P300-based brain–computer interface for environmental control: an asynchronous approach

This article has been downloaded from IOPscience. Please scroll down to see the full text article.
2011 J. Neural Eng. 8 025025
(http://iopscience.iop.org/1741-2552/8/2/025025)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 151.100.59.194
The article was downloaded on 24/03/2011 at 18:39

Please note that terms and conditions apply.
P300-based brain–computer interface for environmental control: an asynchronous approach

F Aloise\textsuperscript{1}, F Schettini\textsuperscript{1,2}, P Aricò\textsuperscript{1,2}, F Leotta\textsuperscript{1,2}, S Salinari\textsuperscript{2}, D Mattia\textsuperscript{1}, F Babiloni\textsuperscript{1,3} and F Cincotti\textsuperscript{1}

\textsuperscript{1}Neuroelectrical Imaging and BCI Lab, Fondazione Santa Lucia IRCCS, Rome, Italy
\textsuperscript{2}Deparment of Computer Science and Engineering, University of Rome ‘Sapienza’, Rome, Italy
\textsuperscript{3}Department of Human Physiology and Pharmacology, University of Rome ‘Sapienza’, Rome, Italy

E-mail: f.aloise@hsantalucia.it

Received 30 November 2010
Accepted for publication 20 January 2011
Published 24 March 2011
Online at stacks.iop.org/JNE/8/025025

Abstract

Brain–computer interface (BCI) systems allow people with severe motor disabilities to communicate and interact with the external world. The P300 potential is one of the most used control signals for EEG-based BCIs. Classic P300-based BCIs work in a synchronous mode; the synchronous control assumes that the user is constantly attending to the stimulation, and the number of stimulation sequences is fixed \textit{a priori}. This issue is an obstacle for the use of these systems in everyday life; users will be engaged in a continuous control state, their distractions will cause misclassification and the speed of selection will not take into account users’ current psychophysical condition. An efficient BCI system should be able to understand the user’s intentions from the ongoing EEG instead. Also, it has to refrain from making a selection when the user is engaged in a different activity and it should increase or decrease its speed of selection depending on the current user’s state. We addressed these issues by introducing an asynchronous BCI and tested its capabilities for effective environmental monitoring, involving 11 volunteers in three recording sessions. Results show that this BCI system can increase the bit rate during control periods while the system is proved to be very efficient in avoiding false negatives when the users are engaged in other tasks.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Brain–computer interfaces (BCIs) provide severely motor disabled people with an alternative channel to communicate and to control simple devices, independent of muscular activity (Wolpaw \textit{et al} 2002). Electroencephalographic (EEG)-based BCIs operate a real-time translation of predefined EEG features into commands that reflect the users’ intent (Farwell and Donchin 1988, Wolpaw and McFarland 2004).

Many of the currently available BCI applications exploit different classes of EEG features as control signals, such as sensorimotor rhythms (Furtscheller \textit{et al} 2006), steady-state visual evoked potentials (SSVEPs Middendorf \textit{et al} 2000) and event-related potentials (ERPs), such as the P300 potential (Donchin and Smith 1970). The P300 potential is typically a large and positive deflection in the EEG activity that reaches a maximum amplitude (about 10–20 μV) over the centro-parietal scalp areas and occurs 250 to 400 ms after a relevant stimulus (Target stimulus) presented within a train of frequent stimuli (No-Target stimuli) is recognized (Fabiani \textit{et al} 1987, Polich and Kok 1995). As such, the P300-based BCI requires a well-defined number of stimulation repetitions to achieve a selection, causing obvious drawbacks. This is the reason for the growing interest regarding the issue of asynchronous (self-paced) BCI control design. In fact, transferring a BCI from the laboratory environment to real-life settings requires a BCI system to recognize the user’s intent without any additional external input (Vaughan \textit{et al} 2006). Moreover, both the user’s
psychophysical state (mood, fatigue, motivation; Nijboer et al 2010) and external factors (environmental conditions, stimulation modality, etc; Allison and Pineda 2003) can affect P300 morphology, and thus influence performance. In other words, the system should be able to recognize when the user intends to exercise his control and then to correctly identify the desired command. Several studies have addressed the issue of the asynchronous BCI in the domain of sensorimotor rhythms (Mason and Birch 2000, Millan and Mouríño 2003, Townsend et al 2004). Recently, Zhang et al (2008) proposed a first computational approach to implement an asynchronous mode of control for the P300-based BCI. Using statistical and probability models about control and no-control user’s state, they developed an algorithm that first recognizes the control state and then looks for the target, giving a classification after at least three stimulation sequences, as a result. They used a stimulation interface containing nine numeric items, and stimuli were provided, intensifying each single button.

In this work, we propose an alternative asynchronous P300-based BCI developed by following a heuristic approach. Stimuli were provided by intensifying rows and columns in a matrix, just like in the classical P300 Speller (Farwell and Donchin 1988). The primary goal of this study was to compare our asynchronous BCI system with a classical synchronous system in terms of bit rate and accuracy. At the same time, we wanted to assess the asynchronous system from its usability and robustness point of view when the user was engaged in real-life activities. To achieve these aims, we first collected and processed data about the users’ Control and No-Control tasks and we introduced some thresholds in the on-line classifier that allowed asynchronous control modality. Secondly, we tested the asynchronous control modality to control a domotic appliance in a real-life setting where a stronger system is required to control errors than traditional systems used for communication only (Cincotti et al 2008).

2. Materials and methods

2.1. Study design and participants

Eleven healthy volunteers (four women, seven men; mean age and std 26.45 ± 4.05 years) were involved in the study. Five out of 11 subjects were naive to the P300-based BCI context.

The acquisition protocol was based on the P3Speller application (Farwell and Donchin 1988) within the BCI2000 framework (Schalk et al 2004). This application was adapted to control a home automation system by using a $4 \times 4$ matrix that allowed a total of $N_t = 8$ stimulation classes. As shown in figure 1, the stimulation interface consisted of 16 black and white icons representing the achievable actions on the environment (Aloise et al 2008). An asynchronous operating mode was implemented within the BCI2000 framework (see below). The fixation cross was static in the middle of the interface and never flashing, also during stimulation.

Scalp EEG potentials were recorded (g.MobiLab, gTec, Austria, sampling rate 256 Hz) from eight positions according to 10–10 standard (Fz, Cz, Pz, Oz, P3, P4, P07 and P08; Krusienski et al 2008). Each channel was referenced to the linked earlobes and grounded to the left mastoid. Stimulation was provided through a 22” LCD monitor; on one half of the screen there was the stimulation matrix, while on the other half there were movies played on a DVD player.

2.2. Data acquisition protocol

Stimuli were provided by the BCI2000 framework through the random intensification of rows and columns in the matrix. Each stimulus was intensified for 125 ms; the inter-stimulus interval (ISI) was set at 125 ms, so that the stimulus onset asynchrony (SOA) interval lasted 250 ms. The EEG signal was reorganized in overlapping epochs lasting 800 ms and following the onset of each stimulus. Epochs were then grouped into sequences. A Sequence consisted in a single flash relative to each row and column on the control interface. A set of sequences in which the target icon was the same composed a Trial. Finally, a Run comprised a trial series.

In the synchronous system after a number of stimulation sequences fixed a priori, a selection was always made. In contrast, the asynchronous mode allowed a selection only when the thresholds were exceeded. When the given thresholds were not reached after a fixed number of stimulation sequences, a new trial began with no selection occurring. During experimentation, the system indicated the Target icon for the next trial, or presented the classification result within the 4 s between two trials. All subjects underwent a total of three recording sessions over 2 weeks (3 or 4 days elapsed between two sessions). We defined the first two sessions as Training sessions and the last one as an On-line session.

2.2.1. Training sessions. The aim of the first two recording sessions was to collect data for the off-line analysis and to extract parameters (features and thresholds) for the On-line session.

Figure 2 illustrates the Training session composition. Each session was composed of eight runs. The first two runs were defined as Control runs and they consisted of eight
Control trials of ten stimulation sequences each. During the Control runs, the subject had to mentally count the occurrences of the Target icon that was cued by its intensification before the beginning of each trial. Throughout the training sessions, all icons were presented as a Target in random order spanning each position on the matrix. We used this dataset both for offline analysis of the synchronous system and to extract control parameters for synchronous control runs in the On-line session. During the Alternate runs, data were acquired under Control and No-Control conditions. In these runs, Control and No-Control trials alternated for a total of ten trials composed of ten stimulation sequences each. During the Control task, the subjects were asked to count Target icon flashing (as in the first two runs of the session). We chose not to set any control parameters during Control trials and thus we did not provide any feedback about the on-line classification results. As for the No-Control condition, subjects attended three different tasks while the stimulation (icon flashing) was running.

- Fixation Cross, Training Session 1, Alternate Runs [3–8]: subjects were instructed to fixate the cross in the center and to ignore the stimulation.
- Watch & Listen, Training Session 2, Alternate Runs [3–5]: subjects were instructed to watch a movie displayed on the half of the screen beside the matrix.
- Computation, Training Session 2, Alternate Runs [6–8]: subjects had to answer simple arithmetic questions posed by the operator while fixating the cross.

By doing so, we ensured that the No-Control dataset would contain trials mimicking a real-life situation wherein the users could direct their attention elsewhere or could interact with other persons. Furthermore, under the No-Control experimental condition, the user’s visual field was not immune to the random stimuli, thus allowing us to test the system robustness to artifacts and to reduce potential visual misclassification.

Each icon on the interface was presented as a Target with the same frequency, both for Control and Alternate runs, over the two training sessions, with the aim of making all icons equally likely for subsequent analysis. Only one icon (the 16th) was never suggested as a Target during Alternate runs because we used it to indicate the No-Control trials to the users.

2.2.2. On-line session. Figure 3 illustrates the adopted scheme for the On-line sessions.

The first four Control runs were performed in order to compare the synchronous versus asynchronous system in terms of the time required to perform a previously defined list of actions. These actions were selected in order to test each device in the environment. The goal of each run was to complete five actions; the Target icons were cued at the beginning of each run and the subjects were also informed about how to correct each potential error by selecting the complementary action. The number of trials was not fixed a priori but depended on a number of errors (synchronous and asynchronous modalities) and abstentions (asynchronous modality). Classification results were shown to the subjects by intensifying a single icon, and through the corresponding device operation. Abstentions with the asynchronous system were fed back by intensifying the icons all together on the interface. The purpose of the two No-Control runs was to assess system reliability avoiding false positives. Each No-Control run took 5 min, during which the stimulation was kept on. During the first and second No-Control runs, the subjects were asked to refrain from the control by watching a movie or by answering arithmetic questions while looking at the fixation cross, respectively.
Finally, the last on-line runs were devoted to test the usefulness of the asynchronous BCI in everyday life. We simulated two different scenarios and we quantified error and false positive occurrences.

- **Scenario 1**, On-line session, Real Life [7]: someone rings the doorbell, thus the user turns on the interphone’s video camera and waits until the image appears (about 30 s). The guest is one of his friends. He decides to open the door and turn on the light for his guest. Subsequently, he turns on the DVD and they watch a video together. After 1 min the user turns off the DVD player (five BCI commands).

- **Scenario 2**, On-line session, Real Life [8]: the user is tired and wants to relax; he lowers the chair and turns off the light. After 1 min the phone rings and he answers. It is his sister and he talks with her for 1 min, then he hangs up, turns on the fan and goes to sleep (five BCI commands).

### 2.3. EEG preprocessing, feature extraction and classification

The EEG signal is divided into 800 ms epochs starting from the onset of each stimulus. We can distinguish Target and No-Target epochs related to Control trials, and No-Control epochs related to No-Control trials. EEG epochs were subsequently reorganized into a three-dimensional (3D) array; each 2D matrix of the array represents a single epoch, where rows stand for acquisition channels and columns correspond to samples of each epoch. Despite sample downsampling (by a factor of 3), the amount of data were still demanding and a further reduction in feature space was performed by using the stepwise linear discriminant analysis (SWLDA; Krusienski et al 2006). We divided the dataset related to Alternate trials into two parts: training and testing. The first contains three runs from the first training session and four runs from the second one. In this way, we included in the training dataset No-Control trials related to all three different No-Control tasks. We used the same runs for each subject to extract thresholds and features for asynchronous runs in the On-line session.

- **Training session 1**: Runs 3–5, Fixation Cross
- **Training session 2**: Runs 3–4, View & Listen; Runs 6–7, Computation.

We ran SWLDA on the training dataset including No-Control trials, assigning a label equal to 0 to the No-Target and No-Control epochs while the label was equal to 1 for Target epochs. For the synchronous system, we used all Control runs of the training sessions to extract the features for the On-line session, while for the Off-line analysis we performed a cross-validation, using two runs to extract significant features and another two to test them. Finally, for both asynchronous and synchronous modalities, the final discriminant function was restricted to contain a maximum of 60 features by SWLDA. Nonzero weights were assigned to these features, \( w \). The score values for each epoch were then calculated as

\[
y_i = \sum_e w \cdot f_e + b, \tag{1}
\]

where \( e \) denotes all features related to single stimulus \( i \) and \( b \) represents the intercept. For the classification, it is assumed that a P300 is elicited for one of the four row/column intensifications during the control period, and that the P300 response is invariant to row/column stimuli, the resultant classification in the synchronous system taken as the maximum of the scored feature vectors for the respective rows, as well as for the columns. The icon that appears at the intersection of the predicted row and column in the matrix is the one chosen.

### 2.4. Extraction of threshold values

The threshold values were chosen through a procedure that relies on the use of ROC curves (Schulzer 1994). We
calculated the score values on the Target, No-Target and No-Control epochs using data from Alternate runs: a training dataset to extract features and testing dataset to estimate score value, as explained in the previous section. Figure 4 shows the trend of these three normalized distributions for a representative subject.

As can be seen in figure 4, a normal distribution fits the score distributions well. In fact, we ran a $t$-test on the three different score distributions for each subject, and its results showed that the hypothesis of the normal distribution is true with a 95% confidence level.

Next, we performed the Kolmogorov–Smirnov test on each pair of samples. The values of the statistic test and the corresponding $p$-value are reported in table 1. The hypothesis of different distributions was confirmed with a 95% confidence level for all subjects except for subject 1. For this reason, it is necessary to take into consideration the No-Control trials to estimate the control parameters and thresholds.

The threshold values were chosen according to the number of stimulation sequences accumulated in the trial. In fact, the scores for the general stimulus $i$ at the sequence $s$ will be defined as

$$y_{acc}^i = \sum_{n=1}^{s} y_{n}^i, \quad i = 1, 2, \ldots, N_s,$$  \hfill (2)

where $y_{n}^i$ is given by (1) for the sequence $n$, $N_s$ is the number of stimulation classes for the domotic interface ($N_s = 8$), and the maximum number of stimulation sequences in a trial was fixed to 10. Figure 5 shows the box plot of score distributions earned on the basis of the accumulated sequence. The greater the number of accumulated sequences, the more the Target distribution deviates from the No-Target and no-control distributions, while the difference between No-Target and No-Control remains much the same with the addition of new sequences.

Afterwards, we looked for the maximum score of the row stimuli and the maximum score of the column stimuli for each sequence; then we assigned to them a label equal to 1 if the maximum scores were relative to a target stimulus and equal to 0 if they referred to No-Target or No-Control stimuli. In this way, we were sure to include the maximum score values.
Figure 6. Threshold extraction process. On the right are ROC curves plotted as a function of the number of sequence elapsed. The ROC curves show a trend closer to the ideal when the number of sequences accumulated in the Trial increases. The box on the left magnifies the area representing the chosen tradeoff between FPR and TPR values. Threshold values are taken at the intersection of the ROC curve to the straight line joining points (0, 1) and (0.05, 0.5).

Table 1. Kolmogorov–Smirnov’s test values comparing pairwise distribution of Target, No-Target and No-Control score distributions. The hypothesis of different distributions was confirmed with the 95% confidence level for all of the subjects except for subject 1.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Target versus No-Target</th>
<th>Target versus No-Control</th>
<th>No-Target-No Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ks-value</td>
<td>p-value</td>
<td>ks-value</td>
</tr>
<tr>
<td>SUBJ 1</td>
<td>0.64</td>
<td>&lt;0.001</td>
<td>0.63</td>
</tr>
<tr>
<td>SUBJ 2</td>
<td>0.58</td>
<td>&lt;0.001</td>
<td>0.53</td>
</tr>
<tr>
<td>SUBJ 3</td>
<td>0.74</td>
<td>&lt;0.001</td>
<td>0.65</td>
</tr>
<tr>
<td>SUBJ 4</td>
<td>0.68</td>
<td>&lt;0.001</td>
<td>0.61</td>
</tr>
<tr>
<td>SUBJ 5</td>
<td>0.66</td>
<td>&lt;0.001</td>
<td>0.57</td>
</tr>
<tr>
<td>SUBJ 6</td>
<td>0.74</td>
<td>&lt;0.001</td>
<td>0.65</td>
</tr>
<tr>
<td>SUBJ 7</td>
<td>0.78</td>
<td>&lt;0.001</td>
<td>0.70</td>
</tr>
<tr>
<td>SUBJ 8</td>
<td>0.62</td>
<td>&lt;0.001</td>
<td>0.52</td>
</tr>
<tr>
<td>SUBJ 9</td>
<td>0.57</td>
<td>&lt;0.001</td>
<td>0.47</td>
</tr>
<tr>
<td>SUBJ 10</td>
<td>0.62</td>
<td>&lt;0.001</td>
<td>0.51</td>
</tr>
<tr>
<td>SUBJ 11</td>
<td>0.67</td>
<td>&lt;0.001</td>
<td>0.67</td>
</tr>
</tbody>
</table>

related to No-Control trials in ROC curve training, so that threshold values took into account possible artifacts that could occur when the subject was not engaged in BCI control. From this point on, ROC curves could be plotted for each sequence using the corresponding scores. In figure 6, there is an example of the ROC curve trend. It is evident that when the number of elapsed sequences in the trial increases, the ROC curves assume an ideal tendency. To choose the threshold value, it is necessary to find a tradeoff between the false positive rate (FPR) and true positive rate (TPR). We have chosen to set the maximum FPR to 0.05 and the lowest TPR to 0.5, so that the threshold will be chosen at the intersection of the ROC curve with the straight line joining points (0, 1) and (0.05, 0.5). The choice of these values was based on empirical considerations. The aim was to control home automation, so we preferred the specificity (low FPR) respecting the sensitivity. In any case, as can be seen from figure 6, this affects threshold values only for the first sequences in the trial, because as the number of sequences accumulated grows, the ROC curves rapidly tend to the ideal trend.

We modified the classification process in the BCI2000 framework. The score values were computed at each new sequence and accumulated to the previous ones in the current trial. Threshold values were related to the number of elapsed sequences in the trial. At the end of each sequence, the maximum row and column values were compared to the specific threshold. If the threshold exceeded simultaneously because of the maximum row and column values, the system classified the icon at their intersection. Conversely, if the threshold values did not exceed throughout the maximum number of stimulation sequence fixed a priori (reset value), the system refrained from making a selection.
Figure 7. Off-line cross-validation for the synchronous system using data from Control trials in the training sessions. Accuracy values are based on the number of stimulation sequences averaged for each subject.

Table 2. Number of stimulation sequences needed to exceed thresholds in the asynchronous modality.

<table>
<thead>
<tr>
<th>NSeq</th>
<th>SUBJ 1</th>
<th>SUBJ 2</th>
<th>SUBJ 3</th>
<th>SUBJ 4</th>
<th>SUBJ 5</th>
<th>SUBJ 6</th>
<th>SUBJ 7</th>
<th>SUBJ 8</th>
<th>SUBJ 9</th>
<th>SUBJ 10</th>
<th>SUBJ 11</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Avg</td>
<td>3.61</td>
<td>4.63</td>
<td>3.08</td>
<td>3.95</td>
<td>3.52</td>
<td>4.70</td>
<td>3.31</td>
<td>4.80</td>
<td>5.56</td>
<td>4.42</td>
<td>3.31</td>
<td>4.08</td>
</tr>
<tr>
<td>Std</td>
<td>2.07</td>
<td>2.31</td>
<td>1.78</td>
<td>2.17</td>
<td>1.98</td>
<td>2.69</td>
<td>1.56</td>
<td>2.12</td>
<td>1.97</td>
<td>2.15</td>
<td>1.72</td>
<td>0.79</td>
</tr>
</tbody>
</table>

* Standard deviation of inter-subject mean values.

3. Results

3.1. Synchronous system off-line performances

The Control runs data from training sessions were used to evaluate synchronous system performances. We operated a cross-validation using two runs to extract significant features and the other two to test them; we then averaged the results of classification for each possible combination of training and testing dataset (a six-round cross-validation). The trials used for both train and test were almost half of the ones used for the asynchronous system, even if 16 trials were enough to train SWLDA for the P300-based BCI (Guger et al 2009). Instead, a greater amount of data was necessary to achieve good resolution in ROC curve plotting; this is why the asynchronous dataset was larger than the synchronous one.

Figure 7 shows the trend of the percentages of correct classification based on the number of stimulation sequences averaged for each subject. The black dotted line represents an accuracy of 95% according to the value of false positives set in the asynchronous system through ROC curves.

These results were used to set the maximum number of sequences within the synchronous runs and the reset value for the asynchronous runs during the On-line session. In the synchronous mode, the number of sequences equal to the number of stimulation sequences needed to reach 95% of accuracy has been set. If it failed, the maximum number of sequences was left to 10. For the asynchronous system, the reset value was chosen according to the maximum number of sequences needed to make a selection during Control periods (see table 2). At the same time, the accuracy values and the number of sequences for both synchronous and asynchronous systems were used for bit-rate valuation in section 3.3.

3.2. Asynchronous system off-line performances

An off-line six-round cross-validation was performed on the data acquired during the Alternate trials in the training session. We used six different training datasets to extract features defining the threshold value (three Fixation Cross, two Watch & Listen and two Computation Alternate runs), and the complementary testing datasets to assess off-line performances for the asynchronous system. We performed a six-round cross-validation in order to match the maximum number of rounds achievable with the Control Runs data. Depending on the user’s state (Control or No-Control), there could be five different classification outcomes.

Control state classification outcomes:
- Correct classification: the target was correctly recognized.
- Wrong classification: there was a target misclassification.
- Missed classification: the thresholds were never exceeded, and the system refrained from taking a decision.

No-Control state classification outcomes:
- Abstention: the system properly refrained from taking decisions.
- Missed abstention: the thresholds were exceeded and the system made a wrong choice.

Figure 8 reports cross-validation results: it can be seen how the system was proved to be robust in avoiding false positives during No-Control trials, in fact abstentions reached
Figure 8. Results of off-line cross-validation for the asynchronous system using data from Alternate sessions. Results refer to Control and No-control trials. Vertical bars on the mean values denote the standard deviation of subjects’ performance. The system demonstrated high reliability during No-Control trials (abstention mean = 98.91%) and at the same time on average 88.73% of Control Trials were correctly classified. The error bars on the mean values denote the inter-subject variability.

Table 3. Information transfer rate (bit min$^{-1}$) for each subject evaluated for the synchronous system and for the asynchronous system. These values refer to off-line analysis.

<table>
<thead>
<tr>
<th></th>
<th>SUBJ 1</th>
<th>SUBJ 2</th>
<th>SUBJ 3</th>
<th>SUBJ 4</th>
<th>SUBJ 5</th>
<th>SUBJ 6</th>
<th>SUBJ 7</th>
<th>SUBJ 8</th>
<th>SUBJ 9</th>
<th>SUBJ 10</th>
<th>SUBJ 11</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYNC</td>
<td>12.39</td>
<td>4.66</td>
<td>16.81</td>
<td>12.01</td>
<td>11.03</td>
<td>16.81</td>
<td>14.01</td>
<td>5.48</td>
<td>3.94</td>
<td>5.63</td>
<td>11.34</td>
<td>10.37</td>
</tr>
</tbody>
</table>

on average 98.61%. On the other hand, the 11.39% on average of missed classification represented the system’s ability to avoid misclassification, because the percentage of wrong classification did not exceed an average of 1.5%. Table 2 reports the maximum, mean and standard deviation values of the number of stimulation sequences needed for each subject to exceed thresholds during Control trials.

3.3. Information transfer rate

In order to assess the efficiency of the two systems in terms of an information transfer rate, we used the definition of bit rate given by Wolpaw et al. (2000) and widely used in BCI community. The latter is based on the definition of information rate proposed by Shannon for noisy channels with some simplifying assumptions; all of the symbols have the same $a$ priori occurrence probability $p = 1/N_s$, the classifier accuracy $P$ is the same for all target symbols and the classification error $1 - P$ is equally distributed amongst all of the remaining symbols:

$$B_{\text{wolpaw}} = \log_2 N_s + P \log_2 P + (1 - P) \log_2 \left[ \frac{(1 - P)}{N_s - 1} \right].$$

This expresses the bit rate or bit/trial for each selection. The information transfer rate (bits per minute) is equal to $B_{\text{wolpaw}}$ multiplied by speed of selection $S$ (selection per minute). In turn, the speed selection for the P300-based system depends on the number of stimulation sequences used, and when we calculated it we took into consideration the 4 s between the two Trials. Table 3 shows the values of the information transfer rate for each subject calculated, using the number of sequences and the percentage of accuracy obtained by off-line analysis. For the asynchronous system, we considered only the results from the Control trials. We ran a $t$-test to evaluate the differences between the two distributions. It did not show any statistical significance between the two distributions ($t = -0.81$, $p$-value = 0.62); however, the asynchronous system exhibited on average an information transfer rate higher than the synchronous one.

3.4. On-line results

As mentioned before, during the On-line session, subjects were asked to manage some devices using the BCI in both synchronous and asynchronous modes.

Figure 9 reports the total time needed by each subject to complete the two Control runs. Results are on average consistent with off-line bit rate values but some subjects (3, 4 and 8) exhibited different performances with respect to those expected from the off-line analysis. The $t$-test does not present a significant difference between the two distributions ($p$-value = 0.74; $t = 0.33$).

Regarding the two No-Control trials in the On-line session, on average we detected 0.26 false positive min$^{-1}$ (std = 0.4). Furthermore, each subject was able to complete the real-life runs. On average, over the two scenarios, there...
Figure 9. Results for comparison in the On-line sessions. The time needed to complete ten different actions with the synchronous and the asynchronous systems is reported. Values refer to the total time needed to complete Control runs 1–2 for the synchronous system, and Control runs 3–4 for the asynchronous system.

were 1.73 false positives (std = 2.14) during No-Control actions, while during Control actions the asynchronous system achieved on average 87.46% (std = 13.34%) of correct classifications.

4. Discussion

Our preliminary findings indicated that the asynchronous mode exhibited a useful level of reliability in avoiding false positives when subjects were engaged in different tasks, in both the off-line and the on-line computations (mean = 0.26 false positive min$^{-1}$). Although this level of reliability awaits a larger group testing, the approach proposed in this study is promising for the increment of robustness as compared to what has been reported previously by Zhang et al (2008). In fact, they reported on average 0.71 false positive min$^{-1}$. From the bit rate point of view, they achieved a mean value of 15 bit min$^{-1}$, which appears higher than that shown with our approach (mean = 11.19 bit min$^{-1}$). It should be considered, however, that a direct comparison between the present study and Zhang et al’s is difficult due to the differences in the stimulation modalities and classification approach. Furthermore, we calibrated our system to control home automation, preferring reliability rather than speed.

Although the synchronous and asynchronous modalities of BCI operation did not statistically differ in terms of speed, the asynchronous mode allowed for a higher information transfer rate with respect to the synchronous operating mode (mean = 10.37 bit min$^{-1}$). These results of a more reliable and stable BCI control under asynchronous mode are encouraging in boosting the benefit of a real-life usage system that does not engage the user in a continuous control task.

The introduction of a threshold-based classification approach might allow the user to divert her/his attention from the control interface at any time and without the use of external inputs. A further advantage consists in increasing the accuracy of the system; an asynchronous BCI may prevent an error through abstentions that in bit-rate terms is less than an error. Altogether, these features would allow for a more independent use of a BCI system by people with severe disabilities. The findings obtained during the scenarios suggest that the presented approach would be suitable to operate in a real-life context. In this regard, further improvements can be made in terms of efficiency; for instance, the threshold values might be updated, depending on the most likely user’s choice or on the current task. Further development could include the introduction of an algorithm that allows the system to automatically recalibrate the control parameters and threshold values based on the results of the current classification.

Acknowledgments

This work was partly supported by the EU grants FP7-224332 ‘SM4ALL’ (smart homes for all, an embedded middleware platform for pervasive and immersive environments for all) project, and FP7-224631 ‘TOBI’ (tools for brain–computer interaction) 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.

References


brain–computer interface system: towards its application as assistive technology Brain Res. Bull. 75 796–803
Donchin E and Smith D B 1970 The contingent negative variation and the late positive wave of the average evoked potential Electroencephalogr. Clin. Neurophysiol. 29 201–3
Fabiani M, Gratton G, Karis D and Donchin E 1987 Definition, identification and reliability of measurement of the P300 component of the event-related brain potential Adv. Psychophysiol. 2 1–78