CrossMark

RESEARCH ARTICLE

An exploration of spatial auditory BCI paradigms with different sounds: music notes versus beeps

Minqiang Huang[1] · Ian Daly[2] · Jing Jin[1] · Yu Zhang[1] ·
Xingyu Wang[1] · Andrzej Cichocki[3,4]

Abstract Visual brain-computer interfaces (BCIs) are not suitable for people who cannot reliably maintain their eye gaze. Considering that this group usually maintains audition, an auditory based BCI may be a good choice for them. In this paper, we explore two auditory patterns: (1) a pattern utilizing symmetrical spatial cues with multiple frequency beeps [called the high low medium (HLM) pattern], and (2) a pattern utilizing non-symmetrical spatial cues with six tones derived from the diatonic scale [called the diatonic scale (DS) pattern]. These two patterns are compared to each other in terms of accuracy to determine which auditory pattern is better. The HLM pattern uses three different frequency beeps and has a symmetrical spatial distribution. The DS pattern uses six spoken stimuli, which are six notes solmizated as "do", "re", "mi", "fa", "sol" and "la", and derived from the diatonic scale. These six sounds are distributed to six, spatially distributed, speakers. Thus, we compare a BCI paradigm using beeps with another BCI paradigm using tones on the diatonic scale, when the stimuli are spatially distributed. Although no significant differences are found between the ERPs, the HLM pattern performs better than the DS pattern: the online accuracy achieved with the HLM pattern is significantly higher than that achieved with the DS pattern ($p = 0.0028$).

Keywords Auditory BCI · Auditory pattern · Spatial cues · Diatonic scale

Introduction

Brain-computer interfaces (BCIs) are communication systems, which can be used to send messages or commands to the external world without the brain's normal output pathways of peripheral nerves and muscles (Freeman 2007a, b; Wang et al. 2014a, b; Wolpaw et al. 2002).The main goal of BCI technology is to improve the quality-of-life of people who are locked-in (e.g., by end-stage amyotrophic lateral sclerosis brainstem stroke, or severe polyneuropathy) or lack any useful muscle control (e.g., due to severe cerebral palsy) (Daly et al. 2013; Farwell 2012; Freeman 2007a, b; Hoffmann et al. 2008). Residual senses, such as vision (Jin et al. 2011, 2014a, b; Wang et al. 2014a, b), audition (Daly et al. 2014), or touch (Kodama et al. 2014) can be utilized to control a BCI. One of the most frequently used brain potentials in BCI is the P300 event-related potential (ERP) (Hwang et al. 2013), which is an enhanced positive EEG amplitude potential with a latency of about 300 ms, that can be evoked by specific events (Farwell and Donchin 1988). Farwell and Donchin (1988) proposed the first visual P300-based BCI, called the P300 speller. Since then, many studies have been done to

✉ Jing Jin
jinjingat@gmail.com

✉ Xingyu Wang
xywang@ecust.edu.cn

Minqiang Huang
wharu@mail.ecust.edu.cn

1 Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, People's Republic of China

2 Brain Embodiment Lab, School of Systems Engineering, University of Reading, Reading, UK

3 Lab for Advanced Brain Signal Processing and BTCC, Riken BSI, Wako-shi, Japan

4 Skolkowo Institute of Science and Technology, Moscow, Russia

improve the performance of visual P300-based BCIs (Li et al. 2016; Yin et al. 2013).

However, people in different situations may experience some limitations in using visual BCI systems. For example, for people who have weak vision or who cannot control their eye movements but keep audition, a visual BCI can prove challenging or unusable, while auditory BCIs can prove to be effective (Guo et al. 2010; Laghari et al. 2013; Ide et al. 2013). Moreover, some amyotrophic lateral sclerosis (ALS) patients can usually still hear clearly (Hayashi and Kato 1989). In view of this, some studies have focused on auditory-based-BCI systems, which can be used by these patients. Although an auditory P300 speller tested with four locked-in patients showed worse performance compared with a visual P300 speller (Kübler et al. 2009), later studies have shown the practicality of the auditory BCI paradigm (Dominguez et al. 2011; Wang and Chang 2008).

A P300 auditory brain-computer interface that utilized a 6 × 6 speller and six different environmental sounds representing different columns or rows was presented by Klobassa et al. (Klobassa et al. 2009). It demonstrated the high speed of the proposed system, with maximum bit rates reaching about 2 bits/min.

Halder et al. (2010) showed that a paradigm with targets varying in pitch achieved the best performance, while a paradigm with targets varying in direction performed relatively worse. Höhne et al. (2011) developed an auditory BCI speller which used three tones varying in pitch and direction and presented via headphones. Similarly, Käthner et al. (2013) proposed an auditory BCI system, which used pitch and direction as cues, and also investigated different inter-stimulus intervals (ISIs). It was found that, for this auditory BCI, the best performance was achieved with an ISI under 400 ms. Auditory spatial paradigms have also been well researched in other studies (Belitski et al. 2011; Lelievre et al. 2013; Rutkowski et al. 2010).

Moreover, Muller-Putz et al. (2012) evaluated a novel auditory single-switch BCI (ssBCI) in nine individuals who were in a minimally conscious state (MCS). The task consisted of a simple and difficult oddball pattern that included two different tone streams with rarely appearing deviant tones. The simple pattern contained one tone stream, while the difficult pattern contained two tone streams played at the same time. The results showed that this auditory BCI could work for MCS patients. With enough training time, an auditory BCI might perform as well as a visual BCI (Nijboer et al. 2008).

Some studies utilized the auditory steady state response (ASSR) in BCI paradigms and showed the usefulness of the ASSR (Higashi et al. 2011; Kim et al. 2011).

Although different kinds of auditory patterns have been investigated, some problems still exist, such as front-back confusion [the phenomenon of a listener giving a response to the opposite location to the stimulus in the vertical plane (Makous and Middlebrooks 1990)] in multi class auditory BCIs which use spatial cues (Schreuder et al. 2011), and the low information content of the binary decision auditory BCI (Schreuder et al. 2010). Therefore, we explored a new auditory BCI system, which was designed to improve performance and avoid the front-back confusion.

In this paper, we compared two patterns, which were named the HLM (High Low Medium) pattern and the DS (Diatonic Scale) pattern respectively. The frequencies of human speech and sung vocalization are in the low frequency range below 1000 Hz (Smith and Price 2014). Thus, the HLM pattern used three different frequency beeps at low (200 Hz), medium (500 Hz), and high (1000 Hz) frequencies, to compare to the six notes of the diatonic scale in the DS pattern, which were pronounced as "do", "re", "mi", "fa", "sol", and "la". It had been shown that a paradigm with multiple spatially distributed speakers performed better than a paradigm using a single speaker (Schreuder et al. 2010). Accordingly, six loudspeakers were arranged to play sounds in both paradigms. In this study, we wanted to compare an auditory BCI paradigm that used beeps with a paradigm that used music notes.

Methods

Participants

Ten healthy volunteers (7 males and 3 females, ages 23–25, mean age 24.3 ± 0.67) participated in this study, which was approved by the Human Research and Ethics Committee of East China University of Science and Technology. Nine of the participants were right handed, and one male was left handed. All the participants were given information about the experiments and signed consent forms beforehand. They reported no problems of hearing or neurological disorders now or previously. Their participation was paid 30 RMB per hour.

Auditory stimuli

The stimuli used in the HLM pattern were three pure tones in different frequencies. The tones were generated with 200 Hz (low), 500 Hz (medium) and 1000 Hz (high) frequencies by Matlab. The frequencies were chosen from a subjective perspective to maximize the differences of the tones. The six notes in the DS pattern were C3, D3, E3, F3, G3, and A3, which were solmizated as "do", "re", "mi", "fa", "sol" and "la". These notes were synthesized using the Luotianyi VOCALOID 3 Editor (Yamaha). All

auditory stimuli were unified to the same loudness and duration using Adobe Audition version 3.0.

In the HLM pattern, a symmetrical play mode was executed. Six speakers were arranged in a semicircle symmetrically around the participant (see Fig. 1). The high tone was presented by speakers located at directions −90° or 90°; the low tone was presented by speakers located at directions −60° or 60°; and the medium tone was presented by speakers located at directions −30° or 30°. In the DS pattern, six notes were presented. The note "do" was presented by the speaker located at −90° and the remaining five notes were distributed to the other speakers in sequence. In both patterns, six kinds of stimuli were presented and the possibility of a given stimuli being the target was 1/6.

Each speaker presented only one kind of sound in both patterns. Each stimulus lasted 200 ms and the stimulus onset asynchrony (SOA) was 400 ms. Stimuli were presented in a random sequence. The next stimulus would not begin until the present stimulus ended. The direction and content of the stimuli were the major features to differentiate them.

Experimental design set up

EEG was recorded with wet active electrodes (Ag–AgCl) via a 64-channel 'g.EEGcap' EEG cap (Guger Technologies, Graz, Austria). We measured the EEG signals with a 'g.HIamp' amplifier (Guger Technologies, Graz, Austria) with a sensitivity of 100 µV, band pass filtered between 0.1 Hz and 100 Hz, and sampled at 512 Hz (Laghari et al. 2013). This device uses wide-range DC-coupled amplifier technology in combination with 24-bit sampling. Data were recorded from a subset of 24 selected electrodes from a larger set of 64 electrodes, which were placed in accordance with the international 10–20 system. These were located at positions Fz, C5, C3, Cz, C4, C6, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, PO7, PO3, POz, PO4,

PO8, O1, Oz, and O2 (see Fig. 2) (Treder et al. 2014). The electrode on the right earlobe was chosen as the reference, and the frontal electrode (FPz) was chosen as the ground. The impedance level of all electrodes was less than 30 KΩ.

The stimuli were played through programs designed in Qt Creator (version 5.3) and controlled by Matlab (version R2012a).

In the proposed design, a circular speaker arrangement was used (Schreuder et al. 2010). However, it was found that severe direction confusion resulted to an unacceptable accuracy with using a circular arrangement in the testing phase. So we changed the circular arrangement to a semi-circle with a radius of 1.5 m round the participant (see Fig. 1), a similar design which was used in Belitski et al. (2011). The angle interval of the speakers on the same side was 30°.

Offline and online experiment task

In order to train the classification models for each participant, an offline experiment was implemented first. Each participant sat in a comfortable chair with a relax posture and was asked to silently count the number of target stimuli in both patterns. Before the experiment, participants were given an oral introduction in Mandarin, as follows: "In the experiments, you will hear an audio-cue, which comes from one of the six speakers; the direction of the cue is the target direction and the sound appearing in the cue is the target sound, these are what you need to focus on. The task is to silently count the number of target sounds from the target direction. Sounds from other speakers are non-target stimuli, which you should ignore as much as possible."

In the offline experiment, there was an audio-cue (lasting about 1.5 s) before each run, which showed the target to participants. The target stimulus was the sound, which came from the same speaker as the audio-cue. For example, if the speaker located at −90° played the audio-cue, it meant the stimuli was the sound that came from the speaker

Fig. 1 The location of each loudspeaker and the parameters of the scene
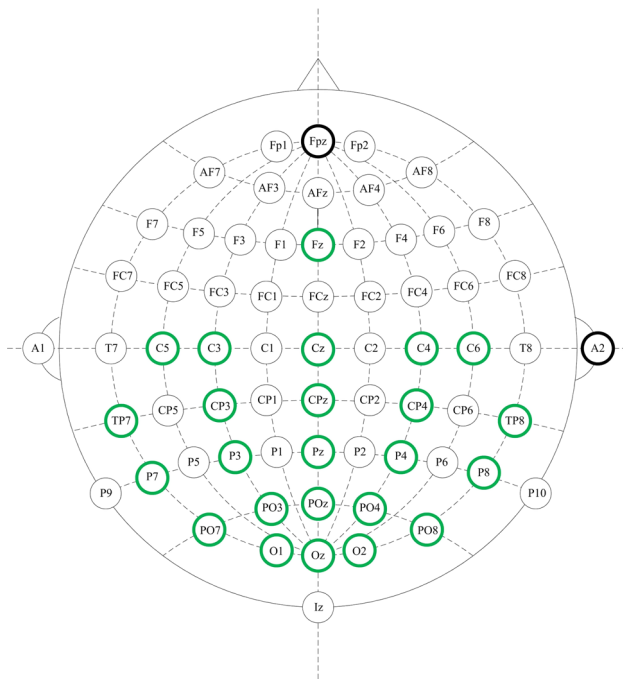
Volume: 60dB
Radius: 1.5m
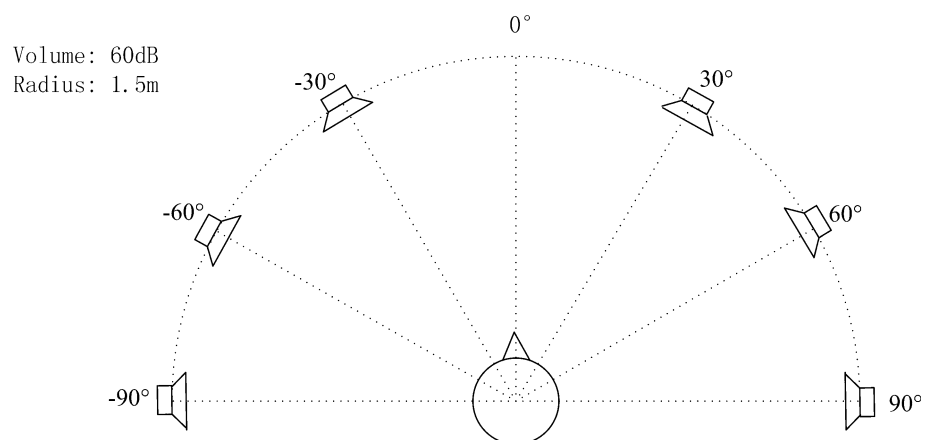
Fig. 2 Configuration of electrode positions. The *green* electrodes indicate recording channels and the *black bold* electrodes (Fpz and A2) indicate the ground and reference channels. (Color figure online)

located at $-90°$ and the content of the audio-cue would be a high beep (1000 Hz) or 'do', depending on which auditory pattern was been executed. In one run, only one sound stimulus was the target and the others were non-target stimuli. Each run contained twelve trials, and each trial contained six stimuli. There were five runs in each block, and there were three blocks in each session. When each run ended, participants were asked to report the number of each target they had counted to ensure they were focused on the task.

After the offline experiment, an online experiment was conducted to evaluate the actual performance of two auditory patterns. The clear differences between offline experiment and online experiment were: (1) During the online experiment feedback was given to the participant after each target selection, while this didn't happen during the offline experiment; (2) The online experiment was faster than the offline experiment because the online experiment needed fewer trials before the selection results could be output. During the online experiment, 24 target selections were carried out for each participant. The 24 target selections were constituted of 6 targets in the offline experiment, which were chosen four times in the online experiment. There were five trials in each run in the online experiment. When participants finished one selection, the selection result was reported by the corresponding speaker. In the online experiment, participants still needed to count

the number of the target stimuli, but they were not asked to report that number. Participants were told that they should ignore the result to avoid affecting their current condition.

Feature extraction procedure

The raw data was filtered with a third order Butterworth band pass filter between 0.1 and 30 Hz to remove high-frequency noise. After filtering, the data was down-sampled from 512 to 64 Hz by selecting every eighth sample from the filtered data. The EEG data corresponding to the first 1000 ms after a stimulus onset was used to extract the feature. The size of each feature matrix was 24×64 (24 channels by 64 time points). 15-fold cross-validation was used to estimate the average classification accuracy. Fourteen fifteenths of the data were used for training a Bayesian linear discriminant analysis (BLDA) classifier and the rest were used to test the classifier.

Classification scheme

We used Bayesian linear discriminant analysis (BLDA) as the classification algorithm, because it can solve the over-fitting problem and get higher classification accuracies compared to Fisher's discriminant analysis (FLDA) (Hoffmann et al. 2008). The details of BLDA can also be found in (Hoffmann et al. 2008).

R-squared value

The r-squared value was calculated as a discriminant index. The r-squared value was equal to the square of $r(x)$, which was defined as (Birbaumer et al. 1999)

$$r(x) = \frac{\sqrt{N_1 N_2}}{N_1 + N_2} \cdot \frac{mean\{x_i | l_i = 1\} - mean\{x_i | l_i = 2\}}{std\{x_i | l_i = 1, 2\}},$$

where N_1 and N_2 denote the number of variables in the class 1 (target) and class 2 (non-target) groups respectively, meanwhile x_i was the ith variable and l_i was the class label of the ith variable.

Online accuracy of each participant

The online accuracy of each participant was calculated as

$$Acc = \frac{T}{F + T},$$

where T and F denote the number of true selection results and the number of false selection results respectively, $F + T$ was always equal to 24 (the number of total target selections in the online experiment). Each selection result was obtained by a BLDA classifier and Acc was the

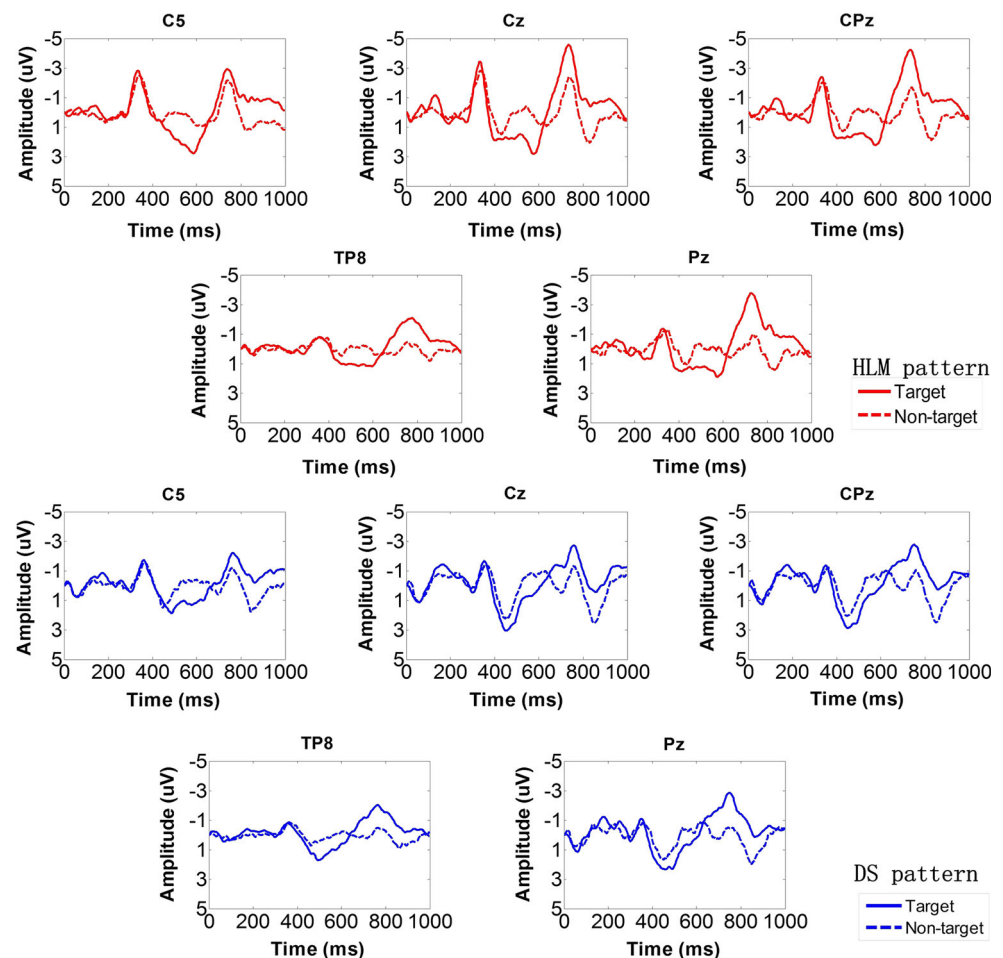accuracy of the BLDA classifier based on the EEG features.

Average accuracy of each direction

To calculate the average accuracy of each direction based on the accuracy above, the number of true and false selection results of each direction for all participants in both patterns was counted. The accuracy was then calculated as:

$$Acc_m = \frac{\sum_{n=1}^{10} T_{mn}}{\sum_{n=1}^{10} (T_{mn} + F_{mn})},$$

where T_{mn} denotes the number of true selection results of the nth participant in the mth direction (for convenience, here directions from left to right were numbered corresponding from 1 to 6 and m denotes the number of the direction), and F_{mn} denotes the number of false selection results of the nth participant in the mth direction.

Questions for feedback

Participants were asked to answer two questions.

1. Do you like these two patterns? Please rate each pattern on a scale of zero to five according to your feeling. The higher the score, the more you like the pattern.
2. Are the two patterns difficult for you? Please rate each pattern on a scale of zero to five according your feeling. The higher the score, the more difficult the pattern.

Experiment results and discussions

Five channels were chosen to show the average amplitudes of target and non-target ERPs in both patterns in Fig. 3. We did not find obvious differences in the early components

Fig. 3 Grand averaged ERP amplitudes across 10 participants in the HLM and DS patterns over five channels: C5, Cz, CPz, TP8, and Pz. *Red solid lines* indicate the target group and *red dashed lines* indicate the non-target group in the HLM pattern. *Blue solid lines* indicate the target group and *blue dashed lines* the non-target group in the DS pattern. (Color figure online)

Fig. 4 The topographic map of r-squared values for the HLM and DS patterns across all 10 participants. The negative class indicates the negative component and the positive class indicates the positive component
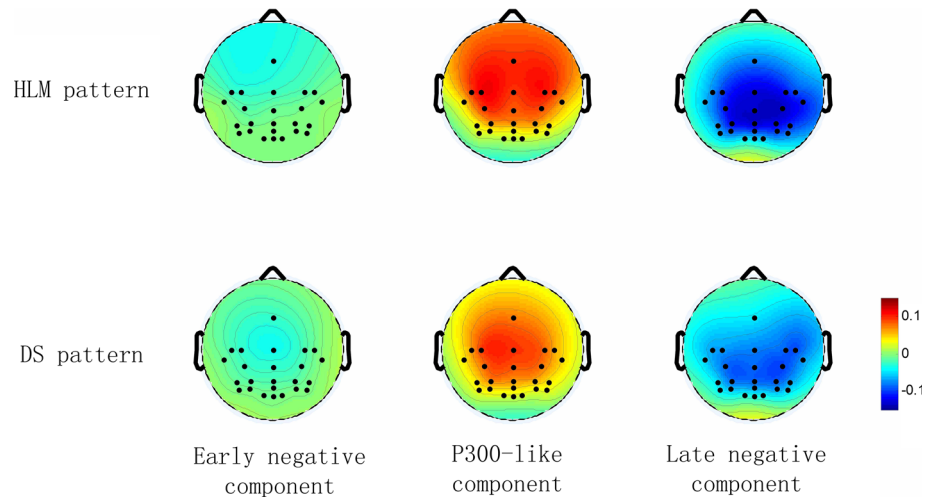
between target and non-target amplitudes in the 0–400 ms time period. In the period from 400 to 600 ms, a more positive component was observed in the target amplitude compared to the non-target amplitude in two patterns. The latency of this component was longer than that of a standard P300 signal and here we considered it as a P300-like component. Following the P300-like component, a distinct negative wave appeared slightly before 800 ms and could be observed in the target group but not in the non-target group. The two components were the major differences between target and non-target groups. The average peak value of the P300-like component across 10 participants was calculated at electrode Cz. Although no significant difference was found between two patterns by a t test, the DS pattern resulted in a higher mean peak value (4.49 ± 1.45) than the HLM pattern (4.11 ± 1.22). In addition, in the non-target condition, the DS pattern was observed to have a higher mean peak value (3.20 ± 1.29) than the HLM pattern (2.58 ± 1.05). The higher peak value in the non-target group might lead to deterioration in classification accuracy.

The same analysis was done on the late negative component. Fz was chosen to compare the difference between the two patterns. The mean peak values of the late negative components were -4.40 ± 1.82 in the HLM pattern and -3.22 ± 1.17 in the DS pattern. Comparing the non-target amplitude in the same period with the late component, a paired t-test was used to show a significant difference between the two patterns ($p = 0.0087$). The mean peak values in the non-target group were -3.43 ± 1.78 in the HLM pattern and -1.98 ± 0.57 in the DS pattern.

To get some insight into the contribution of the EEG features to the classification results, r-squared values were computed to evaluate the contributions to the classification accuracy of the EEG components. Figure 4 shows the r-squared values of the early negative component, the

P300-like component, and the late negative component. The early negative component showed little contribution to the classification accuracy. The P300-like component and late negative component were the critical components to distinguish between targets and non-targets. The topographic map of r-squared values also illustrates the spatial distributions of the EEG features: the P300-like component appeared mostly over the frontal and central region but not over the occipital region; the late component appeared over the central region. This information could provide references to choose appropriate channels without redundancy in a future study.

The online results are shown in Table 1. The average accuracy of the HLM pattern across ten participants was above 70 %, which is regarded by some as the minimum accuracy needed for a useful BCI system (Schreuder et al. 2010), while the DS pattern was considerably worse, with an average accuracy of 60.85 ± 18.9 %. The highest accuracy in the HLM pattern was 91.7 % and three participants' performances reached this level. A t-test was applied to compare the online accuracies and a significant difference was found between the patterns ($p = 0.0028$).

Participant P7 achieved the best accuracy in both two patterns, while participant P8 achieved the worst accuracy. We checked the ERP amplitudes of these two participants at electrode Pz in both patterns and found that the apparent distinction was in the late negative component. Higher amplitudes were observed in the late negative component in participant P7 than P8. Figure 5 presents the target ERPs recorded from participants P7 and P8 at electrode Pz in two patterns. It was apparent that participant P7 had a higher peak than participant P8, especially in the late negative component, which might be one of the reasons that participant P7 achieved better performance than participant P8.

Table 2 shows the average accuracy achieved for each of the six directions by all participants.

Table 1 Online results achieved with the HLM and DS patterns

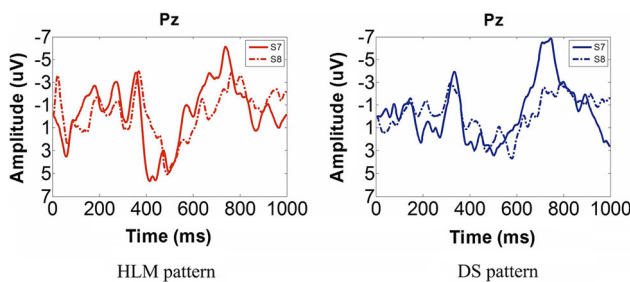| Participant	Age	Sex	Online classification accuracy (%)	
			HLM pattern	DS pattern
P1	25	M	70.8	66.7
P2	23	M	91.7	79.2
P3	24	M	91.7	66.7
P4	24	M	70.8	66.7
P5	25	F	58.3	50
P6	24	F	70.8	54.2
P7	24	F	91.7	83.3
P8	25	M	50	29.2
P9	24	M	87.5	79.2
P10	25	M	75	33.3
Avg			75.83 ± 15.47	60.85 ± 18.9

Fig. 5 The average ERP amplitudes of participants P7 and P8. The *red lines* represent the HLM pattern and the *blue lines* represent the DS pattern. Participant P7 is indicated by the *solid line* and participant P8 is indicated by the *dot-dash line*

Table 2 The average accuracy of each direction of all participants for the HLM and DS patterns

	Accuracy (%)					
	−90°	−60°	−30°	30°	60°	90°
HLM pattern	65	77.5	82.5	80	80	70
DS pattern	47.5	60	65	67.5	57.5	65

The accuracy of each direction was higher when using the HLM pattern than when using the DS pattern. In both patterns, −30° and 30° were the best, while −90° was the worst in terms of accuracy. A *t* test was performed to show that the ERP amplitude of non-target stimuli for the −90° direction at electrode site Pz for the DS pattern was significantly larger than of the corresponding ERP amplitude recorded during the HLM pattern ($t = -3.4$, $p < 0.05$) when the target direction was −90°. The signals induced by the non-target stimuli would lead to a defective influence on the classification accuracy. In this paper, in contrast to the design in the work of Schreuder et al. (2011),
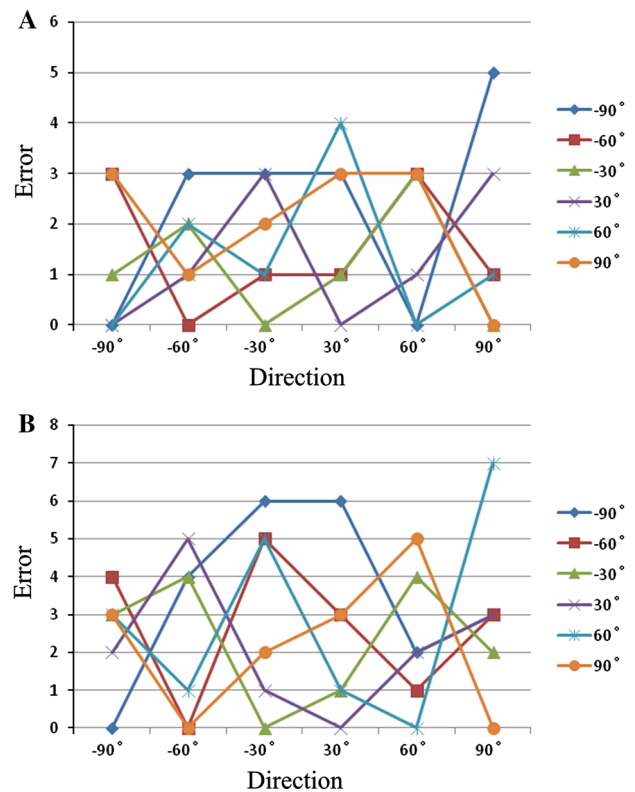

Fig. 6 The *horizontal axis* indicates the six directions of stimuli and the *vertical axis* indicates the number of errors appearing in each non-target direction in the online experiment. **a** The errors in non-target directions during the HLM pattern. **b** The errors in non-symmetrical directions during the DS pattern

we did not have an arrangement of speakers behind the participants, and therefore the front-back confusion was avoided. The errors appeared in the following cases: (1) the symmetrical direction of the target direction like the −90° versus 90°, −60° versus 60° and −30° versus 30°; (2) The neighbors of the target direction, for instance −90° and −60° were neighbors; (3) The directions were not symmetrical and neighboring to the target direction such as −30°, 30° and 60° when −90° was the target direction. The numbers of errors that appeared at symmetrical directions were 17 in the HLM pattern and 10 in the DS pattern. The errors in neighbor directions and non-symmetrical-neighbor directions in the HLM pattern were less than the DS pattern (22 errors for the HLM pattern and 34 errors for the DS pattern of the neighbor directions and 23 errors for the HLM pattern and 52 errors for the DS pattern of non-symmetrical-neighbor directions) (see Fig. 6). The symmetrical sound distribution in the HLM pattern easily led to confusion of the left and right directions for the participants. The fewer sounds and the clearer differences in frequencies between sounds might be the reasons that participants perform better with the HLM pattern than the DS pattern.

Table 3 The scores given by participants for the likability and difficulty of the patterns

	Likability		Difficulty	
	HLM pattern	DS pattern	HLM pattern	DS pattern
P1	4	3	3	4
P2	5	4	1	2
P3	5	3	3	4
P4	4	3	4	2
P5	4	2	2	3
P6	4	5	5	4
P7	5	4	3	4
P8	3	4	3	2
P9	3	4	4	3
P10	3	4	4	4
Avg	4 ± 0.82	3.6 ± 0.84	3.2 ± 1.14	3.2 ± 0.92

The score was scaled from 1 to 5. Higher scores represented higher likability and more difficult patterns

Participant feedback

After all the experiments were completed, we collected feedback from all participants about the two patterns.

The results are presented in Table 3. It may be seen from Table 3 that the two patterns got approximately similar average scores in terms of difficulty, but that the majority of subjects preferred the HLM pattern to the DS pattern. Oral feedback of the participants provided some helpful information to optimize the system. For the HLM pattern, some of the participants noted that the differences between beeps of three frequencies were not clear enough for them to catch targets well, for example, some participants could not discern the medium well. For the DS pattern, some participants expressed the view that there were too many kinds of sounds to remember, leading to confusion. It is important to attempt to improve the auditory stimuli according to this feedback in future work.

In summary, the results demonstrated that the HLM pattern performed better than the DS pattern. However, no significant differences were found between the patterns in terms of brain signals, which meant the music notes might be good auditory stimuli in more appropriate conditions such as changing the way stimuli are presented to make these music notes sound more melodious.

Conclusion

In this paper, an auditory BCI system using different sounds with spatial cues was proposed. The results show that the performance of the HLM pattern was better than the DS pattern. However, the DS pattern showed potential

to become a good music paradigm: No significant difference in the ERPs with HLM pattern meant that the DS pattern was able to induce good ERPs. Our further work will be focused on the optimization of the DS pattern to improve its performance and applicability.

References

Belitski A, Farquhar J, Desain P (2011) P300 audio-visual speller. J Neural Eng 8(2):025022

Birbaumer N, Ghanayim N, Hinterberger T, Iversen I, Kotchoubey B, Kübler A, Perrelmouter J, Taub E, Flor H (1999) A spelling device for the paralysed. Nature 398:297–298

Daly I, Billinger M, Laparra-Hernández J, Aloise F, García ML, Faller J, Scherer R, Müller-Putz G (2013) On the control of brain-computer interfaces by users with cerebral palsy. Clin Neurophysiol 124(9):1787–1797

Daly I, Williams D, Hwang F, Kirke A, Malik A, Roesch E, Weaver J, Miranda E, Nasuto SJ (2014) Investigating music tempo as a feedback mechanism for closed-loop BCI control. Brain Comput Interfaces. doi:10.1080/2326263X.2014.979728

Dominguez LG, Kostelecki W, Wennberg R, Velazquez JLP (2011) Distinct dynamical patterns that distinguish willed and forced actions. Cogn Neurodyn 5(1):67–76

Farwell LA (2012) Brain fingerprinting: a comprehensive tutorial review of detection of concealed information with event-related brain potentials. Cogn Neurodyn 6(2):115–154

Farwell LA, Donchin E (1988) Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr Clin Neurophysiol 70(6):510–523

Freeman WJ (2007a) Definitions of state variables and state space for brain-computer interface. Part 1: multiple hierarchical levels of brain function. Cogn Neurodyn 1(1):3–14

Freeman WJ (2007b) Definitions of state variables and state space for brain-computer interface. Part 2: extraction and classification of feature vectors. Cogn Neurodyn 1(2):85–96

Guo J, Gao S, Hong B (2010) An auditory brain–computer interface using active mental response. IEEE Trans Neural Syst Rehabil Eng 18(3):230–235

Halder S, Rea M, Andreoni R, Nijboer F, Hammer E, Kleih S, Birbaumer N, Kübler A (2010) An auditory oddball brain–computer interface for binary choices. Clin Neurophysiol 121(4):516–523

Hayashi H, Kato S (1989) Total manifestations of amyotrophic lateral sclerosis: ALS in the totally locked-in state. J Neurol Sci 93(1):19–35

Higashi H, Rutkowski TM, Washizawa Y, Cichocki A, Tanaka T (2011) EEG auditory steady state responses classification for the novel BCI. In: Proceedings of the Engineering in Medicine and Biology Society, EMBC (2011) annual international conference of the IEEE

Hoffmann U, Vesin J-M, Ebrahimi T, Diserens K (2008) An efficient P300-based brain–computer interface for disabled subjects. J Neurosci Methods 167(1):115–125

Höhne J, Schreuder M, Blankertz B, Tangermann M (2011) A novel 9-class auditory ERP paradigm driving a predictive text entry system. Front Neurosci. doi:10.3389/fnins.2011.00099

Hwang H-J, Kim S, Choi S, Im C-H (2013) EEG-based brain-computer interfaces: a thorough literature survey. Int J Hum Comput Interact 29(12):814–826

Ide Y, Takahashi M, Lauwereyns J, Sandner G, Tsukada M, Aihara T (2013) Fear conditioning induces guinea pig auditory cortex activation by foot shock alone. Cogn Neurodyn 7(1):67–77

Jin J, Allison BZ, Sellers EW, Brunner C, Horki P, Wang X, Neuper C (2011) An adaptive P300-based control system. J Neural Eng 8(3):036006

Jin J, Allison BZ, Zhang Y, Wang X, Cichocki A (2014a) An erp-based bci using an oddball paradigm with different faces and reduced errors in critical functions. Int J Neural Syst. doi:10.1142/S0129065714500270

Jin J, Daly I, Zhang Y, Wang X, Cichocki A (2014b) An optimized ERP brain–computer interface based on facial expression changes. J Neural Eng 11(3):036004

Käthner I, Ruf CA, Pasqualotto E, Braun C, Birbaumer N, Halder S (2013) A portable auditory P300 brain–computer interface with directional cues. Clin Neurophysiol 124(2):327–338

Kim D-W, Cho J-H, Hwang H-J, Lim J-H, Im C-H (2011) A vision-free brain–computer interface (BCI) paradigm based on auditory selective attention. In: Proceedings of the Engineering in Medicine and Biology Society, EMBC (2011) annual international conference of the IEEE

Klobassa DS, Vaughan T, Brunner P, Schwartz N, Wolpaw J, Neuper C, Sellers E (2009) Toward a high-throughput auditory P300-based brain–computer interface. Clin Neurophysiol 120(7):1252–1261

Kodama T, Makino S, Rutkowski T M (2014) Spatial tactile brain–computer interface paradigm applying vibration stimuli to large areas of user's back. arXiv preprint arXiv:14044226

Kübler A, Furdea A, Halder S, Hammer EM, Nijboer F, Kotchoubey B (2009) A brain–computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients. Ann N Y Acad Sci 1157(1):90–100

Laghari K, Gupta R, Arndt S, Antons J, Schleicher R, Moller S, Falk T (2013) Auditory BCIs for visually impaired users: should developers worry about the quality of text-to-speech readers? In: Proceedings of the international BCI meeting

Lelievre Y, Washizawa Y, Rutkowski TM (2013) Single trial BCI classification accuracy improvement for the novel virtual sound movement-based spatial auditory paradigm. In: Proceedings of the signal and information processing association annual summit and conference (APSIPA) Asia-Pacific

Li Y, Pan J, Long J, Yu T, Wang F, Yu Z, Wu W (2016) Multimodal BCIs: target detection, multidimensional control, and awareness evaluation in patients with disorder of consciousness. In: Proceedings of the IEEE

Makous JC, Middlebrooks JC (1990) Two-dimensional sound localization by human listeners. J Acoust Soc Am 87(5):2188–2200

Muller-Putz G, Klobassa D, Pokorny C, Pichler G, Erlbeck H, Real R, Kubler A, Risetti M, Mattia D (2012) The reviewer's comment. In: Proceedings of the Engineering in Medicine and Biology Society (EMBC), 2012 annual international conference of the IEEE

Nijboer F, Furdea A, Gunst I, Mellinger J, McFarland DJ, Birbaumer N, Kübler A (2008) An auditory brain–computer interface (BCI). J Neurosci Methods 167(1):43–50

Rutkowski T, Tanaka T, Zhao Q, Cichocki A (2010) Spatial auditory BCI/BMI paradigm-multichannel EMD approach to brain responses estimation. In: Proceedings of the APSIPA annual summit and conference

Schreuder M, Blankertz B, Tangermann M (2010) A new auditory multi-class brain-computer interface paradigm: spatial hearing as an informative cue. PLoS ONE 5(4):e9813

Schreuder M, Rost T, Tangermann M (2011) Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI. Front Neurosci. doi:10.3389/fnins.2011.00112

Smith RC, Price SR (2014) Modelling of human low frequency sound localization acuity demonstrates dominance of spatial variation of interaural time difference and suggests uniform just-noticeable differences in interaural time difference. PLoS ONE 9(2):e89033

Treder MS, Purwins H, Miklody D, Sturm I, Blankertz B (2014) Decoding auditory attention to instruments in polyphonic music using single-trial EEG classification. J Neural Eng 11(2):026009

Wang D, Chang P (2008) An oscillatory correlation model of auditory streaming. Cogn Neurodyn 2(1):7–19

Wang H, Li Y, Long J, Yu T, Gu Z (2014a) An asynchronous wheelchair control by hybrid EEG–EOG brain–computer interface. Cogn Neurodyn 8(5):399–409

Wang M, Daly I, Allison BZ, Jin J, Zhang Y, Chen L, Wang X (2014b) A new hybrid BCI paradigm based on P300 and SSVEP. J Neurosci Methods 24:16–25

Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM (2002) Brain–computer interfaces for communication and control. Clin Neurophysiol 113(6):767–791

Yin E, Zhou Z, Jiang J, Chen F, Liu Y, Hu D (2013) A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm. J Neural Eng 10(2):026012