Neuroscience Letters 462 (2009) 94–98

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

How many people are able to control a P300-based brain–computer interface (BCI)?

Christoph Guger [a,*], Shahab Daban [a,b], Eric Sellers [d], Clemens Holzner [a], Gunther Krausz [a], Roberta Carabalona [c], Furio Gramatica [c], Guenter Edlinger [a]

[a] g.tec Medical Engineering GmbH, Herbersteinstrasse 60, 8020 Graz, Austria
[b] University of Applied Sciences Upper Austria, Linz, Austria
[c] Biomedical Technology Department, Santa Maria Nascente Research Hospital, Don Gnocchi Foundation, Milan, Italy
[d] East Tennessee State University, Johnson City, TN, USA

ARTICLE INFO

Article history:
Received 9 April 2009
Received in revised form 16 June 2009
Accepted 16 June 2009

Keywords:
Brain–computer interface (BCI)
Electroencephalogram (EEG)
P300 potential
Rehabilitation
Spelling device

ABSTRACT

An EEG-based brain–computer system can be used to control external devices such as computers, wheelchairs or Virtual Environments. One of the most important applications is a spelling device to aid severely disabled individuals with communication, for example people disabled by amyotrophic lateral sclerosis (ALS). P300-based BCI systems are optimal for spelling characters with high speed and accuracy, as compared to other BCI paradigms such as motor imagery. In this study, 100 subjects tested a P300-based BCI system to spell a 5-character word with only 5 min of training. EEG data were acquired while the subject looked at a 36-character matrix to spell the word WATER. Two different versions of the P300 speller were used: (i) the row/column speller (RC) that flashes an entire column or row of characters and (ii) a single character speller (SC) that flashes each character individually. The subjects were free to decide which version to test. Nineteen subjects opted to test both versions. The BCI system classifier was trained on the data collected for the word WATER. During the real-time phase of the experiment, the subject spelled the word LUCAS, and was provided with the classifier selection accuracy after each of the five letters. Additionally, subjects filled out a questionnaire about age, sex, education, sleep duration, working duration, cigarette consumption, coffee consumption, and level of disturbance that the flashing characters produced. 72.8% ($N = 81$) of the subjects were able to spell with 100% accuracy in the RC paradigm and 55.3% ($N = 38$) of the subjects spelled with 100% accuracy in the SC paradigm. Less than 3% of the subjects did not spell any character correctly. People who slept less than 8 h performed significantly better than other subjects. Sex, education, working duration, and cigarette and coffee consumption were not statistically related to differences in accuracy. The disturbance of the flashing characters was rated with a median score of 1 on a scale from 1 to 5 (1, not disturbing; 5, highly disturbing). This study shows that high spelling accuracy can be achieved with the P300 BCI system using approximately 5 min of training data for a large number of non-disabled subjects, and that the RC paradigm is superior to the SC paradigm. 89% of the 81 RC subjects were able to spell with accuracy 80–100%. A similar study using a motor imagery BCI with 99 subjects showed that only 19% of the subjects were able to achieve accuracy of 80–100%. These large differences in accuracy suggest that with limited amounts of training data the P300-based BCI is superior to the motor imagery BCI. Overall, these results are very encouraging and a similar study should be conducted with subjects who have ALS to determine if their accuracy levels are similar.

A brain–computer interface (BCI) allows people to use electroencephalographic (EEG) activity to control external devices such as robots, virtual environments, or spelling devices [6,16]. It is necessary to train BCI systems on subject-specific EEG data before real-time BCI use is possible. Depending on the type of BCI being used, the amount of training time can vary from minutes to hours. Several different EEG signals can be used for BCI control. For example, slow cortical potentials [1], oscillations in alpha and beta range [6,11], steady-state visual evoked potentials (SSVEP) [2,17] and the P300 event-related potential [13] have all been used successfully for BCI control. BCI systems based on oscillations use mostly motor imagery strategies to generate event-related desynchronization (ERD) and event-related synchronization (ERS) in the alpha and beta frequency ranges of the EEG [9]. This type of BCI is

* Corresponding author at: Sierningstrasse 14, 4521 Schiedlberg, Austria. Tel.: +43 7251 22240 12; fax: +43 7251 22240 39.
E-mail address: guger@gtec.at (C. Guger).
URL: http://www.gtec.at/ (C. Guger).

0304-3940/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.neulet.2009.06.045

mainly used for cursor control on computer screens, for navigation of wheelchairs or in virtual environments [11]. Typically, different motor imagery techniques such as right/left hand movement, foot movement, tongue movement and/or mental counting are used. SSVEP-based BCI systems use flickering lights that induce EEG oscillations with the same frequency as the stimulation source. SSVEP systems have mostly 4–12 different stimulation frequencies to realize, e.g., robot control or mobile phone control [8]. The P300-based BCI systems use the effect that an unlikely event induces the P300 component in the EEG. Such systems are mostly used as spelling devices because a high number of different characters to choose from can enhance the communication speed of the BCI [7].

In the present study, two different paradigms were used to implement the P300 speller: (i) the row/column (RC) speller highlights multiple characters at once and (ii) the single character (SC) speller flashes each character individually. Therefore a higher P300 amplitude and more reliable control should be expected with the SC flasher because it is more unlikely that the target character appears. This was also tested by Guan et al. [3], and the results suggested that the SC paradigm produced larger P300 responses. Sellers found that a 3×3 matrix had higher accuracy than a 6×6 matrix, but a lower communication rate [13]. With an inter-stimulus interval (ISI) of 175 ms and a 3×3 matrix Sellers achieved an accuracy of 88% in the best case [13]. In addition [13] showed that the 6×6 matrix produced larger P300 amplitudes, presumably because the target item was less likely to appear. In a study that included six subjects, Serby et al. [15] achieved mean online accuracy of 79.5%. Sellers et al. [14] showed that accuracy using the standard RC method and a random grouping method of characters both achieved accuracy levels around 90%. Working with four subjects disabled by amyotrophic lateral sclerosis (ALS) Nijboer reported online accuracy of 79% [10]. In addition to using characters and numbers [5,12] showed that images can be used for the P300-based BCI.

In all of these studies different parameters such as matrix size, ISI, type of speller, and other factors were manipulated. Moreover, each study used a relatively small sample size. The primary focus of the current study is to hold all of these variables constant and test a P300-based BCI on a large population ($N = 100$) that is not pre-selected. The current study used a very small amount of calibration data, only 5 min, compared with most other studies [e.g., 7, 10, 13–15]. This allowed us to complete many more sessions in a short amount of time.

In a previous study, we tested a motor imagery-based BCI system with 99 subjects visiting an exhibition in Graz [4]. The subjects were trained for 6 min to imagine left or right hand movement for a few seconds (20 times each) to produce ERD and ERS changes. The BCI system was then trained on the individual EEG data for a subsequent session with visual feedback of cursor movement. The subjects were able to control the cursor to the left or right side of the computer screen. 6.2% of the subjects were able to learn this control with 100% accuracy in this short training session. About 93.3% showed a control above 59% accuracy (50% corresponds to random classification).

Based on experience with the motor imagery study, we sought to replicate the design using a P300-based BCI. Furthermore, we compared SC and RC spellers to examine if one paradigm produces higher rates of accuracy than the other, performance of female and male participants was compared, and subjects were required to complete a short questionnaire about education, sleeping duration, working duration, cigarette and coffee consumption, and disturbance created by the highlighting of the P300 speller characters.

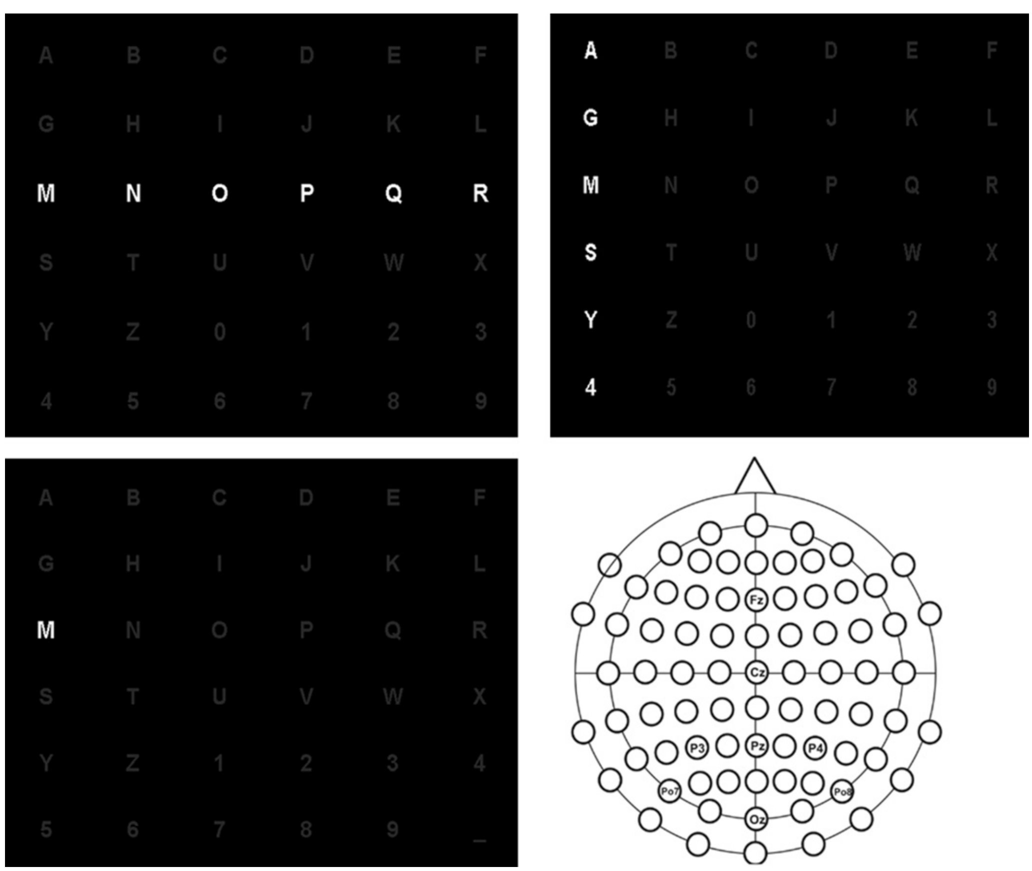

Fig. 1. (Top) The RC speller highlights a whole row or column at once with 6 symbols each. (Bottom, left) The SC speller highlights each character individually. (Bottom, right) Electrode positions according to the international 10/20 electrode system used for the EEG measurements. The head is viewed from above and the nose points to the top of the page.

100 subjects (32 female, 68 male, age: 27.9 ± 10.9) participated in the study. Data were acquired by g.tec Medical Engineering GmbH, Austria (N = 87) and Fondazione Don Gnocchi, Italy (N = 13). The subjects were free of medication and central nervous system abnormalities and had no prior experience with EEG-based communication systems. The experiment was conducted over a period of 6 months.

Subjects sat in front of a laptop computer and were instructed to relax and remain as still as possible. Fig. 1 shows the electrode configuration and the eight electrode locations used for the study. The EEG data were acquired using a g.USBamp (24 Bit biosignal amplification unit, g.tec Medical Engineering GmbH, Austria) at a sampling frequency of 256 Hz. The ground electrode was located on the forehead; the reference was mounted on the right earlobe. EEG electrodes were made of gold or sintered Ag/AgCl material. The first 75 subject's data were recorded with passive electrodes (gold). The final 25 subject's data were recorded with active electrodes (Ag/AgCl); that is, an amplifier was encased within each electrode.

Subjects were given the option to participate in one or both of the spelling experiments: (i) the RC speller and/or the (ii) SC speller. The two paradigms are shown in Fig. 1. Both spellers show 36 characters (A, B, . . ., Z; 0, 1, . . ., 9) on the computer screen. The RC speller highlights a whole column or row for 100 ms while the SC speller highlights each character individually for 60 ms. Between the flashes there is a short time while only the grey matrix items are visible (RC: 60 ms; SC: 40 ms). The subject's task is to attend to (or look at) the character he/she is prompted to spell and count how many times the character is highlighted. Counting is used to help the subject remain focused on the task. Initially, calibration data must be collected for each subject. Therefore, each subject was asked to "select" (or attend to) the word WATER, one letter at a time, without feedback from the BCI. This procedure lasted approximately 5 min. The calibration data are then processed using a linear discriminant analysis (LDA; described below) to derive the EEG weighting parameters. After deriving the LDA, the subject was asked to write the word LUCAS, one character at a time, taking approximately 5 additional minutes. After 15 highlights of each character or row/column, the signal processing unit applied the LDA to the EEG response of each row and column in the RC condition, and to the EEG response of each character in the SC condition. After each character, the BCI provided feedback to the subject indicating whether or not the classifier was correct.

The last 35 subjects filled out a questionnaire about age, sex, education, sleep duration, working duration, cigarette consumption, coffee consumption and disturbance of flashes.

The Simulink model shown in Fig. 2 is used for real-time analysis of the EEG data. The g.USBamp block reads in the EEG of the 8 channels (shown in Fig. 1) with 256 Hz sampling frequency. Then the data are converted to double precision, bandpass filtered between 0.5 and 30 Hz and down-sampled to 64 Hz. The signal processing block buffers the EEG data of each flash and performs the linear discriminant analysis (LDA) using 800 ms of the EEG signal beginning at the onset of highlighted stimulus. Thus, offline (after the calibration data have been collected) single trial epochs of 800 ms are used to derive the LDA weighting coefficients. All time × location features are entered into the LDA. The LDA weights each input parameters according to its importance. The classification result, the sum of the weighted parameters, indicates the class (target or non-target stimuli) to which the character belongs. Hence, the LDA classifier selects the character having the highest sum of the weighted parameters and presents that character on the computer screen as feedback to the subject. Then the process starts again with the next character.

The RowCol Character Speller block (Fig. 2) controls highlighting of the characters and the experiment. The Scope and To File blocks are used to visualize the EEG and to store the data for subsequent off-line analyses. An important point is that the g.USBamp represents the hard real-time clock and drives the whole model with a hardware interrupt. This ensures that the model and therefore all blocks are updated every 1/256 s.

The RowCol Character Speller block (or the Single Character Speller block for the other version) controls the experiment and highlights the corresponding rows and columns randomly. It also sends the ID of the flashing character to the signal processing block. The signal processing block generates a buffer for each character and stores the incoming EEG data around the flash (800 ms epoch). This is done until all 12 RC buffers or all 36 SC buffers are filled with 15 epochs (15 highlights of each character). Finally, the LDA [6] is applied to the EEG data to determine the selected character.

The results of all 100 subjects that participated in the recordings are shown in Table 1. A total of 81 subjects used the RC speller and 38 subjects the SC speller. 19 subjects tested both versions. The most important result is that 72.8% of all subjects were able to control the RC speller and 55.3% of the subjects were able to control the SC speller with 100% accuracy (i.e., all 5 characters of LUCAS

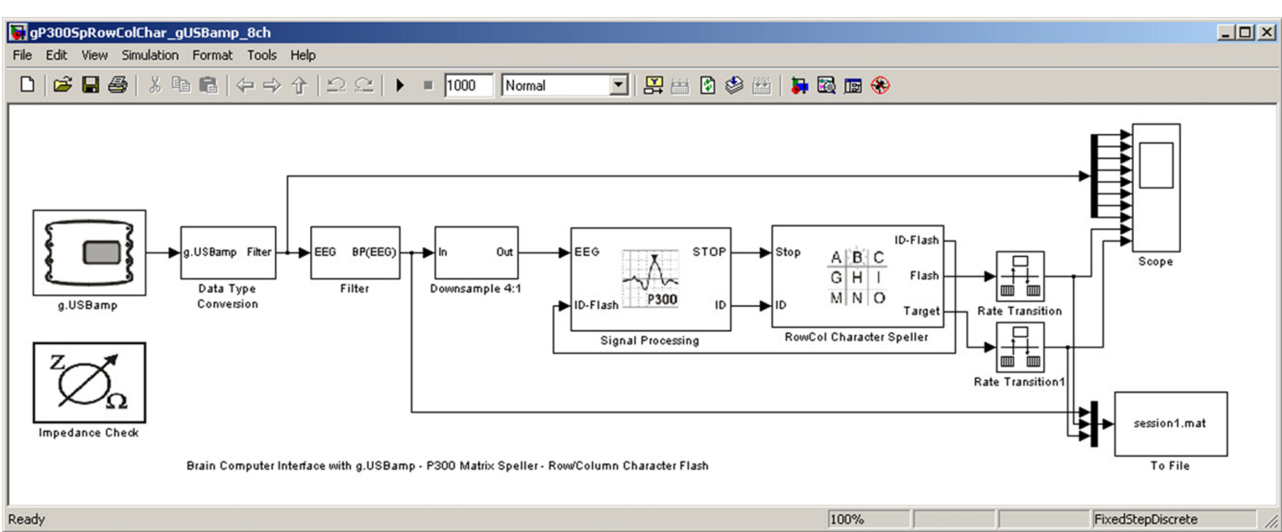

Fig. 2. Simulink model for the real-time analysis of the EEG data. The amplifier g.USBamp reads in the 8 channels of EEG data and passes the data to a bandpass filter and downsamples it. The signal processing block performs the EP analysis and LDA, the RowCol Character Speller block controls the method used to highlight the characters.

Table 1
Percentage of sessions which were classified with certain accuracy. *n* specifies the number of subjects participating.

Classification accuracy in %	Row-column speller: percentage of sessions (N = 81)	Single character speller: percentage in sessions (N = 38)
100	72.8	55.3
80–100	88.9	76.3
60–79	6.2	10.6
40–59	3.7	7.9
20–39	0.0	2.6
0–19	1.2	2.6
Mean accuracy of all subjects	91.0	82.0
Spelling time [s]	28.8	54
Mean accuracy of subjects who participated in RC and SC (N = 19)	85.3	77.9

were correctly selected by the LDA). It must be noted that this is an on-line result and not a cross-validation result. Even 88.9% (RC) and 76.3% (SC) of the subjects were able to make none or only 1 mistake. Moreover, only 1.2% (RC) and 2.6% (SC) of the subjects were not able to spell a single character correctly. For the 19 subjects that participated in both paradigms, the RC speller performed better (85.3%) than the SC speller (77.9%). Nine persons had the same accuracy for SC and RC, for 6 persons RC was better, for 4 persons SC was better. More data would be required to run a statistical test.

The median score of the disturbance of the flashing characters was 1 (1, not disturbing; 5, very disturbing) reported by 35 subjects. Furthermore non-paired accuracy data have been stratified by means of the questionnaire results and analyzed with the Mann–Whitney test. Gender aspects were also controlled. Female participants reached a mean accuracy of 81.9% and male participants reached 90.1%, but the difference was not significant. Furthermore, the level of education, working duration and cigarette and coffee consumption (yes/no question) did not show significant differences. Subjects that slept less than 8 h the night before reached 99.1% accuracy compared to 87% for all others ($p < 0.05$).

Additionally, the maxima of the P300 responses were calculated for all subjects who performed both the RC and SC experiments. The P300 response on electrode position Cz (one of the most important positions) for the RC speller was 7.9 μV and for the SC speller 8.8 μV; the difference was statistically significant (paired t-test, $p < 0.001$; similar differences could be found for other electrode positions).

This study shows that the P300-based BCI can achieve high accuracy after only 5 min of training. 72.8% of the subjects reached 100% accuracy with the RC speller. Moreover, these results show accuracy levels similar to those of other studies that have used much more training data [e.g., 7, 10, 13–15].

The results presented in this paper can be compared to an earlier study performed with 99 subjects using a motor imagery-based BCI in Graz [4]. The subject's task was to imagine left and right hand movement (20 times each) to move a cursor to the corresponding side of a computer monitor. The BCI system was then trained on the EEG data recorded from positions C3 and C4. Training time was approximately 6 min and the recursive least square or bandpower estimation in predefined frequency bands with LDA were used for classification. 6.2% performed at an accuracy level between 90 and 100%, as shown in Table 2. This is well below the P300 results achieved in this study. The motor imagery-based BCI used only 2 bipolar derivations compared to 8 EEG electrodes for the P300 experiment; however, the setup time was almost equal. Thus, the current results strongly suggest the P300-based BCI is superior to a motor imagery system, if the goal is to quickly achieve highly accurate and reliable results.

Table 2
Percentage of sessions which were classified with a certain accuracy for motor imagery classified with the rls algorithm or band power (bp) estimation. *n* specifies the number of subjects.

Classification accuracy in %	RLS + BP percentage of sessions (N = 99)
90–100	6.2
80–89	13.0
70–79	32.1
60–69	42.0
50–59	6.7

The motor imagery experiment used a binary decision between left and right and therefore chance classification corresponds to 50%. The P300 experiment can also be thought of as a binary decision system that discriminates between the target character and non-target characters. However, for the spelling experiment 36 decisions were presented. This is a major advantage because with a single decision step, one of 36 (or more) characters can be selected. This also contributes to a much higher bit rate.

For example, in the motor imagery task, one decision is made every 8–10 s (a random interval after each motor imagery). This provides approximately 6 binary decisions per minute and would yield near 1 character per minute with 100% accuracy. In contrast, the RC speller highlights each column or row for 100 ms and is grey for 60 ms (6 rows × 160 ms + 6 columns × 160 ms × 15 flashes = 28.8 s). This yields one character every 28.8 s. The SC speller flashes each character for 60 ms and is dark for 40 ms (36 characters × 100 ms × 15 = 54 s). This means that the RC flasher is about 2 times faster than the SC flasher and the motor imagery system. However, accuracy of the P300 system is much higher, resulting in fewer spelling errors. It must be noted that the P300-based system must average several flashes for one decision; the motor imagery system uses a single trial but performs averaging over the last samples with a specific window length.

In comparing the performance of the RC and SC paradigms it is clear that the RC speller has a speed advantage because 6 characters are simultaneously highlighted. This results in reduced P300 amplitude in the RC paradigm as compared to the SC paradigm; therefore, it was expected that the SC paradigm would achieve higher overall accuracy than the RC paradigm. However this was not the case. In general, it is possible that the longer character selection time is tiring for many subjects, which increases the variability of the P300 response. It must also be noted that for training of the LDA more non-target characters are used than target characters. Especially for the SC paradigm, only 15 target highlights are used while 35 × 15 highlights are used for non-targets. In contrast, the RC paradigm uses 11 × 15 non-target highlights. This results in a more unbalanced LDA training set for the SC paradigm than the RC paradigm, which could contribute to the lower accuracy observed in the SC paradigm.

For the 19 subjects that used both the RC and SC paradigms, there was a trend toward higher accuracy in the RC paradigm. A larger sample size to investigate this further may be warranted; however, the difference in character selection time may render the RC paradigm superior in either case. Not surprisingly gender aspects and education level did not affect accuracy. Working duration, cigarette consumption, and coffee consumption also had no effect on accuracy. However, an interesting finding was that the system was more accurate for people who slept less the previous night.

Acknowledgement

This work was supported by the EU-IST (FP6-027731) project Presenccia and SM4all.

References

[1] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, H. Flor, A spelling device for the paralysed, Nature 398 (1999) 297–298.
[2] O. Friman, I. Volosyak, A. Graser, Multiple channel detection of steady-state visual evoked potentials for brain–computer interfaces, IEEE Trans. Biomed. Eng. 54 (2007) 742–750.
[3] C. Guan, M. Thulasidas, J. Wu, High performance P300 speller for brain–computer interface, in: Proc. IEEE Int. Workshop on Biomedical Circuits Systems, 2004, S3-5-S316.
[4] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, G. Pfurtscheller, How many people are able to operate an EEG-based brain–computer interface ? IEEE Trans. Neural Syst. Rehab. Eng. 11 (2003) 145–147.
[5] C. Guger, C. Holzner, C. Grönegress, G. Edlinger, M. Slater, Control of a smart home with a brain–computer interface, in: Proceedings of the 4th International Brain–Computer Interface Workshop and Training Course 2008, September 18–21, 2008, pp. 339–342.
[6] C. Guger, A. Schlögl, C. Neuper, D. Walterspacher, T. Strein, G. Pfurtscheller, Rapid prototyping of an EEG-based brain–computer interface (BCI), IEEE Trans. Rehab. Eng. 9 (1) (2001) 49–58.
[7] D. Krusienski, E. Sellers, F. Cabestaing, S. Bayoudh, D. McFarland, T. Vaughan, J. Wolpaw, A comparison of classification techniques for the P300 Speller, J. Neural Eng. 6 (2006) 299–305.
[8] Z. Lin, C. Zhang, W. Wu, X. Gao, Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs, IEEE Trans. Biomed. Eng. 54 (6) (2007) 1172–1176.
[9] C. Neuper, R. Scherer, S. Wriessnegger, G. Pfurtscheller, Motor imagery and action observation: modulation of sensorimotor brain rhythms during mental control of a brain–computer interface, Clin. Neurophysiol. 120 (2) (2009) 239–247.
[10] F. Nijboer, E.W. Sellers, J. Mellinger, M.A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D.J. Krusienski, T.M. Vaughan, J.R. Wolpaw, N. Birbaumer, A. Kübler, A P300-based brain–computer interface for people with amyotrophic lateral sclerosis, Clin. Neurophysiol. 119 (8) (2008) 1909–1916.
[11] G. Pfurtscheller, R. Leeb, C. Keinrath, D. Friedman, C. Neuper, C. Guger, M. Slater, Walking from thought, Brain Res. 1071 (1) (2006) 145–152.
[12] A.D. Santana-Vargas, M.L. Pérez, F. Ostrosky-Solís, Communication based on the P300 component of event-related potentials: a proposal for a matrix with images, Rev. Neurol. 43 (11) (2006) 653–658.
[13] E.W. Sellers, D.J. Krusienski, D.J. McFarland, T.M. Vaughan, J.R. Wolpaw, A P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance, Biol. Psychol. 73 (3) (2006) 242–252.
[14] E.W. Sellers, G. Townsend, C. Boulay, B.K. LaPallo, T.M. Vaughan, J.R. Wolpaw, The P300 brain–computer interface: a new stimulus presentation paradigm, 778.21, Abstract Viewer/Itinerary Planner, Society for Neuroscience, Washington, DC, Online, 2008.
[15] H. Serby, E. Yom-Tov, G.F. Inbar, An improved P300-based brain–computer interface, IEEE Trans. Neural Syst. Rehab. Eng. 13 (1) (2005) 89–98.
[16] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, Brain–computer interfaces for communication and control, Clin. Neurophysiol. 113 (2002) 767–791.
[17] H. Zhang, C. Guan, C. Wang, Asynchronous P300-based brain–computer interfaces: a computational approach with statistical models, IEEE Trans. Biomed. Eng. 55 (6) (2008) 1754–1763.