Prediction of Command Delivery Time for BCI

S. Saeedi, R. Chavarriaga, I. Iturrate, J. d. R. Millán
Chair in Non-Invasive Brain-Machine Interface (CNBI),
Center for Neuroprosthetics, School of Engineering, EPFL
Lausanne, Switzerland
sareh.saeedi@epfl.ch

T. Carlson
Aspire Create,
Institute of Orthopaedics and Musculoskeletal Science
University College London
Royal National Orthopaedic Hospital, Stanmore, HA7 4LP, UK

***Abstract*—One of the challenges in using brain computer interfaces over extended periods of time is the uncertainty in the system. This uncertainty can be due to the user's internal states, the non stationarity of the brain signals, or the variation of the class discriminative information over time. Therefore, the users are often unable to maintain the same accuracy and time efficiency in delivering BCI commands. In this paper, we tackle the issue of variation in BCI command delivery time for a motor imagery task with the aim of providing assistance through adaptive shared control. This is important mainly because having long delivery of mental commands leads to uncertainty in the user's intent classification and limits the responsiveness of the system. In order to address this issue, we separate the trials into long and short groups so that we have the same number of trials in each group. We demonstrate that using only a few samples at the beginning of the trial, we are able to predict whether the current trial will be short or long with high accuracies ($70\% - 86\%$). Eventually, this prediction enables us to tune the shared control parameters to overcome the issue of uncertainty.**

***Index Terms*—Brain computer interface (BCI), EEG, Adaptive shared controls**

I. Introduction

A Brain-Computer Interface (BCI) monitors the user's brain activity and translates his/her intentions into commands to an external device, such as a wheelchair or a prosthetic device [1]. There are several challenges for determining the user's intention using an uncertain channel, such as BCI. These challenges are mostly associated with relatively low accuracies, low temporal precision, low information transfer rates, measurement noise, and uncertainty in the system. There are several sources of uncertainty in the system, such as the users internal state (attention, etc.), non stationarity of the brain signal, or variation of the class discriminative information within and between users [2]. As a result, capabilities of the users in delivering a mental command are not constant over time. This change can be reflected in accuracy or duration of the time required to deliver a BCI command. In this respect, the use of shared control systems, in which both the user and the system contribute to the control process are been shown to be beneficial [3], [4].

Nevertheless, the shared control techniques applied for BCI usually have predefined settings based on the task and the environment. In consequence, they provide a constant level of assistance based on these settings. Yet, given the mentioned variations in performance, the level of assistance they provide should be adaptive so as to properly complement the user's capabilities at each time [5]. This allows the user to always remain in control of the brain-actuated device [6], [7]. To provide such an adaptive shared controller, there is a need to assess the reliability of the BCI output. Figure 1 shows such a system which works in parallel with BCI and modulates the action generation by estimating the performance.

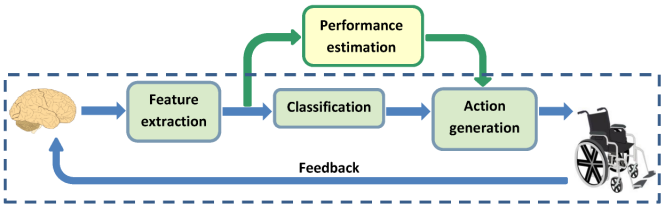

Fig. 1. General block diagram of a BCI in dashed box. Brain signals are decoded into the control commands to the device after processing and classification stages. The goal is to design a system which works in parallel with the BCI and can assess the performance of the BCI system.

Long delivery of mental commands is usually frustrating for the subjects and reduces their engagement in performing the mental task. Also, it can increase the workload. For example, in the task of controlling a brain actuated telepresence robot, the user drives the robot to the right or to the left while the robot moves forward in the absence of mental commands [8]. If the user wants the robot to enter a doorway on the right side, he/she needs to be fast enough to deliver the right command at the proper moment. Otherwise, he/she will lose the chance to turn the robot into the doorway and will need to deliver additional commands to bring the robot back to the same position. In such cases, having a system that can predict if a command is going to take a long time can be helpful. This prediction enables providing adaptive assistance to the user. For instance, the shared controller can reduce the speed of the robot so that the user has enough time to enter the doorway.

The main goal of this study is to assess the reliability of a motor imagery (MI) based BCI output by predicting the time required to deliver a command. In order to do this prediction, we have considered the position of the features at each time point with respect to the classifier properties. In other words, we evaluate the uncertainty of the signal by comparing it to the characteristics of the signal at the calibration phase (the signal used for classification). This prediction allows us to regulate the level of assistance provided for the user based on their capabilities.

In this paper, we will begin with describing our MI BCI system. Then, we will continue by methods used for classification of short and long trials. This will be followed by the results and discussion.

II. Methods

A. Motor Imagery (MI) BCI

In a synchronous MI BCI paradigm, the users learned to voluntarily modulate EEG oscillatory rhythms by performing motor imagery tasks, e.g. movement imagination of right hand, left hand, or feet. Two female and two male healthy subjects (mean age 27, range 24-30 years old) participated in this study. One of the subjects had experience with MI BCI. All the experiments were conducted in the laboratory conditions to minimise additional sources of noise. EEG was recorded using 16 electrodes over the sensorimotor cortex at $512Hz$ and band-pass filtered between $0.1Hz$ and $100Hz$. Laplacian spatial filtering was then applied on the signal. Then, feature extraction and classification (detailed below) were executed to decode the user's intention, i.e. moving the cursor to the right or to the left.

1) Experimental protocol: First, subjects underwent a training phase, where they were asked to imagine the movement of their right hand, left hand, and feet following the relevant cue. The training phase was done in a session comprising four '*offline*' runs, which were used to train the classifier. The runs consisted of 15 trials of each mental task which were randomly organized. Timing of trials is depicted in Figure 2. First, a cross appears on the screen showing that the subject should get ready to execute the task. Then, a cue (arrows to the right, left, or up) is shown, based on which the subject needs to do the instructed mental task for a period of four seconds. During these four seconds, they see the gray bar moving in the direction indicated by the cue with a constant speed. A feedback is then given to them showing that the trial has ended. The recorded EEG signal was assessed using the feature selection and classification methods (discussed in the following sections). Then, in case of achieving a certain level of classification accuracy, the two most separable mental tasks were chosen to be used for online BCI control.

In the following sessions, the subjects were recorded in a two-class motor imagery task (e.g. hand/feet), in which they were asked to do the relevant mental task following a cue on the screen while receiving a visual feedback from the classifier outputs ('*online*' runs). In fact, the classifier outputs were translated into the movement of the gray bar at each time point. The gray bar continued to move until the classifier output surpassed a subject-specific threshold, at which point the corresponding BCI command was 'delivered' and the subject had a brief rest (random between 2 and $3s$). In this way, the users were able to learn from the congruent feedback and adjust their techniques of performing the mental tasks accordingly. The experiment was done in a session comprising six *online* runs. The runs consisted of 15 randomly organized trials of each mental task. As mentioned in the previous section, the command delivery time is not the same for all the *online* trials as the movement of the feedback bar is directly controlled by the classifier output [9].

It is worth mentioning that all the four subjects performed right hand movement imagination as class 1. As class 2, subject1 and subject4 performed left hand movement imagination while subject2 and subject3 performed feet movement imagination. All the three subjects went through one offline session. The number of online sessions was 1, 2, 2 for subjects 1, 2, and 3 respectively. Subject4 had experience with MI BCI.

2) EEG decoding: Decoding of the user's intention from EEG was carried out in the following steps:

Feature extraction/selection: The brain correlates associated to motor imagery appear as a decrease/increase in the band power of the EEG signal [10] in specific frequency bands (typically μ, $8-14Hz$ and β, $18-24Hz$). Therefore, the power spectral density (PSD) of the signal (over the last one second) was calculated with the resolution of $2Hz$. The PSDs were estimated every $62.5ms$ (i.e., 16 times per second) using Welch method with 3 overlapped (50%) Hamming windows of $500ms$. Given the number of channels (16) and the number of frequency components (23), each EEG sample comprises 368 features.

After extracting the features, we performed a feature selection process to find for each subject those features that maximized the separability of the two mental tasks. Canonical variate analysis (CVA) was used to project PSD samples onto the canonical space [11]. Subsequently, the features were ranked based on their correlation with the projected ones. The final feature selection was done manually considering this rank and the neurophysiological evidence on the cortical areas/frequency bands, which are expected to contribute to the mental task [12]. In this study, we selected 7 to 13 features per subject.

Classification: Data from the training period was used to fit a Gaussian Mixture model (GMM) classifier of four prototypes per class. Then, in the online runs, real-time classification of selected features was done at each time point. Classes were assumed to have equal priors as well as common diagonal covariance matrices. The activation of j^{th} prototype of class i, with center μ_{ij} and covariance matrix Σ_i is given by:

$$a_{ij} = \frac{1}{\prod_k \sum_{ik}} exp(-\frac{1}{2}\sum_k \frac{(x_k - \mu_{ijk})^2}{\sum_{ik}}) \quad (1)$$

where x is a sample with k elements (selected features, i.e. EEG channels and frequency bands).

The posterior probability p_t of class c is derived as a function of the total activation of the classifier (A) and the activation of class c (a_c):

$$p_t = \frac{a_c}{A} \quad (2)$$

$$A = \sum_{i=1}^{N_c}\sum_{j=1}^{N_p} a_{ij} \quad (3)$$

$$a_c = \sum_{j=1}^{N_p} a_{cj} \quad (4)$$

Where N_c is the number of classes and N_p is the number of prototypes for each class.

Evidence accumulation: The GMM classifier provides a discrete posterior probability distribution over the two mental tasks ($p_t = P(c_i|x_t)$) given the feature vector x_t extracted from the EEG signal at time t. In order to tackle the uncertainty of the single sample classification and to provide smooth feedback to the user, we introduced memory to the system by incorporating the past evidence. That is, the feedback to the user (the movement of the bar) is updated based on an integrated probability $P^*_{(t)}$, which is computed as:

$$P^*_{(t)} = \alpha P^*_{(t-1)} + (1-\alpha)p_t \quad (5)$$

Where $\alpha \in [0, 1]$ and it was set as 0.96 or 0.97 in our experiments based on previous experience. α is one of the user specific values which affects the speed of the feedback bar (since this represents the integrated classifier output) and consequently the delivery time. Delivery of a BCI command is performed when the integrated probability reaches a decision threshold (th_d).

B. Real-time prediction of long/short delivery of mental commands

The uncertainty in BCI systems in general and the evidence accumulation strategy in our system lead to (sometimes high) variations in command delivery time. The distribution of command delivery time over different trials is depicted in Figure 3 for the four subjects and confirms this variation between different trials in the experiment. Our main goal in this study is to estimate reliability of commands by predicting the expected delivery time in online runs based on the initial samples. As the BCI chain has different modules (Figure 1), one may think of doing this assessment in different levels, such as the EEG signal, the PSD features, and the classifier output.

Previously, we used Entropy as a measure of information content of the EEG signal in order to evaluate how reliable the BCI command is [13]. However, this method is challenging for many reasons: firstly, the window of data required for a reliable entropy estimation is longer than some of the trials. Secondly, in order to perform well in real-time, a simple and fast method of estimating entropy should be applied which may not necessarily lead to accurate estimations. Thirdly, some of the preprocessing steps, like binning the data before entropy estimation, requires the data of each trial to be normalised which may mask some modulations in the signal.

In order to overcome these issues, in this study, we focus on the feature level and the classifier output. To do so, we have explored five different cases for conducting the long vs. short classification of trials. In order to do the prediction, we considered a window W at the beginning of a trial (the green window in Figure 2) which is shorter than the shortest trial (W is $1.2s$ to $1.5s$ for all the subjects). In this way, we can predict for all the trials whether they will be short or long. Also, we separated the trials into '*long*' and '*short*' ones based on the median of command delivery time (MDT) in a session of online experiment. These analyses were conducted separately for different mental tasks (right hand, left hand, or feet movement imagination). A linear discriminant analysis (LDA) classifier with five-fold cross validation was implemented for classification of short vs. long trials in all cases (I to V).

(I) **Classifier output:** The goal is to assess if the posterior probabilities at the beginning of a trial reflect the level of time efficiency in that trial. To do so, the average of the posterior probabilities within W was calculated in order to classify long vs. short trials.

PSD features: As it was mentioned, the posterior probability of a feature vector 'x_t' belonging to class C_i is an exponential function of the distance between the feature vector and the center of the prototypes of the class. That is, the closer the sample is to the center of prototypes of C_i, the higher the probability of that sample belonging to C_i is. Given the feature vectors extracted from EEG within W, we define a distance measure for class i and prototype j as:

$$Dist_{ij} = \frac{1}{N_w} \sum_{l=1}^{N_w} \sum_{k=1}^{N_f} \frac{(x_{lk} - \mu_{ijk})^2}{\sum_{ik}} \quad (6)$$

Where N_w is the number of feature vectors within the window, and N_f is the number of features in the feature vector x_l. Therefore, for each trial we have $2 \times N_p$ distances to the prototypes of the two classes. The defined distance measure is similar to the one used for calculating the posterior probability of the GMM classifier. However, in order to compute the posterior probabilities, the distances are normalized using an exponential function which smooths the differences. Also, posterior probabilities are sum of the activation of prototypes while some of the prototypes may be more influential than others in the differences between long and short trials. Therefore, four different sets of features have been considered:

(II) For class i, the prototype with the smallest distance to the feature vector x_l contributes more to the posterior probability. In this case, the minimum distance to prototypes of the two classes were considered as the features to be used for long/short classification.

$$f_{sl2} = [\min_{j=1:N_p} Dist_{1j}, \min_{j=1:N_p} Dist_{2j}] \quad (7)$$

(III) A measure of how a feature vector x_l is close to the prototypes of one class and far from the others can be derived by subtracting the average of distances to the prototypes of the two classes.

$$f_{sl3} = (\sum_{j=1}^{N_p} Dist_{1j} - \sum_{j=1}^{N_p} Dist_{2j})^2 \quad (8)$$

(IV) The distances of a feature vector x_l to all the eight prototypes.

$$f_{sl4} = Dist_{ij}, i = 1 : N_c, j = 1 : N_p \quad (9)$$

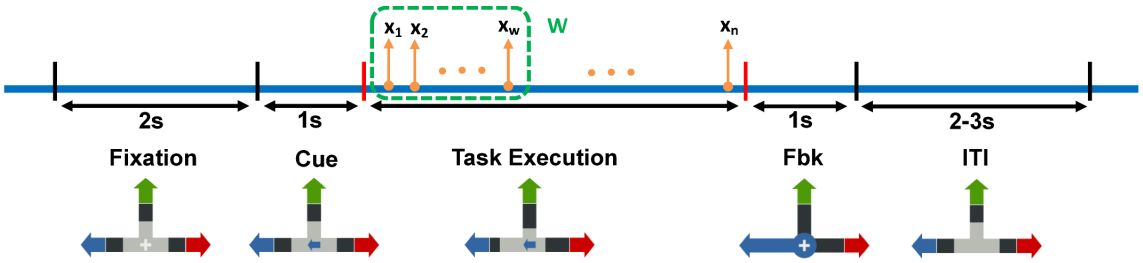

Fig. 2. Timing of a trial in a MI BCI experiment. Offline trials: First, a cross is shown on the screen so that the user prepares for the task. Then the cue appears and the subject needs to execute the relevant mental task for 4s. A feedback is then given to them showing that the trial has ended. This is followed by an inter trial interval (ITI) period. In the task execution part, PSD features ($x_i, i = 1 : n$) are extracted 16 times per second and a GMM classifier is trained based on a subset of them. In these trials, the user executes three different mental tasks. Online trials: all the steps are the same, except for the duration of the task execution, which depends on the classifier output at each time and can have different lengths across trials. Also, two out of three mental tasks are used. The green square shows the window at the beginning of a task execution which is used for real-time prediction of long/short trials.

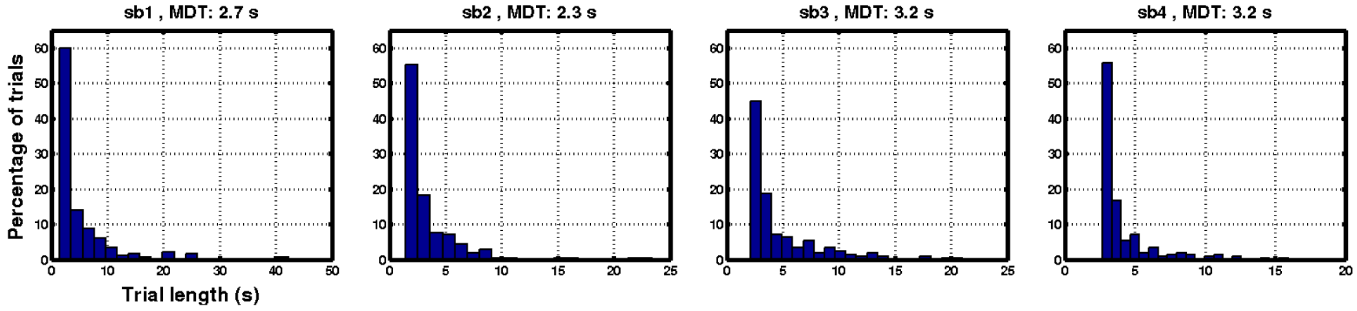

Fig. 3. Distribution of mental command delivery time is shown for the four subjects, which confirms relatively high variations across different trials. The median delivery time (MDT), which is the treshold for separating the long and short trials, is defined on top of the figures for each subject.

In this case, CVA has been used to select the features in f_{sl} which contribute more to the classification of long vs. short trials.

(V) Combining f_{sl3} and f_{sl4} can provide additional information for the classification. In this case also CVA was used as a feature selection method.

III. Results

All subjects performed well in the MI experiment with performances of 98%, 99%, 88%, and 95% (subjects 1 to 4, respectively). Figure 4 illustrates the average distance of the PSD feature vectors (within the window W) to the prototypes, averaged over all the online trials (subject 2). Solid lines show the distance to the prototypes of the desired class, i.e. the mental task that the subject is doing with respect to the cue. The distance between the feature vectors and the prototypes of the desired class does not show any significant difference between short and long trial. Dashed lines show the distance to the prototypes of the other class. According to this Figure, the distances between the feature vectors and the prototypes of the other class is higher for short trials compared to long ones. This suggests that, for short trials, even at the beginning of the trial, the PSD features are close to the prototypes of the desired class and far from the other one, whereas the PSD features of long trials may fall somewhere in between the prototypes of the two classes. Similar results have been observed for other three subjects.

As described in section II, we used these distances in four different ways, as well as the posterior probabilities in order to classify short vs. long trials. Figure 5 shows the average accuracy of classification (over 5 folds). According to the results, case III has the lowest accuracy among all, which suggests that by calculating the mean of the distances we are losing some information.

Case I (i.e. the posterior probabilities) results in a good accuracy most of the times, but the accuracy is not consistent for all subjects. Besides, the variation of accuracies (among folds) is quite high compared to the other cases. Case II also does not result in a good accuracy for all the subjects. However, case IV, i.e. using the distances to a selected subset of prototypes shows a higher and more consistent accuracy of classification (more than 70%) for all the subjects.

Figure 6 shows the receiver operating characteristic curve for classification in case IV for each subject. The x axis and the y axis denote the false positive rate (FPR) and the true positive rate (TPR), respectively (long trials are considered as positive). Table I illustrates area under the ROC curve for cases I to V. According to this table, only cases IV and V (which use more or less the same features) show consistent AUC's higher than 0.70 for all the subjects.

IV. Discussion

One of the challenges in using BCI's over extended periods of time is that they are subject to performance variations. This is usually reflected in the accuracy in performing and decoding a task and the time it takes to do so. Shared control techniques have addressed to some extent this variability in the user's

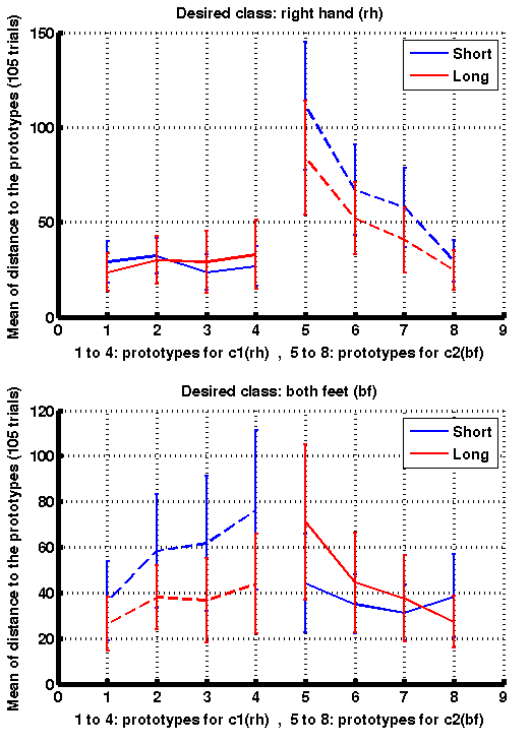

Fig. 4. The average distances of the PSD features (within window W) to the prototypes of both classes, averaged over trials (subject2). Solid lines show the distance to the prototypes of the desired class, i.e. the mental task that the subject is doing with respect to the cue. Dashed lines show the distance to the prototypes of the other class. Short trials show smaller distance to the prototypes of the desired class and larger distance to the prototypes of the other one.

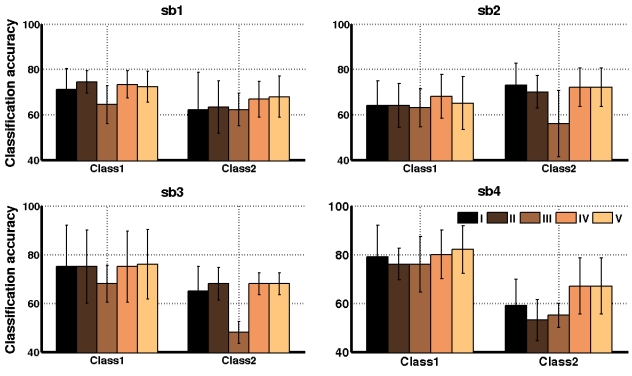

Fig. 5. Average of accuracy of classification over 5 folds, in the five cases. The first case is when we use the average of posterior probabilities for classification and the rest are when we use different features based on the distances between feature vectors and classifier prototypes.

control capabilities over time [3]. However, these techniques usually do not take into account the sources of uncertainty in the system, such as the user's internal states at each time [7]. One of the issues is the variation in the trial lengths across trials of the online runs. To tackle this issue, we have designed a classifier to predict (within around one second) if a trial will be long or short in a MI BCI.

The classifier that is used for our MI BCI system is a

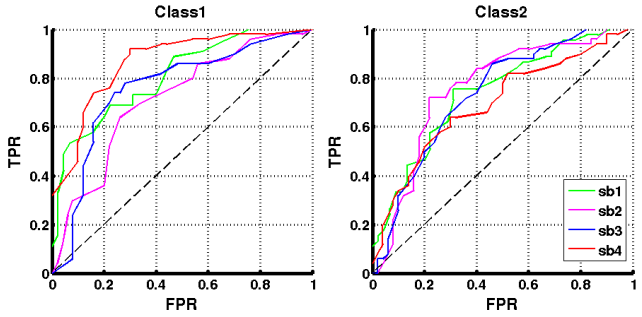

Fig. 6. ROC curve for classification in case IV for each subject. The x axis and the y axis denote the false positive rate (FPR) and the true positive rate(TPR), respectively (long trials are considered as positive). The dashed line shows the random case.

	Class1				Class2			
	sb1	sb2	sb3	sb4	sb1	sb2	sb3	sb4
I	0.79	0.70	0.74	0.83	0.64	0.76	0.70	0.58
II	0.83	0.72	0.78	0.86	0.70	0.79	0.72	0.57
III	0.76	0.70	0.70	0.81	0.66	0.50	0.47	0.61
IV	0.81	0.71	0.76	0.86	0.74	0.76	0.74	0.70
V	0.80	0.71	0.78	0.87	0.73	0.76	0.73	0.69

TABLE I
AREA UNDER THE ROC CURVE FOR EACH CASE.

GMM classifier, in which the posterior probability of a sample belonging to a class is an exponential function of the distance between that sample and the center of the prototypes of the class. The distance between the samples (in the beginning of a trial) and the prototypes of both classes show different patterns for long and short trials (Figure 4). In both cases, there is more or less the same distance to the desired class, but the short trials have higher distance to the other class than the long ones. This suggests that in short trials, even the few first samples are close to prototypes of the desired class and quite far from the prototypes of the other. This is also reflected in the posterior probabilities. That is, for short trials there is a higher certainty of samples belonging to the desired class.

Five different types of features have been considered: the first based on the posterior probabilities and the rest based on the distance between the features and the classifier prototypes. Among all, cases I (where posterior probabilities were considered) and IV (where a subset of the distances to all prototypes were considered) showed the higher and more consistent accuracies for all the subjects (Figure 5). This suggests that not only the closest prototypes (case II), but a subset of them (as chosen by CVA) contribute to the differences in long and short trials. There is a small difference between the accuracies in these two cases which can be due to the fact that for the calculation of posterior probabilities (case I), all the distances are considered, whereas in case IV, the redundant information is discarded by selecting a subset of features that carry more information. Besides, to compute the posterior probabilities, the distances are normalized using an exponential function which smooths the differences.

Comparing the results of long vs. short classification of

trials with the results in [13] highlights the advantages of choosing features from the PSD feature level or classifier output rather than using entropy of the EEG signal. Firstly, when using the distances between the features and the prototypes or the posterior probabilities, there is no need of additional preprocessing of the signal. Secondly, the window required for making prediction about the trial delivery time is shorter than the shortest trial. That is, the prediction can be reliably executed for all trials in around $1s$.

In classification of long vs. short trials, an important factor is the false positive rate (FPR), which shows the percentage of the long trials which are misclassified. The AUC's for cases I to V are compared in Table I. These results confirm that classification of long vs. short in case IV is more reliable and consistent for all subjects. In other words, the trials can be reliably classified as long/short ones, considering only a window of $1.2 - 1.5s$ at the beginning of the trial (i.e. 900ms or more before the actual median delivery time of the subject). According to this table and Figure 6, the results of classification for all subjects are better for the first class (right hand movement imagination). This is probably due to the fact that they are right handed and the movement of right hand is more natural for them compared to left hand or feet.

In conclusion, we have proposed a method for real-time classification of long and short trials in a MI BCI. According to the results, this method allows us to make a reliable prediction of how fast the user will deliver a command within a few seconds. This prediction is essential for regulating the level of assistance in shared control systems. That is, we can provide an adaptive shared controller to overcome some aspects of uncertainty in the BCI systems.

ACKNOWLEDGMENT

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

REFERENCES

[1] J. d. R. Millán, R. Rupp, G. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K. Müller, and D. Mattia, "Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges," *Frontiers in Neuroscience*, vol. 4, p. 161, 2010.

[2] B. Allison, S. M. Dunne, R. Leeb, J. d. R. Millán, and A. Nijholt, *Towards Practical Brain-computer Interfaces : Bridging the Gap from Research to Real-world Applications*, ser. Biological and medical physics, biomedical engineering. Heidelberg: Springer Verlag, 2012.

[3] H. K. Kim, S. J. Biggs, D. W. Schloerb, J. M. Carmena, M. A. Lebedev, M. A. L. Nicolelis, and M. A. Srinivasan, "Continuous shared control for stabilizing reaching and grasping with brain-machine interfaces." *IEEE Trans Biomed Eng*, vol. 53, no. 6, pp. 1164–1173, Jun 2006.

[4] D. Vanhooydonck, E. Demeester, M. Nuttin, and H. V. Brussel, "Shared control for intelligent wheelchairs: an implicit estimation of the user intention," in *Proceedings of the ASER '03 1 st International Workshop on Advances in Service Robotics*, pp. 13–15.

[5] T. Carlson, R. Leeb, R. Chavarriaga, and J. d. R. Millán, "Online Modulation of the Level of Assistance in Shared Control Systems," in *Proceedings of the IEEE International Conference on Sytems Man and Cybernetics (SMC 2012)*, 2012, pp. 3321–3326.

[6] J. Philips, J. d. R. Millán, G. Vanacker, E. Lew, F. Galán, P. Ferrez, H. V. Brussel, and M. Nuttin, "Adaptive Shared Control of a Brain-Actuated Simulated Wheelchair," in *Proceedings of the 10th IEEE International Conference on Rehabilitation Robotics*, 2007, pp. 408–414.

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

[8] L. Tonin, T. Carlson, R. Leeb, and J. d. R. Millán, "Brain-controlled telepresence robot by motor-disabled people," in *Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC 2011*, 2011, pp. 4227–4230.

[9] 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.

[10] E. Niedermeyer and F. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 5th ed. Lippincott Williams & Wilkins, Nov. 2004.

[11] W. J. Krzanowski, Ed., *Principles of Multivariate Analysis: A User's Perspective*. New York, NY, USA: Oxford University Press, Inc., 1988.

[12] 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 *IEEE Int Symp Intelligent Signal Processing*, 2007.

[13] S. Saeedi, R. Chavarriaga, M. C. Gastpar, and J. d. R. Millán, "Real-time Prediction of Fast and Slow Delivery of Mental Commands in a Motor Imagery BCI: An Entropy-based Approach," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 42, no. 3, 2012, pp. 3327–3331.