Asynchronous P300-Based Brain-Computer Interface to Control a Virtual Environment: Initial Tests on End Users

Fabio Aloise, Francesca Schettini, Pietro Aricò, Serenella Salinari, Christoph Guger, Johanna Rinsma, Marco Aiello, Donatella Mattia and Febo Cincotti

Key Words
Asynchronous Control
Brain-Computer Interface (BCI)
Domotic Appliances
Electroencephalography
P300

ABSTRACT
Motor disability and/or ageing can prevent individuals from fully enjoying home facilities, thus worsening their quality of life. Advances in the field of accessible user interfaces for domotic appliances can represent a valuable way to improve the independence of these persons. An asynchronous P300-based Brain-Computer Interface (BCI) system was recently validated with the participation of healthy young volunteers for environmental control.

In this study, the asynchronous P300-based BCI for the interaction with a virtual home environment was tested with the participation of potential end-users (clients of a Frisian home care organization) with limited autonomy due to ageing and/or motor disabilities. System testing revealed that the minimum number of stimulation sequences needed to achieve correct classification had a higher intra-subject variability in potential end-users with respect to what was previously observed in young controls. Here we show that the asynchronous modality performed significantly better as compared to the synchronous mode in continuously adapting its speed to the users’ state. Furthermore, the asynchronous system modality confirmed its reliability in avoiding misclassifications and false positives, as previously shown in young healthy subjects.

The asynchronous modality may contribute to filling the usability gap between BCI systems and traditional input devices, representing an important step towards their use in the activities of daily living.

INTRODUCTION
Environmental control is an important challenge for people who have (partially) lost motor ability. The opportunity to independently perform simple everyday actions, such as turning on/off lights, opening a door or changing TV channels, might represent a significant improvement in the quality of life, by reducing the level of dependency from the caregivers. In this respect, the recent advance in the field of domotics (a set of methods and techniques for the automation of the home) has remarkably augmented the potential to interact with the environment and thus, the management of everyday life activity in case of disability (as a forefront example see SMALL - Smart Homes for All European project1). The access to environmental control interface might be, however, limited when disability is severe and progressive.

Brain-computer interface (BCI) systems utilize neurophysiological signals originating in the brain to activate external devices or a computer.2,3 Nowadays, non-invasive BCIs exploiting electroencephalo-
The asynchronous modality was first implemented for a P300-based BCI by Zhang et al., and was based on a statistical method. In Aloise et al., a different asynchronous P300-based BCI system was implemented to control real domotic appliances, and it was proved to be reliable in avoiding false positives when the healthy young subjects (n=11) diverted their attention from the stimulation interface. However, no significant improvement in selection speed could be determined with respect to a synchronous system.

The aim of the present study was to validate our asynchronous P300-based BCI system for domotic appliance control with the participation of a sample of potential end-users who might benefit from an enhanced environmental control because of their moderate to severe disability due to ageing and/or neurological disorders. Two main questions were addressed: 1) whether the intra-subject variability expected to be higher in potential end-users (with respect to previously reported data obtained from young controls) would have a direct effect on the stability of the number of stimulus sequences needed to achieve a correct classification; 2) whether such a system would be robust to false positives and to misclassifications avoidance.

**MATERIALS AND METHODS**

**Participants**

Seven elderly (4 males, 3 females; mean age=64.85 ± 5.81 years) individuals who are clients of the Frisian home care organization (THFL) joined the recording protocol. Four of them suffered from chronic neurological disorders: one affected by amyotrophic lateral sclerosis (ALS), two by multiple sclerosis (MS) and one by stroke. The degree of functional motor impairments was defined on the basis of the ALS Functional Rating Scale-Revised (ALSFRS-R), the Kurtzke Expanded Disability Status Scale (EDSS), the Rankin Scale for Stroke Disability (RSSD). In addition, the Barthel Index (BI) was administered to all participants to estimate a global degree of independence in performing daily activities. Functional scores revealed a moderate to severe motor disability (EDSS score = 3 for MS patients; ALSFRS-R score = 10 for ALS patient; RSSD score = IV for the stroke patient; BI 85.71±19.24; n=7). Cognitive functions were preserved in all of them as indicated by the MMSE scores (between 27-30). Each participant had previous experience in exploring the virtual environment (see below) by means of classical input devices such as mouse or joystick. All of them were naive to the BCI control.

**Experimental setup**

The domotic environment to be controlled was based on a virtual reconstruction of a real apartment that is built at the premises of the Fondazione Santa Lucia in Rome (see Figure 1A). The apartment consisted of four rooms: two bedrooms, a kitchen and a living room, and the devices operable by means of the BCI were lights, doors, curtains, windows, bed, TV, air conditioning, SOS and “sleep macro.” The “sleep macro” arranges the environment in sleep modality (e.g., turns off the lights, closes the curtains). Detailed description of the environment simulation is reported in Kaldeli et al.. In our experiments, a projector was used to present the apartment in a bird-eye’s view to the participant; a laptop processed the instructions coming from the BCI software. Figure 1B shows a moment of the actual experimentation.

The acquisition protocol was based on the Speller paradigm, implemented in the P3Speller application within the BCI2000 framework. Stimuli were provided by row and column flashing in a matrix and a static fixation cross was placed in the middle of the interface. The first recording session consisted in a text input task, and for this reason we used a 6 by 6 matrix containing alpha numeric items. For the second recording session, which concerned control of the domotic appliances the Speller was reduced to a 4 by 4 matrix and letters were replaced with the icons representing the devices available in the virtual apartment. In order to perform the asynchronous control, we used a modified version of the P3Speller application. When the BCI system recognized a Target, a message was sent to an external application that parsed the information and generated a call request using the XML based Simple Object Access Protocol (SOAP) protocol to the virtual environment. Finally, the software provided a feedback to the user through the animation of the selected device in the virtual environment. For instance, commanding the closing of a curtain resulted in the actual movement of the curtain fabric in the virtual environment.

Scalp EEG potentials were recorded from 16 positions according to the 10-10 standard (Fz, FCz, Cz, CPz, Pz, O2, F3, C3, C4, CP3, CP4, P3, P4, PO7, PO8) with g.LADYbird active ring electrode using the g.USBaamp amplifier (g.tec medical engineering GmbH, Austria). The EEG was sampled at 256 Hz. Each channel was referenced to the left earlobe and grounded to the right mastoid. Stimulation was provided to the subjects using a 17”LCD monitor placed at 1 meter of distance from him/her.

**Recording protocol**

The purpose of the first recording session was to investigate the subjects’ ability to control a P300-based BCI and to allow them to familiarize with the system.
Each participant underwent two recording sessions on different days. In the first recording session (Speller), we asked the subjects to perform 5 runs. During the second session (Domotic), in order to train and test the asynchronous classifier, subjects performed also No-Control runs for a total of 10 runs. A run consisted of 5 trials and every trial was composed of 12 stimulation sequences (each composed by the single flash of each row and column on the interface). Each stimulus was intensified for 125ms, the inter-stimulus interval (ISI) was set at 125ms.

**Speller session**

The Speller session consisted of 5 Copy mode runs in synchronous mode, i.e., the user was cued with a target letter to concentrate on, and a fixed-length train (12 sequences) of random stimuli was presented. Subjects were asked to spell five common words: WATER, WATER, KOPJE, BROOD and KLEIN (these words are in Dutch and mean water, water, cup, bread and small). The word WATER was repeated twice as previously reported by Guger et al. In order to extract the most significant features we applied a Stepwise Linear Discriminant Analysis (SWLDA) on the first two Copy runs, during which subjects did not receive any feedback regarding classification outcome. During the last three runs, the selected letter was presented to the subject at the end of each trial. Since on-line results referred to 12 stimulation sequences, we performed an off-line cross-validation to provide accuracy as a function of stimulation sequences. In particular we performed 10 rounds cross-validation using all possible combinations of 2 runs to extract control parameters, whereas the remaining 3 runs were used for testing them.

**Domotic session**

The Domotic session was in two parts; synchronous and asynchronous applications.

The synchronous part consisted of 4 Copy mode runs and 2 No-Control runs. During each Copy mode run the subjects had to operate five different devices in the virtual environment. The system suggested the Target icon at the beginning of each trial in a pseudo random order, ensuring that all the items of the matrix were presented at least once. We used the first two runs to extract control parameters through SWLDA. The parameters estimated from the Speller session were not used since the matrix was different in size, and thus implied changes in the Target to Target Interval (TTI). The TTI can affect P300 morphology and, as a consequence, the control parameters. From the third run, we provided feedback to the user about the outcome of classification: at the end of every trial the selected icon was intensified and the subjects could see the corresponding device change its state in the virtual environment. This occurred within 5 seconds after each trial. We extended the inter-trial time to 10 seconds in order to avoid artifacts due to subject’s movements looking at the image of the virtual apartment. As for the data acquired in the Speller session, we performed on the Domotic session data a 6 rounds cross-validation: data relating to the copy mode runs were divided into a training data set composed of two runs and a testing data set including the remaining runs. During No-Control runs no target icons were provided to the subjects who were required to ignore the stimulation and execute two different No-Control Tasks: during the No-Control Task 1 the subject had to gaze at a fixation cross in the middle of the interface while the stimulation was on, during the No-Control Task 2 the subject had to gaze at the fixation cross and talk with the operator (if he/she was able to verbally communicate) while the stimulation was on.

The asynchronous part was composed of 1 Control run and 2 No-Control runs operated by the asynchronous classifier. The asynchronous system is based on the introduction of threshold values in the on-line classifier: at the end of each sequence, the maximum row and column score values were compared to the specific threshold. If the threshold was exceeded because of the maximum row and column values, the system classified the icon at their intersection. Conversely, if the threshold values did not exceed the maximum number of stimulation sequence fixed a priori (reset value), the system refrained from making a selection. After the reset, the system set to zero the scores values accumulated for each stimulus class, and a new trial began. We extracted features and thresholds using data acquired during the synchronous part. In order to make the system robust to possible artifacts that can occur during No-Control tasks, the SWLDA was applied also on the No-Control data which were labeled as a Non-Target. Threshold extracting relies on ROC curve plotting of score values, and the latter were dependant on the number of stimulation sequences accumulated in a trial. During the Control run, subjects were asked to operate five different devices in the environment using the BCI. At the beginning of each trial, the operator suggested the icon on which the user had to focus. There could be three different classification outcomes: 1) correct classification: the system correctly recognized the target icon; 2) misclassification: the system selected an unwanted item; 3) abstention: the thresholds were not exceeded throughout the reset value of stimulation sequences, so a new trial began without selections. Subjects had 10 minutes to complete the task, otherwise the task was considered incorrect. If an unwanted
abstention or a misclassification occurred, the operator invited the subject to again select the desired icon. This run allowed quantification of the accuracy of the asynchronous system when the subject intends to operate a control on the environment. In order to quantify the robustness of the asynchronous system with respect to false positives, subjects performed two No-Control runs lasting 2.5 minutes, during which subjects repeated the two No-Control tasks.

**Intra-subject variability analysis**

In order to quantify intra-subject variability of the stimulation sequences needed to achieve a correct classification for end-users, the data acquired during the domotic session were used. In particular, from the 6 rounds off-line cross-validation, we collected the number of stimulation sequences at which the correct classification was achieved. We compared the results obtained from data acquired in this study with those obtained from the dataset collected from 11 healthy young subjects for a previous study. This latter Control dataset related to 4 Copy mode runs of 8 trials each in which subjects were engaged in a similar experimental task. The stimulation modalities were the same as the actual study: stimuli were provided through a 4 by 4 matrix containing 16 black and white icons (Stimulus Duration = 125ms, ISI = 125ms), representing some device that can be really operated by BCI. The icons presented in this interface are slightly different with respect to the icons used in the previous study. In order to make the two datasets comparable, we reduced the end-users’ dataset to 8 EEG acquisition channels (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8) and we considered only the first 10 stimulation sequences as it was in the Control dataset. Also for the latter we used only the first 5 trials of each run.

**RESULTS**

**On-line results**

Figure 2.A illustrates online results of the Speller and Domotic sessions. The bars represent the mean accuracy reached during the runs in abstention or a misclassification occurred, the operator invited the subject to again select the desired icon. This run allowed quantification of the accuracy of the asynchronous system when the subject intends to operate a control on the environment. In order to quantify the robustness of the asynchronous system with respect to false positives, subjects performed two No-Control runs lasting 2.5 minutes, during which subjects repeated the two No-Control tasks.

**Intra-subject variability analysis**

In order to quantify intra-subject variability of the stimulation sequences needed to achieve a correct classification for end-users, the data acquired during the domotic session were used. In particular, from the 6 rounds off-line cross-validation, we collected the number of stimulation sequences at which the correct classification was achieved. We compared the results obtained from data acquired in this study with those obtained from the dataset collected from 11 healthy young subjects for a previous study. This latter Control dataset related to 4 Copy mode runs of 8 trials each in which subjects were engaged in a similar experimental task. The stimulation modalities were the same as the actual study: stimuli were provided through a 4 by 4 matrix containing 16 black and white icons (Stimulus Duration = 125ms, ISI = 125ms), representing some device that can be really operated by BCI. The icons presented in this interface are slightly different with respect to the icons used in the previous study. In order to make the two datasets comparable, we reduced the end-users’ dataset to 8 EEG acquisition channels (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8) and we considered only the first 10 stimulation sequences as it was in the Control dataset. Also for the latter we used only the first 5 trials of each run.

**RESULTS**

**On-line results**

Figure 2.A illustrates online results of the Speller and Domotic sessions. The bars represent the mean accuracy reached during the runs in
other tasks. If thresholds were not passed, 2 abstentions could be collected in a minute; on the contrary, if the thresholds were incorrectly exceeded after the first stimulation sequence a maximum of 30 false positives could occur in a minute not considering the inter-trial interval. We detected on average 0.225 false positives/minute. This value is comparable to the 0.26 false positives/minutes achieved in our previous study with the control subjects.23

Off-line results

Figure 3.A and 3.B show the outcome of the 10 and 6 rounds cross-validation performed for Speller and Domotic datasets, respectively. A t-test was used on the accuracy values reached at the 12th stimulation sequence obtained on-line and off-line for the Speller and Domotic sessions. It did not show statistical difference between the on-line and the off-line accuracy (Speller: t-value = 0.082, p-value = 0.93; Domotic: t-value = 0.90, p-value = 0.38).

The box plots in Figure 4.A and 4.B represent the distribution of the number of stimulation sequences needed to achieve the correct classification for end- and control-users. Distributions are the outcomes of the 6 rounds cross-validation. It can be seen that end-users exhibited a higher intra-subject variability of stimulation sequences with respect to the control-users. This was confirmed performing a t-test (x = 0.05): the two standard deviations distributions of the number of stimulation sequences for end-users and control-users were statistically different (p-value = 0.04, t-value = 2.22); in particular, variability in the end-users distribution (mean value = 2.1 sequences) was higher than in the control (mean value = 1.53 sequences).

Figure 4.C illustrates the mean value and the standard deviation of the number of stimulation sequences to achieve correct classifications for each subject during the online asynchronous control run. The standard deviation shows that the asynchronous system was able to adapt its speed of selection to the intra-subject variability, which typically is in the range of two sequences.

DISCUSSION

One of the most important features of an asynchronous system is the capability to estimate from the ongoing EEG when the user intends to exert his control on the system, thus avoiding misclassifications when attention is focussed elsewhere. The finding of 0.225 false positives/min on average obtained from the asynchronous modality testing with potential end-users is indeed comparable with the rate of 0.26 false positives/min on average previously reported in control volunteers.23 However, in our previous study the asynchronous system modality did not show a significant improvement in the bit-rate as compared with the synchronous operational modality of the P300-based BCI. In this regard, it should be considered that the previous study involved young subjects (mean age 26.45±4.05). We hypothesize that ageing (even normal) would play a role in determining the level of intra-subject variability which in turn might affect the number of stimulation sequences needed to achieved a correct classification. In the case of young volunteers with a lower intra-subject variability, the feature of the asynchronous system to adapt its speed of selection to the current user state was not enhanced. In the current study, aged potential end-users showed a higher intra-subject variability in terms of time (number of stimuli) to achieve correct classification. Under this condition, a synchronous system would cause obvious uncertainty in deciding the number of sequences to be used for the online control of domotic and/or communication appliances. Choosing a higher value can improve performance in terms of system accuracy, but would also lead to a slower system. The asynchronous system can provide a solution by continuously adapting its speed to the most effective number of stimulation sequences, thus, maintaining high accuracy without lowering the system’s bit-rate.

The present findings reinforce the usability and reliability of an asynchronous BCI system for environmental control, indicating how these systems could be considered as input devices to interact with the external world and to restore the personal independence of people with severe motor disabilities.

CONCLUSIONS

An asynchronous P300-based BCI system has some advantages as compared with a synchronous modality, both in terms of usability and efficiency. First, it is able to automatically suspend the control when the user does not attend to the stimulation, therefore, avoiding the need for an explicit “pause” command the user should otherwise issue. Second, the system feature of “abstention” could avoid misclassification when the EEG feature is not sufficiently reliable. Finally, it can continuously adapt the time required for each classification to the changes of user state, finding an optimal trade-off between speed and accuracy.

In this study, we showed that this latter system feature might play a relevant role in enhancing system usability in a large population of potential end-users with high intra-subject variability (old in age with or without neurological disorders). They can benefit from the system’s adaptation to the number of stimuli needed to achieve correct classification. Furthermore, under this condition, the system’s robustness to false positives and misclassifications is maintained.

Although a definitive system validation requires further testing on a large sample size, the advantage in usability and efficiency of an asynchronous system paves the way towards the applicability of BCI systems as assistive device.

ACKNOWLEDGMENT

The work is partly supported by the EU grant FP7-224332 “SM4ALL” (Smart Homes for All an embedded middleware platform for pervasive and immersive environments for all) project. This paper only reflects the authors’ views and funding agencies are not liable for any use that may be made of the information contained herein. We also thank Francesco Leotta, Eirini Kaldeli, Alexander Lazovik and Ehsan Warriach for software development, and Dinie Sieswerda, Willy van der Brink, Allard Naber and Ehsan Warriach for their precious support during the experimentation. Finally, we are grateful to all the subjects that participated in the study.

DISCLOSURE AND CONFLICT OF INTEREST

F. Aloise, F. Schettini, P. Aricò, S. Salinari, C. Guger, J. Rinsma, M. Aiello, D. Mattia and F. Cincotti have no conflicts of interest in relation to this article.
REFERENCES


