Spatial Location and Sound Timbre as Informative Cues in Auditory BCI/BMI - Electrodes Position Optimization for Brain Evoked Potential Enhancement

Zhenyu Cai*, Hiroko Terasawa*, Shoji Makino*, Takeshi Yamada* and Tomasz M. Rutkowski*†
* TARA Center, University of Tsukuba, Tsukuba, Japan
E-mail: tomek@tara.tsukuba.ac.jp
Tel/Fax: +81-29-853-6566
† RIKEN Brain Science Institute, Wako-shi, Japan

***Abstract*—The paper introduces a novel auditory BCI/BMI paradigm based on combined sound timbre and horizontal plane spatial locations as informative cues. The presented concept is based on responses to eight-directional audio stimuli with various tonal and environmental sound stimuli. The approach is based on a monitoring of brain electrical activity by means of the electroencephalogram (EEG). The previously developed by the authors spatial auditory stimulus is extended to varying in timbre sound stimuli which feature helps the subjects to attend to the targets. The main achievement discussed in the paper is an offline BCI analysis based on an optimization of electrode locations on the scalp and evoked response latency for further classification results improvement. The so developed new BCI paradigm is more user-friendly and it leads to better results comparing to previously utilized simple tonal or steady-state stimuli.**

I. Introduction

A concept of a spatial auditory stimulus creates a very interesting possibility to target "the less crucial" auditory activity. We propose to utilize spatial audio stimuli design with a target application in a new BCI/BMI paradigms where users consciously direct their attention to different locations in surround sound environment with various tonal frequency stimuli [1], as depicted in Figure 1. Contemporary applications limit their scope to frontal surround sound loudspeakers [2], while our proposal includes also rear loudspeakers sound presentation allowing for eight commands BCI/BMI applications (full octagonal surround sound loudspeakers setup).

In the approach first proposed in [3] it was shown that responses in a spatial tonal stimuli within the 7.1 channels surround sound system (subjects were positioned in the middle of the loudspeakers systems and requested to direct attention to single direction loudspeakers) were distinguishable in EEG for targets and non-targets interfacing commands. The target and non-target direction sequences were presented randomly. The current proposal extends the design to fully octagonal loudspeakers setup with stimuli direction sequences presented also randomly.

Within this framework, the subjects are asked to focus their attention to a direction of the tonal or environmental sound. The EEG responses are recorded with an EEG amplifier. Additionally vertical and horizontal eye-movements are recorded in order to have a reference signal indicting potential muscle activity used later in artifacts removal algorithm.

In the presented study we decided first to process only "artifact-free-data" (eye blinks, facial muscle and head movements, etc. trials were discarded) in order to validate the proposed spatial auditory stimuli paradigm. To reduce computational complexity of the interfacing approach we propose channel and event-related-potentials (ERP) response samples selection in order to optimize classification results.

This approach allows for EEG channels selection optimization and ERP regions-of-interest (ROI) for target and non-target stimuli optimization by adaptively finding only the components carrying brain activities maximizing the contrasts between responses when the subjects attend or ignore the spatial stimuli. The so obtained brain activity spatial patterns clearly follow the expectations of stronger activities in parietal and temporal cortical areas, which are known for spatial stimuli processing [4].

The paper is organized as follows. In the next section the experimental paradigm is described together with EEG preprocessing steps. Next the channel selection and ERP response period optimization procedures for each subject are described. Finally classification results and discussion conclude the paper.

II. Methods

The experiments to validate the proposed spatial auditory BCI/BMI paradigm were conducted in a Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute in Wako-shi, Saitama, Japan with agreement of the institute's ethical committee guidelines. All experimental procedures and this study targets were explained to the subjects who agreed to participate voluntarily by signing consent forms. EEG signals were recorded with 64 channels active electrodes EEG caps with BIOSEMI *ActiveTwo* system. The stimuli sounds were played through loudspeakers positioned in octagonal setting around a head of a subject. All experiments

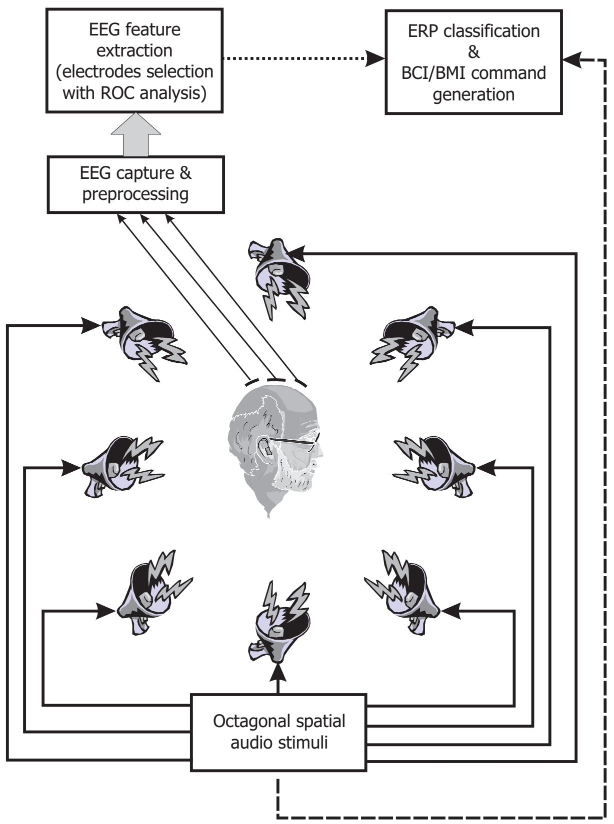

Fig. 1. Spatial auditory BCI/BMI paradigm concept with octagonal loudspeaker arrangement.

were conducted in a silent and low reverberation room in order to limit an interference of "environmental noise" in this preliminary study in order to validate usability of all front- and rear-head sound usability (in contrast to a study conducted in a noisy environment which failed to utilize back loudspeakers for BCI/BMI application [2] - only psychoacoustic experiment for rear-head loudspeakers succeeded there).

A diagram of stimuli and further described in this section steps of EEG processing, electrodes positions optimization, feature extraction and final classification are depicted in Figure 1.

A. Spatial Auditory BCI/BMI Paradigm

We propose to utilize a concept of a horizontal two-dimensional sounds spatial sources arrangement of stimuli paradigm in order to evoke P300 ("aha response") discriminating subject's intentionally attended target and non-target directions. We test the hypothesis whether the evoked in EEG ERP signals are feasible for a future multi-command (eight commands in the proposed case) BCI/BMI application similarly as it was developed in visual domain [5].

B. Experiment Description

EEG recording experiment for offline BCI/BMI paradigm testing were conducted with six healthy subjects (five males and one female; age range $20-50$ years). The subjects were instructed to sit in a comfortable chair in center of eight octagonally positioned loudspeakers. The elevation of the loudspeakers was fixed to the subjects' ear level. Instructions which target direction to attend was given visually on a display located in front of them. A visual fixation cross was also presented on that display to avoid unnecessary eye movements.

Two types of auditory stimuli were presented spatially to the subjects in the octagonal speaker system through only single speaker at a time. The first stimuli was composed of a 400ms long (the longer stimuli comparing to results from [2] was chosen to create more spatial localization realistic situation) and 440Hz sinusoidal tone with 10ms linear raise and decay to avoid "click-effect." The low frequency tone was chosen to evaluate feasibility of an inter-aural-time-delay (ITD) brain auditory localization principle only [6]. The second stimuli was a sound of a car horn recorded on a noisy street and presented also in a form of 400ms long waveforms with 10ms linear raise and decay periods. This second stimuli represented a "broadband" acoustic waveform targeting both brain auditory mechanisms of inter-aural-level-difference (ILD) and ITD. The sound directions were presented in random order to avoid habituation effects.

III. Analysis

The EEG analysis leading to final eight-directions-spatial auditory classification for target and non-target locations is composed of three steps as follows:

- a EEG signal preprocessing (filtering and artifact removal);
- an informative electrodes selection and ERP's ROI optimization for further classification outcome maximization;
- a final classification of evoked responses within each of ten chosen "best channels."

A. EEG Preprocessing

The recorded raw EEG 64 channel signals with BIOSEMI *ActiveTwo* system have to be first referenced by removing mean values of all channels. Next a notch filter is applied at 50Hz center frequency to remove power line noise interference. Two Butterworth 5^{th} order low-pass and high-pass filters are applied next with cutoff frequencies at 0.5Hz and 25Hz respectively to remove low frequency and DC-shift interferences. The low-pass filtering removes possible muscle frequency artifacts. Next the signals are segmented creating event related epochs starting at 0ms of the stimulus onset and ending at 500ms after it (see Figure 2). In a next step eye movement artifacts rejection is carried out. Spatial stimuli are known to cause uncontrolled eye movements which in the current approach are removed with a threshold value set at 80μV (signal voltage above EEG activity level). The EEG conversion from the original BIOSEMI BDF format and the above preprocessing steps were conducted within SPM8 package [7]. The rejected epochs are not processed further since in current approach an emphasis is put forward on the spatial paradigm validation. An example with artifact cleaned and

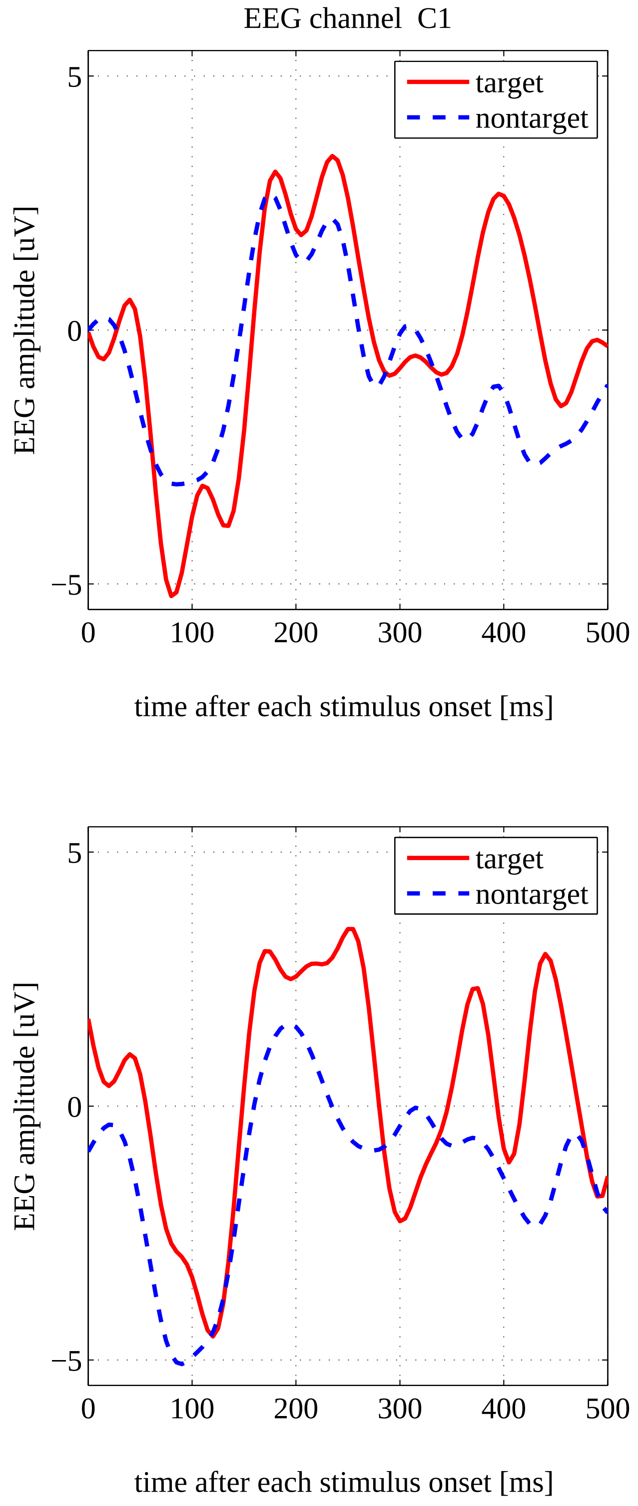

Fig. 2. Results of ERP P300 response for frontal (upper panel) and rear (lower panel) speakers confirming the feasibility of the proposed approach. The zero stands for stimuli onsets. The blue/dashed lines depict non-target (no P300 response after 300ms) responses and red/solid traces visualize attended spatial targets (obvious positive EEG response deflections after in $300-500$ms range).

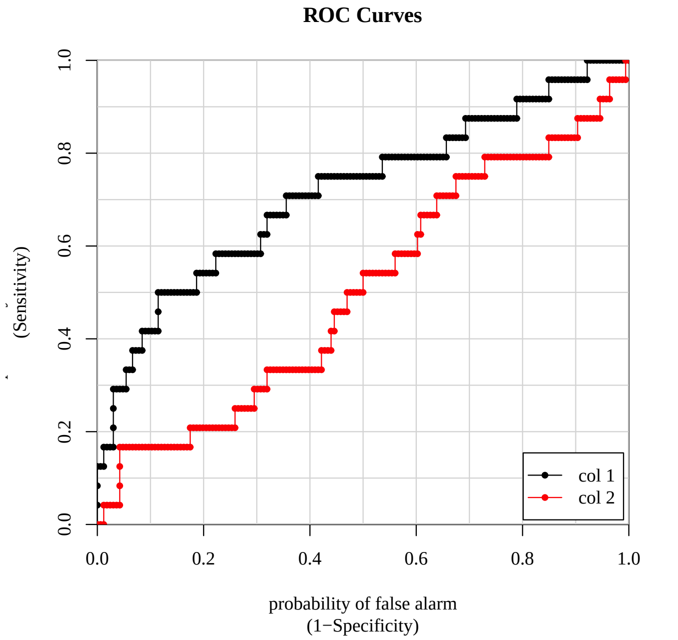

Fig. 3. ROC analysis results of a good channel candidate (col1: black/upper trace) discriminating and a “chance level” one (col2: red/lower trace). Analysis conducted with [10], [11].

averaged epochs separately for frontal and rear loudspeakers is presented in Figure 2. This confirms a feasibility to utilize all frontal and rear channels, since in averaged ERPs those responses are clearly easy to distinguish in $300-500$ms time range.

B. Electrodes and ERP Features Selection for Classification

In order to find ten electrodes from the original 64 channels recordings which for each subject would discriminate the ERP responses for target and non-target responses, we propose to test the two measures. The first one is based on a classical linear-discriminant-analysis (LDA) [8] classification applied to all channels separately. The best classification results set of ten channels from 64 available in a training set would be later applied to test sets. The second proposed measure is based on a receiver operating characteristic (ROC) [9] applied to quantify the separability of two single-trial response distributions for each sample point of ERPs (see Figure 2 with averaged ERP responses for targets and non-targets suggesting that ROI there shall be chosen from around $350-500$ms). While LDA is a standard classification method the ROC related measures are usually used to evaluate the performance of classifiers, they can also quantify the discriminability of feature distributions leading to classifiers optimization. An advantage over other methods for feature selection is that ROC analysis does not rely on the assumption that the distributions are Gaussian [9]. The ROC curve derived from perfectly mixed distributions is the diagonal line (a no-discrimination line as presented approximately in Figure 3 for not chosen channel). The numbers along the major diagonal of ROC graph represent

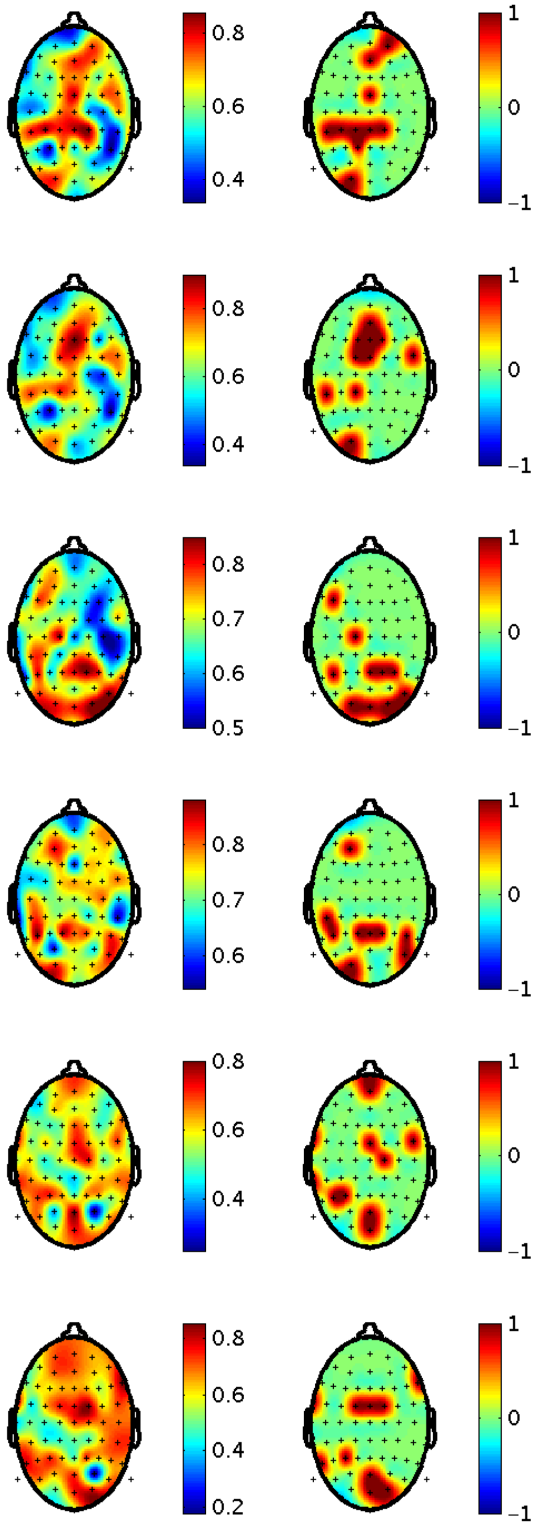

Fig. 4. Results of classical LDA application to binary classification of a spatial tonal 440Hz stimuli applied to all electrodes for a single subject in six cross-validation trials are visualized in the left column while the best resulting electrodes are presented in the right column for each subject. All graphs have the same scaling.

the correct decisions made, and the numbers off this diagonal represent the errors of the confusion between the classes. The sensitivity (also called a true positive rate) of a classifier is calculated as:

$$sensitivity = \frac{PCC}{TP}; \tag{1}$$

and the classifier specificity is as:

$$specificity = \frac{TN}{TP + TN}; \tag{2}$$

where PCC stands for *positives correctly classified*; TP for *total positives*; TN for *true negatives*; and FP for *false positives* respectively. The results of ROC analysis for the chosen and discarded EEG channels are presented in Figure 3.

In order to choose channels with EEG ERP features leading to best classification results we utilize a separability index which is calculated as an area under the curve (AUC) between the ROC curve and the no-discrimination line (diagonal) multiplied by two to relate it to the Wilcoxon test of ranks or the Gini coefficient [9]. We decided to choose ten channels scoring with highest AUC for each subject and within each experimental paradigm (440Hz tone and car horn in the current study). Additionally within each of the chosen channels the best discriminable two areas were chosen taking only 50ms regions around AUC maxima derived from $0 - 200$ms and $200 - 500$ms regions. Those two vectors of single-trial ERP subregions formed features used in subsequent LDA classification within each channel.

Finally the EEG ERP responses were classified into targets and non-targets using two approaches to validate the proposed electrodes and ROI estimation with ROC together with classical LDA applied to all electrodes and the P300 response region. For classical LDA the results are visualized in Figure 4, where classification outcomes are shown for all electrodes together with the ten best electrode candidates. For the proposed approach of ROC based channel and ERP ROI selection the best electrode candidates are visualized in Figure 5, where the ten best electrodes are indicated in red for each subject and condition. The detailed results are discussed in the next section.

IV. Results

The results of the proposed approach to compare target and non-target evoked potentials have been summarized in Figures 6 and 7. In Figure 6 it has been shown that the proposed approach to identify the ten best electrodes and ERP response ROI based on ROC analysis had allowed for a gain of classification results ranging from an increase of 8% boost at the best for subject $\#1$ and tonal stimuli of 440Hz comparing to classical application of LDA analysis to P300 response area for all electrodes and the whole ERP region. In case of a *car horn* stimuli, the best classification increase has been obtained also at the level of 5% for the same subject. Figure 7 presents a comparison of target vs. non-target classification results for frontal and rear loudspeaker sound directions confirming the hypothesis of a possibility

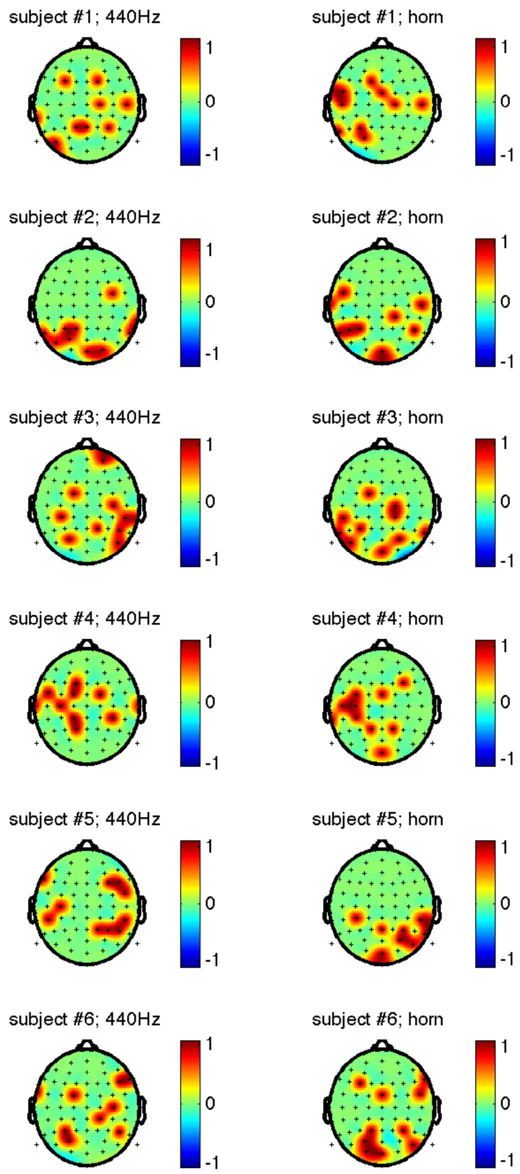

Fig. 5. Results of the proposed ROC analysis based channel selection for all six subjects and two stimuli cases revealing the temporal and parietal scalp regions as best candidates for spatial stimuli P300 responses identification.

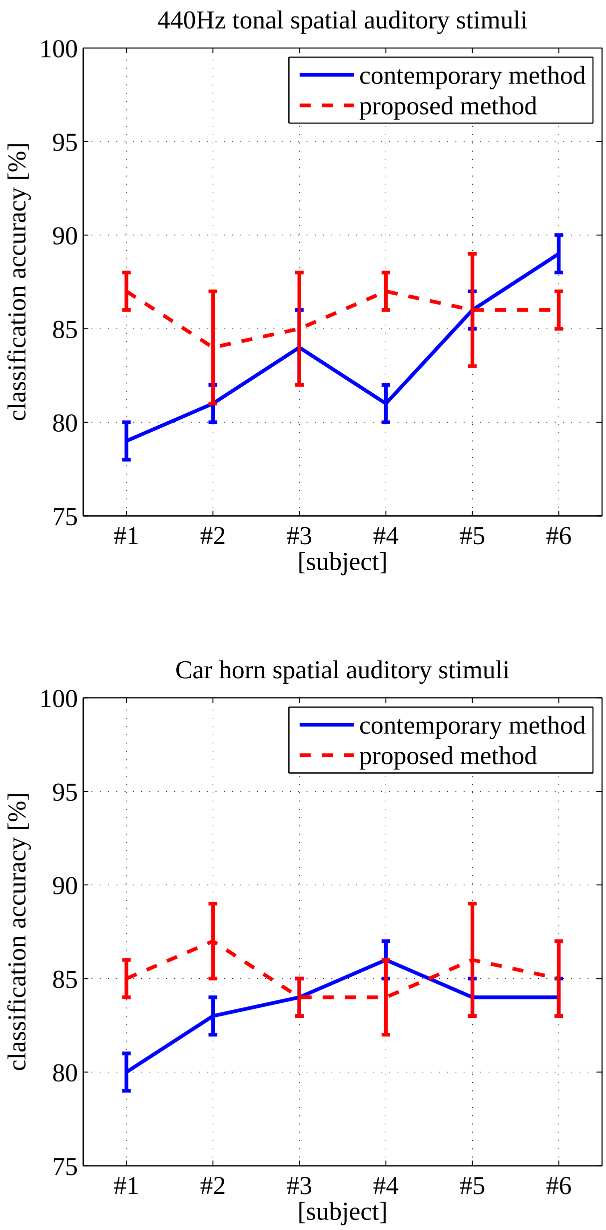

Fig. 6. Percentage of correctly classified spatial audio BCI/BMI P300 (“*aha*”) responses with LDA classifiers derived from three experimental sessions for each subject (blue/solid lines) and based on ten best electrode results (red/dotted lines). Chance level is 50%. LDA analysis conducted with [10], [12], [13].

to utilize those direction modalities despite of the known in psychoacoustics “front–back–confusion” effect.

A variability of the results for various subjects and conditions calls still for further research in this area which our group will continue.

V. Conclusions

In the paper it has been shown that in contrary to the contemporary results with the spatial auditory BCI/BMI paradigms, which fail to utilize rear-head loudspeakers, it is possible to achieve good results for a fully surround sound octagonal loudspeakers setup.

The developed by the authors approach to select the optimal ten channels and ERP ROI intervals resulted with very good classification results for all eight sound stimuli directions. This has been achieved for two types of 400ms long acoustical stimuli targeting ITD and ILD auditory spatial localization shall be also constructed with virtual-acoustical-space simulation via headphones warn by the subject.

ACKNOWLEDGMENTS

Authors would like to thank Sungyoung Kim and Masahiro Ikeda of YAMAHA Corporation for their technical support and rental of experimental spatial auditory equipment used for the experiments.

This research was supported in part by KAKENHI, the Japan Society for the Promotion of Science grant no. 21360179.

REFERENCES

[1] 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), p. P5, November 11-13, 2009.

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

[3] T. M. Rutkowski, T. Tanaka, Q. Zhao, and A. Cichocki, "Spatial auditory BCI/BMI paradigm - Multichannel EMD approach to brain responses estimation," in *Proceedings of the Second APSIPA Annual Summit and Conference (APSIPA ASC 2010)*, (Biopolis, Singapore), pp. 197–202, APSIPA, December 14-17, 2010.

[4] J. Lewald, K. A. J. Riederer, T. Lentz, and I. G. Meister, "Processing of sound location in human cortex," *European Journal of Neuroscience*, vol. 27, no. 5, pp. 1261–1270, 2008.

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

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

[7] Wellcome Trust Centre for Neuroimaging, "Statistical parametric mapping - SPM8 package." http://www.fil.ion.ucl.ac.uk/spm/, 2010.

[8] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Acedemic Press, fourh ed., 2009.

[9] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861 – 874, 2006. ROC Analysis in Pattern Recognition.

[10] R Development Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. ISBN 3-900051-07-0.

[11] J. Tuszynski, *caTools: Tools: moving window statistics, GIF, Base64, ROC AUC, etc.*, 2011. R package version 1.12.

[12] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S*. New York: Springer, fourth ed., 2002. ISBN 0-387-95457-0.

[13] C. Vidaurre, T. H. Sander, and A. Schlögl, "Biosig: The free and open source software library for biomedical signal processing," *Computational Intelligence and Neuroscience*, vol. 2011, pp. Article ID 935364, 12 pages, 2011.

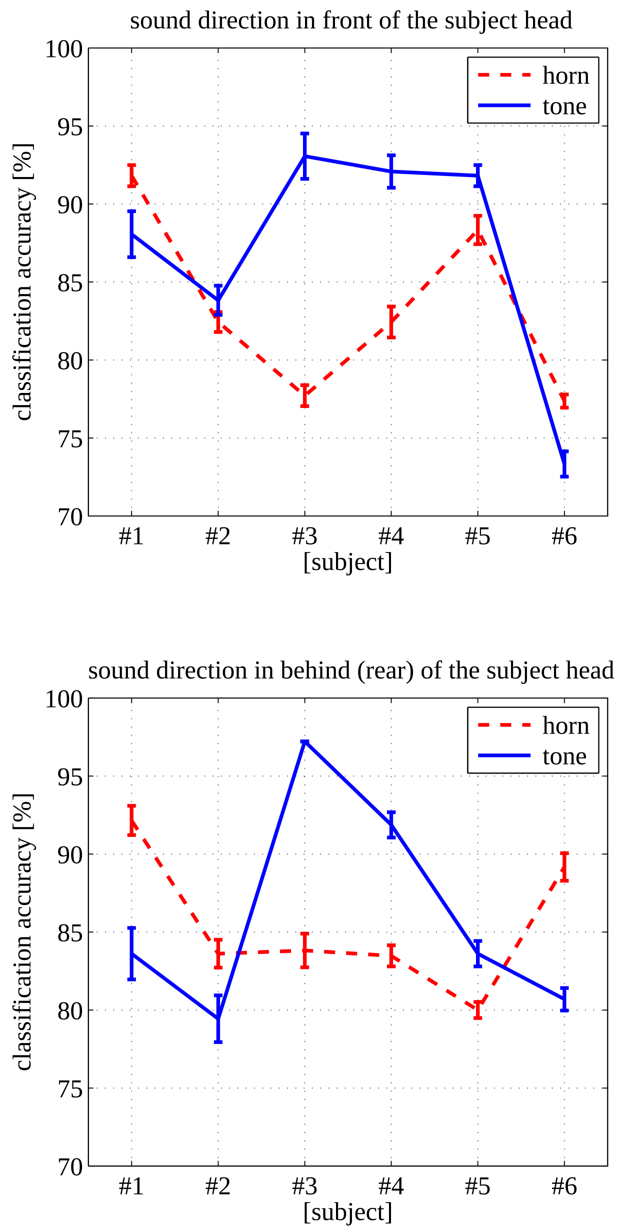

Fig. 7. Comparison of classification responses to frontal (top panel) and rear/back (bottom panel) loudspeakers stimuli directions for six subjects and two stimuli conditions. The results confirm only slight subjects' preferences to frontal stimuli directions except of subject #6 who had better results for rear sound directions.

features. In both cases the classification results obtained with classical LDA and with the proposed combination of ROC-based channel and ERP's ROI selection, which boosted the classification outcomes, were significantly over a chance level of the P300 response based binary paradigm.

Still the results in the offline BCI/BMI mode are not fully satisfactory calling for further research in order to achieve better classification outcomes for an online application which