 9 healthy subjects. Moreover, we assessed the feasibility of providing online adaptive assistance using the performance estimator in a BCI game by comparing two conditions: (i) allowing a 'fixed timeout' to deliver each command or (ii) providing 'adaptive assistance' by giving more time if the performance estimator detects a long CDT. The results revealed that providing adaptive assistance increases the ratio of correct commands significantly ($p < 0.01$). Moreover, the task load index (measured via the NASA TLX questionnaire) shows a significantly higher user acceptance in case of providing adaptive assistance ($p < 0.01$). Furthermore, the results obtained in this study were used to simulate a robotic navigation scenario, which showed how adaptive assistance improved performance.

*Corresponding authors: Sareh Saeedi (Email: sareh.saeedi@epfl.ch), José del R. Millán (Email: jose.millan@epfl.ch)

Online resources:

Supplementary materials can be found in

`http://infoscience.epfl.ch/record/212925/files/IEEECIM_supplementary.pdf`

This document provides a description of the ITR derivation (Section 2.1.3) as well as a second set of experiments where we use the results obtained in this article in a robotic navigation scenario.

1. Introduction

Brain-computer interfaces (BCIs) aim at offering an interaction modality for people with severe motor disabilities. A common approach relies on the decoding of sensorimotor rhythms (SMRs) measured using electroencephalography (EEG). Despite promising advances, BCIs are still confronted with multiple challenges in determining user's intentions reliably, mainly due to high performance variations among and within subjects [1].

Several studies have addressed the issue of performance variations in SMR-based BCIs. Most of these studies focus on inter-subject variability from a physiological [2–6], anatomical [7, 8], or psychological [9, 10] perspectives. Although precise distinction between user-related and system-related causes of performance variations may not be simple [11], these studies provide a better understanding of these causes. To tackle the system-related issues and to boost reliability, some studies have used adaptive machine learning techniques [12], while others proposed methods for removing signal contamination (e.g. due to muscular or ocular artifacts) [13].

Other studies have investigated the issue of intra-subject performance variability. In a longitudinal study, motivational factors were shown to be correlated with BCI performance on participants with amyotrophic lateral sclerosis (ALS) [14]. Other studies have focused on neurophysiological markers. However, most of these studies are limited to the analysis of a single session using offline recordings (where no feedback was provided to the subject). For instance, it is suggested that classification certainty for each subject is positively correlated with the power of EEG oscillations in the gamma band ($55-85$ Hz) [3, 6]. Similarly, trial-by-trial classification performance was found to correlate with high-frequency gamma oscillations ($70-80$ Hz) prior to the beginning of each trial [15], while others found correlations with a weighted combination of the theta ($3-8$ Hz), alpha ($8-13$ Hz), and beta ($16-24$ Hz) oscillations in frontal, pariatel and central areas of the brain, respectively [16]. Another study made a step forward by conducting both offline and online recordings (where subjects received feedback on their performance) in a single session. In this study, pre-cue SMRs in subject-specific electrodes and frequency bands were found to be positively correlated with single-sample classification performance [17].

The mentioned studies provide a good insight into the correlates of intra-subject performance variability. However, they do not account for performance changes over extended periods of time (i.e., over several sessions/days), neither do they exploit this information online. Thus, there are still several open research questions around this issue. First, the main goal of BCIs is to provide online decoding of users' intentions. It follows that it is essential to evaluate performance changes in online sessions, as distribution of data normally differs between offline and online sessions because of the feedback subjects receive in the latter. Second, one of the challenges for a BCI is to cope with performance variations over extended periods of time (e.g., over different sessions/days). Hence, a method capable of evaluating these variations over different sessions would improve usability and reliability. Third, it is crucial to predict performance on a short-time basis so as to compensate for the user's varying capabilities while using a brain-controlled device.

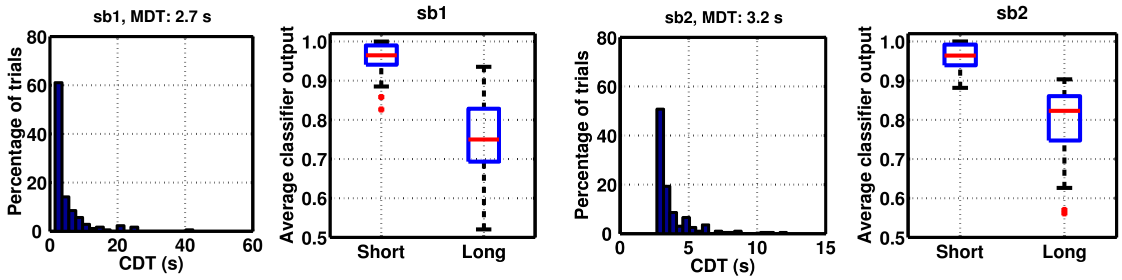

Figure 1. Distribution of command delivery time (CDT) for two subjects (sb1 and sb2) in an MI online protocol. The comparison of short and long trials (with respect to the median delivery time, MDT) reveals significant differences in the average classifier output ($p < 0.001$).

BCIs based on motor imagery (MI-BCIs) typically combine the outcome of classification for samples in a given time window in order to improve reliability. Some implementations make a decision after a fixed amount of time [2]. Others accumulate evidence over time by integrating the classifier outputs until it reaches a certain threshold, at which point the command is delivered [18]. The latter approach results in different command delivery time (CDT) across trials [19, 20]. Figure 1 illustrates such variations in a BCI session for two subjects. Since the integrated probability should reach a threshold for a command to be delivered, long CDT should be due to lower average classifier output for samples of the trial. Indeed, after separating the trials into short and long ones based on the median delivery time (MDT), we observe that the average classifier output across samples is higher for short trials compared to long ones ($p < 0.001$ in Wilcoxon rank sum test, Figure 1).

Having long CDT is usually frustrating for subjects and may reduce their engagement in performing the mental task. In addition, it can increase the workload or affect the performance of the system if the BCI application has to meet temporal constraints. For example, consider the task of controlling a brain-actuated robot [19]. In this case, the user delivers right or left commands, while the robot moves forward if the subject voluntarily decides not to deliver any command. To make the robot cross a doorway on the right

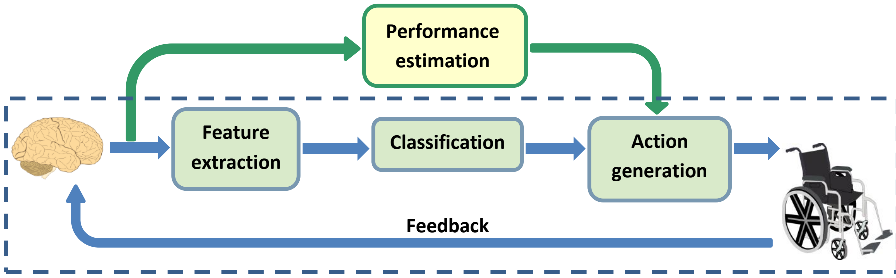

Figure 2. Common block diagram of a BCI in the dashed box. The goal is to design a performance estimator, which works in parallel with the BCI and can be used to provide online adaptive assistance for the user. In this study, performance estimation is achieved by a trial-by-trial prediction of command delivery time (CDT).

side, the user needs to be fast enough to deliver the corresponding command at the proper moment. Otherwise, the robot will miss the doorway and the user will need to deliver additional commands to bring it back. In such cases, predicting that a command is going to take a longer time to be delivered would be extremely helpful, as it would enable providing adequate assistance to the user. For instance, the speed of the robot can be reduced so that the user has enough time to make it cross the doorway.

As a result, adding a performance estimator to the BCI system may allow to compensate for performance variations by adapting the interaction to the user's current capabilities (e.g., through the use of shared control) [21]. Figure 2 shows such a system that works in parallel with the BCI and modulates action generation based on an estimation of the performance. The goal of this study is to propose a method capable of making a trial-by-trial prediction of the performance in an MI-BCI in terms of the time it takes to deliver a command. This predictor estimates the reliability of a command (i.e., having 'short' or 'long' CDT) based on the input features to the BCI. For this approach to be useful, the estimation has to be performed at the very beginning of a mental task execution. Here, we show that the information within 1.5 s from the beginning of the MI execution is sufficient to make such a prediction. Importantly, the experiments were done in an online setting lasting several sessions (2-3 depending on the subject) so as to account for the effect of feedback on the modulation of brain signals. Finally, the feasibility of using this method for online adaptation of assistance has been studied in a BCI game. In the Supplementary Materials, we also investigate a robot navigation scenario.

2. Materials and methods

2.1. Experimental protocol

Figure 3 illustrates the structure of the experimental protocol, as well as the number of sessions for each condition. First, in the MI training phase, we built the MI classifier based on an offline session and then tested it in an online session. Second, the subjects played an MI-BCI game. One to two sessions were required to tune the parameters and to re-train the classifier in case a drop of performance was observed (Calibration). Third, 2 to 3 sessions of the MI-BCI game were conducted with the same classifier and fixed parameters (Evaluation), which were used to build the performance estimator (short vs. long classifier). Finally, a single session of the MI-BCI game was performed in order to apply online adaptive assistance in the game based on performance estimation. The number of subjects (N_{sb}) participating in each phase is displayed on the left side of this figure.

2.1.1. MI training

Nine healthy subjects (sb1-sb9; five females, age 26.5 ± 4.2 years old) participated in a synchronous MI-BCI experiment. All participants gave written informed consent and the protocols were approved by the local ethics committee. Two of the subjects (sb2, sb5) had previous experience with the MI-BCI. EEG was recorded using 16 electrodes over the sensorimotor cortex at 512 Hz and band-pass filtered between 0.1 Hz and 100 Hz. Laplacian

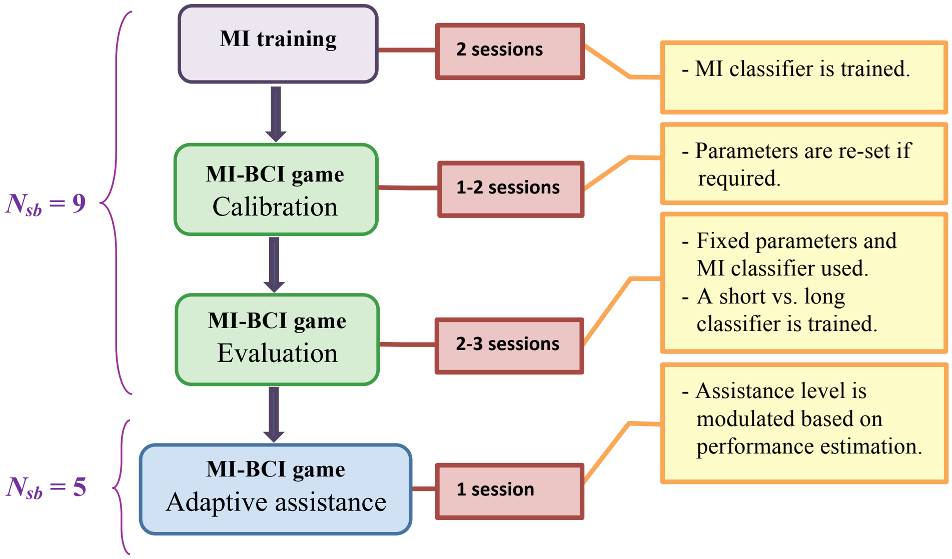

Figure 3. Experimental protocol: MI training was performed to calibrate the MI classifier and to allow the subjects to practice the mental tasks. Then, subjects played an MI-BCI game for several online sessions. Finally, subjects carried out another session of the game where they received online adaptive assistance.

spatial filtering was applied to the signal.

MI training was performed as described in [18]. First, offline calibration recordings were conducted in order to train an MI classifier to discriminate between two mental tasks (e.g., imagination of left or right hand movements). Then, the participants performed one online MI session (i.e., experiments performed on a day), comprising 4 to 6 runs in order to get familiar with the task. Each run consisted of 15 trials per mental task. The timing of one trial is depicted in Figure 4(a). The user was required to perform a mental task following the cue while receiving visual feedback from the classifier outputs (movement of the grey bar). The command was delivered as soon as the integrated classifier output surpassed a subject-specific threshold (th_d). Feedback was then given to the user showing that the trial has ended, which was followed by a brief rest period (random between 2 and 3 s). A key point in this experiment was that the subjects had unlimited amount of time during a trial to deliver a command. This can be observed in the distributions of CDT shown in Figure 1. The threshold th_d was set based on previous experience so that more than 75% accuracy was achieved in the command delivery of both classes (right and left) in each run while having the majority of the CDT lower than 8 s.

2.1.2. MI-BCI game

An MI-BCI game was designed for this study to provide a more engaging environment for users (Figure 4(b)). In this game, subjects were asked to rescue a parachutist by moving a platform right or left (defined by the cue). There was the same number of trials for the two classes (right and left) in each run. The speed of the parachutist was set so that it lands at second 8. Depending on the performance, there were three types of command delivery: 'hit' (when the platform reaches the correct side before the parachutist touches the ground), 'miss' (when the platform reaches the wrong side), and 'timeout' (when the platform does not reach either side on time). In the first case (i.e., hit), the platform color changes to green; otherwise to red (i.e., miss or timeout). Moreover, in order to increase the quality of feedback to the users, as recommended in [22], their progress (number of correct commands and average CDT) was displayed at the end of each run.

Participants initially performed 1 to 2 sessions of the BCI game (Calibration in Figure 3) in order to tune subject-specific parameters of the BCI classifier. The threshold, th_d, was set using the same criteria as in the MI training. Also, the MI classifier was re-trained if the performance dropped below 75%. Once the optimal parameters and classifier were found, 2 to 3 additional sessions (8 runs each) were conducted (Evaluation in Figure 3). No re-training or parameter update was performed either in this phase, or in the next phase (online adaptive assistance). The data of these Evaluation sessions was used to train the performance estimator. Finally, the performance estimator was applied in order to provide

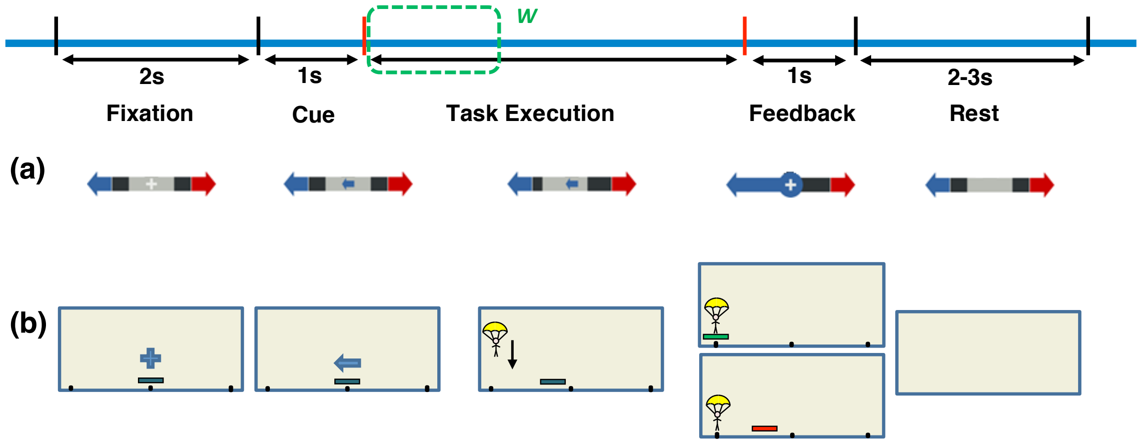

Figure 4. Timing of an online trial in the MI-BCI experiments: **(a) MI online recordings:** First, a cross is shown on the screen so that the user prepares for the task. Then the cue appears and the subject needs to execute the relevant mental task. The duration of the task execution is different across trials, as it depends on the integrated classifier output. Feedback is given to the user at the end of the trial and a rest period follows. **(b) MI-BCI game:** The subject controls the platform performing the two selected mental tasks so that the parachutist lands on it. The parachutist lands at second 8 as default, but it can have different speeds depending on the level of assistance. If the platform reaches the target on time, it changes color to green and otherwise to red.

adaptive assistance during the game as being described in the next section.

2.1.3. Online adaptive assistance in the BCI game

Five subjects (sb1, sb3, sb4, sb7, sb8) participated in this part of the experiment (1 recording session; blue block in Figure 3). In order to assess the benefit of adaptive assistance based on the proposed performance estimator, we compared two conditions:

(i) Fixed timeout: The parachutist reaches the ground at a fixed timeout. That is, the user always has a constant time (fixed timeout) to deliver a command. This fixed timeout (t_{sl}) was set to the value of the threshold used to separate trials into short and long groups (detailed in Section 2.2.2).

(ii) Adaptive assistance: In this condition if the performance estimator classifies a command as short, the speed of the parachutist is set so that it lands at t_{sl}. Otherwise, the parachutist slows down and reaches the ground at 8 s. That is, the user has by default t_{sl} to deliver a command unless the command is predicted as long, in which case she/he has 8 s to do it. The timeout of 8 s in case of long CDT was selected in order to be consistent with the previous MI-BCI game, where 8 s was sufficient for all subjects to deliver a command.

Subjects carried out 3 runs per condition, each run having 15 trials per class (right and left). The two conditions were performed sequentially, starting from the 'fixed timeout' for two subjects (sb1 and sb3) and from 'adaptive assistance' for the other three. In order to measure the perceived workload, participants filled in the NASA-TLX (task load index) questionnaire [23] at the end of each condition.

Apart from this subjective measure, the two conditions were compared based on the success rate (the ratio of correct command delivery) and information transfer rate (ITR) measures. For computing the latter, we assume that the 'timeout' commands are equivalent to 'no decision'; cf., Supplementary Materials. Throughout this paper, we have implemented the Wilcoxon paired signed-rank test to investigate statistical difference of the results in different conditions, as the data exhibit non-Gaussian distributions.

2.2. EEG decoders

2.2.1. Classification of motor imagery

The procedure for decoding of users' intention from EEG is detailed in [18]. Power spectral densities (PSDs) were used as features for classification of different motor imagery tasks. Features were ranked based on their relevance to the mental tasks using Canonical Variate Analysis (CVA) [24]. Considering this ranking and the neurophysiological evidence on the cortical areas/frequency bands contributing to a specific mental task, 5 to 7 features (in the alpha and beta range) were selected manually for each subject. These features were the input to a Gaussian classifier with four prototypes per class. In the online sessions, the single sample classification was integrated over time.

2.2.2. Single-trial prediction of short vs. long CDT

The performance estimator (whether the CDT of a trial will be short or long) uses the same selected PSD features as the MI classifier. We hypothesise that the characteristics of these control features in the beginning of task execution reflect the reliability of the command and, hence, the CDT.

To design the performance estimator, we separated the trials into short and long ones based on a subject-specific temporal threshold (t_{sl}). In order to select t_{sl}, we compared short vs. long classification performance over the Evaluation sessions for different percentiles of the CDT. That is, if the threshold is chosen as 35th percentile of CDT, then 35% of the trials are considered as short and the remaining 65% as long ones. The accuracy of the designed performance estimator was assessed using 10-fold cross validation for different percentiles (from 35 to 65 with a step size of 5), so as to find the optimal threshold for each subject.

A short window of samples, W, in the beginning of a trial was considered for predicting the CDT (the green window in Figure 4). The length of W was the shortest possible below the threshold t_{sl} that yielded a reliable prediction. The median delivery time (MDT) and the length of this window is shown in the 'Results' section in Table 1 for all subjects.

Features: For estimating the trial performance given the feature vectors ($\mathbf{x}$) extracted from EEG within W, we define a distance measure for class i and prototype j as:

$$Dist_{ij} = \frac{1}{N_w} \sum_{t=1}^{N_w} \sum_{k=1}^{N_f} \frac{(x_{tk} - \mu_{ijk})^2}{\sum_{ik}}, \tag{1}$$

where $\mathbf{x}$ is the feature vector with N_f selected features (i.e. EEG channels and frequency bands), μ_{ij} is the center of the j^{th} prototype of class i, Σ_{ik} is the variance of feature k for class i, and N_w is the number of feature vectors within the window. The defined distance measure is similar to the one used for calculating the posterior probability of the Gaussian classifier. However, we consider that some prototypes may be more influential than others for discriminating between short and long trials. Given that we have N_p (equal to 4) prototypes per class, the feature vector f_{sl} for estimating the performance is composed of the distance of sample x_t to every prototype ($f_{sl} = Dist_{ij}, i = 1 : N_c, j = 1 : N_p$).

Finally, since these features are not independent, CVA was used to select the features in f_{sl} that better discriminate between short and long trials.

Classification: As the characteristics of the commands were different for the two different mental tasks (right and left)[1], we built a separate short vs. long classifier for each of the two classes per subject. The threshold t_{sl} was chosen based on the classification performance (Area under the ROC curve, AUC ≈ 0.8) in the Evaluation phase (10-fold cross validation with a linear discriminant analysis, LDA, classifier). Moreover, a single t_{sl} was chosen for both classes (right and left), which may result in a different percentile of CDT for the two classes. The output of the resulting classifier was used for providing adaptive assistance.

3. Results

3.1. Performance in MI-BCI

Table 1 shows the success rate and the MDT for all subjects over the Evaluation sessions of the game. All subjects had a rather good accuracy, higher than 0.7 (a value often considered sufficient for BCI operation [25]).

[1] Most of the subjects are usually more efficient in performing one task than the other. Thus, they may generate more consistent brain patterns across trials while doing that task.

Table 1. Performance assessment of the Evaluation sessions in the MI-BCI game. 'MDT': median delivery time, 'W': length of the window used for performance estimation, 'Success rate': ratio of correct command delivery over all trials, 't_{sl}': temporal threshold for classifying short and long trials. The underlined subject IDs indicate those who participated in the online adaptive assistance session.

	sb1	sb2	**sb3**	**sb4**	sb5	sb6	**sb7**	**sb8**	sb9
MDT (s)	2.69	2.29	1.91	1.83	3.07	2.38	2.00	3.96	2.09
W **(s)**	1.0	1.0	1.0	1.0	1.0	1.0	0.625	1.5	0.625
Success rate	0.85	0.88	0.91	0.92	0.89	0.75	0.82	0.75	0.96
t_{sl} **(s)**	2.8	2.2	2.4	2.0	3.0	2.5	2.0	3.7	2.1

3.2. Single-trial prediction of short vs. long CDT

Figure 5 shows the average AUC for 10-fold cross validation of short vs. long classification for all subjects. These results were obtained offline using the data from the MI-BCI game Evaluation sessions. As mentioned before, different percentiles of CDT have been used to separate the trials into short and long groups. The values of t_{sl} (percentile of CDT) are displayed in the x-axis of each plot, as well as the average AUC for both classes: C1 (right command, in red) and C2 (left command, in blue). As the figure illustrates, a reliable classification result can be achieved in most cases.

The AUC tends to decrease as higher thresholds are considered. Based on the short vs. long classification performance for each subject, optimal t_{sl}, equal for the two classes, was individually chosen to be used in the adaptive assistance phase. Table 1 reports the selected t_{sl} for all subjects.

3.3. Online adaptive assistance in the MI-BCI game

As depicted in Figure 6, providing adaptive assistance leads to a significantly higher success rate with respect to having a fixed timeout equal to t_{sl} ($p < 0.01$). The former resulted in an

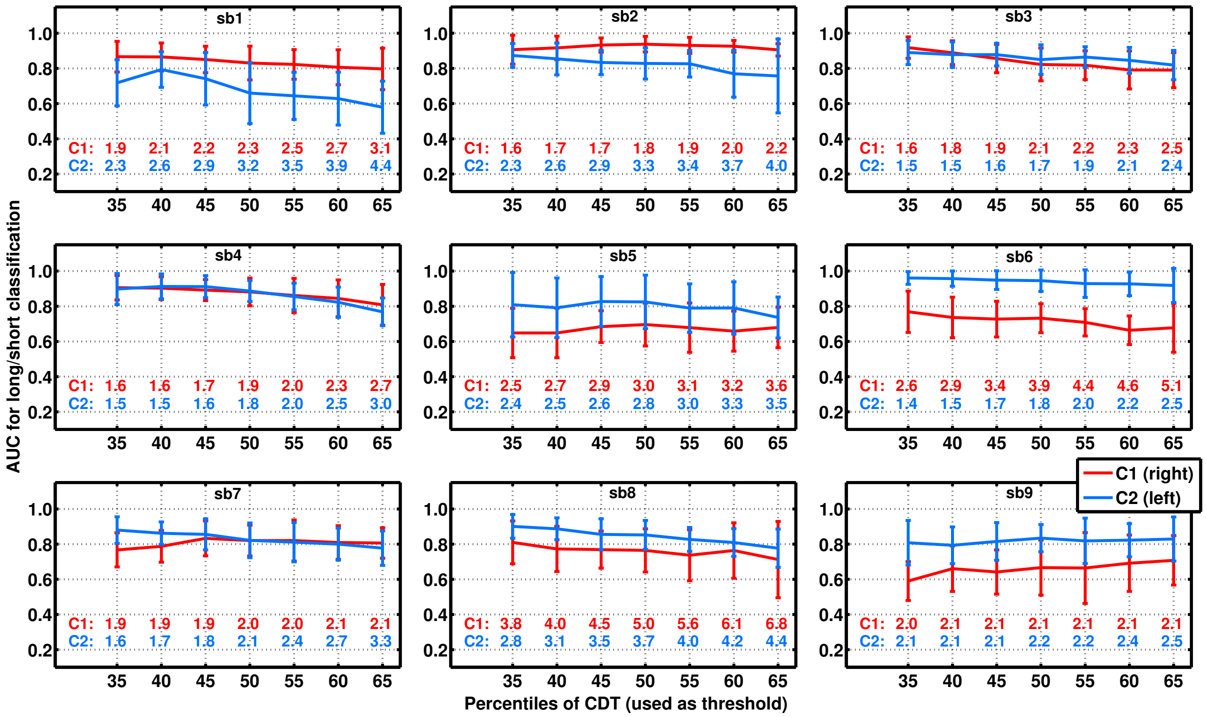

Figure 5. AUC of the performance estimator for all subjects and for the two classes, C1 (right, red) and C2 (left, blue), using different percentiles of CDT. The figure shows the time threshold corresponding to each percentile.

average success rate of 0.76 ± 0.04 and the latter an average success rate of 0.53 ± 0.05 across the 5 subjects (underlined subject IDs in Table 1). However, considering timeout trials as no decision, the two cases have a comparable ITR with no significant difference.

In addition, based on the users' evaluations, there is a significant reduction in the NASA TLX scores when implementing the adaptive assistance ($p < 0.01$) as depicted in Figure 7. The main contributing factors were the temporal demand ($p < 0.01$), frustration ($p < 0.05$), and performance ($p < 0.05$).

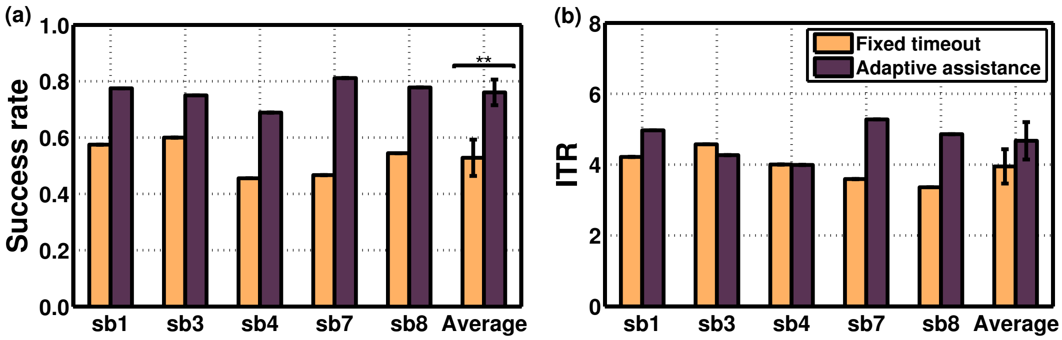

Figure 6. Comparison between fixed timeout and adaptive assistance conditions for the five subjects who participated in this phase: (a)Success rate ($**p < 0.01$), (b)ITR assuming that timeout trials are equivalent to no decision.

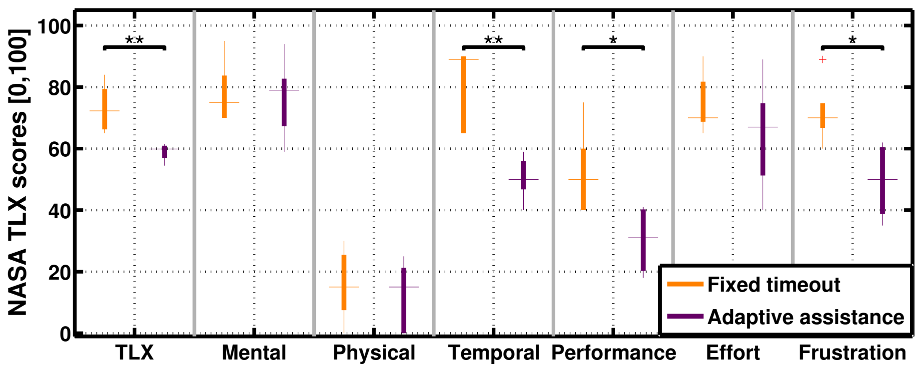

Figure 7. The NASA TLX: the main contributing factors to the users' perceived workload were the temporal demand, frustration, and performance ($* * p < 0.01$, $*p < 0.05$). For all factors (including performance), a lower score is better.

4. Discussion

One of the challenges in using BCI systems over extended periods of time is the performance variation across different sessions for the same subject. One way of tackling this issue is to adapt the level of assistance the system provides based on the users' needs. For instance, adaptive assistance based on EEG has been shown to improve perceived performance (NASA TLX) in an air traffic control task [26]. Here, we addressed the issue of performance variation in an MI-BCI by designing a performance estimator whose output is used to provide online adaptive assistance to the user. This estimator predicts CDT on a trial-by-trial basis. The results on 9 healthy subjects reveal that a subject-specific estimator can be found to reliably classify single-trials into short and long ones (according to the distribution of CDT) within a short time window at the beginning of each trial. This classification was done separately for the two classes (right and left) due to different characteristics of the mental tasks.

The proposed performance estimator takes into consideration some critical issues in studying intra-subject variabilities. Firstly, it is designed based on the recorded data in online sessions, where the users received feedback on their performance and could adapt their strategies accordingly. Secondly, its design takes into account performance variations across several sessions (2 to 3) and not only across trials in a single session. Thirdly, it is capable of detecting performance variations on a single-trial basis based on a short time window. Thus, it can be beneficial for monitoring the performance of the subjects who exhibit variations on a short time interval (e.g., across trials within the same run).

In addition, the feasibility of implementing this performance estimator for providing adaptive assistance was investigated for 5 subjects in an MI-BCI game. The results show that adaptive assistance (regulating the timeout for delivering a command in the MI-BCI game) improves the success rate significantly, compared to a fixed timeout. Although the bit rate for a fixed timeout might be comparable to the one in the adaptive assistance case, the latter showed significantly higher user acceptance as indicated by the NASA TLX score.

Intra-subject variabilities in BCI systems have not been extensively studied so far. Existing studies have some limitations. For instance, they focus on a single experimental session and do not take into consideration the variability across sessions, which is usually more critical. Besides, a performance measure is usually defined based on offline recordings, which might not be an ideal measure, as the subjects did not receive any feedback on their performance. Moreover, most of them focus on finding correlates of high or low performances. Thus, they are not capable of making decisions on a trial-by-trial basis. A study that investigated the feasibility of performance estimation in a short-time basis revealed that the proposed method can be beneficial for subjects who show performance variations on a time scale of several minutes [15], different from our approach that works on a time scale of seconds. Finally, to the best of our knowledge, none of these studies explored the provision of adaptive assistance based on performance estimation.

In order for subjects to be able to use an SMR-based BCI in real applications, e.g., driving a wheelchair, several training sessions are required [18]. At this stage, BCI performance tends to fluctuate substantially due to variations in the SMR correlates, which hinders training and requires offline recalibration of the BCI classifier. Thus, implementing a performance estimator for regulating the level of assistance according to the users' performance may largely improve training.

The MI task in this study was designed as a game in order to make it more engaging. Based on the data collected during the game, mainly CDT, we have simulated a navigation task. Description of the task and results are detailed in the Supplementary Materials. Our simulation results show that providing adaptive assistance (adjusting the speed of the robot according to the CDT) reduces the time to finish the task as compared to delivering BCI commands after a fixed period. Our approach is also superior in dealing with large variations of BCI performance across sessions.

Several improvements to the proposed framework can be developed in future work. First, the proposed performance estimator relies on the data at the beginning of a trial, which may not be compatible with asynchronous BCIs (where no cue is used and the subject can start the mental task whenever she/he intends). Extensions of this approach where the performance is estimated on the ongoing patterns (i.e., inferring the probability of delivering a command in the near future) should be developed and validated. Second, due to different characteristics of the two classes (right and left), two separate performance estimators were built. Therefore, the online use of the performance estimator requires class labels, which are rarely available. Semi-supervised approaches, using environmental information [27, 28] or error potentials [29] could be used to infer this information.

Acknowledgment

This work is supported by the Swiss National Center of Competence in Research (NCCR) Robotics and the Hasler Foundation.

References

[1] M. Grosse-Wentrup and B. Schölkopf, "A review of performance variations in SMR-based brain-computer interfaces (BCIs)," in *Brain-Computer Interface Research*, C. Guger, B. Z. Allison, and G. Edlinger, Eds. Springer Berlin Heidelberg, 2013, pp. 39–51.

[2] B. Blankertz, C. Sannelli, S. Halder, E. M. Hammer, A. Kübler, K.-R. Müller, G. Curio, and T. Dickhaus, "Neurophysiological predictor of SMR-based BCI performance," *NeuroImage*, vol. 51, no. 4, pp. 1303–1309, 2010.

[3] M. Grosse-Wentrup, B. Schölkopf, and J. Hill, "Causal influence of gamma oscillations on the sensorimotor rhythm," *NeuroImage*, vol. 56, no. 2, pp. 837–842, 2011.

[4] C. Sannelli, M. Braun, M. Tangermann, and K.-R. Müller, "Estimating noise and dimensionality in BCI data sets: Towards illiteracy comprehension," in *Proceedings of the 4th International BrainComputer Interface Workshop and Training Course*, 2008, pp. 26–31.

[5] M. Ahn, H. Cho, S. Ahn, and S. C. Jun, "High theta and low alpha powers may be indicative of BCI-illiteracy in motor imagery," *PLoS ONE*, vol. 8, no. 11, p. e80886, 2013.

[6] M. Ahn, S. Ahn, J. H. Hong, H. Cho, K. Kim, B. S. Kim, J. W. Chang, and S. C. Jun, "Gamma band activity associated with BCI performance: Simultaneous MEG/EEG study," *Frontiers in Human Neuroscience*, vol. 7, no. 848, 2013.

[7] S. Halder, D. Agorastos, R. Veit, E. M. Hammer, S. Lee, B. Varkuti, M. Bogdan, W. Rosenstiel, N. Birbaumer, and A. Kübler, "Neural mechanisms of brain–computer interface control," *NeuroImage*, vol. 55, no. 4, pp. 1779–1790, 2011.

[8] S. Halder, B. Varkuti, M. Bogdan, A. Kübler, W. Rosenstiel, R. Sitaram, and N. Birbaumer, "Prediction of brain-computer interface aptitude from individual brain structure," *Frontiers in Human Neuroscience*, vol. 7, no. 105, 2013.

[9] E. M. Hammer, S. Halder, B. Blankertz, C. Sannelli, T. Dickhaus, S. Kleih, K.-R. Müller, and A. Kübler, "Psychological predictors of SMR-BCI performance," *Biological Psychology*, vol. 89, no. 1, pp. 80–86, 2012.

[10] A. Vuckovic and B. A. Osuagwu, "Using a motor imagery questionnaire to estimate the performance of a brain-computer interface based on object oriented motor imagery," *Clinical Neurophysiology*, vol. 124, no. 8, pp. 1586–1595, 2013.

[11] M. Ahn and S. C. Jun, "Performance variation in motor imagery brain-computer interface: A brief review," *Journal of Neuroscience Methods*, vol. 243, pp. 103–110, 2015.

[12] C. Vidaurre, C. Sannelli, K.-R. Mller, and B. Blankertz, "Machine-learning-based coadaptive calibration for brain-computer interfaces." *Neural Computation*, vol. 23, no. 3, pp. 791–816, 2011.

[13] S. K. Goh, H. Abbass, K. C. Tan, and A. A. Mamun, "Artifact removal from EEG using a multi-objective independent component analysis model," in *Neural Information Processing*, C. Loo, K. Yap, K. Wong, A. Teoh, and K. Huang, Eds. Springer International Publishing, 2014, vol. 8834, pp. 570–577.

[14] F. Nijboer, N. Birbaumer, and A. Kübler, "The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis - A longitudinal study," *Frontiers in Neuroscience*, vol. 4, no. 55, July 2010.

[15] M. Grosse-Wentrup and B. Schölkopf, "High gamma-power predicts performance in sensorimotor-rhythm brain-computer interfaces," *Journal of Neural Engineering*, vol. 9, no. 4, p. 046001, 2012.

[16] A. Bamdadian, C. Guan, K. K. Ang, and J. Xu, "The predictive role of pre-cue EEG rhythms on MI-based BCI classification performance," *Journal of Neuroscience Methods*, vol. 235, pp. 138–144, 2014.

[17] C. L. Maeder, C. Sannelli, S. Haufe, and B. Blankertz, "Pre-stimulus sensorimotor rhythms influence brain–computer interface classification performance," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 5, pp. 653–662, 2012.

[18] R. Leeb, S. Perdikis, L. Tonin, A. Biasiucci, M. Tavella, M. Creatura, A. Molina, A. Al-Khodairy, T. Carlson, and J. d. R. Millán, "Transferring brain–computer interfaces beyond the laboratory: Successful application control for motor-disabled users," *Artificial Intelligence in Medicine*, vol. 59,

no. 2, pp. 121–132, 2013.

[19] R. Leeb, L. Tonin, M. Rohm, L. Desideri, T. Carlson, and J. d. R. Millán, "Towards independence: A BCI telepresence robot for people with severe motor disabilities," *Proceedings of the IEEE*, vol. 103, no. 6, pp. 969–982, 2015.

[20] T. Carlson and J. d. R. Millán, "Brain-controlled wheelchairs: A robotic architecture," *IEEE Robotics and Automation Magazine*, vol. 20, no. 1, pp. 65–73, 2013.

[21] S. Saeedi, T. Carlson, R. Chavarriaga, and J. d. R. Millán, "Making the most of context-awareness in brain-computer interfaces," in *Proceedings of IEEE International Conference on Cybernetics*, 2013, pp. 68–73.

[22] F. Lotte, F. Larrue, and C. Mühl, "Flaws in current human training protocols for spontaneous brain-computer interfaces: Lessons learned from instructional design," *Frontiers in Human Neuroscience*, vol. 17, no. 568, 2013.

[23] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Advances in Psychology*, vol. 52, pp. 139–183, 1988.

[24] F. Galán, P. W. Ferrez, F. Oliva, J. Guàrdia, and J. d. R. Millán, "Feature extraction for multi-class BCI using canonical variates analysis," in *Proceedings of the IEEE International Symposium on Intelligent Signal Processing*, 2007, pp. 1–6.

[25] A. Kübler, N. Neumann, B. Wilhelm, T. Hinterberger, and N. Birbaumer, "Predictability of brain-computer communication," *Journal of Psychophysiology*, vol. 18, no. 2-3, pp. 121–129, 2004.

[26] H. A. Abbass, J. Tang, R. Amin, M. Ellejmi, and S. Kirby, "Augmented cognition using real-time EEG-based adaptive strategies for air traffic control," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1. SAGE Publications, 2014, pp. 230–234.

[27] A. L. Orsborn, S. Dangi, H. G. Moorman, and J. M. Carmena, "Closed-loop decoder adaptation on intermediate time-scales facilitates rapid BMI performance improvements independent of decoder initialization conditions." *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 4, pp. 468–477, July 2012.

[28] T. Gürel and C. Mehring, "Unsupervised adaptation of brain-machine interface decoders," *Frontiers in Neuroscience*, vol. 6, no. 164, 2012.

[29] R. Chavarriaga and J. d. R. Millán, "Learning from EEG error-related potentials in noninvasive brain-computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18,

no. 4, pp. 381–388, 2010.