Controlling a Smartphone with Brain-Computer Interfaces: A Preliminary Study

Víctor Martínez-Cagigal[1], Eduardo Santamaría-Vázquez[1], and Roberto Hornero[123]*
victor.martinez@gib.tel.uva.es · eduardo.santamaria@gib.tel.uva.es · robhor@tel.uva.es

[1] Biomedical Engineering Group, E.T.S.I. de Telecomunicación, Universidad de Valladolid, Valladolid (Spain)
[2] IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid (Spain)
[3] INCYL, Instituto de Neurociencias de Castilla y León, Salamanca (Spain)

Abstract. Nowadays, smartphones are essential parts of our lives. The wide range of functionalities that they offer to us, from calling, taking photos, sharing information or contacting with people, have contributed to make them a useful tool. However, its accessibility remains restricted to disabled people that are unable to control their motor functions. In this preliminary study, we have developed a Brain-Computer Interface system that allows users to control two main functionalities of our smartphones using their own brain signals. In particular, due to the importance of the socializing apps in today's world, the system includes the control of social networking and instant message services: Twitter and Telegram, respectively. The system has been tested with 10 healthy subjects, who were asked to perform several tasks, reaching an average accuracy of 92.3%. Preliminary results show that users can successfully control the system, bridging the accessibility gap in smartphone applications.

Keywords: Brain-Computer Interfaces (BCI), smartphones, electroencephalogram, P300 evoked potentials, Event-Related Potentials (ERP).

1 Introduction

Brain-Computer Interfaces (BCIs) have been originally developed for improving the quality of life of severely motor-disabled people. The facility of these systems to create a communication system between our brains and the environment makes them a suitable alternative to bypass diseases that impairs the neural pathways that control muscles [1–3]. For instance, BCIs have been successfully applied with users that suffer from traumatic brain injuries, muscle disorders, ataxia, cerebral palsy, multiple sclerosis, among others [1, 2, 4]. In order to perform such bypassing, brain signals should be monitored. This is commonly

* ✉ Corresponding author

achieved by recording the electroencephalogram (EEG) of the user, due to its non-invasiveness and ease of use [3].

Owing to the range of capabilities and the continuous Internet connection that offer the smartphones nowadays, these devices have become an indispensable part of people's lives. In fact, the market penetration of the smartphones reaches the 66%, with a 4.9 billion of unique mobile users [5]. Although their main functionalities cover from taking photos, reading news, watching videos or playing games, more that the 56% of the time spent with these devices is dedicated to socializing (i.e., social media and instant messaging) [6]. Nevertheless, their access is still restricted to disabled people that are unable to use accurately their hands and fingers.

In relation to that accessibility, despite the growing popularity of smartphones, there are very few attempts in the literature that have tried to integrate a BCI system for controlling their main functionalities. These studies are limited to accept incoming calls [7], dial numbers [8], select contacts [8, 9] or open pre-installed apps [10]. However, none of those studies have been focused on providing a high-level control of the smartphones, nor controlling anything related to the socializing category, the most popular one, both in everyday and work environments [6].

The main objective of this study is to design, develop and test a BCI system that allows users to control socializing-related functionalities of the smartphones with their own brain signals. In particular, the system should provide a complete control of Twitter and Telegram, a social network and an instant messaging app that currently have more than 317 and 100 millions of mobile active users, respectively [5].

2 Subjects and methods

The application has been tested with 10 healthy subjects (9 males and 1 female) with a mean age of 26.2% $\pm$ 3.45 years. All the subjects gave their informed consent for participating in the study, composed of 3 different sessions (2 for calibration and 1 for testing).

As can be noticed in the Fig. 1, the system is composed of three main stages that communicate among themselves: (1) acquisition, which involves the EEG signal recording and monitoring; (2) processing, which applies the real-time methods to determine the command that the user wants to select; and (3) application, intended to interpret those commands and provide visual feedback. These stages are detailed below.

2.1 Acquisition

The acquisition stage is intended to record, monitor and send the EEG signal of the user to the processing stage in real-time. EEG signals were recorded with a g.USBamp amplifier using a 8-channel cap. BCI2000 platform was used to record, display and process the data [11]. Active electrodes were placed on Fz,

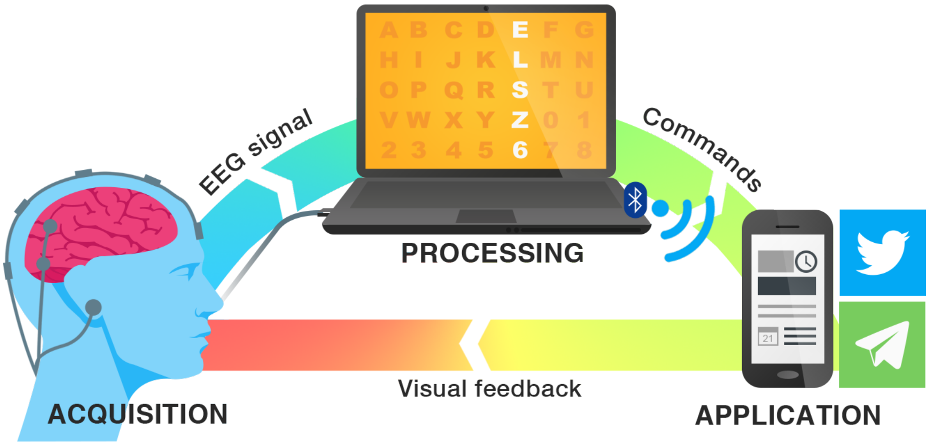

Fig. 1. Structure of the BCI smartphone system. The acquisition stage records the EEG signal and sends it to the laptop, whose main task is to apply the signal processing methods to decode the user's intentions. Finally, the selected commands are sent in real-time via Bluetooth to the final device, which runs the application and provides visual feedback to the user.

Cz, Pz, P3, P4, PO7, PO8 and Oz, according to the International 10–20 System distribution [12], and referenced to the earlobe, using the FPz as a ground. Moreover, notch (50 Hz), bandpass (0.1 – 60 Hz) and common average reference filters were applied as a pre-processing stage.

2.2 Processing

P300 evoked potentials were selected as control signals, due to its exogenous nature and the large amount of commands that can be selected by the user [3, 4, 13]. These potentials, defined as voltage deflections that appear in parietal and occipital cortex in response to infrequent and significant stimuli about 300 ms after their onset, are elicited using an *oddball* paradigm [3]. In this paradigm, a target infrequent stimulus, which has to be attended, is presented among other distracting stimuli that have to be ignored. Thus, a P300 potential is generated when the user receives an unexpected target stimulus.

In this study, row-col paradigm (RCP) matrices, an extension of the *oddball* paradigm, have been used to determine the user's intention [14]. As shown in Fig. 2, a matrix that contains the application commands is displayed. The user just need to focus attention on one of these commands, while the matrix rows and columns are randomly intensified. Whenever the target's row or column is flashed, a P300 potential is produced in the scalp of the user. Therefore, the desired character can be determine by computing the intersection where those potentials were found [3, 4, 13, 14]. In particular, two switchable RCP matrices have been used: (i) navigation matrix, a small one intended to provide efficient navigation; and (ii) keyboard matrix, intended to write texts and fill out forms.

In order to determine the command that the user is looking at, it is required to perform a signal processing stage, composed by feature (i) extraction, (ii) selection and (iii) classification. The signal processing pipeline that has been followed in this study is the most common one in the P300-based BCI literature, which applies: down-sampling to 20 Hz as feature extraction, step-wise (SW) regression (max. of 60 features, $p_{in} = 0.1$, $p_{out} = 0.15$) as feature selection, and linear discriminant analysis (LDA) as feature classification [4, 15–18]. As a result, the likelihood of selecting each matrix command is returned, and the final selected command will be the one that provides the maximum probability (i.e., $p_{sel} = \max \boldsymbol{p}$).

RCP paradigm is a synchronous process. This implies that a set of probabilities will be always returned and thus, the system will select a command even if the user is not paying attention to the flashings [4, 19, 20]. In order to avoid this problem, we have applied an asynchrony management method based on thresholding [4]. The algorithm is simple: (1) EEG signals of the user paying attention (i.e., control state) and ignoring the stimuli (i.e., non-control state) are recorded in a calibration session; (2) probability scores are stored in control $\boldsymbol{p}_c$, and non-control $\boldsymbol{p}_n$ vectors; (3) these vectors are fed as different classes into a ROC curve; and (4) threshold is obtained by maximizing the sensitivity-specificity tuple. Hence, when a selection occurs, p_{sel} is compared to the threshold T: if $p_{sel} > T$, the selected command is sent to the application stage; otherwise, the application shows a warning message that encourages the user to pay more attention to the stimulation.

2.3 Application

As previously indicated, the developed system allows users to control both Twitter and Telegram, and switch freely between both functionalities. Therefore, the application stage receives the commands selected by the user via Bluetooth and interprets them, controlling both functionalities and providing real-time feedback. Fig. 2 shows several snapshots of the final application.

Twitter. Popular social networking service where users post small messages (up to 140 characters), known as "tweets" Moreover, its activity is not only limited to personal computers, but also to mobile phones, where the number of active users reaches more than 317 million [5]. Our BCI application implements the entire set of Twitter functionalities, including the possibility of interacting with tweets: writing, answering, mark as favorite or "retweet" them; and accounts: surfing among profiles, or sending private messages.

Telegram. Non-profit cloud-based instant messaging service where users can send messages and exchange files of any type. Although it also have a desktop version, it has more than 100 million of mobile active users, becoming the most popular instant message app in several countries [5]. The developed BCI application covers its main functionalities, including the possibility to create new chats with any contact stored in the phone; and interacting with chats, groups and channels, sending messages and receiving them in real-time.

2.4 Evaluation procedure

During the assessment, participants were comfortably seated in front of a panoramic screen that displayed the current RCP command matrix, connected to a laptop (Intel Core i7 @ 2.6 GHz, 16GB RAM, Windows 10) that executed the processing stage; as well as in front of a smartphone (Samsung Galaxy S7, 4GB RAM, Android 7.0) on a small tripod, which run the application. Each user carried out a total of 3 sessions (2 calibration sessions and 1 evaluation session), detailed below:

Calibration 1 The first session was indented to calculate the optimal parameters for each user, such as the classifier weight vector, the optimal number of sequences (i.e., repetitions of the stimuli) and the optimal asynchronous threshold. First, users were ask to sequentially pay attention to 6 items in 4 trials (i.e., spelling 4 words of 6 characters with the keyboard matrix) while the matrix was flashing. For this task, 15 sequences where used and thus, each character cell was highlighted 30 times. In order to keep the attention on the task, they were recommended to count how many times the target command was illuminated. Then, SW and LDA were performed for determining the optimal weights and number of sequences of each user. From here onwards, the customized classifier and number of sequences were used. Finally, the first session of threshold calibration was performed. Composed of 8 trials with 6 items, users were asked to pay attention to 4 trials (i.e., control state), and to ignore the remaining 4 (i.e., non-control state).

Calibration 2 The second session was intended to record additional data with the objective of creating a more robust asynchronous threshold [4]. Hence, users were asked to pay attention to 4 trials and to ignore 4 trials more, all of them composed by 6 items. It is noteworthy to mention that these trials were performed using the navigation matrix, which reduces the average duration time of the session.

Evaluation The last session was intended to assess the usefulness and the performance of the developed BCI application. Users were asked to complete 6 different tasks, whose difficulty increased progressively. The duration of each task varied among users due to their optimal number of sequences. However, the optimum number of selections, the mean average time, and its standard deviation are provided below:

- **Task 1) Toggling between Twitter and Telegram.** The first and easiest task is intended to introduce the system to the user. In this task, users had to scroll up and down the Twitter timeline and toggle to Telegram (3 items, $1:10 \pm 0:25$ min).
- **Task 2) Retweeting a *tweet*.** Using Twitter, users had to scroll down the timeline, select one *tweet* and retweet it (4 items, 1:50 ± 0:55 min).
- **Task 3) Writing a new *tweet*.** This was the first task that involved the use of both matrices, increasing the duration time and the difficulty to

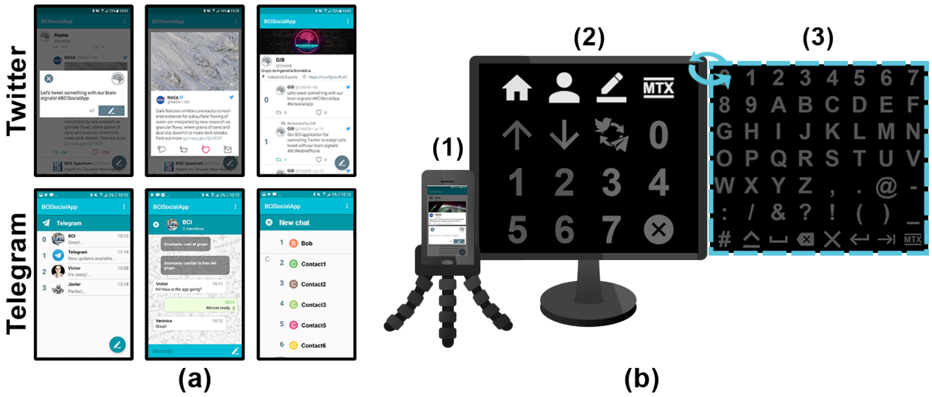

Fig. 2. (a) Snapshots of the developed BCI application whilst controlling Twitter and Telegram. (b) Evaluation setup: (1) smartphone; and panoramic screen with (2) navigation matrix or (3) keyboard matrix. Note that the first row of the navigation matrix is currently flashed.

finish it. Using Twitter, users had to open the form to write a new *tweet* and spell “hello”(7 items, 3:54 ± 1:39 min).

Task 4) Checking the profile and answering a *tweet*. Using Twitter, users had to visit their profile, select the last written *tweet* and answer it by “great!” (11 items, 5:53 ± 2:00 min).

Task 5) Create a new chat. Using Telegram, users had to select one contact and create a new chat, spelling “how are you?” (11 items, 6:15 ± 2:10 min).

Task 6) Chating with someone. Using Telegram, users had to select one chat from the list, in which the interlocutor had inquiring: ”hi! how are you?”, and reply with: “fine, and you?” (12 items, 7:31 ± 2:48 min).

3 Results

The results of the evaluation session are shown in the Table 1 and the Fig. 3, where accuracies and the required time to accomplish each task are provided for each participant. Accuracy is calculated as $1 - N_e/N_t$, where N_e is the number of errors and N_t is the total number of selections. Note that selections that have not overcome the asynchronous threshold are not considered errors, since they have not been sent to the final device. As previously mentioned, the duration of each task depends on the number of sequences of each user (i.e., N_s) in a large extent. Owing to that fact, N_s is provided as well.

Users	Task 1		Task 2		Task 3		Task 4		Task 5		Task 6		N_s	Average accuracy
	Tim.	Acc.	Tim.	Acc.	Tim.	Acc.	Tim.	Acc.	Tim.	Acc.	Tim.	Acc.		
C01	01:42	100%	02:16	100%	05:38	100%	07:47	90.9%	08:05	90.9%	09:12	91.7%	11	93.8%
C02	00:56	100%	01:14	100%	03:04	85.7%	04:40	100%	04:51	100%	05:31	100%	6	97.9%
C03	02:01	100%	04:02	83.3%	07:43	85.7%	10:07	100%	10:30	100%	13:01	92.3%	13	94.2%
C04	01:05	100%	02:10	66.7%	03:35	100%	05:27	81.8%	05:39	100%	09:43	73.3%	7	85.2%
C05	00:47	100%	01:02	100%	02:33	100%	03:54	90.9%	04:03	100%	04:36	100%	5	98.0%
C06	00:56	100%	01:14	100%	03:04	71.4%	06:14	100%	08:05	100%	09:12	66.7%	8	86.7%
C07	01:14	100%	02:04	60.0%	03:35	57.1%	05:27	81.8%	06:53	91.7%	07:22	81.8%	8	79.6%
C08	00:37	100%	00:50	100%	02:03	100%	03:07	100%	03:14	90.9%	03:41	91.7%	4	95.8%
C09	01:14	100%	01:39	100%	04:06	100%	06:38	91.7%	05:39	100%	06:26	100%	8	98.0%
C10	01:05	100%	01:49	80.0%	03:35	100%	05:27	100%	05:27	90.9%	06:26	91.7%	7	93.9%
Mean	**01:10**	**100%**	**01:50**	**89.0%**	**03:54**	**90.0%**	**05:53**	**93.7%**	**06:15**	**96.4%**	**07:31**	**88.96%**	**7.7**	**92.3%**
SD	00:25	0.0%	00:55	15.6%	01:39	15.1%	02:00	7.5%	02:10	4.6%	02:48	11.5%	2.7	6.3%

Table 1. Evaluation session results for each participant.

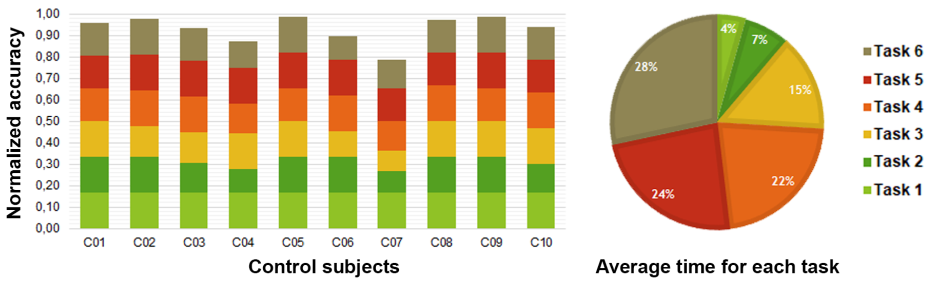

Fig. 3. Results of the assessment sesion. (Left) Stacked normalized accuracy of the participants for each task. (Right) Average duration percentage of each evaluation task.

4 Discussion

Results show that the developed application, tested with 10 healthy subjects, can be successfully controlled using only the brain signals of the users, reaching an average accuracy of 92.3% in the evaluation session. Its standard deviation is kept low, since 7 participants were able to reach an average accuracy greater than 90%, and only 3 have obtained values below the average. Moreover, it is noteworthy to mention that all of them were able to finish all the tasks.

With regard to the difficulty of these tasks, the average durations shown in Fig. 3 reinforce the fact that it was increased progressively. In fact, the last two tasks took more than the half of the duration of the session, while the sum of the 4 first ones took only over a quarter. However, the average reached accuracies of the tasks does not show a constant decreasing, which could be expected in order to follow that difficulty. The first task was easily finished by all the participants, obtaining a perfect score (i.e., $100\% \pm 0.0\%$), which means that they were able to complete it without a single mistake. Despite of the difference of the required time to finish them, the second and the third tasks have obtained similar average accuracies, $89.0\% \pm 15.6\%$ and $90.0\% \pm 15.1\%$, respectively. This may be because several participants, such as C02 or C10, demonstrated that they were more proficient controlling the keyboard matrix (i.e., present in tasks 3–6) than the navigation matrix (i.e., present in all tasks). After the first one, the fourth and the fifth tasks have achieved high accuracies, reaching $93.7\% \pm 7.5\%$ and $96.4\% \pm 4.6\%$, respectively. Finally, the sixth and last task has obtained the lowest performance, with an average accuracy of $88.96\% \pm 11.5\%$, possibly because of its difficulty, where 10 out of 12 selections had to be performed using the keyboard matrix.

As pointed earlier, there are vey few BCI-based studies that have attempted to control any functionalities of a smartphone. These studies are limited to dial numbers in cell phones [8], accept incoming calls [7], perform calls [8, 9], open the photo gallery [10] or playing simple games [21]. None of them have been focused to provide a high level control of a smartphone, nor implement any socializing feature. In addition, none of them have been tested with disabled users. Moreover, within these attempts, only two of them are P300-based [9, 10], while the rest uses steady-state visual evoked potentials [8], or MindSet concentration features [7,21]. Regarding the P300-based studies, Katona *et al.* developed an application to answering or rejecting incoming calls, reaching an average accuracy of 75% with 5 healthy subjects; while Elsawy *et al.* developed an application that allowed users to open pre-installed apps and visualize images from the gallery, obtaining mean accuracies of $79,17\%$ and 87.5% for both features, respectively, using 6 healthy subjects. As can be noticed, not only our application provides a higher level of control of an smartphone, but also it have reached higher accuracies than these previous attempts.

Even though the results show that the developed application allow users to successfully control the socializing features of a smartphone, we can point out several limitations. Firstly, the application has been only tested with healthy subjects, and not with motor-disabled people, who are the target of this kind of BCI systems. In addition, the signal processing stage is executed in a laptop, making the system dependent on a computer and thus, impairing its portability. In order to overcome these limitations, we contemplate the following future research lines: (i) testing the system with motor-disabled people, in order to assess its actual usefulness to improve their quality of life, and (ii) encapsulating the signal processing stage inside the final device, improving the portability and its application in a real world scenario.

5 Conclusion

An asynchronous P300-based BCI system to control socializing apps of a smartphone has been designed, developed and tested. The system uses the P300 potentials of the user, generated by two exchangeable RCP matrices that are displayed on a panoramic screen, to determine the command that the user wants to select in real-time. These commands are then sent to the smartphone app via Bluetooth, and visual feedback is presented to the user. The application has been tested with 10 healthy subjects, who were asked to perform a total of 6 different tasks with increased difficulty, reaching an overall accuracy of 92.3%. Although the preliminary results shows that the developed system can be successfully controlled with the brain signals of the users, care must be taken to generalize them for disabled people.

Acknowledgments

This study was partially funded by projects TEC2014-53196-R of 'Ministerio of Economía y Competitividad' and FEDER, the project "Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer" (Inter-regional cooperation program VA Spain-Portugal POCTEP 2014-202) of the European Commission and FEDER, and project VA037U16 of the 'Junta de Castilla y León' and FEDER. V. Martínez-Cagigal was in receipt of a PIF-UVa grant of the University of Valladolid. The authors declare no conflict of interest.

References

1. Kübler, A., Nijboer, F., Birbaumer, N.: Brain-Computer Interfaces for communication and motor control – perspectives on clinical application. In: Toward Brain-Computer Interfacing. 1 edn. MA: The MIT Press (2007) 373–391
2. Kübler, A., Birbaumer, N.: Brain-computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? Clinical Neurophysiology **119**(11) (2008) 2658–2666
3. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer interfaces for communication and control. Clinical Neurophysiology **113**(6) (2002) 767–91
4. Martínez-Cagigal, V., Gomez-Pilar, J., Álvarez, D., Hornero, R.: An Asynchronous P300-Based Brain-Computer Interface Web Browser for Severely Disabled People. IEEE Transactions on Neural Systems and Rehabilitation Engineering **25**(8) (2017) 1332–1342
5. Kemp, S.: Digital in 2017: Global Overview. Technical report (2017)
6. Ipsos MORI, Google: Something For Everyone: Why the Growth of Mobile Apps is Good News for Brands. Technical report (2017)
7. Katona, J., Peter, D., Ujbanyi, T., Kovari, A.: Control of incoming calls by a Windows Phone based Brain Computer Interface. 15th IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2014) (2014) 121–125
8. Wang, Y.T., Wang, Y., Jung, T.P.: A cell-phone-based brain-computer interface for communication in daily life. Journal of neural engineering **8** (2011) 025018
9. Campbell, A., Choudhury, T., Hu, S., Lu, H., Mukerjee, M.K., Rabbi, M., Raizada, R.D.: NeuroPhone: Brain-Mobile Phone Interface using a Wireless EEG Headset. Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds - MobiHeld '10 (2010) 3–8
10. Elsawy, A.S., Eldawlatly, S.: P300-based Applications for Interacting with Smart Mobile Devices. In: 7th Annual International IEEE EMBS Conference on Neural Engineering. (2015) 166–169

11. Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R.: BCI2000: A general-purpose brain-computer interface (BCI) system. IEEE Transactions on Biomedical Engineering **51**(6) (2004) 1034–1043
12. Jasper, H.H.: The ten twenty electrode system of the international federation. Electroencephalography and clinical neurophysiology **10** (1958) 371–375
13. Farwell, L.A., Donchin, E.: Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and Clinical Neurophysiology **70**(6) (1988) 510–523
14. Townsend, G., LaPallo, B.K., Boulay, C.B., Krusienski, D.J., Frye, G.E., Hauser, C.K., Schwartz, N.E., Vaughan, T.M., Wolpaw, J.R., Sellers, E.W.: A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns. Clinical Neurophysiology **121**(7) (2010) 1109–1120
15. Corralejo, R., Nicolás-Alonso, L.F., Álvarez, D., Hornero, R.: A P300-based brain-computer interface aimed at operating electronic devices at home for severely disabled people. Medical & Biological Engineering & Computing **52**(10) (2014) 861–872
16. Martinez-Cagigal, V., Hornero, R.: Selección de Canales en Sistemas BCI basados en Potenciales P300 mediante Inteligencia de Enjambre. Revista Iberoamericana de Automática e Informática Industrial **14**(4) (2017) 372–383
17. Krusienski, D., Sellers, E., McFarland, D., Vaughan, T., Wolpaw, J.: Toward enhanced P300 speller performance. Journal of Neuroscience Methods **167**(1) (2008) 15–21
18. Krusienski, D.J., Sellers, E.W., Cabestaing, F., Bayoudh, S., McFarland, D.J., Vaughan, T.M., Wolpaw, J.R.: A comparison of classification techniques for the P300 Speller. Journal of neural engineering **3**(4) (2006) 299–305
19. Pinegger, A., Faller, J., Halder, S., Wriessnegger, S.C., Müller-Putz, G.R.: Control or non-control state: that is the question! An asynchronous visual P300-based BCI approach. Journal of neural engineering **12**(1) (2015) 014001
20. Aloise, F., Schettini, F., Aricò, P., Leotta, F., Salinari, S., Mattia, D., Babiloni, F., Cincotti, F.: P300-based braincomputer interface for environmental control: an asynchronous approach. Journal of Neural Engineering **8**(2) (2011) 025025
21. Wu, G., Xie, Z., Wang, X.: Development of a mind-controlled Android racing game using a brain computer interface (BCI). 2014 4th IEEE International Conference on Information Science and Technology (2014) 652–655