Spatial Auditory BCI with ERP Responses to Front–Back to the Head Stimuli Distinction Support

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

***Abstract*—This paper presents recent results obtained with a new auditory spatial localization based BCI paradigm in which ERP shape differences at early latencies are employed to enhance classification accuracy in an oddball experimental setting. The concept relies on recent results in auditory neuroscience showing the possibility to differentiate early anterior contralateral responses to the spatial sources attended to. We also find that early brain responses indicate which direction, front or rear loudspeaker source, the subject attended to. Contemporary stimuli–driven BCI paradigms benefit most from the P300 ERP latencies in a so-called "aha-response" setting. We show the further enhancement of the classification results in a spatial auditory paradigm, in which we incorporate N200 latencies. The results reveal that these early spatial auditory ERPs boost offline classification results of the BCI application. The offline BCI experiments with the multi-command BCI prototype support our research hypothesis with higher classification results and improved information transfer rates.**

I. Introduction

A brain computer interface (BCI) utilizes human neurophysiological signals to communicate with an external computer or a machine, without depending on muscle activity [1]. Particularly, in the case of patients suffering from amyotrophic lateral sclerosis (ALS), this could help them to communicate or to complete various daily tasks (control a computer or type messages on a virtual keyboard, etc). This would create a very good option for ALS patients to communicate with their families, friends or caretakers by using only their brain waves. Hitherto, many approaches have focused on visual modality BCI applications. A visual modality BCI cannot be used by ALS patients who, in the advanced stages of the disease, often suffer from limited or lost sight. In this paper, we present the concept of an auditory BCI based on spatial sound stimuli, which we refer to in brief as saBCI (spatial auditory BCI). The saBCI concept is based on a basic feature of the human auditory brain pathway which is very sensitive to the localization of changing spatial auditory sources [2]. The auditory pathway also has a very good temporal resolution, which is an additional feature we would like to utilize in the saBCI design. This will allow the decrease of the inter–stimulus interval (ISI) of the presented sound stimuli in comparison to vision based applications [3].

The majority of contemporary BCI applications aim at the $P300$ response latency without consideration of the remaining event related potential (ERP) ranges. In this paper, we compare and discuss $N200$ response [4] suitability, and we show that its utilization improves the final classification results.

As a first step, we propose to utilize the early ERP latency based modulation related to the so-called "$N200$ anterior contralateral" ($N2ac$) response [4]. This response is characterized by different shapes in the brain ipsilateral and contralateral to targets ERPs. Our previous research also confirmed in [5] that the $N2ac$ setting improved the BCI classification accuracy. Using this setting allowed us to classify targets from left and right side loudspeaker sound sources, respectively. Psychophysical experiments also confirmed that the subjects had no discrimination problems related to auditory *front–back confusion*. Another previous study, conducted by the authors, further confirmed the feasibility of utilizing front and rear loudspeaker directions for the saBCI paradigm [6].

In this paper, we report on the new finding that the different ERP shape at the $N200$ latency supports the discrimination between auditory stimulus targets attended to, originating from front and rear loudspeakers respectively. This new finding allows us to identify the auditory stimulus direction to which the subject attended (front or rear). We call this new finding "a $N200$ front–rear" ($N2fr$) response.

We also report on the design of a new saBCI experimental paradigm based on the auditory spatial localization principle as the informative cue with support of both the $N2ac$ and $N2fr$ ERP components elicited in the new experimental setup, as depicted in Figure 1. Our hypothesis is that the new ERP component shall improve the classification results and the final information transfer rate (ITR) leading to better BCI usage comfort in general.

Within the novel saBCI paradigm framework, the subjects are asked, as in the usual oddball paradigm [1], to attend to and count the target stimuli from the instructed or intended directions, while ignoring the others. The EEG signals are recorded with a g.MOBIlab+ EEG amplifier by g.tec Medical Engineering GmbH, Austria. We use dry g.SAHARA electrodes by the same producer, which further improve the interfacing comfort, since there is no need to apply a conductive gel to the subject's scalp. In order to decrease the unnecessary and signal quality degrading muscular movement related electromyographic (EMG) noise on ERP responses, the subjects are requested to minimize their eye, facial and body movements in general during the experiments.

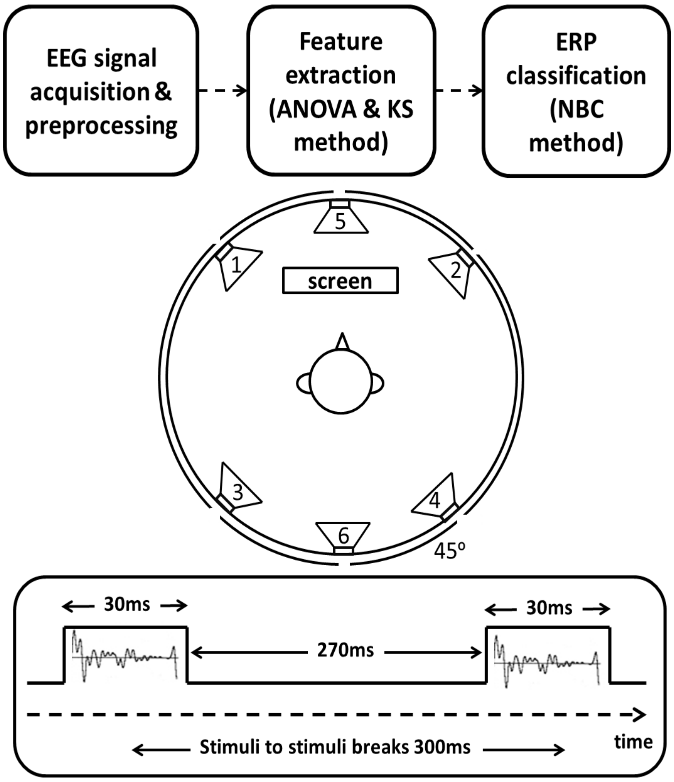

Fig. 1. The top panel depicts the novel localization paradigm for auditory sources to the front and rear of the subject's head, based on spatial sound stimuli. The bottom panel shows our stimulus presentation concept illustrated in the time domain. Each stimulus was presented in our experiments for 30 ms with 270 ms silent breaks, with an inter–stimulus interval (ISI) of 300 ms.

The remainder of the paper is organized as follows. In the next section, the experimental setup and the novel paradigm are described, together with EEG signal pre–processing steps. Next, analysis and optimization procedures of the ERP at $N200$ and $P300$ response latencies for all experimental subjects are described. Finally, classification and discussion of ITR results conclude the paper, together with future research directions.

II. Methods

The EEG experiments to validate the proposed spatial auditory BCI paradigm utilizing the $N200$ and $P300$ latency responses were conducted in the Multimedia Laboratory of the Life Science Center of TARA at the University of Tsukuba, Tsukuba, Japan. All the details of the experimental procedure and the research targets of this approach were explained to the seven human subjects, who agreed voluntarily to take part. The EEG BCI experiments were conducted in accordance with *The World Medical Association Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects*. The experimental procedures were designed in accordance with ethical committee guidelines of the institutions to which the authors of this paper are affiliated. The EEG signals were recorded by a g.MOBIlab+ EEG amplifier with six dry g.SAHARA electrodes. The sampling frequency was set to 256 Hz, with a notch filter to reject the 50 Hz power line noise, set to remove the EEG signal band in a range of $48 \sim 52$ Hz.

The auditory stimulus was presented through six loudspeakers distributed at an equal radius of one meter around the subject's head, as depicted in Figure 1. Three speakers at equal distances were positioned to the front and rear of the head. Two short *white* and *pink* noise stimulus bursts were used, as described in the following section. All the experiments were conducted in a silent, low reverberation room in order to limit environmental noise interference.

A. The Offline saBCI Experimental Protocol

The experimental hypothesis was that we would be able to distinguish from the ERP shape which direction left, right, front or rear to the head, the subject attended to based on the $N2ac$ and $N2fr$ responses.

To test the hypothesis, we conducted a series of EEG recording experiments in the offline BCI mode [1] (no instant feedback or classification results given to the subjects). The experiments were performed with the seven healthy subjects (six males and one female; age range $21 - 42$ with a mean of 26.4 years). The experimental procedure was explained in detail to each subject and her/his written consent was obtained. The subject was seated in the center of the experimental studio and the dry EEG electrodes were attached to the scalp. The subject's chair was positioned in the middle of the six surrounding loudspeakers. The elevation of the loudspeakers was set at the subject's ear level. A computer display with instructions about the experiment was set in front of the subject. The six loudspeakers were distributed in a circle with three loudspeakers $(1, 2, 5)$ positioned to the front at a $45°$ angular distance. The remaining three loudspeakers $(3, 4, 6)$ were located to the rear at the same angular distance (see Figure 1). Four loudspeakers $(1, 2, 3, 4)$ were used to test the $N2ac$ effect, and the other two loudspeakers $(5, 6)$ were used to confirm our hypothesis about $N2fr$ response.

The sound stimuli were presented in a random order, one at a time from a single loudspeaker (one trial consisted of the delivery of a single *target* and five *non-targets*). We decided to use two broadband noise stimulus types in order to utilize the two spatial localization mechanisms of the human auditory pathway (the inter-aural time and level differences – ITD/ILD) [2]. The *white* and *pink* noise stimuli of 30 ms lengths with 5 ms linear attack and sustain periods were chosen. The ISI was set to 300 ms. A single session consisted of the six single trials (6 *targets* from each direction, accompanied by 30 *non–targets*). The *target* direction in each trial was presented randomly together with five *non–targets*. For each subject and each stimuli, we performed 15 sessions (all together 90 *targets* and 450 *non-targets* were delivered). The target direction instruction was presented visually on a computer display and the auditory stimulus was given from the loudspeaker which the subject was later to attend to. Before each experiment the subjects were allowed a short practice session to familiarize themselves with the spatial auditory conditions.

III. Analysis of ERP Responses in Offline BCI Paradigm

In many current auditory BCI applications, the focus is put on a binary classification of brain evoked responses to *targets* versus *non–targets* [7], [6], [8], [9]. The majority of contemporary BCI applications aim at the $P300$ response latency without consideration of the remaining ERP ranges. Only one of the papers recently published has mentioned the $N200$ latency range as possibly useful in supporting classification [9], but there is no comparison made so far with the $P300$ only related results, which we present in this paper.

Basically, the concept of adding the early latency $N2ac$ or $N2fr$ responses is based on the concept contained in our previous research [6] and recent publications by other groups [4] about this ERP range modulation by *ipsilateral* vs. *contralateral* stimulus spatial locations, which results in the different ERP shapes. This difference confirms the feasibility of utilizing the early $N200$ response latency to improve the *target* vs. *non–target* classification accuracy. In order to analyze precisely the impact of the early ERP responses on the saBCI paradigm classification, we propose to conduct three separate analyses that compare results from:

- the $P300$ only based BCI classification;
- the $N2ac$ response based improvement;
- the further $N2fr$ related classification boosting.

A. EEG Preprocessing

The EEG signals captured by the dry electrodes were first filtered digitally with the two 5^{th}–order Butterworth high– and low–pass filters with cutoff frequencies of 0.5Hz and 25Hz, respectively.

The high–pass filtering removed very slow baseline drift related artifacts, as well as slow eye movements related EMG interference. The low–pass filter limited higher frequency EMG artifacts related to subject body muscle movements.

Next the EEG signals were segmented creating the ERP related *epochs*. Each *epoch* started 100 ms before each stimuli onset and ended 700 ms after it.

In the next step, the eye movement artifact rejection was carried out. Auditory spatial stimuli have been known to cause uncontrolled eye movements in subjects [10], which in the current approach were removed with a threshold value set at 80μV (signal amplitude level above the usual EEG activity). The rejected epochs were not further processed, since in the current approach, the emphasis was on the spatial paradigm validation. An example in Figure 2 shows the averaged and artifact–removed $P300$ responses to *target* and *non-target* sound stimuli with standard error bars respectively.

B. The Optimization of the EEG Electrode Locations and ERP Feature Extraction

In the previously reported research on the $N2ac$ phenomenon [4], the anterior cluster of electrodes sites $F3$, $F7$, $C3$, $T7$, $F4$, $F8$, $C4$, and $T8$ was used, as in the 10/20 *international system* [11]. In our experimental setup, we select the $F5$, $F6$, $C3$, $C4$, $P5$, and $P6$ electrodes in order to have additional responses from parietal cortices known to generate ERPs related to spatial and $P300$ responses [12]. Additionally, we show that the $P5$ and $P6$ sites are also useful to differentiate the responses to lateral stimuli similarly to the

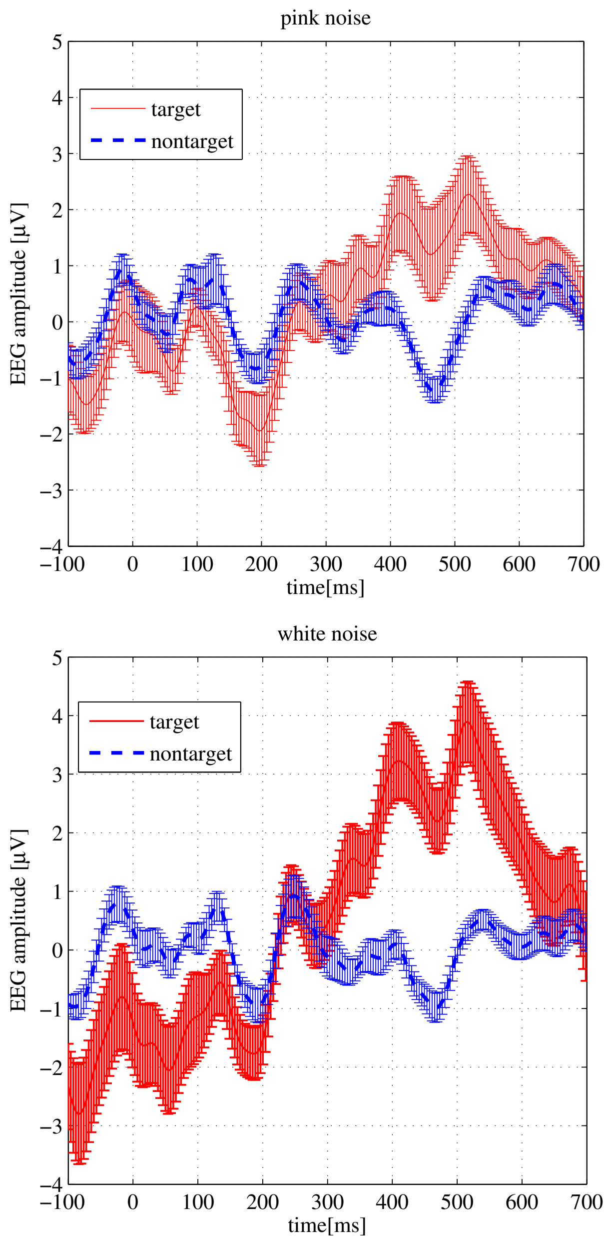

Fig. 2. Grand mean averaged ERP responses of the seven subjects. The upper panel presents the grand mean averaged ERP responses to pink noise. The lower panel depicts respective results obtained with white noise stimulus. The red lines represent *targets* and the blue lines *non–targets*. All the results are presented together with standard error bars. The differences between *targets* and *non–targets* are obvious in the range of 300 $\sim$ 600ms (the so-called “aha” or $P300$ response).

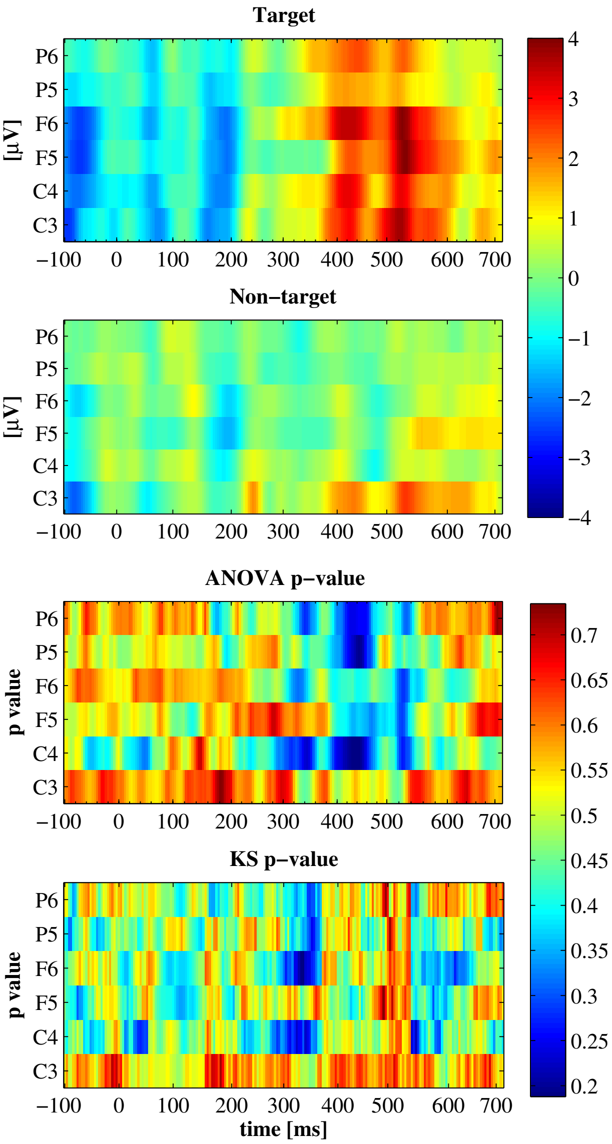

Fig. 3. Grand mean averaged ERP responses of all seven subjects and the six electrodes are plotted separately in the top two panels. The top panel shows the *target*, and the second from the top the *non–target* averaged responses, respectively. The significant differences between the responses can be seen, as visualized by color coding of the p-values obtained from the *ANOVA test* (statistical significance at $p < 0.05$) in the third panel from the top. The bottom panels present the *Kolmogorov–Smirnov test*. The EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiments.

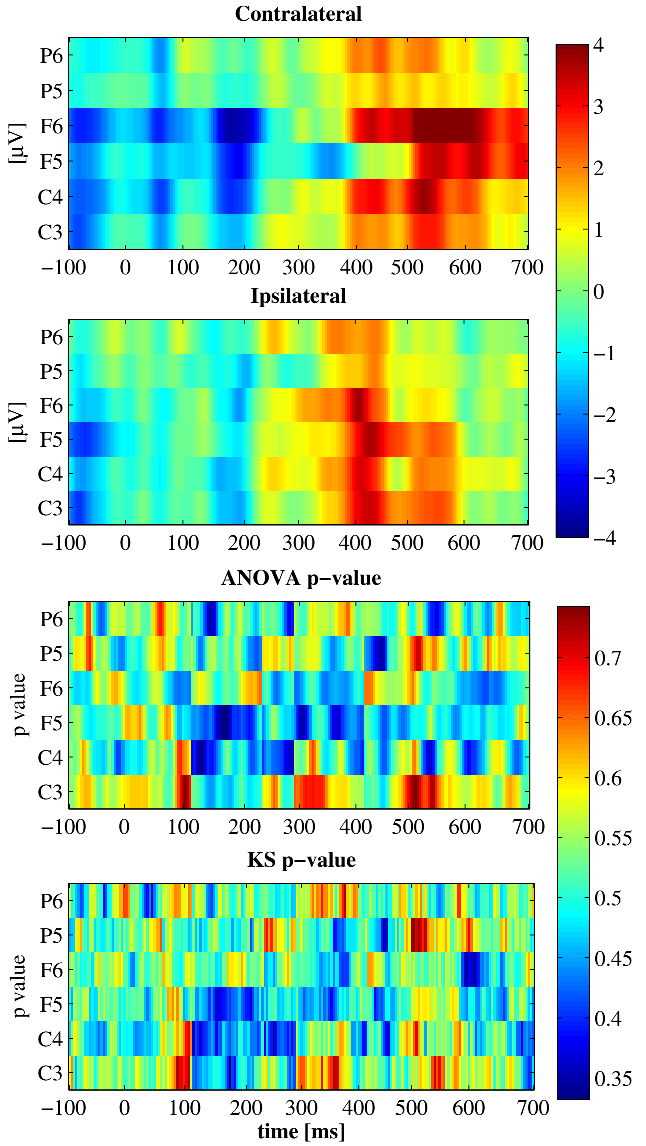

Fig. 4. Grand mean average ERP responses of all seven subjects and the six electrodes plotted separately in each panel in *contralateral* (top panel) vs. *ipsilateral* (second from the top panel) stimulus direction presentation settings. The significant differences between the two responses can be seen, as visualized by the color with p-values of *ANOVA test* and *KS test* results (statistical significance for $p < 0.05$) in the third panel and fourth panel, respectively. EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiment.

left vs. right only comparison revealed by $N2ac$.

An example in Figure 4 shows the averaged and artifact–removed $N2ac$ responses to *ipsilateral* and *contralateral* sound stimuli, as confirmed by our experiments. The $N200$ area responses presented are elucidated for *ipsilateral* and *contralateral* targets.

In order to validate statistically the differences between *target* and *non–target* responses, we conducted an analysis of variance test (ANOVA) [13] and Kolmogorov Smirnov test (KS) (to compare the two class ERP distribution similarities). We also used this method to analyze the two class ERP distributions in the *ipsilateral* vs. *contralateral*, as well as

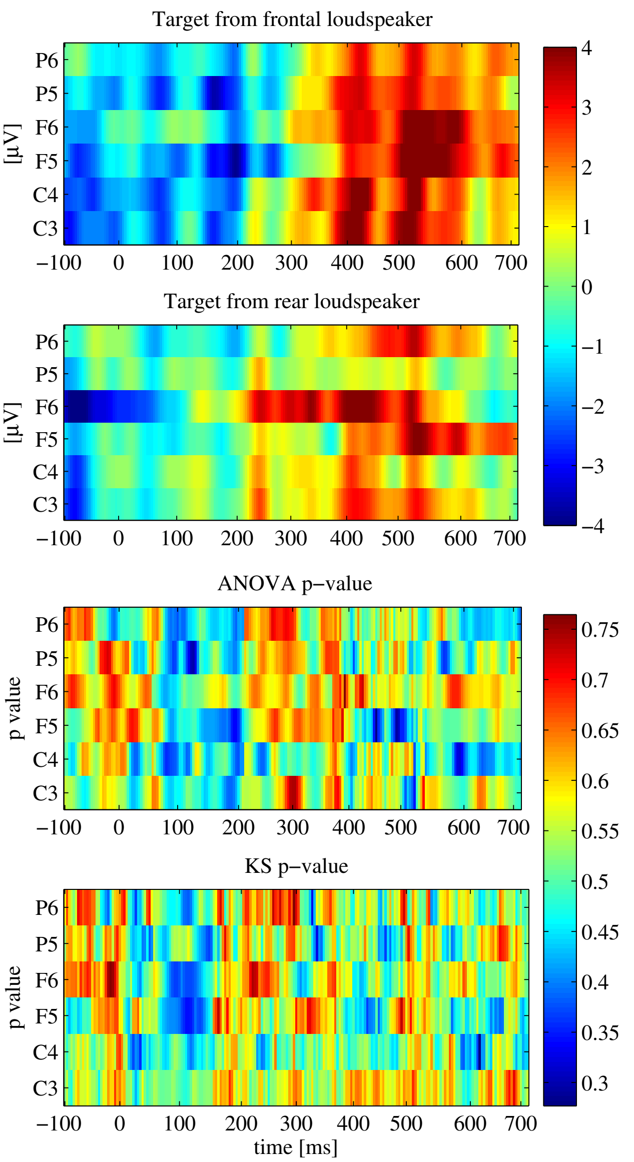

Fig. 5. Grand mean averaged ERP responses of all subjects and the six electrodes plotted separately in each panel for *the front* (top panel) and *the rear loudspeaker* (second panel from the top) ERP responses. The significant differences between the two responses can be seen, as visualized by the color with p-values of *ANOVA test* and *KS test* results (statistical significance at $p < 0.05$) in the third and fourth panels from the top, respectively. EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiment.

targets originating from the *the front* vs. *rear loudspeakers*, respectively. The ANOVA and KS test methods were applied to compare the differences in response distributions in single trials for each sample point of the collected ERPs. As a result we were able to extract discriminative information (in $N200$ and $P300$ latencies) leading to later classification optimization. The results of target and non-target analyses are depicted in Figure 3. The bottom panels in the above figure visualize the ANOVA and KS test $p-$value results, which are the probability of the null hypothesis rejections that the distributions from both the ERPs compared are significantly different (usually in life sciences, $p < 0.05$ is considered as a significant value). The ANOVA test results panel in Figure 3 show that the $P300$ related significant responses are located in the range from 300ms to 500ms. The ERP analysis results of ipsilateral vs. contralateral responses are shown in Figure 4. The $p-$values obtained from ANOVA and KS tests and presented in Figure 4 reveal that the $N200$ significant responses are located in the range from 100ms to $300ms$, similarly to the previously published $N2ac$ phenomenon. This finding confirms our hypothesis that the early $N200$–range latencies are related to spatial localization processes in the human brain and that the parietal electrodes also contribute to the result. The ERP analysis results of the $N2fr$ experimental setting are depicted in Figure 5. The $p-$values resulting from the ANOVA and KS tests confirm that the significant $N200$ response is located in the range from 100ms to 200ms. This difference is a source of BCI classification accuracy improvement, discussed below.

In this paper, three types of binary classification related feature extraction problems have been outlined so far:

1) First, we have evaluated ERP features separability in the $P300$ latency for classic *target* vs. *non-target* classification purposes.
2) Second, we have shown that the $N2ac$ latency has further contributed to the ERP features separability in the *ipsilateral* vs. *contralateral* stimulus presentation setting.
3) Finally, we have postulated and validated the new concept of *front* vs. *rear loudspeaker* stimulus response differences, which we named $N2fr$.

We propose to identify the most discriminable features from ERP responses using the ANOVA and KS tests described above, evaluating the statistical significance of the latency differences. Next, a "hand picking" procedure could be applied to only those ERP latencies within each subject's ERP p–values that are smaller than a reasonable threshold of 0.20 as depicted by blue shades in the third panels from the top and the bottom panels in Figures 3, 4 and 5.

In the next section we show that the proposed method of utilizing ANOVA and KS tests for $p < 0.20$ improves the saBCI classification results when utilizing the $N200$ latency responses.

C. The Offline saBCI Classification

In order to test the proposed feature extraction method, the offline saBCI mode of classification was performed for each subject separately, meaning that all the procedures were conducted after each data collection experiment, without any online feedback to the subjects. The classification procedure was performed in a so-called binary task paradigm of *targets* vs. *non–targets*, *contralateral* vs. *ipsilateral*, or *front* vs. *rear loudspeaker* targets respectively.

In each classifier training and testing step, 90 *targets* and a random subset of 90 *non-targets* were selected (from the 450 available) to have balanced numbers in the each class set. The resulting chance level was 50%.

For the *contralateral* vs. *ipsilateral* response classification, we selected 30 *contralateral* and 30 *ipsilateral* events.

To classify the *targets* from front vs. rear loudspeakers, we selected 15 *targets* from each direction respectively.

Based on our previous classification trials reported in [5], we decided to continue to use a Bayesian classifier, which for the datasets in our case outperforms the linear discrimination analysis methods. Despite its simplicity, the NBC (naive Bayesian classifier) approach often outperforms more sophisticated classification methods [14], on the assumption that the individual ERP features are statistically independent. It turns out that the NBC can be very robust, even to violations of the independence assumption [14].

The results of the successful application of the NBC technique are presented in the next section.

IV. Results

As the outcome of the presented research we have obtained results showing that for both the experimental settings of the saBCI offline paradigm, the classic $P300$ latency could be improved with the $N2ac$ and $N2fr$ features identified with the $p-$values calculated using the ANOVA or KS tests for significance. We summarize below the results obtained.

A. The Classification Results from the $P300$ ERP Latencies in the Classic Oddball Paradigm Setting

The first summary of classification results is presented in Table I, where classification accuracies for the features drawn from $P300$ are shown. All of the subjects performed above the chance level of 50% for single feature latencies of $P300$ based on ANOVA and KS tests $p-$values. The proposed "hand–picked" feature set identification using the ANOVA and KS tests of significant ERP samples allowed us to boost the classification results (in both cases with the maximum classification boost of 29%) using *leave–one–out* cross validation [14] for the NBC technique.

B. The Classification Results from the $N2ac$ ERP Feature in the Ipsilateral vs. Contralateral Settings

The results of the approach to compare *ipsilateral* and *contralateral* to target evoked potentials are summarized in Table II, based on the ERP features drawn from results of the ANOVA and KS tests depicted in Figure 4. The classification accuracy results were boosted 22% in the best case with ANOVA test based feature extraction. With the KS test, the classification accuracy results were even 28% enhanced in the best case. In both the enhancement cases, the same method of the NBC *leave–one–out* cross validation classification was applied.

C. The Classification Results of the $N2fr$ ERP Feature Extraction Method

The classification results of the proposed approach to compare target stimuli originating from *front* and *rear loudspeaker* are summarized in Table III, based on the ERP features selection from results of the ANOVA and KS tests analysis depicted in Figure 5. The classification accuracy results in this case were also 26% boosted.

TABLE I
THE CLASSIFICATION RESULTS FOR ERP LATENCIES IN $P300$ RESPONSES FOR THE *target* VS. *non–target* PARADIGM. THE THREE SETS (WHOLE ERP, $P300$ LATENCIES OPTIMIZED BY THE ANOVA METHOD AND KS METHOD) CLASSIFICATION RESULTS ARE COMPARED.

subject	noise stimulus type	conventional whole ERP [%]	ANOVA P300 [%]	KS P300 [%]
#1	pink	61	71	71
	white	65	68	67
#2	pink	61	85	78
	white	63	72	71
#3	pink	58	62	65
	white	55	59	67
#4	pink	49	68	53
	white	55	84	79
#5	pink	48	69	82
	white	56	83	81
#6	pink	55	70	66
	white	57	75	64
#7	pink	54	82	83
	white	65	81	72

TABLE II
THE CLASSIFICATION RESULTS FOR ERP LATENCIES IN $N200$ RESPONSES FOR THE *ipsilateral* VS. *contralateral* PARADIGM. THE THREE FEATURE SETS (WHOLE ERP ($0ms - 700$ MS), $N200$ LATENCIES OPTIMIZED BY THE ANOVA METHOD AND KS METHOD) CLASSIFICATION RESULTS ARE COMPARED.

subject	noise stimulus type	conventional whole ERP [%]	ANOVA N200 [%]	KS N200 [%]
#1	pink	51	58	57
	white	35	51	56
#2	pink	46	57	58
	white	38	64	59
#3	pink	51	61	57
	white	53	60	65
#4	pink	52	67	59
	white	51	67	71
#5	pink	55	60	58
	white	38	60	52
#6	pink	41	51	52
	white	43	48	57
#7	pink	53	57	65
	white	45	65	73

D. Analysis of Information Transfer Rate Improvement Results

The amount of information carried by every selection in the BCI application is usually quantified by the ITR, which is calculated based on bits per selection R, defined as in [7]:

$$R = \log_2 N + C \cdot \log_2 C + (1 - C) \cdot \log_2 \left(\frac{1 - C}{N - 1} \right), \quad (1)$$

TABLE III

THE CLASSIFICATION RESULTS FOR ERP LATENCIES IN $N200$ RESPONSES FOR THE *targets from front loudspeaker* VS. *targets from rear loudspeaker* PARADIGM. THE THREE FEATURE SETS (WHOLE ERP ($0-700$ MS), $N200$ LATENCIES OPTIMIZED BY THE ANOVA METHOD AND KS METHOD) CLASSIFICATION RESULTS ARE COMPARED.

subject	noise stimulus type	conventional whole ERP [%]	ANOVA N200 [%]	KS N200 [%]
#1	pink	46	52	50
	white	51	63	77
#2	pink	40	66	53
	white	36	51	56
#3	pink	52	65	58
	white	51	55	67
#4	pink	63	73	70
	white	53	60	57
#5	pink	43	57	56
	white	46	57	51
#6	pink	37	56	58
	white	52	62	56
#7	pink	57	68	65
	white	44	51	55

TABLE IV

THE ITR OF TARGET VS. NON-TARGET, SEE EQUATIONS (1) AND (2), FOR THE THREE ERP SETS RELATED CLASSIFICATION APPROACHES USING $wholeERP$, $P300$ EXTRACTED WITH ANOVA AND KS.

subject	noise stimulus type	Whole ERP [bit/min]	ANOVA P300 [bit/min]	KS P300 [bit/min]
#1	pink	3.25	13.13	13.13
	white	6.59	9.56	8.51
#2	pink	3.52	39.02	23.98
	white	4.93	14.45	13.13
#3	pink	1.85	4.19	6.59
	white	0.72	2.35	8.51
#4	pink	0.00	9.56	0.26
	white	0.72	36.57	25.85
#5	pink	0.00	10.68	31.99
	white	1.04	34.23	29.85
#6	pink	0.72	11.87	7.52
	white	1.42	18.87	5.73
#7	pink	0.46	31.99	34.23
	white	6.59	29.85	14.45

TABLE V

THE ITR OF N2AC, SEE EQUATIONS (1) AND (2), FOR THE THREE ERP SETS RELATED CLASSIFICATION APPROACHES USING $wholeERP$, $N200$ EXTRACTED WITH ANOVA AND KS.

subject	noise stimulus type	Whole ERP [bit/min]	ANOVA N200 [bit/min]	KS N200 [bit/min]
#1	pink	0.03	1.85	1.42
	white	0.00	0.03	1.04
#2	pink	0.00	1.42	1.85
	white	0.00	5.73	2.35
#3	pink	0.03	3.52	1.42
	white	0.26	2.90	6.59
#4	pink	0.12	8.51	2.35
	white	0.03	8.51	13.13
#5	pink	0.72	2.90	1.85
	white	0.00	2.90	0.12
#6	pink	0.00	0.03	0.12
	white	0.00	0.00	1.42
#7	pink	0.26	1.42	6.59
	white	0.00	6.59	15.85

TABLE VI

THE ITR OF TARGETS FROM FRONT LOUDSPEAKER AND TARGETS FROM REAR LOUDSPEAKER, SEE EQUATIONS (1) AND (2), FOR THE THREE ERP SETS RELATED CLASSIFICATION APPROACHES USING $wholeERP$, $N200$ EXTRACTED WITH ANOVA AND KS.

subject	noise stimulus type	Whole ERP [bit/min]	ANOVA N200 [bit/min]	KS N200 [bit/min]
#1	pink	0.00	0.12	0.00
	white	0.03	4.93	22.19
#2	pink	0.00	7.52	0.26
	white	0.00	0.03	0.72
#3	pink	0.12	6.59	1.85
	white	0.03	0.72	8.51
#4	pink	4.93	15.85	11.87
	white	0.26	2.90	1.41
#5	pink	0.00	1.42	1.04
	white	0.00	1.42	0.03
#6	pink	0.00	1.04	1.85
	white	0.12	4.19	1.04
#7	pink	1.42	9.56	6.59
	white	0.00	0.03	0.72

where C is the classification accuracy and N is the number of classes ($N = 2$ in this paper). The final *bit per minute rate* B is obtained after multiplication by a classification speed V, resulting in bits per minute [bit/min] as:

$$B = V \cdot R \quad (2)$$

The ITR results are summarized in Tables IV, V, and VI respectively. For all the cases of the $P300$, the $N2ac$, and the $N2fr$, there were significant increases of the ITRs for the majority of subjects.

V. CONCLUSIONS

In this paper, we have presented three approaches leading to the improvement of the classification accuracy and ITR in the offline saBCI paradigm. The improvement was obtained by introducing the novel ERP feature extraction methods in the $P300$ latency, the $N2ac$ and the $N2fr$.

The first improvement has been presented in the form of a comparison of classification rates for the three ERP feature sets of the $P300$ latencies processed separately with ANOVA and KS tests, versus the classic whole ERP features. The classification accuracy was increased for all the subjects at a maximum of 29% boost (ITR improvement at a maximum of 34 bit/min). This is a very good result, giving the possibility to further improve the auditory based BCI paradigm.

The second improvement step was based on the $N2ac$ concept. The classification obtained and ITR improvement are also very encouraging.

The third improvement step was based on the $N2fr$ responses in the front–back to the subject head stimulus presentation. The classification and ITR improvement confirm that our new findings can improve the final classification results.

The three main achievements reported in this paper allow us to boost the accuracy of the novel saBCI paradigm in offline mode, which is a step forward in non–vision based interfacing strategies.

The results obtained reveal that not only the cortical auditory information processing centers related to the cognitive streams can be utilized for BCI purposes. Also the differences in ERPs at early latencies before 300 ms are useful and they guarantee good classification results and resulting ITR scores. These results reveal that very early spatial auditory ERPs are potentially interesting for faster BCI applications.

AUTHOR CONTRIBUTIONS

Performed the EEG experiments and analyzed the data: ZC, TMR. Conceived the concept of the spatial auditory BCI and designed the EEG experiments: TMR. Supported the project: SM. Wrote the paper: ZC, TMR.

ACKNOWLEDGEMENTS

This research was supported in part by the Strategic Information and Communications R&D Promotion Program (SCOPE) 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] J. Schnupp, I. Nelken, and A. King, *Auditory Neuroscience - Making Sense of Sound*. MIT Press, 2010.

[3] 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.

[4] M. L. Gamble and S. J. Luck, "N2ac: An ERP component associated with the focusing of attention within an auditory scene," *Psychophysiology*, vol. 48, pp. 1057–1068, 2011.

[5] Z. Cai, S. Makino, T. Yamada, and T. M. Rutkowski, "Multichannel analysis of ERP at early and P300 latencies in application for spatial auditory BCI improvement," in *Proceedings of the Fourth APSIPA Annual Summit and Conference (APSIPA ASC 2012)*, (Xi'an, China), p. paper 355 (7 pages), APSIPA, December 3 - 6, 2012.

[6] 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), p. paper 241 (6 pages), APSIPA, October 18 - 21, 2011.

[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, p. e9813, 04 2010.

[8] M. Schreuder, T. Rost, and M. Tangermann, "Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI," *Frontiers in Neuroscience*, vol. 5, p. Article 112, 2011.

[9] J. Hoehne, M. Schreuder, B. Blankertz, and M. Tangermann, "A novel 9-class auditory ERP paradigm driving a predictive text entry system," *Frontiers in Neuroscience*, vol. 5, p. Article 99, 2011.

[10] T. M. Rutkowski, "Auditory brain-computer/machine interface paradigms design," in *Haptic and Audio Interaction Design* (E. Cooper, V. Kryssanov, H. Ogawa, and S. Brewster, eds.), vol. 6851 of *Lecture Notes in Computer Science*, pp. 110–119, Springer Berlin / Heidelberg, 2011.

[11] 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.

[12] E. Niedermeyer and F. L. Da Silva, eds., *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 5 ed., 2004.

[13] H. F. van Emden, *Statistics for Terrified Biologists*. Blackwell Publishing, 2008.

[14] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic Press, fourth ed., 2009.