Spatial Auditory BCI Paradigm based on Real and Virtual Sound Image Generation

Nozomu Nishikawa*, Shoji Makino*, Tomasz M. Rutkowski*,†,▽
* TARA Center, University of Tsukuba, Tsukuba, Japan
E-mail: tomek@tara.tsukuba.ac.jp (▽corresponding author)
Website: http://bci-lab.info
† RIKEN Brain Science Institute, Wako-shi, Japan

***Abstract*—This paper presents a novel concept of spatial auditory brain–computer interface utilizing real and virtual sound images. We report results obtained from psychophysical and EEG experiments with nine subjects utilizing a novel method of spatial real or virtual sound images as spatial auditory brain computer interface (BCI) cues. Real spatial sound sources result in better behavioral and BCI response classification accuracies, yet a direct comparison of partial results in a mixed experiment confirms the usability of the virtual sound images for the spatial auditory BCI. Additionally, we compare stepwise linear discriminant analysis (SWLDA) and support vector machine (SVM) classifiers in a single sequence BCI experiment. The interesting point of the mixed usage of real and virtual spatial sound images in a single experiment is that both stimuli types generate distinct event related potential (ERP) response patterns allowing for their separate classification. This discovery is the strongest point of the reported research and it brings the possibility to create new spatial auditory BCI paradigms.**

I. Introduction

The many researchers studying the brain computer interface (BCI) aim to provide handicapped people with various services. Since the interface utilizes only captured brain wave patterns to control a computer, patients are able to successfully move a cursor or a wheelchair without any muscle activity [1].

The audio modality BCI is an interface that makes use of brain waves captured in response to sound stimuli. Since totally locked–in syndrome (TLS) patients cannot control their own eyes, it is envisaged that an auditory BCI system which reads human intentions from brain responses without any eye movements shall provide a solution to this problem. Figure 1 presents the proposed BCI system prototype, performing the following functions:

- Firstly, it creates a spatial auditory stimulus using a sound generation system developed by the authors;
- Secondly, it captures and analyzes EEG event related potential (ERP) responses to the spatial auditory stimuli attended to and ignored;
- Finally, it classifies the responses in order to generate BCI commands for the user application.

In the project described in this paper, we focus on the fact that given a sound stimulus, the human brain generates an ERP pattern [2]. It is a well-known fact that attentional modulation adds a characteristic deflection (the so-called “aha–response”), which could be utilized for auditory BCI development. This ERP feature, which depends on whether the sound is attended to (a target) or not (a non-target), allows further classification of the subject’s choices [3], [4]. When compared with the non–target sound, “the aha–response,” to a target sound being attended to by the subject has a characteristic positive deflection 300 ms after the onset, and thus it is called a $P300$ response. Figure 2 presents an example of averaged $P300$ responses (red lines – targets) from our experiments, together with ignored sounds (blue lines – non–targets).

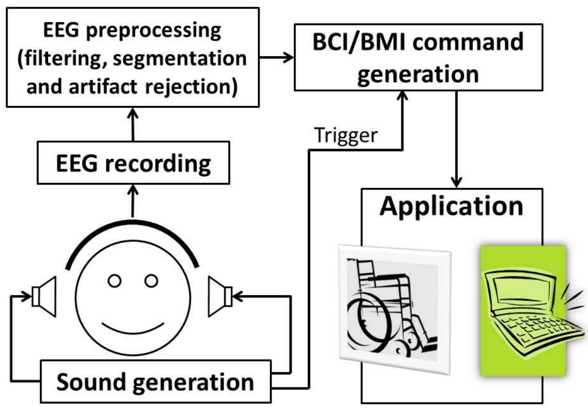

Fig. 1. Schematic diagram of the spatial auditory BCI (saBCI) application. EEG signals are processed by the BCI software processing stages in synchrony with presented auditory stimuli in order to generate interactive commands (the so-called BCI–online–mode).

When the P300 response is properly classified by a machine learning approach, it allows the BCI application user to select a command corresponding to a sound stimulus. In this paper, we compare stepwise linear discriminant analysis (SWLDA) [5] and a linear (SVM) [6] as the machine learning applications.

In our research, we aim at the practical realization and development of the sound generation system for the spatial auditory BCI (saBCI) [7], [8]. The saBCI is based on a sound localization principle. In the saBCI, the application user pays attention to a sound stimulus localized in a certain (intended) direction. Therefore, it is necessary to generate as many easy to discriminate spatial sound stimulus patterns as possible. However, in contemporary solutions [7], [8], the number

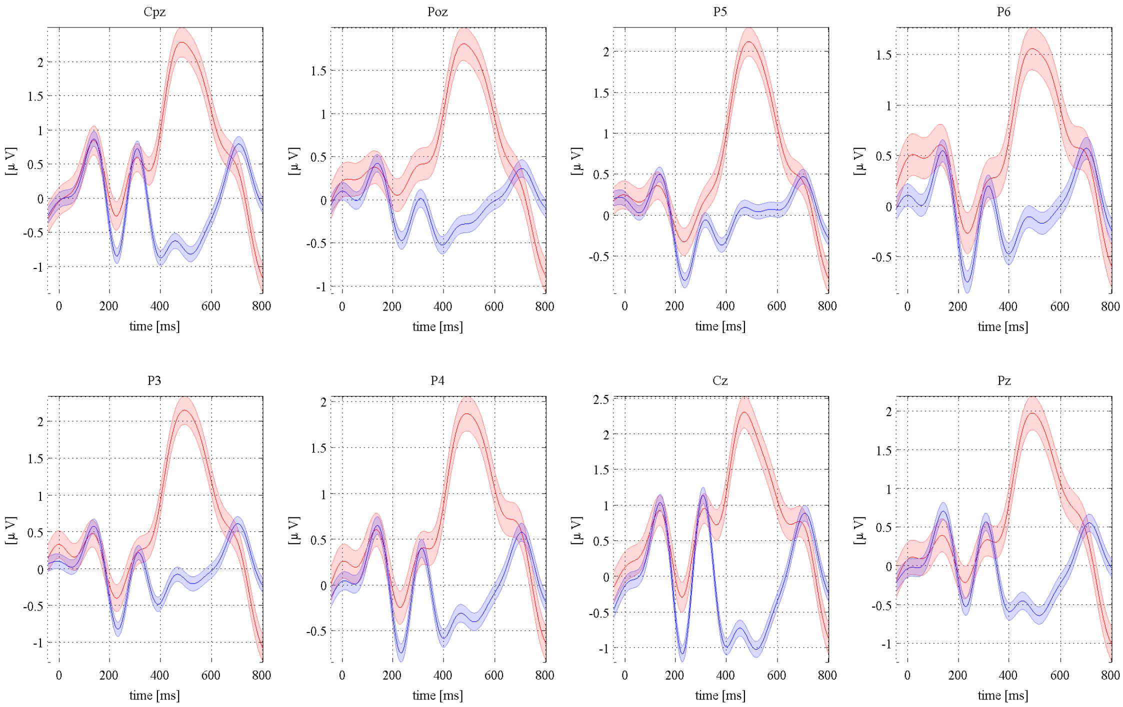

Fig. 2. An example of grand mean averaged ERP brain responses of 27 trials (3 trials $\times$ 9 subjects) to spatial real targets (red line) and non–targets (blue line), together with standard errors. The difference in ERP waves of all electrodes can be observed around 300 ms, which constitutes a $P300$ (positive ERP deflection at this latency) or the so-called "aha–response." The zero seconds time point indicates the stimulus onset.

of commands depends on the available real loudspeakers, since they generate the sound images. This is why a novel sound generation system not depending on the number of loudspeakers is sought and developed by our team.

We propose a combination of a real and virtual sound generation system for a novel spatial auditory BCI. By employing a concept of virtual sound sources or virtual sound images, our saBCI will make it possible to decrease the number of necessary loudspeakers and to increase the number of commands in the current setup. The virtual sound images are generated using multiple loudspeakers (at least two) simultaneously. We have already reported previously [8] that it is possible to evoke a P300 response with virtual sound sources, similarly to with real sound images. The method to create the virtual sound images is known as vector based amplitude panning (VBAP) [9], [10]. It generates the virtual sound images as described in the following sections, where the real–and–virtual spatial sound generation system for the saBCI is discussed.

We conduct psychophysical and EEG experiments to evaluate the real–and–virtual spatial sound generation system by comparing the results obtained with the classic solution of real–only sound image generation. In our experiments, we measure the accuracy rate of psychophysical and EEG responses using real–and–virtual and real–only generation systems. The comparison confirms our hypothesis that the real–and–virtual method is suited for the saBCI paradigm. A detailed evaluation and accuracy rates are described in the following sections, together with the experimental procedures and conditions. Conclusions and future research directions conclude the paper.

II. Methods

This section describes real–only and real–and–virtual spatial sound generation methods for the eight command saBCI application. The real–only method is a reference to evaluate the proposal based on real–and–virtual sound images. With both methods, sound source stimuli are created using a MAX environment [11].

A. *Real–only Method with Eight Loudspeakers*

In the real–only saBCI paradigm, the user receives only real sound source stimuli that are generated using a surround sound environment consisting of the eight loudspeakers distributed octagonally, as shown in Figure 3. The eight spatially distributed loudspeakers are controlled by the MAX environment in real time. Each loudspeaker can generate a single spatial real sound stimulus or a command separately.

B. *Real-and-Virtual Spatial Sound Generation System with the VBAP Method*

In the real–and–virtual saBCI paradigm, the user receives the spatial real and virtual sound source stimuli which were also created using the MAX environment, similarly to the real–only solution described above. The real–and–virtual method consists of four loudspeakers distributed at $-90°, 0°, 90°, 180°$ positions around the subject's head, as shown in Figure 3.

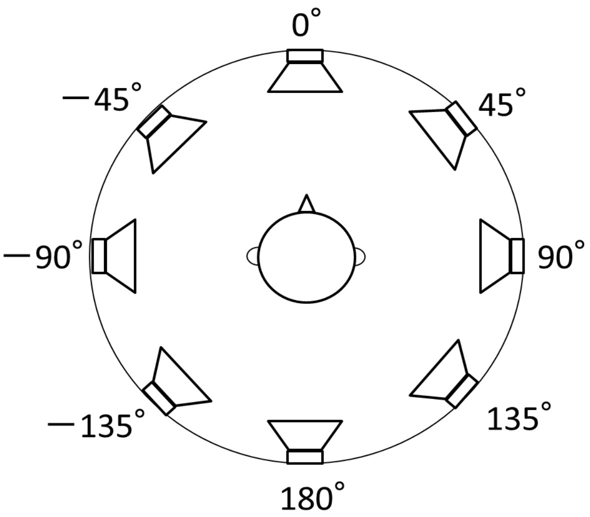

Fig. 3. The locations of the eight loudspeakers for the real–only method in our experiments. A loudspeaker frontal to the subject is situated at 0° angular position in relation to the front of the subject's head.

We implemented a custom VBAP method patcher in the MAX environment that, additionally to the above-mentioned real locations, controlled the four spatially distributed loudspeakers. Thus, it created together eight directions of spatial sound stimuli in real time. The VBAP patcher could be used to position a virtual sound source with the pairs of neighboring loudspeakers. For example, to position a virtual sound image at a direction of $45°$, the VBAP patch created the sound stimulus using only two loudspeakers placed at $0°$ and $90°$ at the same time. The proposed real–and–virtual method could position virtual sound images at the directions of $-135°, -45°, 45°, 135°$ in the same way as in the above example. The classic real sound images were positioned in the directions of $-90°, 0°, 90°, 180°$.

III. Experiments

This section describes the psychophysical and the EEG in saBCI paradigm experiments. In the series of experiments, we compare the behavioral and brain signal responses between real–only and real–and–virtual spatial sound generation systems. These experiments were conducted with nine subjects, and the experimental procedure for each single trial was as follows:

1) An instruction was given to the subject about which spatial sound stimuli direction to attend to in the form of an acoustic prompt followed by a short pause.
2) Next, the subject listened to the eight randomly ordered directions with the same stimulus timbre played from each.
3) The subject responded immediately (as fast as possible) by pressing a computer keyboard button (in the psychophysical experiment), or by counting the number of the target sound stimuli (in the EEG experiment) after the target direction sound occurred.

TABLE I
The psychophysical experiment conditions.

Condition	Parameters
Subjects	9
Sound stimuli	G–chord played using the guitar (MIDI)
Sound sources	4 - and 8 - channel loudspeaker system
Response input	computer keyboard
Sound directions	8 ($-135°$, $-90°$, $-45°$, $0°$, $45°$, $90°$, $135°$, $180°$)
Sound stimuli	40 targets and 280 non–targets in each trial
Trials	2 of real–only and real–and–virtual sounds
Stimulus length	300 ms
ISI	500 ms

TABLE II
The EEG experiment conditions

Condition	Parameters
Subjects	9
Sound stimuli	G–chord played using the guitar (MIDI)
Sound sources	4− or 8−channels loudspeaker system
Response input	Dry 8-electrodes EEG system
Sound directions	8 ($-135°$, $-90°$, $-45°$, $0°$, $45°$, $90°$, $135°$, $180°$)
Sound stimuli	40 targets and 280 non–targets in each trial
Trials	6 (3 trial × 2 real-only/real-and-virtual -sounds)
Classifier method	SWLDA and linear SVM
Stimulus length	300 ms
ISI	300 ms

4) The above procedure was repeated for all eight target directions within a single trial.

In the EEG experiment, the single trial consisted of five sequences, and each of the sequences was defined by eight randomly ordered spatial sound stimuli. The detailed experimental settings are summarized in following subsections.

A. Psychophysical Experiment to Evaluate Real-and-Virtual Sound Localization Method

In order to evaluate the real–and–virtual method for difficulty levels, we first performed a psychophysical experiment to measure the behavioral responses before the EEG evaluation. We conducted two trials for each method, and each trial consisted of eight spatial sound sources (chance level $= 1/8 = 12.5\%$) with a MIDI sound simulating a G–chord sound played by a synthetic guitar sound generator. Table I summarizes the details of the psychophysical experiment settings.

The results of the psychophysical experiment are presented in Figure 4. A comparison of the results reveals that the real–only sound generation system performed better than the real–and–virtual one. Additionally, a comparison of accuracy rates of psychophysical responses within only the real–and–virtual method resulted in the real sound stimulus performing better than the virtual one in the spatial sound localization task. Even though the averaged accuracy rates of the real–and–virtual method in our experiments resulted only in the median of 73.6%, the outcome is much higher than the results previously reported in [8] of 44.7% accuracy for a virtual–only method.

B. EEG Experiment Conditions, Processing and saBCI Classification

In this EEG experiment, we verified whether the real–and–virtual method is competent for the successful saBCI

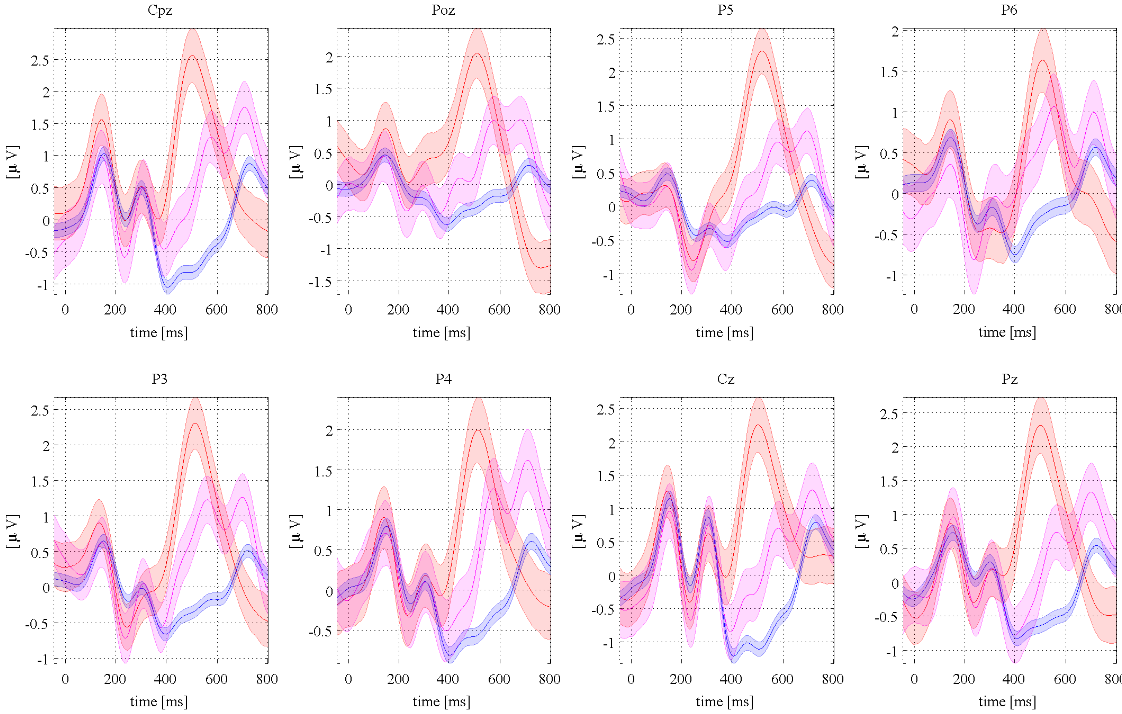

Fig. 5. ERP to real and virtual sound image stimuli for all EEG electrodes used in the experiments conducted. The red lines represent the grand mean averaged ERP responses of 27 trials (3 trials $\times$ 9 subjects) to the real target sounds, while the magenta lines indicate responses to the virtual targets, and the blue lines responses to the non-targets, respectively. Error bars depicting standard errors are also drawn around each of the averaged responses.

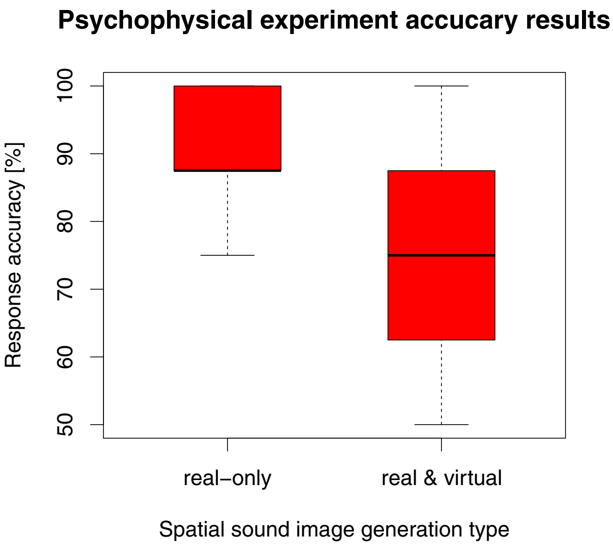

Fig. 4. Accuracy rates of psychophysical responses using real-only and real-and-virtual methods. The chance level was 12.5%.

in the oddball experimental paradigm. The subjects were the same as in the psychophysical experiments. The number of trials was changed because the EEG experiment required more data for an averaged ERP classification. The EEG signals were captured with an eight dry electrode portable wireless acquisition system, and the electrodes were attached at *Cpz, Poz, P3, P4, P5, P6, Cz,* and *Pz* positions, as in the $10/10$ extended international system [12]. In the next step, the captured EEG signals were processed by the in–house enhanced BCI2000 [13] application using the SWLDA and SVM classifiers with features drawn from $0-800$ ms ERP intervals. The sampling frequency was set at 256 Hz, the high pass filter at 0.1 Hz, and the low pass filter at 40 Hz, with a power line interference notch filter set in the $48-52$ Hz band. Each target was presented five times in a single sequence, and the averages of five ERPs were later used for the classification of a single letter. Table II summarizes the details of the EEG experiment settings. The following three experiment sessions were conducted for each subject in the EEG experiments:

1) The first session was recorded for the machine learning approach. The captured EEG data were used to train the classifier used in the following (second) session.
2) The second session was defined as the first evaluation of the classifier in the online BCI–speller test setting.
3) The third session was the second classifier evaluation. The classifier was trained using the EEG data from the second session above.

The single sequence classification results of the EEG experiments conducted are presented in Table III. The real–only method resulted in better outcomes compared with the real–and–virtual method using SWLDA and SVM classifiers. However, in the real–and–virtual sound images experiment, the virtual directions were classified with better scores, which are denoted as ARP(V), compared with the real ones, denoted as ARP(R). These results suggest that improvement is still necessary in this field of research. The grand mean averaged target

TABLE III
BCI INTERACTION SINGLE SEQUENCE CLASSIFICATION ACCURACY RATE OF EEG RESPONSES USING THE REAL–ONLY AND REAL–AND–VIRTUAL METHOD (ARC = ACCURACY RATE WITH REAL–ONLY METHOD; ARP = ACCURACY RATE WITH REAL–AND–VIRTUAL METHOD; R = REAL SOUND IMAGE; V = VIRTUAL SOUND IMAGE). SWLDA AND SVM CLASSIFIERS ARE COMPARED WITH BETTER RESULTS OBTAINED WITH SVM METHOD.

SWLDA classifier results				
Subject	ARC	ARP (R+V)	ARP (R)	ARP (V)
#1	50.0%	0.0%	0.0%	0.0%
#2	37.5%	25.0%	0.0%	25.0%
#3	12.5%	25.0%	12.5%	12.5%
#4	12.5%	12.5%	12.5%	0.0%
#5	25.0%	25.0%	0.0%	25.0%
#6	25.0%	12.5%	0.0%	12.5%
#7	25.0%	12.5%	0.0%	12.5%
#8	12.5%	0.0%	0.0%	0.0%
#9	25.0%	25.0%	12.5%	12.5%
Average:	**25.0%**	**15.3%**	**4.2%**	**11.1%**
Linear SVM classifier results				
Subject	ARC	ARP (R+V)	ARP (R)	ARP (V)
#1	13.0%	19.0%	12.0%	7.0%
#2	47.0%	23.0%	11.0%	12.0%
#3	23.0%	23.0%	8.0%	15.0%
#4	30.0%	24.0%	9.0%	15.0%
#5	27.0%	27.0%	10.0%	17.0%
#6	20.0%	19.0%	10.0%	9.0%
#7	23.0%	20.0%	7.0%	13.0%
#8	47.0%	19.0%	9.0%	11.0%
#9	19.0%	23.0%	10.0%	13.0%
Average:	**27.7%**	**21.9%**	**9.6%**	**12.4%**

ERP responses to real and virtual sound stimuli separately are shown in Figure 5, together with the non–target responses. The illustrated ERP responses, together with standard errors to real target sound images (red lines), virtual targets (magenta lines), and non–targets (blue lines), confirm the hypothesis of the different $P300$ shapes generated by real and visual sound images.

IV. CONCLUSIONS

We conducted a series of psychophysical and EEG experiments in order to evaluate the spatial real and virtual sound images generation system usability for a novel saBCI paradigm. In the psychophysical experiment, the real–and–virtual method was inferior to the real–only sound images directions localization.

The EEG experiment results also show that the real–and–virtual method was inferior to the real–only one. In addition, we were able to obtain an interesting result showing the difference between the target ERP responses to real and virtual sound stimuli created by the real–and–virtual method. This result suggests a great potential in virtual sound image generation for the saBCI.

The results presented confirm the hypothesis of the usability of virtual spatial sound images for novel saBCI paradigms.

AUTHOR CONTRIBUTIONS

Performed the EEG experiments and analyzed the data: NN, TMR. Conceived the concept of the spatial auditory BCI based on real and virtual sound images generation: TMR. Designed the EEG experiments: NN, TMR. Supported the project: SM. Wrote the paper: NN, TMR.

ACKNOWLEDGEMENTS

This research was supported in part by the Strategic Information and Communications R&D Promotion Program no. 121803027 of The Ministry of Internal Affairs and Communication in Japan, and by KAKENHI, the Japan Society for the Promotion of Science, grant no. 12010738 and 24700154. We also acknowledge the technical support of YAMAHA Sound & IT Development Division in Hamamatsu, Japan.

REFERENCES

[1] J. Wolpaw and E. W. Wolpaw, Eds., *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 2012.

[2] D. L. Schomer and F. H. Lopes da Silva, Eds., *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 6th ed. Wolters & Kluwer - Lippincott Williams & Wilkins, 2011.

[3] T. M. Rutkowski, "Auditory brain-computer/machine interface paradigms design," in *Haptic and Audio Interaction Design*, ser. Lecture Notes in Computer Science, E. Cooper, V. Kryssanov, H. Ogawa, and S. Brewster, Eds. Springer Berlin / Heidelberg, 2011, vol. 6851, pp. 110–119. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-22950-3_12

[4] Z. Cai, H. Terasawa, S. Makino, T. Yamada, and T. M. Rutkowski, "Sound timbre and spatial location as informative cues in auditory BCI - brain evoked potential enhancement and involuntary eye movements artifacts suppression approach," in *Proceedings of the Third APSIPA Annual Summit and Conference (APSIPA ASC 2011)*. Xi'an, China: APSIPA, October 18 - 21, 2011, p. paper 241 (6 pages). [Online]. Available: http://www.apsipa.org/proceedings_2011/pdf/APSIPA241.pdf

[5] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A comparison of classification techniques for the P300 speller," *Journal of Neural Engineering*, vol. 3, no. 4, p. 299, 2006. [Online]. Available: http://stacks.iop.org/1741-2552/3/i=4/a=007

[6] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *Journal of Machine Learning Research*, vol. 9, pp. 1871–1874, Jun. 2008. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/liblinear

[7] 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*, Miyagi-Zao Royal Hotel, Sendai, Japan, November 11-13, 2009, p. P5.

[8] N. Nishikawa, Y. Matsumoto, S. Makino, and T. M. Rutkowski, "The spatial real and virtual sound stimuli optimization for the auditory BCI," in *Proceedings of the Fourth APSIPA Annual Summit and Conference (APSIPA ASC 2012)*. Hollywood, CA, USA: APSIPA, December 3-6, 2012, p. paper #356. [Online]. Available: http://www.apsipa.org/proceedings_2012/papers/356.pdf

[9] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=7853

[10] N. Nishikawa, "Spatial sound stimuli optimization for auditory localization based brain machine interface," Bachelor Degree Thesis, School of Informatics, University of Tsukuba, February 2012.

[11] "Max 6," 2012. [Online]. Available: http://cycling74.com/

[12] V. Jurcak, D. Tsuzuki, and I. Dan, "10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems," *NeuroImage*, vol. 34, no. 4, pp. 1600 – 1611, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811906009724

[13] Schalk LAB, 2013. [Online]. Available: http://www.schalklab.org/research/bci2000