2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Auditory Paradigm for a P300 BCI system using Spatial Hearing

Francesco Ferracuti, Alessandro Freddi, Sabrina Iarlori, Sauro Longhi, Paolo Peretti

Abstract— **The present paper proposes an auditory BCI paradigm for systems based on P300 signals which are generated by auditory stimuli characterized by different sound typologies and locations. A Head Related Transfer Function approach is adopted to virtualize auditory stimuli. When virtualized audio is used, the user has to focus the attention both on the type and location of the stimulus, thus generating P300 signals whose amplitude is higher than that generated without audio virtualization. Classification is performed by Support Vector Machines in which gaussian radial basis functions are used as kernel functions. The system has been validated with 14 users, who were asked to choose one among five common spoken words, previously virtualized and transmitted to stereophonic headphones. Classification results prove that the proposed auditory BCI system performed similarly to common visual BCI P300 systems, representing then an alternative to visual BCI for users with visual impairments.**

I. INTRODUCTION

Brain-Computer Interfaces (BCIs) are devices which translate the brain activity of the user into specific signals, which may be used for communicating or controlling external devices [1], [2] without the use of peripheral nerves and muscles [3]. BCIs represent an interesting option to people affected by neuromuscolar disorders, but whose brain activity is normal, such as in patients affected by Amyotrophic Lateral Sclerosis (ALS).

In the literature three different types of stimuli are commonly adopted to drive a BCI: visual stimuli, tactile stimuli and auditory stimuli. Visual stimuli were the first to be studied, and typically lead to the best classification results [4], [5]. Visual stimuli, however, can not be used when the user's sight has been compromised (e.g. limited horizontal eyes movement, incapability to focus the gaze, etc ...), which is the most critical problem faced by both visual BCI and non-BCI systems (such as Eye Gaze systems [6]). In these cases tactile stimuli and auditory stimuli can be adopted instead. Tactile BCI proved to be a good choice for navigation purposes [7]–[9], but only recently it has been used as a communication device [10]. Different typologies of ElectroEncefaloGraphic (EEG) signals have been used in the literature for developing auditory BCIs (e.g. cortical potentials, sensorimotor rhythm, steady state evoked potential and P300), and auditory BCIs represent at the moment the most suitable alternative to visual BCI.

Slow Cortical Potential (SCP) signals were studied in [11]. The users involved in the experiment received either visual, auditory or combined visual/auditory feedback of their SCPs. Results showed that even if the visual feedback led to the highest number of correct answers, auditory stimuli could be used as well. In [12], instead, the authors adopted an auditory BCI driven by the Sensory Motor Rhythm (SMR) signal. Experimental results showed that auditory stimuli led to similar final results as visual stimuli, even if in the first case the training time was longer. A different approach to auditory paradigm exploits the Steady-State Auditory Evoked Potentials (SSAEP). These are elicited by click-trains, amplitude/frequency modulated tones. A steady-state response is represented by a significant amount of power at the modulation amplitude/frequency of a stimulus [13]. Many of the auditory BCIs available in the literature, however, are based on the P300 component of the Event Related Potential (ERP). In [14] P300 responses to two simultaneous auditory stimulus streams were classified. The users had to choose among one of the two streams and focus their attention by counting the target stimulus. The outcome of the experiment was that a user could possibly direct his/her attention using auditory stimuli only. In [15] a four-choice BCI was tested with both healthy users and patients affected by ALS. The users were presented auditory and visual stimuli and they had to choose the words "yes" or "no" among "yes", "no", "pass" and "end", according to a classic Oddball Paradigm (OP). The results showed that a target probability of 25% was enough to elicit a reliable P300 signal both in healthy users and ALS patients. A P300 speller driven by auditory stimuli was first presented in [16]. The authors created a 5×5 letter matrix similar to that adopted in common visual P300 BCI spellers (see [17]). Column and row flashes were replaced with auditory stimuli that were coded to particular columns and rows in the matrix (i.e. spoken number of column and row). Even if the presence of a visual support matrix was still needed, more than half of the users were able to focus their attention so that the auditory stimulus could be correctly detected and classified, even if with an average accuracy and bit rate lower than those achievable through visual BCIs. Similar results were obtained in [18], where the authors extended the letter matrix to 36 characters and added visual cues early in the training phase. A larger amount of choices did not compromise the classification performances, while the addition of visual cues allowed for a better accuracy during the online phase.

The above mentioned articles prove that auditory BCI is a possible alternative to visual BCI, however at the cost of lower classification scores and average bit rates. An alternative method to improve performances has been presented in [19], where the authors adopted spatial auditory

F. Ferracuti, A. Freddi, P. Peretti, S. Iarlori and S. Longhi are with Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, 60131 Ancona (Italy). E-mail: {f.ferracuti, a.freddi, s.iarlori, s.longhi, p.peretti}@univpm.it

stimuli. Users had to sit in the middle of a room surrounded by five speakers with $45°$ angle between them. All speakers were given a unique complex audio stimulus, so that the discriminating cue was both the physical property and spatial location of the stimulus. The results showed an increment in the classification score w.r.t. the case where a single speaker only was adopted. Moreover by increasing the number of runs (times that the audio stimuli were repeated) it was also possible to achieve results similar to that of visual BCIs, however impacting negatively on the bit rate.

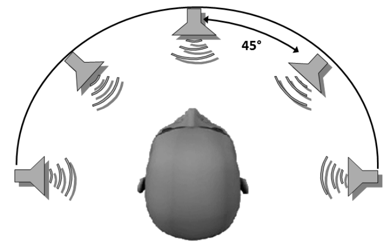
The main drawback of the proposed solution was that the user had to stand still in the middle of a room surrounded by speakers. The present paper tries to overcome this obstacle by using a single stereo headphone where audio stimuli are virtualized. Sound virtualization has already been studied in [20] to show that spatial location can be a cue determining factor for BCI applications. The auditory paradigm presented in this work aims to give the user the opportunity to choose one between five different audio stimuli, retaining at the same time the possibility for the user to be moved within the home environment. Moreover the audio stimuli presented to the users are simple words referred to common daily life activities, rather than audio tones set at specific frequencies [11], [14], [19], [20], numbers [16], or instrumental sounds [12], [18]. It is the authors claim that the use of words of the common language in auditory BCIs can lead to a straightforward communication paradigm, reducing at the same time the training time needed to use the BCI correctly.

The article is divided as follows. Section II describes how spatial hearing have been adopted for developing the proposed auditory BCI. The methodology, the hardware and the software adopted to perform the tests are described in Section III. In section IV the results are presented and discussed. Final remarks and possible future research prospectives are given at the end of the paper.

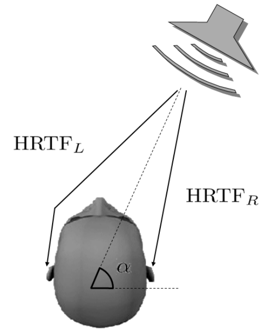
II. Spatial audio

Given an audio source in a room, the human ear can perceive mainly two information: the sound and the position of the source. In anechoic chamber, in case of source in front of the listener, the human auditory system can recognize variation of sound source direction of about $1°$ on the horizontal plane. In case of source behind or beside the listener, the sensibility significantly decreases to about $10°$. On the vertical plane there are no differences between sources in front of and behind the listener and also in this case the order is about $10°$ [21]–[23].

In order to obtain spatial audio, one of the most used technique is the binaural recording: the aim is to get a very realistic recording of a sound event, which takes place in a real environment, through a single pair of microphones, placed on an artificial head at the ears. In this paper we are interested in obtaining spatial audio which can later be used as auditory stimulus directly fed into the user's headphone: the binaural recording thus represents a natural approach to obtain highly realistic sound images. In this context the Head Related Transfer Function (HRTF) assumes a great



(a) The five audio stimuli directions, played by headphones, with an off-set of $45°$.



(b) Schematic of left and right HRTF relative to a sound source coming from a well-defined direction α.

Fig. 1. Spatial hearing

importance. HRTF is an impulse response that describes how a sound coming from a well-defined direction is perceived by the human ear. With a set of two HRTFs, one for each ear, any direction of sound source propagation can be synthesized (Fig. 1(b)).

Therefore, given the left and right HRTFs relative to a desired sound direction α, a mono signal s becomes directive through the operation of convolution:

$$\text{out}_R = s * \text{HRTF}_R(\alpha) \quad (1)$$

$$\text{out}_L = s * \text{HRTF}_L(\alpha) \quad (2)$$

Database of HRTFs for several sound directions in anechoic environment can be found in the literature: the one used in this work has been realized by MIT Media Lab [24].

Five audio signals, namely the words "bathroom", "bedroom", "kitchen", "help" and "stop" have been virtualized through the use of ten different HRTFs, i.e. five different sound directions per ear with an off-set of $45°$ (Fig. 1(a)).

III. Testing Methodologies

A. Participants

Fourteen healthy subjects (10 males, 4 females, mean age 25.4, standard deviation $\pm$ 2.85, range $22-33$) participated in the study. All subjects were volunteering group members and

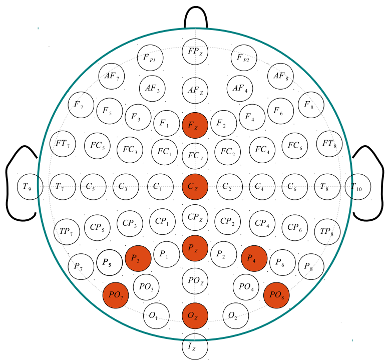


Fig. 2. Electrode set for recording and analysis. Eight data channels are according to the International 10-20 electrodes system; the reference and ground electrodes are selected as the left earlobe and the left mastoid, respectively.

had some previous experience with visual BCI, mainly based on imagined movement and P300 tasks. No one had previous experiences with auditory BCI. The lack of experience is not a main issue: the proposed BCI system, based on auditory stimuli represented by common spoken words, is simpler to use than auditory BCI systems in which stimuli are represented by tones or instrumental sounds, thus requiring short training phase.

B. Data acquisition

The EEG was recorded monopolarly using an electrode cap with 8 active high-purity gold (Au) electrodes (g.tec medical engineering GmbH) following the American Electroencephalographic Society modified version of the 10−20 system [25]. These are located at positions Fz, Cz, Po7, P3, Pz, P4, Po8, and Oz (see Fig. 2). Channels are referenced to the left earlobe and grounded to the left mastoid. Signals were acquired and amplified using a g.MobiLab+ (g.tec medical engineering GmbH, Germany). Data collection and stimulus presentation were controlled by the BCI2000 software package [26].

C. Procedure

Prior to recording periods, participants were asked to minimize eye movements and muscle contractions during the experiment. Each participant was equipped with stereophonic headphones, and was requested to repeatedly fulfill the following auditory task: listen to a sequence of five words and focus his/her attention when the target word was played (i.e. mentally counting how many times the target word was listened to). Each run contained 1 target word and 4 non-target words: both the sequence of the five words and their spatial orientation were randomly chosen. The users were not requested to consciously identify the word spatial orientation, however this association is unconsciously made by the users, thus increasing the P300 activity as already shown in [19]. Each run was repeated 150 times, for a total of 750 audio stimuli of which 150 were target stimuli and 600 non-target stimuli. A ratio of 1 to 5 between target and non-target stimuli has been shown to be rare enough to produce a P300 response [15]. A stimulus duration of 1500 ms and an Inter Stimulus Interval (ISI) of 250 ms were chosen. Electrooculogram (EOG) was not recorded, then the artifact rejection was not considered, but the artifact reduction was implemented using the following filters: a high pass filter at 0.1 Hz, a low pass filter at 30 Hz and a notch filter at 50 Hz. A Common Average Reference (CAR) spatial filter was applied to the temporal filtered signals [27]. Acquired signals were segmented into epochs of 800 ms starting at the onset of a stimulus. The data, that was originally sampled at a rate of 256 Hz, was decimated and moving average filtered by a frequency of 20 Hz. This resulted in 150 target trials (i.e. number of audio stimuli listened) and 600 non-target trials.

A Support Vector Machine (SVM) was used for data classification [28], [29], with the following gaussian radial basis function used as kernel function:

$$\phi(\|x_i - x_j\|) = e^{-a\|x_i - x_j\|}, \tag{3}$$

where x_i, x_j, are the i-th and j-th data sample. The kernel function parameter a is chosen as the value that maximizes the average between the target and non-target classification accuracy. To increase sensitivity, outcomes of multiple runs for the same task can be averaged. In this way, the influence of single trials can be decreased and the selection score can be more robust. One possibility is to average the raw trials timeseries for each task and classify them as a single trial. Another option is to classify each original trial individually and average over the classifier scores: which implies the use of two or more iterations (i.e. number of runs repeated before the classificator generates the output). We opted for this second approach, since it showed better performances.

Datasets from the BCI experiments contained four times more non-target stimuli than targets. Although the classification task is essentially binary, chance level for classification is 80%, which could potentially be obtained by simply assigning all samples to the non-target group. Therefore, to evaluate the performances different type of classification accuracy indexes are considered.

- **The classification accuracy**, which refers to the binary classification and is defined as the percentage of trials in which both the target or non-target stimuli are correctly scored.
- **The target accuracy**, which is defined as the percentage of trials in which a target stimulus is correctly scored.
- **The non-target accuracy**, which is defined as the percentage of trials a non-target stimulus is correctly scored.
- **The selection accuracy**, which denotes the percentage of trials in which the BCI system returns the target action thought by the user.

The selection accuracy index is evaluated for all iterations, therefore it is the average of the classifier scores for each trial. In order to have a single target output from the BCI system, just the target which has the largest classification output is chosen thus multiple targets are not allowed and one target always exists.

D. Information transfer rate

The Information Transfer Rate (ITR) measures the amount of information carried by every selection and, is a performance index for the evaluation of BCI systems. The ITR facilitates the performance comparison with other BCI applications and it is calculated in bits per selection with the following formula [30]:

$$B = \log_2 N + P \log_2 P + (1 - P) \log_2 \left(\frac{1 - P}{N - 1} \right), \quad (4)$$

where N represents the number of classes (five in the present case of study) and P is the selection accuracy. The ITR in bits per minute was obtained by multiplying the bit rate B by the classification speed V, that is the average number of selections per minute, as follows:

$$ITR = B \cdot V. \quad (5)$$

Eq. (4) shows that even though the selection accuracy may increase when using two or more iterations, the ITR may stay the same or even decrease when V decreases, that is to say when selection takes more time. This is typically the case of our auditory BCI, which requires audio stimuli of long duration based on words of common language rather than digital tones.

IV. Results

A. Classification performance

Table I gives the classification, target and non-target accuracy for the BCI experiment when the SVM is required to perform a classification within a single run. In this case only one subject reaches 70% of target accuracy, while the remaining subjects scored a target accuracy below the 70% limit, which is assumed to be the minimal limit for useful BCI operations [31]. Please note that target accuracy being lower than non-target accuracy is considered normal: whenever the ratio between target and non-target words is small, the classificator tends to weight non-target words more than target ones. When using multiple iterations, instead, the score went up quickly for most of the subjects, as shown in Fig. 3, which summarizes the selection accuracy in function of iterations required by the SVM to perform classification, for users 1, 5, 6, 8, 9 and 13.

The average value of iterations to reach the 70% selection score is 5, as shown in Fig. 4. Mean selection score for a single run is about 50%, as shown in Fig. 4 and table I. The participants reached the 80% selection score after ten iterations and 90% after fourteen iterations.

Fig. 5 shows the boxplot of the selection score for all iterations. On each box, the central mark is the median, the edges of the box are the lower and higher quartiles. When the lower quartile is considered, seven subjects are over the 70% selection accuracy. Considering the median, six subjects are over the 70% selection accuracy. If the higher quartile is considered, instead, eleven participants are over

TABLE I

Classification accuracy, target accuracy and non-target accuracy for auditory stimuli (stimulus duration 1500 ms, ISI 250 ms) within a single run. Peak amplitude for the auditory condition is determined as the maximum amplitude in the range from 0 to 800 ms.

Participant	Classification accuracy (%)	Target accuracy (%)	Non-target accuracy (%)
1	76,8	42,0	85,5
2	78,8	44,0	86
3	84,0	51,6	90,7
4	82,9	48,4	90,0
5	84,0	52,0	90,4
6	78,0	33,9	86,8
7	86,2	58,1	92,0
8	82,5	47,3	89,5
9	78,9	36,7	87,3
10	87,3	61,3	92,7
11	82,5	47,3	89,5
12	80,2	40,6	88,1
13	90,6	71,0	94,7
14	85,1	54,8	91,3
Mean	82,5	49,1	89,6
SD	3,9	9,7	2,6

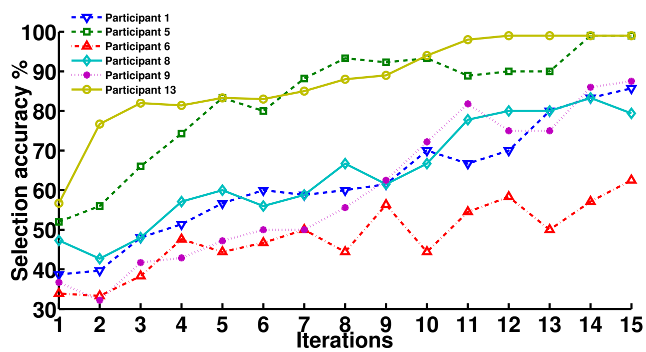


Fig. 3. Selection scores, for auditory stimuli (stimulus duration 1500 ms, ISI 250 ms), plotted as a function of the number of iterations for the users 1, 5, 6, 8, 9 and 13.

the selection score limit, and only subjects number 6 and 11 are below. Subject 6, with the minimum selection value, does not reach the 70% selection score. The maximum selection score achieved is 99% and the minimum is 32.3%. Selection scores are comparable to those achievable with visual and auditory P300 spellers [16].

B. ITR performance

ITR performances are shown in Fig. 6 for six participants. When using multiple iterations, ITR for most subjects went down quickly as shown in Fig. 6. This is a consequence of the classification speed (V) reduction: since ISI is 250 ms and stimulus duration is 1500 ms, each additional iteration increases the classification time of 1750 ms by n^o trials within each run (i.e 7.5 s).

The worst ITR is 0.2 $bits/min$, the best result is 3.8 $bits/min$, which is achieved by the seventh participant. The average value of iterations needed to reach the 70% selection score is 5. At this iteration value, ITR is 1.3 $bits/min$ as

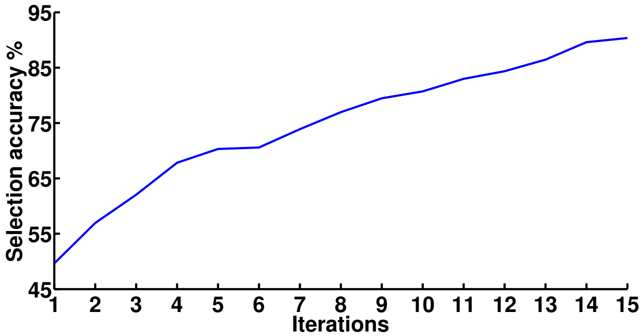


Fig. 4. Mean selection accuracy, for auditory stimuli (stimulus duration 1500 ms, ISI 250 ms), plotted as a function of the number of iterations for fourteen participants.

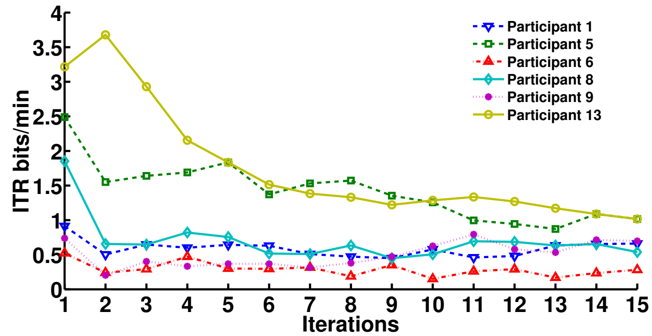


Fig. 6. ITR for auditory stimuli (stimulus duration 1500 ms, ISI 250 ms), plotted as a function of the number of iterations for the subjects 1, 5, 6, 8, 9 and 13.

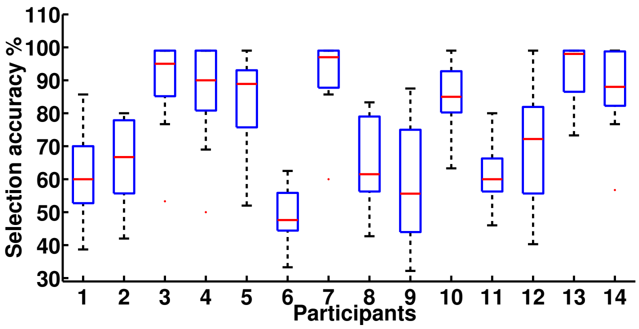


Fig. 5. Selection accuracy boxplot, for auditory stimuli (stimulus duration 1500 ms, ISI 250 ms), of all participants. Boxplot is evaluated with all iterations.

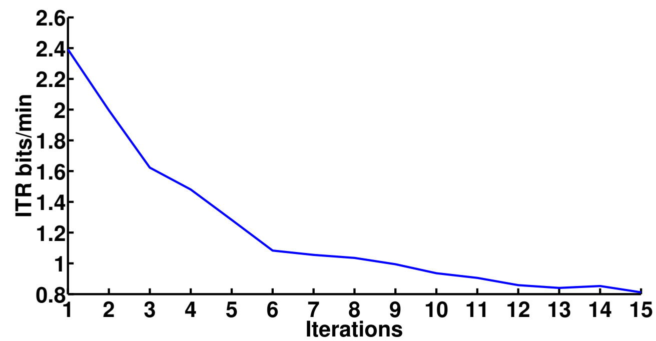


Fig. 7. Mean ITR, for auditory stimuli (stimulus duration 1500 ms, ISI 250 ms), plotted as a function of the number of iterations for fourteen participants.

shown in Fig. 7. Mean ITR for one run is 2.4 $bits/min$, as shown in Fig. 7. When the participants reach the 80% selection score after ten iterations, the ITR is 1 $bits/min$. After fourteen iterations, which corresponds to 90% selection accuracy, ITR is 0.9 $bits/min$. Fig. 8 shows the boxplot of the ITR for all iterations. The boxplot shows that seven subjects are, for all iterations, below 1 $bits/min$ and seven subjects are above this value. The best ITR median is 1.8 $bits/min$, which is achieved by subject 13. Subject 6 shows the worst ITR performances.

ITR are not high compared to visual and auditory P300 spellers [16]. ITR performances, as shown in Eq. (4) and (5), depend from speed and selection accuracy. Speed is related to stimulus duration and ISI. In the present study, the time interval, between the onset of one stimulus until the next, is 1.75 s, that is much higher respect to visual and auditory P300 speller based systems. This high time interval entails a lower ITR but a more natural way of communicating with the user, because the subject has not to pay attention to different tones, timbres or pitches but to single words only. ITR results are comparable to those achievable by auditory P300 in BCI [18] and [19].

V. Conclusion

Visual BCI systems have been intensively researched in the literature, however they can not be adopted by users suffering of visual impairments. Auditory BCI systems represent a valid alternative, even if they yield to lower classification scores. Our contribution was that of developing an auditory BCI system which tries to overcome this obstacle by using a stereo headphone where audio stimuli are virtualized. Sound spatial location can be a cue determining factor for BCI applications. When the user has to focus the attention both on the type and location of the stimulus, then generated P300 signals have an higher amplitude than without audio virtualization.

The proposed sound virtualization procedure is based on Head Related Transfer Function: an impulse response that describes how a sound coming from a well-defined direction is perceived by the human ear. With a set of two of these transfer functions, one for each ear, any direction of sound source propagation can be synthesized. The classification is realized using Support Vector Machines based on gaussian radial basis functions used as kernel functions.

The results obtained with a group of 14 users show that classification and selection accuracy are comparable to those achievable with visual BCI systems. However the average bit rate is lower than that of visual BCI and BCI using tonal audio stimuli. This is due to the typology and duration of the chosen audio stimuli, namely the following spoken words: “bathroom”, “bedroom”, “kitchen”, “help” and “stop”. The use of words of the common language in

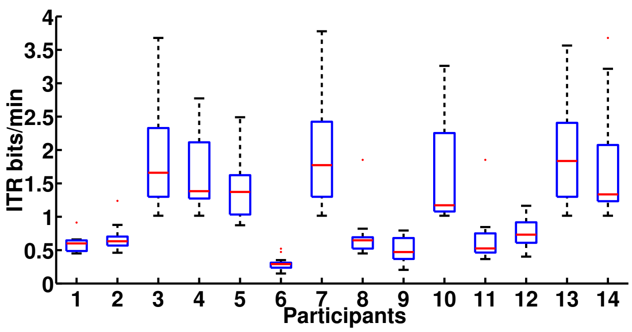


Fig. 8. ITR boxplot for auditory stimuli (stimulus duration 1500 ms, ISI 250 ms), of all participants. Boxplot is evaluated for all iterations.

auditory BCIs can lead to a straightforward communication paradigm, reducing at the same time the training time needed to use the BCI correctly. The authors preferred usability over bit rate since auditory BCIs are typically designed for people with severe physical impairments, which cause the impossibility of moving the eyes (e.g. patients affected by Amyotrophic Lateral Sclerosis). In these cases short training times and system simplicity may be more important than communication data rate.

The authors are currently considering two possible future developments for the auditory BCI paradigm. The first is related to study all the possible experiment variations (e.g.: to focus the attention both on the target word and its spatial location, to increase/decrease the number of sound directions, etc . . .) in order to evaluate both the classification performances and usability of the different solutions. The second one is related to the testing phase and to obtain a more accurate set of indices by increasing the number of users. Both the aspects are under investigation and should provide an improvement to the system and a better comparison of the system to those which have already been presented in the BCI literature.

REFERENCES

[1] N. Birbaumer and L. Cohen, "Brain-computer interfaces: communication and restoration of movement in paralysis," *The Journal of physiology*, vol. 579, no. 3, pp. 621–636, 2007.

[2] D. J. McFarland and J. R. Wolpaw, "Brain-computer interfaces for communication and control," *Communications of the ACM*, vol. 54, no. 5, pp. 60–66, 2011.

[3] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 6, no. 113, 2002.

[4] A. Lenhardt, M. Kaper, and H. Ritter, "An adaptive P300-based online brain-computer interface," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 16, no. 2, pp. 121–130, 2008.

[5] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger, "How many people are able to control a P300-based brain-computer interface (BCI)?" *Neuroscience letters*, vol. 462, no. 1, pp. 94–98, 2009.

[6] A. Calvo, A. Chiò, E. Castellina, F. Corno, L. Farinetti, P. Ghiglione, V. Pasian, and A. Vignola, "Eye Tracking Impact on Quality-of-Life of ALS Patients," 2008, pp. 70–77.

[7] J. Van Erp, "Presenting directions with a vibrotactile torso display," *Ergonomics*, vol. 48, no. 3, pp. 302–313, 2005.

[8] M. Thurlings, J. Erp, A. Brouwer, and P. Werkhoven, "EEG-based navigation from a human factors perspective," *Brain-Computer Interfaces*, pp. 71–86, 2010.

[9] A. Brouwer and J. Van Erp, "A tactile P300 brain-computer interface," *Frontiers in neuroscience*, vol. 4, 2010.

[10] M. van der Waal, M. Severens, J. Geuze, and P. Desain, "Introducing the tactile speller: an ERP-based brain-computer interface for communication," *Journal of Neural Engineering*, vol. 9, no. 4, 2012.

[11] T. Hinterberger, N. Neumann, M. Pham, A. Kübler, A. Grether, N. Hofmayer, B. Wilhelm, H. Flor, and N. Birbaumer, "A multimodal brain-based feedback and communication system," *Experimental Brain Research*, vol. 154, pp. 521–526, 2004.

[12] F. Nijboer, A. Furdea, I. Gunst, J. Mellinger, D. J. McFarland, N. Birbaumer, and A. Kübler, "An auditory brain-computer interface (BCI)," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 43–50, 2008.

[13] D.-W. Kim, H.-J. Hwang, J.-H. Lim, Y.-H. Lee, K.-Y. Jung, and C.-H. Im, "Classification of selective attention to auditory stimuli: Toward vision-free brain-computer interfacing," *Journal of Neuroscience Methods*, vol. 197, no. 1, pp. 180 – 185, 2011.

[14] N. Hill, T. Lal, K. Bierig, N. Birbaumer, and B. Schölkopf, "An auditory paradigm for brain–computer interfaces," *Advances in neural information processing systems*, pp. 569–576, 2004.

[15] E. W. Sellers and E. Donchin, "A P300-based brain-computer interface: Initial tests by ALS patients," *Clinical Neurophysiology*, vol. 117, no. 3, pp. 538–548, 2006.

[16] A. Furdea, S. Halder, D. Krusienski, D. Bross, F. Nijboer, N. Birbaumer, and A. Kübler, "An auditory oddball (P300) spelling system for brain-computer interfaces," *Psychophysiology*, vol. 46, no. 3, pp. 617–625, 2009.

[17] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.

[18] D. Klobassa, T. Vaughan, P. Brunner, N. Schwartz, J. Wolpaw, C. Neuper, and E. Sellers, "Toward a high-throughput auditory P300-based brain-computer interface," *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, vol. 120, no. 7, p. 1252, 2009.

[19] M. Schreuder, B. Blankertz, and M. Tangermann, "A new auditory multi-class brain-computer interface paradigm: spatial hearing as an informative cue," *PLoS One*, vol. 5, no. 4, p. e9813, 2010.

[20] R. Sonnadara, C. Alain, L. Trainor, *et al.*, "Effects of spatial separation and stimulus probability on the event-related potentials elicited by occasional changes in sound location," *Brain research*, vol. 1071, no. 1, pp. 175–185, 2006.

[21] W. M. Hartmann, "Localization of Sound in Rooms," *J. Acoust. Soc. Amer.*, vol. 74, no. 5, pp. 1380–1391, 1983.

[22] J. Blauert, *Spatial Hearing - Revised Edition: The Psychophysics of Human Sound Localization*. MIT Press, 1996.

[23] J. S. Bradley and G. A. Soulodre, "The Influence of Late Arriving Energy on Spatial Impression," *J. Acoust. Soc. Amer.*, vol. 97, no. 4, pp. 2263–2271, 1995.

[24] M. Gardner and K. Martin, "HRTF Measurements of a KEMAR Dummy-Head Microphone," MIT Media Lab, Tech. Rep., 1994.

[25] F. Sharbrough, G. E. Chatrian, R. P. Lesser, H. Luders, M. Nuwer, and T. W. Picton, "American Electroencephalographic Society guidelines for standard electrode position nomenclature," *J. Clin. Neurophysiol.*, vol. 8, pp. 200–202, 1991.

[26] G. Schalk, D. McFarland, T. Hinterberger, N. Birbaumer, and J. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[27] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for EEG-based communication," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.

[28] V. N. Vapnik, *Statistical learning theory*, 1st ed. Wiley, 1998.

[29] V. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, 1999.

[30] J. Wolpaw, H. Ramoser, D. McFarland, and G. Pfurtscheller, "EEG-based communication: improved accuracy by response verification," *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 3, pp. 326 –333, 1998.

[31] A. Kübler, N. Neumann, B. Wilhelm, T. Hinterberger, and N. Birbaumer, "Predictability of brain-computer communication," *Journal of Psychophysiology*, vol. 18, no. 2, pp. 121–129, 2004.