DOI:10.3217/978-3-85125-378-8-20

Head–related Impulse Response–based Spatial Auditory Brain–computer Interface

Chisaki Nakaizumi[1,*], Toshie Matsui[1], Koichi Mori[2,*], Shoji Makino[1,*]
and Tomasz M. Rutkowski[1,3,*,†]

[1] Life Science Center of TARA, University of Tsukuba, Japan
tomek@bci-lab.info & http://bci-lab.info/
[2] Research Institute of National Rehabilitation Center for Persons with Disabilities, Japan
[3] RIKEN Brain Science Institute, Japan

Abstract

This study provides a comprehensive test of the head–related impulse response (HRIR) to an auditory spatial speller brain–computer interface (BCI) paradigm, including a comparison with a conventional virtual headphone–based spatial auditory modality. Five BCI–naive users participated in an experiment based on five Japanese vowels. The auditory evoked potentials obtained produced encouragingly good and stable P300–responses in online BCI experiments. Our case study indicates that the auditory HRIR spatial sound paradigm reproduced with headphones could be a viable alternative to established multi–loudspeaker surround sound BCI–speller applications.

1 Introduction

A brain-computer interface (BCI) is capable of providing a speller for disabled people with conditions such as amyotrophic lateral sclerosis (ALS). Although the currently successful visual modality may provide a fast BCI speller, patients at an advanced stage who are in a locked–in state cannot use the modality because they lose all intentional muscle control, including even blinking and movements of the eyes. An auditory BCI may be an alternative method because it does not require good eyesight. However, the modality is not as precise as the visual.

We propose an alternative method to extend the previously published spatial auditory BCI (saBCI) paradigm [1] by making use of a head–related impulse response (HRIR) for virtual sound image spatialization with headphone–based sound reproduction. Our research goal is a virtual spatial auditory BCI using HRIR–based spatialized cues in the part of the non–invasive, stimulus–driven, auditory modality which does not require long–term training. Experiments were conducted to reproduce and provide a comparison with previously reported vector–based amplitude panning (VBAP)–based spatial auditory experiments [1]. The more precise HRIR–based spatial auditory BCI stimulus reproduction was used to simplify previously reported real sound sources generated with surround sound loudspeakers [2].

HRIR appends interaural intensity differences (IID), interaural time differences (ITD), and spectral modifications to create the spatial stimuli, while VBAP appends only IID. HRIR allows for more precise and fully spatial virtual sound image positioning, even without utilizing the user's owner HRIR measurements [3].

The next section of this paper describes the experiment set–up and the HRIR–based saBCI paradigm, together with EEG signal acquisition, pre–processing and classification steps. In the

---

[*] The author was supported by the Strategic Information and Communications R&D Promotion Program (SCOPE) no. 121803027 of The Ministry of Internal Affairs and Communications in Japan.

[†] Corresponding author.

Published by Graz University of Technology Publishing House

third section, the event related potentials (ERP), and especially the $P300$ response latencies are described, with a classification and discussion of the HRIR–based saBCI paradigm information transfer rate (ITR) results, including a comparison with the conventional method. Finally, the conclusions and future research directions are indicated.

2 Methods

All of the experiments were performed at the Life Science Center of TARA, University of Tsukuba, Japan. Five paid BCI-naive users participated in the experiments. The average age of the users was 21.6 years (standard deviation 0.547 years; five females). The psychophysical and online EEG BCI experiments were conducted in accordance with *The World Medical Association Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects.* The experiment procedures were approved and designed in agreement with the ethical committee guidelines of the Faculty of Engineering, Information and Systems at the University of Tsukuba, Japan. Five Japanese vowels (*a, i, u, e, o*) were used in this experiment. The vowels were taken from a sound dataset of female voices [4]. The monaural sounds were spatialized using the public domain CIPIC HRTF DATABASE provided by the University of California, Davis [5]. Each Japanese vowel was set on a horizontal plane at azimuth locations of $-80°$,$-40°$, $0°$, $40°$, $80°$ for the vowels *a, i, u, e, o*, respectively. The psychophysical experiments were conducted to investigate the response time and recognition accuracy. The users were instructed to respond by pressing the button as soon as possible after they perceived the *target* stimulus, as in a classical oddball paradigm [6]. In a single psychophysical experimental run, 20 *targets* and 80 *non-targets* were presented. An online EEG experiment was conducted to investigate the P300 response with BCI–naive users. The brain signals were collected with a biosignal amplifier system g.USBamp by g.tec Medical Engineering GmbH, Austria. The EEG signals were captured by sixteen active gel–based electrodes attached to the following head locations *Cz, Pz, P3, P4, Cp5, Cp6, P1, P2, Poz, C1, C2, FC1, FC2,* and *FCz*, as in the extended 10/10 international system.The ground electrode was attached on the forehead at the *FPz* location, and the reference on the user's left earlobe. BCI2000 software was used for the saBCI experiments to present stimuli and display online classification results. A single experiment was comprised of five runs which contained 10 *target* and 40 *non-target* stimuli. Each run contained five selections. The stimulus duration was set to 250 ms, the interstimulus interval (ISI) to 150 ms, and brain signal ERPs were averaged 10 times for each vowel classification. In brief, the single experiment was comprised of 25 selections. The EEG sampling rate was set to 512 Hz, and a 50 Hz notch filter to remove electric power line interference was applied in a rejection band of $48 - 52$ Hz. The band pass filter was set with 0.1 Hz and 60 Hz cut-off frequencies. The acquired EEG brain signals were classified online by the BCI2000 application using a stepwise linear discriminant analysis (SWLDA) classifier with features drawn from the $0 \sim 800$ ms ERP interval.

3 Results

This section presents and discusses results obtained from the psychophysical and EEG experiments conducted with five users, as described in the previous section. In the psychophysical experiment, the accuracy rates for all stimuli were above 94%. The majority of responses were concentrated at the 350 ms latency. There were no significant differences in the response times between the target stimuli as tested by ANOVA ($p < 0.05$). The results of the EEG experiment are depicted in Figure 1. The left panel shows the grand mean averaged ERP results at four

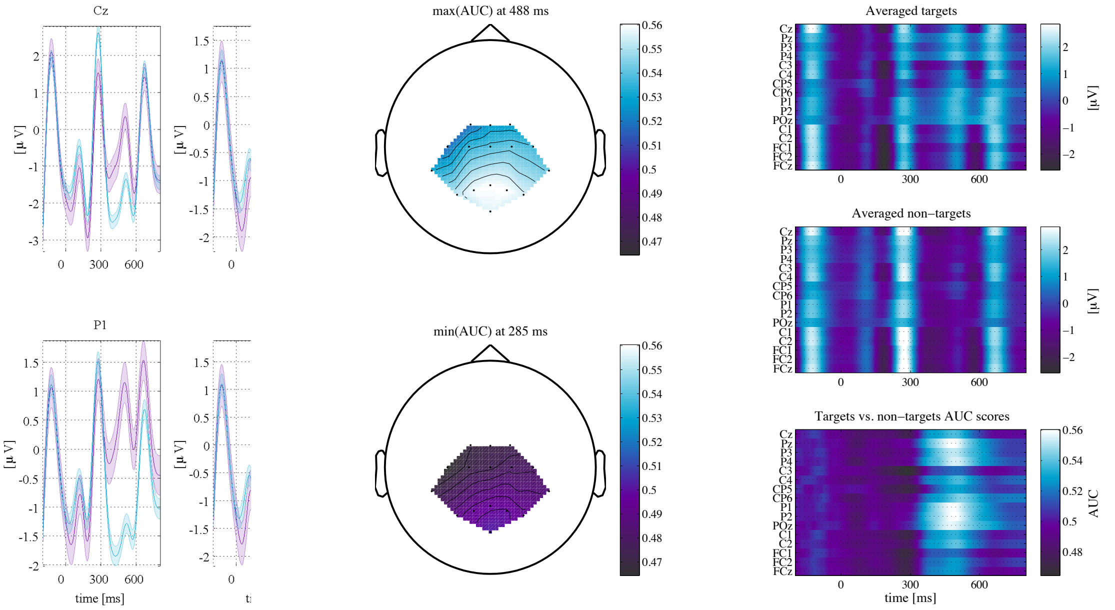

Figure 1: Grand mean averaged ERP and AUC scores leading to final classification results for the participants. The left panel shows the averaged ERP for all participants. The purple line shows the brain waves for *targets*, and blue line is for *non-targets*. The centre panel presents the head topographies at the maximum and minimum AUC scores as obtained from the right bottom panel. The top right panel presents averaged ERP responses to the *target* and the middle panel to the *non–target* stimuli. The right bottom panel visualizes the AUC analysis results of *target* versus *non–target* response distribution differences.

representative electrodes. The centre panel provides the results as scalp topographies at the maximum and minimum area under curve (AUC) of a receiver operating with characteristic values [7] for *target* vs. *non–target* latencies. It also demonstrates the EEG electrode positions used in the experiments. The top right panel indicates the averaged ERP responses of all electrodes to the *target*, and the second panel shows responses to the *non–target* stimuli. The bottom panel indicates the AUC of *target* versus *non-target* responses, clearly confirming the usability of $400 \sim 600$ ms latencies for the subsequent classification. Table 1 presents the classification accuracies of the P300 responses as obtained with the SWLDA classifier and the ITR scores. The average score was obtained as a mean value calculated from $1 \sim 5$ runs (the training run was not included in the calculation of accuracies). The ITR is a major comparison measure [7] among the BCI paradigms. All five users scored above the five vowel sequences spelling chance levels of 20%. There was one user who achieved 100% accuracy, which was the best in the experiments reported. We also compared the ITR scores with a VBAP–based spatial auditory BCI, which is regarded as a conventional method [1]. The VBAP experiment was conducted in 2 runs and with 16 BCI–naive users in [1]. The electrode positions were the same as in our current experiments. The sound stimuli were presented with small ear–fitting headphones in both the modalities. The ISI was set to 500 ms in the VBAP experiment, and to 150 ms in the HRIR experiment. In the VBAP modality, the average ITR score was 1.05 bit/min and the best was 1.78 bit/min. In the HRIR modality, the average ITR was 1.35 bit/min and the best was 2.40 bit/min. The ITR scores of the HRIR experiment were recalculated for 2 runs,

	Run							ITR [bit/min]	
User	1	2	3	4	5	Average	Best	Average	Best
#1	60%	80%	40%	20%	40%	48%	80%	2.26	9.60
#2	0%	40%	100%	80%	100%	64%	100%	5.27	18.58
#3	20%	20%	0%	40%	80%	32%	80%	0.46	9.60
#4	20%	20%	20%	20%	40%	24%	40%	0.06	1.21
#5	0%	40%	40%	80%	60%	44%	80%	1.70	9.60

Table 1: Vowel spelling accuracies and ITRs of each user obtained in the EEG experiments

the same as for the VBAP experiment. HRIR based modality produced better results than the VBAP based modality for both the average and the best score.

4 Conclusions

The EEG results presented confirm the P300 responses of BCI–naive users. The mean accuracy was not very good owing to the short ISI, but the accuracy tends to improve when the number of run increases. Therefore, more attention training or interface using practice may be necessary for BCI–naive users. The ITR scores were higher compared (no significance analysis due to different user groups) with our previous study using HRIR stimuli, and also compared with the previously reported VBAP–based spatial auditory BCI. Nevertheless, the current study is not able to compete with the faster visual BCI spellers. Furthermore, it is necessary to improve the ITR for a more comfortable spelling. We plan to continue research with larger numbers of sound stimuli, a better suited ISI, and more complex spatial sound patterns.

References

[1] M. Chang, N. Nishikawa, Z. R. Struzik, K. Mori, S. Makino, D. Mandic, and T. M. Rutkowski. Comparison of P300 responses in auditory, visual and audiovisual spatial speller BCI paradigms. In *Proceedings of the Fifth International Brain-Computer Interface Meeting 2013*, page Article ID: 156, Asilomar Conference Center, Pacific Grove, CA USA, June 3-7, 2013. Graz University of Technology Publishing House, Austria.

[2] T. M. Rutkowski, A. Cichocki, and D. P. Mandic. Spatial auditory paradigms for brain computer/machine interfacing. In *International Workshop On The Principles and Applications of Spatial Hearing 2009 (IWPASH 2009) - Proceedings of the International Workshop*, page P5, Miyagi-Zao Royal Hotel, Sendai, Japan, November 11-13, 2009.

[3] J. Schnupp, I. Nelken, and A. King. *Auditory Neuroscience - Making Sense of Sound.* MIT Press, 2010.

[4] S. Amano, S. Sakamoto, T. Kondo, and Y. Suzuki. Development of familiarity-controlled word lists 2003 (FW03) to assess spoken-word intelligibility in Japanese. *Speech Communication*, 51(1):76–82, 2009.

[5] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano. The CIPIC HRTF database. In *2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pages 99–102. IEEE, 2001.

[6] J. Wolpaw and E. Winter Wolpaw, editors. *Brain-Computer Interfaces: Principles and Practice.* Oxford University Press, 2012.

[7] M. Schreuder, B. Blankertz, and M. Tangermann. A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue. *PLoS ONE*, 5(4):e9813, 04 2010.