P300 Harmonies: A Brain-Computer Musical Interface

Zacharias Vamvakousis
Universitat Pompeu Fabra
zacharias.vamvakousis@upf.edu

Rafael Ramirez
Universitat Pompeu Fabra
rafael.ramirez@upf.edu

ABSTRACT

We present P300 harmonies: a P300-based Brain-Computer Musical Interface. Using a commercial low-cost EEG device, the user can voluntarily change the harmony of an arpeggio by focusing and mentally counting the occurrences of each note. The arpeggio consists of 6 notes separated by an interval of 175ms. The notes of the arpeggio are controlled through 6 switches, where each switch has two possible states: up and down. When a switch is in the up-state the note produced by this switch is one tone or semitone -depending on the switch- higher than when in the down-state. By focusing on each of the notes of the arpeggio, the user may change -after 12 repetitions- the state of the corresponding switch. The notes of the arpeggio appear in a random order. The state of each switch is shown on a screen. Each switch flashes when the corresponding note is heard. The user can either focus exclusively on the auditory presentation or make use of the visual presentation as well. The interface was presented in a live performance, where the user was able to successfully change the state of all switches with 100% accuracy. An additional preliminary evaluation was performed with 3 more users, in which the selection accuracy was 83.33%.

1. INTRODUCTION

A Brain-Computer Interface (BCI) works by capturing the user's brain activity and converting it to meaningful information in order to control a computer. Most BCIs are built using the the electroencephalogram (EEG). An EEG device captures the electromagnetic activity of the brain's cortex, using electrodes in touch with the skin of the user's scalp. During the last decade a few commercial, low-cost EEG devices have made EEG technology more accessible (Emotiv EEG system, Neurosky). The target group the could benefit the most from the development of BCIs is this of people with severe physical disabilities, such as patients with locked-in syndrome.

Using existing BCI applications someone can perform various tasks, such as controlling a wheel chair, writing, drawing, browsing the internet, playing computer games or controlling musical parameters of an interface [1]. Several Brain-Computer Interfaces for controlling musical parameters have been proposed in previous research. The first proposed musical Interface was *Music for Solo Performer* [2], in the 1960s by Alvin Lucier. The amplified EEG signals were driven to loudspeakers. The vibrations caused were triggering sounds though a set of percussive instruments attached in the loudspeakers. This case though, is better described as a sonification of the brain activity, rather than a BCI. An attention-based BCI was first proposed by David Rosenboom [3]. In this interface EEG components, related with the shifts in the selective attention of the user, were introduced as parameters in a generative music system. It is uncertain though whether the features used were indeed related to the user's selective attention. Many approaches propose the direct mapping of certain EEG bands to musical parameters [4, 5]. In these approaches though, the amount of control the user has over the interface is questionable. It would require extensive training for a user to be able to manipulate his own brain's activity. The limits between a biofeedback interface and an interface where the user voluntarily controls its functions are not always clear.

Probably the most robust way of building a voluntarily controlled BCI that wouldn't require almost any training on behalf of the user, is through the P300 potential. The P300 potential is a positive deflection of the captured electromagnetic activity, 300ms after a rare or unexpected event is perceived, centred around the vertex of the cortex and spread all over the cortex. In a multi-class P300-based BCI, a number of stimuli are presented to the user in a random order and the user draws his attention to a specific stimulus (usually by mentally counting its occurrences). After a number of repetitions of each stimulus, the system is able to predict on which stimulus the user was focusing on. The nature of the stimulus might be visual, auditory, tactile or combination of these. By altering his attention to different stimulus the user is able to perform different actions.

The most well-known P300-based multi-class BCI is the P300 speller proposed in 1988 by Farwell and Donchin [6]. In the typical P300-speller paradigm the user stares at a screen where the characters are placed on a grid. As the characters are flushing in a random order, the user focuses on the character he/she wants to spell. Every time the attended character flashes, a P300 potential is generated. After a number of repetitions, the character that causes the stronger P300 peaks is classified by the system as the attended character. Implementations of the P300-speller have also been proposed using auditory instead of visual stimuli[7, 8].

Copyright: ©2014 Zacharias Vamvakousis et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Apart from typing, a big variety of P300-based BCIs -targeted mainly for locked-in patients- has been proposed, such as controlling the mouse cursor [9], controlling an internet browser [10], controlling a wheelchair [11], painting [12], or controlling musical interfaces [13, 14].

In ICMC 2008, it was presented a P300-based BCI where the user selects the midi-note number placed on a grid, in a similar way a user spells letters in the P300 speller. The maximum speed achieved among 5 subjects was one note every 7 seconds.

Another P300 based BCI proposed [14], integrates the idea of the P300 speller in a music 8x8 step sequencer. The notes of the sequencer are flashing in a random order, and the user selects them as he/she would select letter in the speller. At the same time the melody produced by the sequencer is played back.

These last two proposed interfaces use visual stimuli for controlling the musical interface. In the current paper we propose a P300-based Brain-Computer Musical Interface (BCMI) where the produced musical outcome is at the same time the stimuli that evokes the P300 potentials. This interface can be controlled using just the auditory modality.

2. MATERIALS AND METHODS

2.1 Materials

The Emotiv Epoc [1] 14-channel EEG commercial device was used for capturing the brain activity. The Emotiv Epoc EEG device is targeted for gaming purposes. It is proven though that it is capable of capturing reliable P300 potentials [15, 16]. The signal processing and classification process were performed using OpenVibe software [17]. Using the VRPN server object, stimulations are sent from OpenVibe to a c++ application implemented in openframeworks toolkit [2]. The openframeworks application was used to visualize the interface and send midi messages through LoopBe [3] virtual midi port to propellerhead Reason 5.0 [4] for sonifying a synthesizer. The system was tried on a laptop with a 2.53GHz i5 460M processor with 4GB of RAM running windows 7 OS, using the laptop's internal Realtek ALC269 sound card. The resulting latency of the sound stimuli was 46ms.

2.2 The Interface

The interface consists of an arpeggio of six notes that is continuously being played back. The notes of the arpeggio sound in a random order. The arpeggio consists of 6 notes separated by an interval of 175ms. The notes of the arpeggio are controlled through 6 switches, where each switch has two possible states: up and down. When a switch is in the up-state the note produced by this switch is one tone or semitone -depending on the switch- higher than when in the down-state. By focusing on each of the notes of the arpeggio, the user may change -after 12 repetitions- the state of the corresponding switch. The state of each switch is shown on a screen (see figure 1). Each switch flashes when the corresponding note is heard. The user can either focus exclusively on distinguishing the desired sound or focus as well on the flashings of its corresponding switch. When all notes of the arpeggio have sounded 12 times, the background colour of the screen changes, indicating that the user can then focus on the next sound he desires to change.

Figure 1. From each switch the user can select between two possible notes. The selected note of each switch is highlighted in blue color. When the program starts, all switches are placed down.

In figure 1 are shown the notes assigned to each switch. When all switches are placed in the down-position, the resulting arpeggio consists of the notes G3 (sol in the 3rd octave), B3, D4, F#4, B4, D5, resulting in a G Major seventh chord, while when all switches are in the up-state, the arpeggio consists of the notes A3, C4, E4, G4, C5, E5, resulting in a A minor/minor seventh chord. Stereo spatialization is applied to the notes: the low pitch notes are placed to the left while as the pitch goes higher, the spatialization moves to the right.

The interface has been tried so far with a sound of a harp. By switching his attention to the notes of the arpeggio , the user can build a big variety of possible harmonies. The advantage of the proposed interface, when compared to previously proposed P300-based Musical Interfaces is that it can depend only on the auditory modality: the users changes the music, only by listening to it. Moreover, there is no time interval between the trials, resulting in a continuous musical outcome.

2.2.1 Classification Process

Before using the interface, the xDawn algorithm for Enhancing Evoked Potentials and a 2-class Linear Discriminate Analysis classifier have to be trained. The user is comfortably seated in a chair, in from of a a screen. He/she is asked to remain still and avoid as much as possible swallowing or moving any facial muscle. The brain signals are captured and transmitted wirelessly using the Emotiv Epoc headset. At the beginning of a training session one of the 6 notes of the arpeggio is played back to the user. After a small interval of 3 seconds, the stimuli are presented in a random order, under the constraint that at least one note interferes between two occurrences of the same note. The user is asked to mentally count the occurrences of the presented target-stimulus. A stimulus consists of the sound of the note, along with a blink on the screen of duration 100ms of the corresponding switch. The Inter-Stimulus-Interval (ISI) is set to 175ms. All stimulus are presented 12 times, until the next target stimulus is presented to the user. This process is repeated 6 times -one for each stimulus-.

---

[1] `http://emotiv.com/`

[2] `http://www.openframeworks.cc/`

[3] `http://www.nerds.de/en/loopbe1.html`

[4] `http://www.propellerheads.se/products/reason/`

As a result, the training data consist of 432 epochs, 72 of which are target epochs.

An epoch consists of the 14-channel recording of the time interval 250 to 750ms after the presentation of a stimulus. The signal is downsampled to 32Hz and band-pass filtered to 1-12Hz. Using the xDAWN [18] Spatial Filter Trainer in Openvibe, a 14 to 3 channels spatial filter is acquired. The 48 resulting values per epoch are then used to train a two-class Linear Discriminate Analysis Classifier (LDA) to distinguish target from non-target epochs (figure 2. Once the spatial filter and the LDA classifier parameters are acquired, the use might start using the interface.

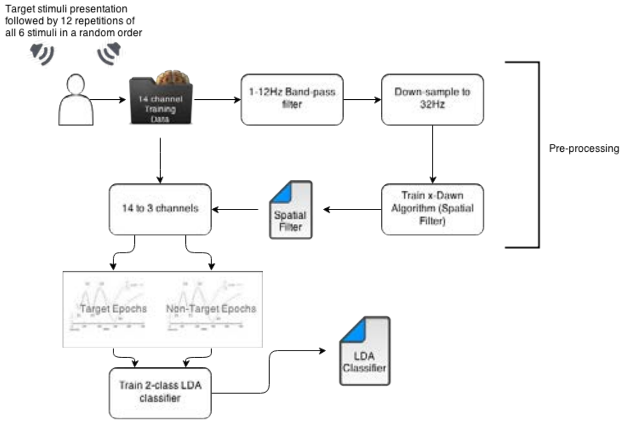

Figure 2. Aqcuiring the Spatial Filter and LDA classifier.

During the on-line session, the features per epoch, are being produced as in the training session. Then, for each stimulus a voting classifier computes the sum of the hyperplane distances -given by the LDA classifiers-, and outputs as the attended stimulus the one with the lowest sum 3.

An evaluation of the accuracy of the described classification process is being reported in a previous publication [19].

2.2.2 Controlling the Interface

In the initial state all switches are placed down. Once the arpeggio starts being reproduced, every 72 notes (12 occurrences of each one of the 6 stimuli), the background colour of the screen changes, indicating that the user might then attend the next note he/she wishes to change. After about 1 second the voting classifier outputs the detected target stimulus, changing the state of the corresponding switch. As a result a different harmony is being produced by the arpeggio. This process, allows a continuous playback of the musical outcome of the interface. The number of trials -that determines the duration of the performance- has to be determined at the beginning of the session.

3. EVALUATION

The interface was evaluated with 4 subjects (3 male). After training the system -as described in paragraph 2.2.1- they were asked to move all switches up, starting from the leftmost one and moving to the one in the right. The average age of all subjects was 35 years. The only female subject performed the task in an exhibition setting, using loudspeakers for sound generation achieving 100% accuracy[5]. The 3 remaining subjects were asked to perform the same task in an office environment, using in-ear headphones. The accuracy was 6/6, 4/6 and 5/6. All subjects used both the visual and auditory modality of the interface to control the interface.

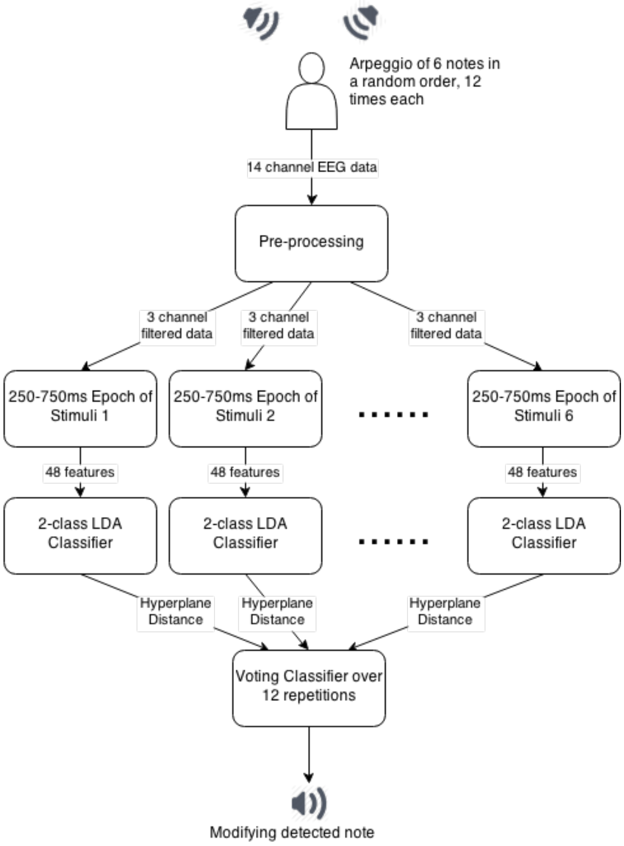

Figure 3. Classification in the on-line session by summing hyper-plane distances of the LDA classifiers of each stimulus.

4. DISCUSSION AND FUTURE WORK

We presented a first prototype of a P300-based BCMI. The novelty of the proposed interface lies in the fact that the user voluntarily interacts with the music while listening to it. The idea of incorporating the stimuli presentation of a re-active BCI with the produced musical outcome could create interesting attention-based musical compositions. In such interfaces the stimuli presentation should be part of the produced music. The limitation that a P300-based auditory BCMI introduces is that the stimuli should be presented in a random order. Even with this limitation though, interesting musical interfaces can be designed. Such interfaces could be useful for some cases of locked-in patients.

In the proposed BCI, the stimuli presentation of a trial starts before presenting the outcome of the preceding. Due to this fact, the Information Transfer Rate of the system increases when compared to a system where a time interval is introduced between the trials. The average ITR among

---

[5] Video of the performance at: `https://vidd.me/roV`

all 4 subjects was 7.37 bits/min, while in the case of the 2 subjects that performed with 100% accuracy the achieved ITR was 12.31 bits/min[6]. If this idea is combined with a algorithm that detects and corrects the possible mistakes, the information transfer rate of a system could rise. For example, in the case of a speller, a misspelled letter could be automatically replaced by the correct one. In this case the user wouldn't have to cancel a wrong choice. As a result the continuous stimulus presentation could be used, increasing the spelling speed.

The performer of the proposed BCMI has control over the chords produced by an arpeggio. Similar interfaces could be designed, in which the system responds in different ways to the user's selective attention. For example when attending a note instead of changing just this note, the whole harmony could change. The system could propose 6 (if 6 is the number of stimuli) different chord on each step, depending on the previous chords. In such a system, the person using the BCMI could accompany another musician that would make a solo on the performed harmonies. In the future, the proposed system should be tested using only the auditory modality in the stimuli design.

5. REFERENCES

[1] B. Blankertz, M. Tangermann, and K.-R. Müller, "Bci applications for the general population," *Brain-Computer Interfaces - Principles and Practice*, pp. 363–372, 2012.

[2] A. Lucier, "Statement on: Music for solo performer," *Biofeedback and the Arts: Results of Early Experiments*, 1976.

[3] D. Rosenboom, "The performing brain." *COMP. MUSIC J.*, vol. 14, no. 1, pp. 48–66, 1990.

[4] T. Hinterberger and G. Baier, "Poser: Parametric orchestral sonification of eeg in real-time for the self-regulation of brain states," *IEEE Trans. Multimedia*, vol. 12, p. 70, 2005.

[5] S. Mealla, A. Väljamäe, M. Bosi, and S. Jordà, "Listening to your brain: Implicit interaction in collaborative music performances," in *The International Conference on New Interfaces for Musical Expression (NIME)*, 2011.

[6] 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*, 1988.

[7] M. Schreuder, T. Rost, and M. Tangermann, "Listen, you are writing! speeding up online spelling with a dynamic auditory bci," *Frontiers in neuroscience*, vol. 5, p. 112, 2011.

[8] I. Käthner, C. A. Ruf, E. Pasqualotto, C. Braun, N. Birbaumer, and S. Halder, "A portable auditory p300 brain–computer interface with directional cues," *Clinical neurophysiology*, vol. 124, no. 2, pp. 327–338, 2013.

[9] L. Citi, R. Poli, C. Cinel, and F. Sepulveda, "P300-based bci mouse with genetically-optimized analogue control," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 16, no. 1, pp. 51–61, Feb 2008.

[10] E. Mugler, C. Ruf, S. Halder, M. Bensch, and A. Kubler, "Design and implementation of a p300-based brain-computer interface for controlling an internet browser," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 18, no. 6, pp. 599–609, Dec 2010.

[11] B. Rebsamen, C.-L. Teo, Q. Zeng, V. Ang, E. Burdet, C. Guan, H. Zhang, and C. Laugier, "Controlling a wheelchair indoors using thought," *Intelligent Systems, IEEE*, vol. 22, no. 2, pp. 18–24, March 2007.

[12] B. van de Laar, I. Brugman, F. Nijboer, M. Poel, and A. Nijholt, "Brainbrush, a multimodal application for creative expressivity," in *ACHI 2013, The Sixth International Conference on Advances in Computer-Human Interactions*, 2013, pp. 62–67.

[13] M. Grierson, "Composing with brainwaves: minimal trial p300b recognition as an indication of subjective preference for the control of a musical instrument," in *Proceedings of International Cryogenic Materials Conference (ICMC08)*, 2008.

[14] Y. C. D. Chew and E. Caspary, "Museegk: Design of a bcmi," in *Proceedings of the 8th ACM conference on Creativity and cognition*. ACM, 2011, pp. 325–326.

[15] M. Duvinage, T. Castermans, M. Petieau, T. Hoellinger, G. Cheron, and T. Dutoit, "Performance of the emotiv epoc headset for p300-based applications," *Biomedical engineering online*, vol. 12, no. 1, p. 56, 2013.

[16] N. A. Badcock, P. Mousikou, Y. Mahajan, P. de Lissa, J. Thie, and G. McArthur, "Validation of the emotiv epoc® eeg gaming system for measuring research quality auditory erps," *PeerJ*, vol. 1, p. e38, 2013.

[17] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer, "Openvibe: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments," *Presence: teleoperators and virtual environments*, vol. 19, no. 1, pp. 35–53, 2010.

---

[6] The ITR value (bits/min) is computed using the formula:

$$ITR = S \cdot \left[log_2(N) + P \cdot log_2(P) + (1-P) \cdot log_2 \left(\frac{1-P}{N-1} \right) \right]$$

Where, S represents the number of selections per minute, N represents the number of possible targets and P represents the probability that they are correctly classified

[18] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, “xdawn algorithm to enhance evoked potentials: application to brain–computer interface,” *Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 8, pp. 2035–2043, 2009.

[19] V. Zacharias and R. Ramirez, “A high-throughput auditory p300 interface for everyone,” in *12th European AAATE Conference*, IOS Press. Portugal, Volamoura: IOS Press, 19/09/2013 2013, pp. 478–482.