Single Trial BCI Classification Accuracy Improvement for the Novel Virtual Sound Movement–based Spatial Auditory Paradigm

Yohann Lelièvre*†§, Yoshikazu Washizawa‡§, and Tomasz M. Rutkowski†§▽
*Department of Neurosciences and Neuropsychopharmacology
University of Bordeaux, Bordeaux, France
† Life Science Center of TARA, University of Tsukuba, Tsukuba, Ibaraki, Japan
‡ The University of Electro-Communications, Tokyo, Japan
§ RIKEN Brain Science Institute, Wako-shi, Saitama, Japan
E-mail: tomek@tara.tsukuba.ac.jp (▽corresponding author)

***Abstract*—This paper presents a successful attempt to improve single trial P300 response classification results in a novel moving sound spatial auditory BCI paradigm. We present a novel paradigm, together with a linear support vector machine classifier application, which allows a boost in single trial based spelling accuracy in comparison with classic stepwise linear discriminant analysis methods. The results of the offline classification of the P300 responses of seven subjects support the proposed concept, with a classification improvement of up to 80%, leading, in the best case presented, to an information transfer rate boost of 28.8 bit/min.**

I. Introduction

Severe motor disabilities limit patients' ability to communicate, especially in cases of amyotrophic–lateral–sclerosis (ALS), severe cerebral palsy, head trauma, multiple sclerosis, and muscular dystrophies. Such people are incapable of conveying their intentions (locked–in syndrome (LIS) [1]) to their external environment. Given that ALS is the third most common neurodegenerative disease, with an incidence in Japan of 5 in $100,000$ cases per year [2] and that this illness mostly occurs in adulthood, an improvement in the patients' dependence on the health care system is a major issue for rapidly aging societies.

Over recent decades, numerous research projects have been undertaken in order to develop novel communication techniques which could rehabilitate or bypass the peripheral nerves and muscles destroyed by disease degenerative processes. A promising method to create or establish new communication abilities from the central neural system (the brain) to the external environment utilizes electroencephalography (EEG) in a brain–computer interface application (BCI) [3]. The most successful BCI paradigms utilize the so-called exogenous mode, which requires the user's sensory ability to be involved in receiving a stimulation from the external environment to induce sensory neurophysiological responses possible to be captured in noninvasive EEG recordings.

So far, the primary choice of interaction modality has been vision, relying on the subject's ability to control eye movements, which might become impossible to maintain in the case of patients suffering from LIS [1]. Therefore, other modalities have recently been explored, such as hearing [4] and touch [5] in order to create vision independent BCIs. Furthermore, a very recent report [6] confirmed the superiority of the tactile BCI in comparison with visual and auditory modalities tested with an LIS patient.

The first auditory paradigm BCIs were developed based on binary decisions allowing for lower information transfer rates (ITR) [7] compared with multi–class interfacing solutions. In order to address the need for a multi–class BCI, a spatial auditory paradigm has been proposed [8], [7]. This paradigm uses spatially distributed, auditory cues. However this paradigm still does not result in ITR scores leading to a fast BCI application that could rely solely on the auditory modality. The ITR results are still inferior compared with the visual paradigms. Moreover, the problem of so-called lower audible angles in spatial auditory perception limits the accuracy of static sound based BCI applications. Therefore, we propose to the use moving sound stimuli to create cues that are richer and easier to perceive [9]. The moving stimuli allow for the creation of sound cues traveling in multiple angular directions, bypassing possible lower angles in the subject's spatial audition.

Our research hypothesis is summarized as follows. In the spatial auditory BCI, the use of moving sound stimuli resolves the problem of lower audible angles and it should lead to an increase in accuracy compared with classic static cases. The concept has already been tested in [10], [11] without fully satisfactory results, due to the necessary averaging of EEG event related potentials (ERP) and the low classification accuracy obtained with stepwise linear discriminant analysis (SWLDA) [12] classifiers. In order to improve the classification accuracy and ITR, we propose to use a linear support vector machine (SVM) classifier as implemented by [13] in a single trial (no response averaging) scenario.

In order to test the hypothesis, a five vowel spatial speller task is proposed, as first developed for the static case in [9]. We test and compare the results offline with the linear SVM single trial classification concept, using the experimental data collected within the project [10].

II. Methods

To realize the experiments, seven healthy subjects were involved (mean age of 30.71 with a standard deviation of 7.70; two females and five males). All the experiments were performed at the Life Science Center of TARA, University of Tsukuba, Japan. The 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 two protocols were designed to reproduce the auditory experiments as proposed in [9] with the addition of the moving sound modality.

A. BCI Static and Moving Sound Experimental Protocols

The 200 ms long spatial unimodal (auditory) stimuli were presented from five distinct virtual spatial locations through the use of headphones. All the subjects conducted a psychophysical test with a button press response to confirm understanding of the experimental setup for the two protocols. These tests resulted in average response delays with means of around 450 ms, suggesting that the tasks were well understood and the mental load differences among the stimuli were not significant. The stimuli sounds used were the Japanese vowels *a, i, u, e, o*, represented in *hirigana*.

The subjects were instructed to mentally attend to the targets presented and spell the five vowel random sequences, which were presented audibly in each session. The target was presented with 20% occurrence probability. It has been shown that such an occurrence probability is rare enough to produce a clear P300 response [14]. Each target was presented ten times in a single spelling session and the single ERP trial (a set of a single target and four non–targets set) responses were later used for the classification in order to improve the BCI interfacing speed. The inter–stimulus interval (ISI) was set to 500 ms and each stimuli length was 200 ms.

In the static sound, the auditory source image of the vowel was virtually positioned approximately at one meter distance from the subject's head using a vector based amplitude panning (VBAP) approach [15]. The five vowels were angularly separated by 45° from right to left as depicted in Figure 1.

In the moving sound protocol, the spatial angle origin positions were the same as in the static example. The virtual sound moving effects were applied in order to simulate five movement trajectories, as depicted in Figure 2. The sound movement effect was generated using a combination of binaural loudness and Doppler effects, as proposed and described in [10].

Both the static and moving sound stimuli oddball auditory BCI–spelling sequences were conducted three times by each subject, of which the first sequence was considered as training, the second was used for the classifier setup and the third for the BCI–speller accuracy tests reported in Table I.

A subjective comfort preference questionnaire between static and moving sound protocols was collected from the subjects after the experiments.

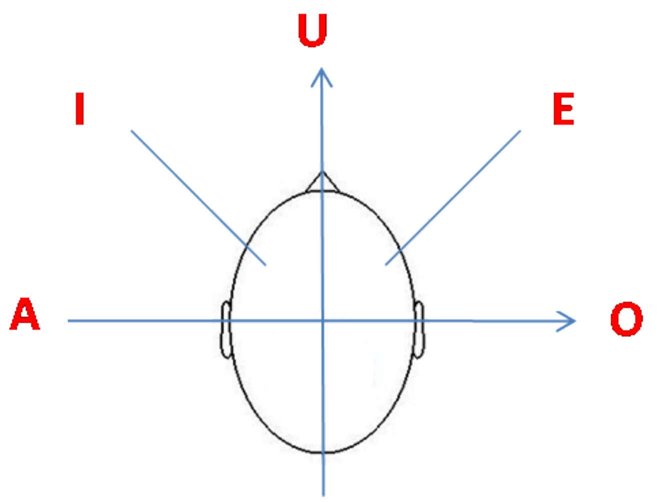

Fig. 1. Auditory source image directions visualization in the static spatial sound BCI protocol. The five vowels were angularly separated by 45° starting from 0° at the right ear level.

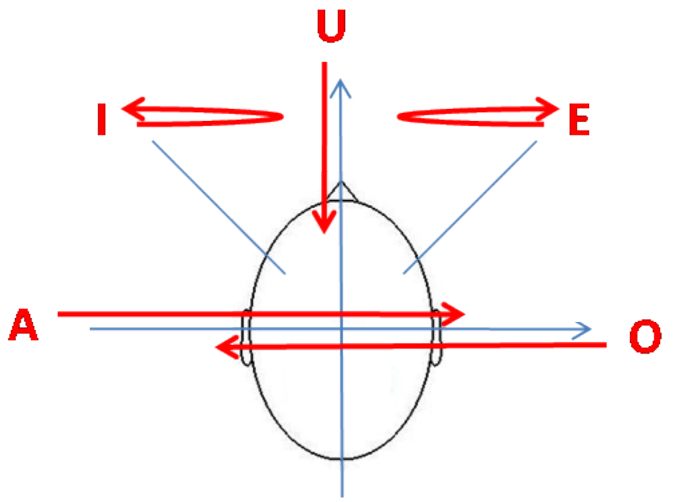

Fig. 2. Auditory source image directions and movement trajectories visualization in the moving sound BCI protocol. The five sounds were virtually moved spatially, as depicted by the arrows in the figure. The vowel sounds of *a* and *o* were moved using the sound motion virtualization method from left to right and right to left accordingly. The vowel *u* was moved frontally toward the subject's head, while the vowels *i* and *e* oscillated around 45° and 135° angular directions.

B. EEG Recording and Processing Steps

During the original online BCI experiments, in a project reported in [10], [11], the EEG signals were captured with eight active dry g.SAHARA electrodes connected to the g.MOBIlab+ EEG amplifier, both by g.tec Medical Instruments GmbH, Austria. The electrodes were attached to the following head locations $Cz, CPz, P1, P2, P3, P4, Cp5$, and $Cp6$ as in the $10/10$ extended international system [16]. The ground and reference electrodes were attached to the left and right mastoids respectively. The sampling frequency was set to 256 Hz, with a high pass filter at 0.1 Hz, the low pass filter at 40 Hz, with a power line interference notch filter set in the

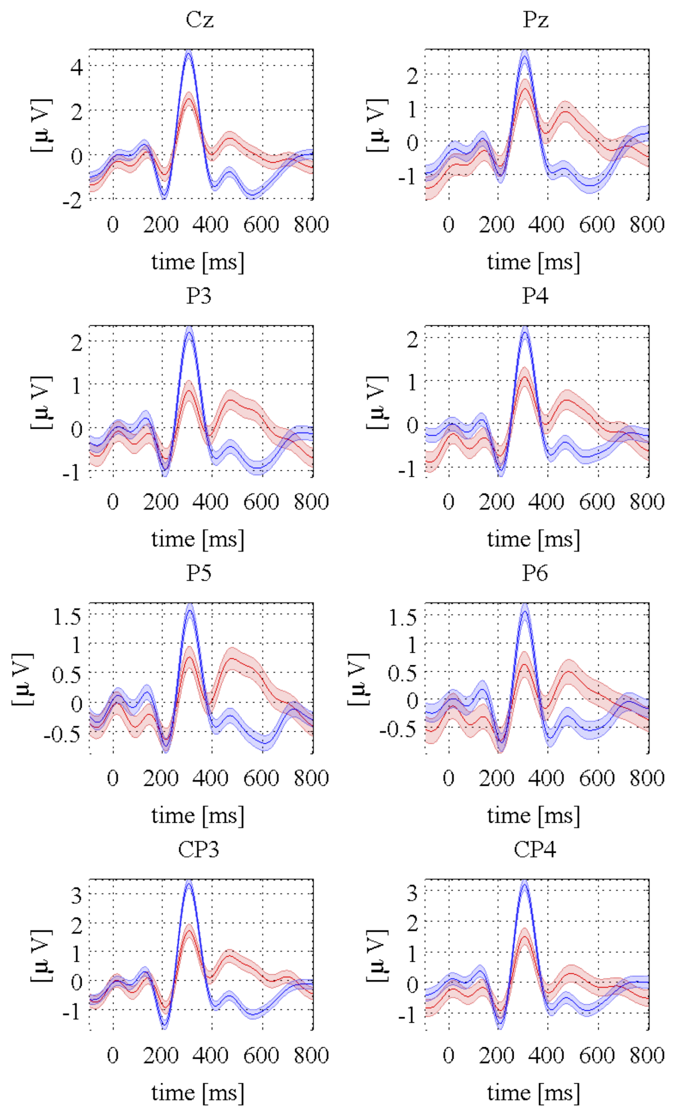

Fig. 3. Grand mean auditory ERP visualization for all subjects, together with P300 responses obtained in the static sound spatial auditory BCI experiment. The eight EEG electrodes are depicted separately. The blue lines represent non–targets, and the red lines represent the targets. The P300 response is clearly visible in the 400 − 700 ms latency ranges. The interesting P200 modulation (lower amplitudes for targets) is also depicted.

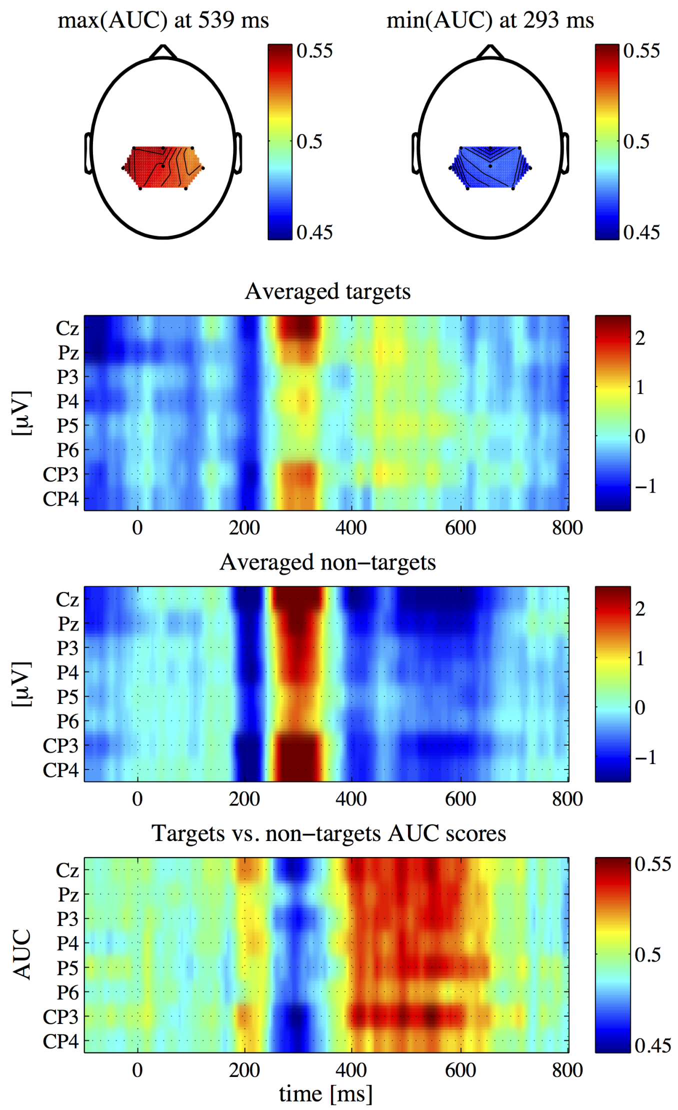

Fig. 4. AUC analysis result of target and non–target response discrimination in the static sound spatial auditory BCI experiment. The top panels present two topographic head plots with electrode positions and spatial maps of the response at maximum and minimum AUC values as obtained from the bottom panel time series depicting this quantity for each electrode separately. The middle two panels depict the ERP responses to targets and non–targets together. A resulting AUC analysis of the two middle panel plots is presented in the bottom panel of the figure.

$48-52$ Hz band to avoid spurious subharmonic and possible amplifier saturation effects.

The recorded EEG signals were then processed offline and classified using the SWLDA [17] and linear SVM [13] classifiers with features drawn from $0-800$ ms ERP intervals, as explained in detail in the following section.

III. EEG Responses Analysis and BCI Classification

The segmented EEG data were first bandpass filtered within a band of $0.1-25$ Hz, and responses with an amplitude larger than 80 μV were rejected from further analysis.

The P300 grand mean average for all subjects response analysis results are depicted in Figures 3 and 4 for static and in Figures 5 and 6 for moving sound protocols, respectively. The top panels in Figures 4 and 6 present the area under the curve (AUC) values projected on a head topographic plot with EEG electrode locations used for maximum and minimum results. The AUC method allows for ERP response discrimination analysis for a subsequent classification, as also depicted in the three lower panels in both the figures for target (note P300 responses), non–target and AUC time series. The separated ERP responses for targets and non–targets recorded for each electrode are depicted in Figures 3 and 5.

The best mean discrimination between target and non-target stimuli was obtained at around 539 ms after stimulus emission

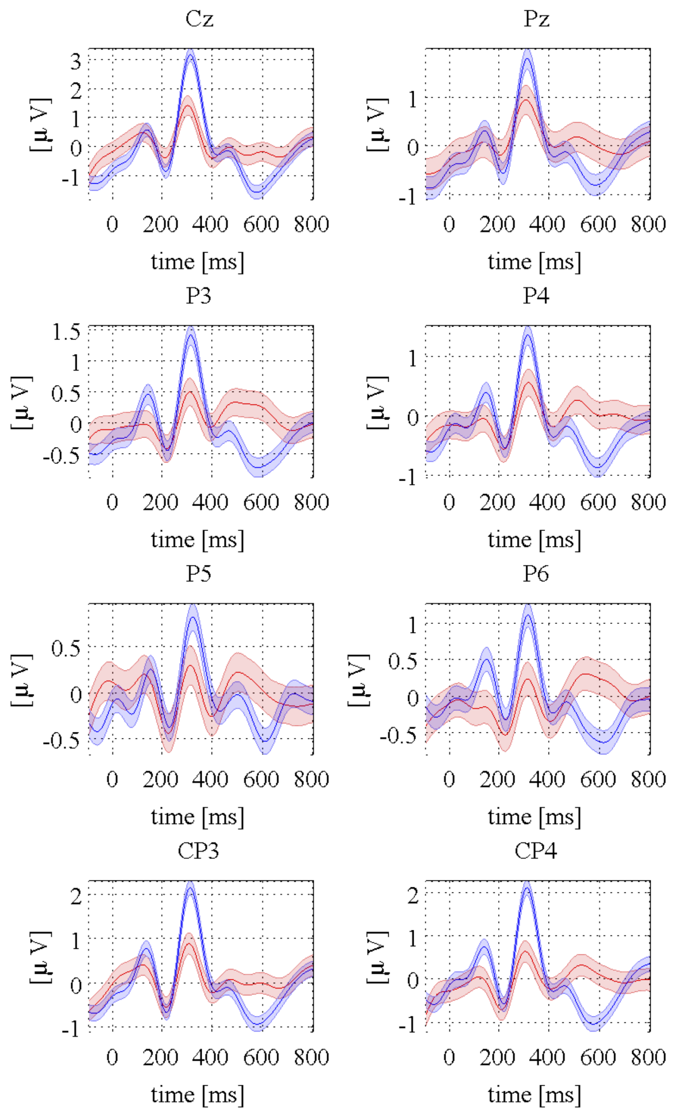

Fig. 5. Grand mean for all subjects auditory ERP visualization together with P300 responses obtained in the moving sound spatial auditory BCI experiment. The eight EEG electrodes are depicted separately. The blue lines represent non–targets, and the red represent the targets. The P300 response is clearly visible in the 400 − 700 ms latency range. The interesting P200 modulation (lower amplitudes for targets), similar to in the static case presented in Figure 3, is also depicted.

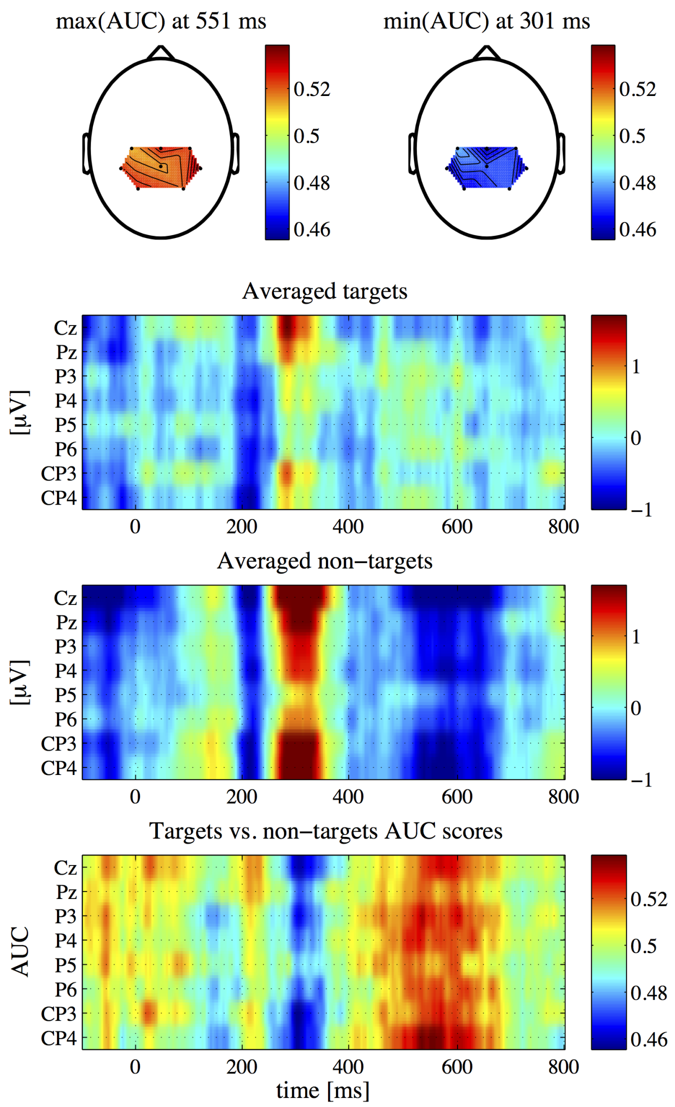

Fig. 6. AUC analysis result of target and non–target response discrimination in the moving sound spatial auditory BCI experiment. The top panels present two topographic head plots with electrode positions and spatial maps of the response at maximum and minimum AUC values as obtained from the bottom panel time series depicting this quantity for each electrode separately. The middle two panels depict the ERP responses to targets and non–targets together. A resulting AUC analysis of the two middle panel plots is presented in the bottom panel of the figure.

for the static and at 551 ms after stimulus emission for the moving sounds, as in the grand mean averaged responses for all the subjects in the study. These periods fit well with the rise in P300 in the case of auditory stimuli [18]. The AUC scores for static sound resulted in between 0.52 and 0.55, which allows for the features' subsequent classification. For the moving sound, the scores similarly resulted in between 0.51 and 0.54. The three bottom panels in Figures 4 and 6 represent the grand mean average responses of all subjects to target and non–target stimuli, and the AUC scores over time for each electrode. For the static protocol, we can see that the differences in electrical activity between target and non–target, as depicted in Figure 3, were the most significant for the $P200$ [19] around 320 ms, and for the $P300$ around 540 ms. Yet the AUC value is significant (above 0.5) only for the $P300$ event. Similar responses were obtained for the moving sound protocol, as depicted in Figure 5, which is proof of the usability of the proposed concept for a novel spatial auditory BCI–speller paradigm.

The results of the SWLDA and linear SVM classifiers in the single trial P300 classification are summarized in Table I. The application of the linear SVM classifier from the package [13] allowed for a boost in classification results of up to 71.4% on average for moving and up to 62.9% for static sounds. This is in comparison with the SWLDA classifiers previously reported

TABLE I

SINGLE TRIAL BASED BCI–SPELLER ACCURACY (NOTE, THIS IS NOT A BINARY P300 CLASSIFICATION RESULT, BUT THE ENSUING SPELLING RESULT WITH A THEORETICAL CHANCE LEVEL OF 20%) IN SPATIAL STATIC AND MOVING SOUND SPELLING TASK USING THE CLASSIC SWLDA AND LINEAR SVM CLASSIFIERS.

The proposed spatial moving sound paradigm		
Subject number	SWLDA result	Linear SVM result
#1	40%	80%
#2	0%	80%
#3	20%	60%
#4	40%	100%
#5	40%	80%
#6	20%	40%
#7	40%	60%
Average:	**28.6%**	**71.4%**
The classic spatial static sound paradigm		
Subject number	SWLDA result	Linear SVM result
#1	20%	60%
#2	20%	80%
#3	20%	60%
#4	40%	80%
#5	0%	40%
#6	20%	60%
#7	20%	60%
Average:	**20%**	**62.9%**

TABLE II

SINGLE TRIAL BASED SPELLING ACCURACY (SEE TABLE I) BASED ITR RESULTS. FOR A COMPARISON, THE MAXIMUM ITR REPORTED IN THE CONTEMPORARY STATIC SPATIAL BCI PARADIGM [7] WAS 25.20 BIT/MIN (AVERAGE 17.39 BIT/MIN).

The proposed spatial moving sound paradigm		
Subject number	SWLDA based ITR	Linear SVM based ITR
#1	3.62 bit/min	28.80 bit/min
#2	0.00 bit/min	28.80 bit/min
#3	0.00 bit/min	13.22 bit/min
#4	3.62 bit/min	55.73 bit/min
#5	3.62 bit/min	28.80 bit/min
#6	0.00 bit/min	3.62 bit/min
#7	3.62 bit/min	13.22 bit/min
Average:	**2.06** bit/min	**24.60** bit/min
The classic spatial static sound paradigm		
Subject number	SWLDA result	Linear SVM result
#1	0.00 bit/min	13.22 bit/min
#2	0.00 bit/min	28.80 bit/min
#3	0.00 bit/min	13.22 bit/min
#4	3.62 bit/min	28.80 bit/min
#5	0.00 bit/min	3.62 bit/min
#6	0.00 bit/min	13.22 bit/min
#7	0.00 bit/min	13.22 bit/min
Average:	**0.52** bit/min	**16.30** bit/min

in [10], which in the same case resulted in an average of 28.6% and 20.0% spelling accuracies respectively, which is close to the theoretical chance level.

We conclude that the analysis of EEG data demonstrates much better results offline with the SVM classifier from the package [13] than the linear SWLDA [17] in the single trial scenario, and also than averaged responses on the same dataset reported in [10], [11]. As a result of the single trial ERP classification in the spatial sound auditory BCI pilot study, a better accuracy was obtained for the moving sound stimuli compared with the classic static stimuli.

The results of the ITR calculation presented in Table II support the concept presented, with much better scores compared with the contemporary static spatial BCI paradigm [7], where 25.20 bit/min at best (average 17.39 bit/min) was reported. Our moving sound stimuli cue resulted in a best score of 55.73 bit/min (average 24.60 bit/min).

Figure 7 reports the subject preference questionnaire results. The moving stimulus had in the greatest preference at 43%, compared with the static at 29%, and no preference with a 28% score. The preference analysis also supports the proposed concept of the moving sound stimulus validity for the spatial auditory BCI–speller utilization.

IV. DISCUSSION

The aim of this study was to enhance single trial classification results in the proposed moving sound sources spatial auditory BCI paradigm in comparison with the classic static spatial protocols. Offline results obtained with the SWLDA classifier have been enhanced by the proposed linear SVM approach, which boosts the BCI spelling results up to 71.4% on average for moving and up to 62.9% for static sounds (chance level of 20%). The resulting ITR scores at a maximum of 55.73 bit/min (24.60 bit/min on average) are also very promising for future online application with patients suffering from LIS.

The increase in accuracy of the moving sound in comparison to the static allows subjects to avoid the problem of too low audible angles in auditory spatial cognition. These are the novel and attractive points of the proposed new paradigm.

The moving sound protocol proved to be more comfortable than the static one. This characteristic is of interest because auditory paradigms tend to be more boring than other modalities.

The approach presented shall help, if not to reach the goal, to get closer to our objective of the design of a more user friendly BCI. Thus, we can expect that patients suffering from LIS will be able to use an appropriate BCI interface more efficiently and comfortably to restore their basic communication abilities.

V. CONCLUSIONS

The results of this study shall lead directly to a new improved spatial auditory BCI paradigm and restoration of the communication ability of LIS patients. We have demonstrated that the use of moving sound stimulus ameliorates the spelling accuracy results and improves the interfacing comfort based on subject preference reports. Our research shows that a moving sound stimulus is proven to be more accurate, as shown by P300 response classification with a linear SVM in the single trial cases, and more comfortable than the classic static cases.

The proposed moving sound protocol helps to avoid the problem of lower audible angles in spatial auditory perception. Still there remains a long way to go before providing an

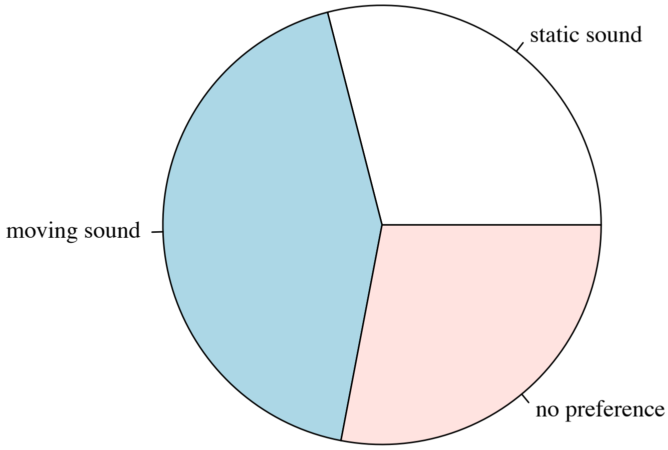

Fig. 7. Stimulus preference and comfort analysis results for all subjects between static and moving protocols. The subject responses resulted in 43% for moving sound stimuli, 29% for the classic static, and 28% without any preference indication.

efficient and comfortable auditory BCI, but our research has progressed toward this goal.

ACKNOWLEDGMENTS

We acknowledge for their essential support the International FidEx program of the University of Bordeaux France, the Student Exchange Support Program of the Japan Student Services Organization (JASSO), the Strategic Information and Communications R&D Promotion Program (SCOPE) of the Ministry of Internal Affairs and Communications (MIC) of Japan, and Dr. Andrzej Cichocki from the RIKEN Brain Science Institute, Japan.

We thank also the YAMAHA Sound & IT Development Division for their precious technical support and the members of the Multimedia Laboratory of the Life Science Center of TARA at the University of Tsukuba, Japan, for their kind cooperation.

REFERENCES

[1] F. Plum and J. B. Posner, *The Diagnosis of Stupor and Coma*. Philadelphia, PA, USA: FA Davis, 1966.

[2] T. Kihira, S. Yoshida, M. Hironishi, H. Miwa, K. Okamato, and T. Kondo, "Changes in the incidence of amyotrophic lateral sclerosis in Wakayama, Japan," *Amyotrophic Lateral Sclerosis*, vol. 6, no. 3, pp. 155–163, September 2005. [Online]. Available: http://informahealthcare.com/doi/abs/10.1080/14660820510030031

[3] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: a review of the first international meeting." *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 8, no. 2, pp. 164–173, Jun. 2000. [Online]. Available: http://view.ncbi.nlm.nih.gov/pubmed/10896178

[4] N. J. Hill, T. N. Lal, K. Bierig, N. Birbaumer, and B. Schoelkopf, "An auditory paradigm for brain–computer interfaces," in *Advances In Neural Information Processing Systems*, L. Saul, Y. Weiss, and L. Bottou, Eds., vol. 17. Cambridge, MA, USA: MIT Press, December 2005. [Online]. Available: http://eprints.pascal-network.org/archive/00000407/01/pdf2482.pdf

[5] G. Muller-Putz, R. Scherer, C. Neuper, and G. Pfurtscheller, "Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces?" *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 14, no. 1, pp. 30–37, March 2006. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1605261

[6] T. Kaufmann, E. M. Holz, and A. Kuebler, "Comparison of tactile, auditory and visual modality for brain-computer interface use: A case study with a patient in the locked-in state," *Frontiers in Neuroscience*, vol. 7, no. 129, 2013. [Online]. Available: http://www.frontiersin.org/neuroprosthetics/10.3389/fnins.2013.00129/abstract

[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*, vol. 5, no. 4, p. e9813, 04 2010. [Online]. Available: http://dx.doi.org/10.1371%2Fjournal.pone.0009813

[8] 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.

[9] 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*. Asilomar Conference Center, Pacific Grove, CA USA: Graz University of Technology Publishing House, Austria, June 3-7, 2013, p. Article ID: 156. [Online]. Available: http://arxiv.org/ftp/arxiv/papers/1301/1301.6360.pdf

[10] Y. Lelièvre, "Moving sound stimulus study for the novel spatial auditory brain–computer interface paradigm," Master's Thesis, Department of Neurosciences and Neuropsychopharmacology, University of Bordeaux, Bordeaux, France, June 2013.

[11] Y. Lelièvre and T. M. Rutkowski, "Novel virtual moving sound-based spatial auditory brain-computer interface paradigm," in *Proceedings of the 6th International IEEE EMBS Conference on Neural Engineering*, EMBS. Sheraton San Diego Hotel & Marina, San Diego, CA, USA: IEEE Press, November 6-8, 2013, pp. (accepted, in press). [Online]. Available: http://arxiv.org/abs/1308.2630

[12] 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

[13] 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

[14] E. W. Sellers and E. Donchin, "A p300-based brain–computer interface: Initial tests by ALS patients," *Clinical Neurophysiology*, vol. 117, no. 3, pp. 538 – 548, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1388245705004608

[15] 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.

[16] 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

[17] C. R. M. Stocks, "Py3GUI," 2011. [Online]. Available: https://github.com/collinstocks/Py3GUI

[18] S. Halder, E. M. Hammer, S. C. Kleih, M. Bogdan, W. Rosenstiel, N. Birbaumer, and A. Kübler, "Prediction of auditory and visual p300 brain-computer interface aptitude," *PLoS ONE*, vol. 8, no. 2, p. e53513, 02 2013. [Online]. Available: http://dx.doi.org/10.1371%2Fjournal.pone.0053513

[19] E. J. Golob and J. L. Holmes, "Cortical mechanisms of auditory spatial attention in a target detection task," *Brain Research*, vol. 1384, no. 0, pp. 128 – 139, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0006899311001971