Spatial Auditory BCI Paradigm Utilizing N200 and P300 Responses

Zhenyu Cai*, Shoji Makino*, Takeshi Yamada*, Tomasz M. Rutkowski*†▽
* TARA Center, University of Tsukuba, Tsukuba, Japan
E-mail: tomek@tara.tsukuba.ac.jp (▽corresponding author)
Tel: +81-29-853-6261
† RIKEN Brain Science Institute, Wako-shi, Japan

Abstract—The paper presents our recent results obtained with a new auditory spatial localization based BCI paradigm in which the ERP shape differences at early latencies are employed to enhance the traditional P300 responses in an oddball experimental setting. The concept relies on the recent results in auditory neuroscience showing a possibility to differentiate early anterior contralateral responses to attended spatial sources. Contemporary stimuli–driven BCI paradigms benefit mostly from the $P_{300}$ ERP latencies in so called “aha-response” settings. We show the further enhancement of the classification results in spatial auditory paradigms by incorporating the $N_{200}$ latencies, which differentiate the brain responses to lateral, in relation to the subject head, sound locations in the auditory space. The results reveal that those early spatial auditory ERPs boost online classification results of the BCI application. The online BCI experiments with the multi-command BCI prototype support our research hypothesis with the higher classification results and the improved information-transfer-rates.

I. INTRODUCTION

The brain computer interface (BCI) utilizes the human neurophysiological signals to control an external computer or a machine [1]. BCI does not depend on muscle or peripheral nervous system activities. Particularly, in case of amyotrophic lateral sclerosis (ALS) suffering patients, it 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 the caretakers by only using their brain waves. So far, many approaches are focusing on the visual modality BCI applications. The visual modality BCI can not be used by ALS patients who often suffer in the advanced disease stages from limited or lost sight. We present in this paper a concept of an auditory BCI based on spatial sound stimuli, which we call shortly saBCI (spatial auditory BCI). The saBCI concept is based on a basic feature of the human auditory pathway which is very sensitive to localization of changing spatial auditory sources [2]. The auditory pathway has also very good temporal resolution, which is an additional feature we would like to utilize in the saBCI design. This will allow to decrease stimuli onset asynchrony (SOA) of the presented sound stimuli in comparison to vision based applications [3].

The contemporary stimuli–driven BCI approaches are mostly based on P300 responses to distinguish targets and non–targets from series of event related potential (ERP) responses [4]. Recently a new result [5] was published elucidating the $N_{200}$-anterior-contralateral ($N_{2ac}$) component at the early latency (around 200ms) of an auditory ERP. The $N_{2ac}$ was obtained in an experiment using two 750ms long sound stimuli which were presented simultaneously from a different loudspeaker each. Subjects were requested to attend to the instructed target sound that could occur from any loudspeaker.

We propose to design the new saBCI experimental paradigm based on the auditory spatial localization principle as the informative cues with support of the $N_{2ac}$ component elicited in the new 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 a better BCI usage comfort.

Within the novel saBCI paradigm framework, the subjects

![Fig. 1. The novel $N_{2ac}$ paradigm based on spatial sound stimuli](image-url)
are asked, as in usual oddball paradigm, to attend and count the target stimuli from the instructed or intended direction, while ignoring the other. The EEG signals are recorded with g.MOBlab+ EEG amplifier by g.tec. We use the novel dry g.SAHARA electrodes by the same producer which further improve the interfacing comfort, since there is no need to apply a conductive gel. In order to decrease the unnecessary and signal quality degrading muscular movement related electromyography (EMG) noise on ERP responses, the subjects are asked to minimize their eye, facial and body in general movements during the experiments.

In the previous paper [6] we proposed a channel and ERP–latency selection in order to improve the classification results. At that time the $P_{200}$ response (so-called “aha–response” at the latency around 300 ms elicited to the expected/instructed target stimulus [7]) was the major feature used for the classification of the attended targets of the oddball paradigm. In this paper, we introduce the early latencies around 200 ms ($N_{200}$), and the $P_{300}$ which precede the $P_{300}$. They shall improve the final classification rates of the saBCI application.

The objective of this paper is to test and confirm our working hypothesis that the auditory evoked response based on $N_{200}$ paradigm should improve the saBCI application classification rates based on the new lateral to the subject head stimuli responses analysis.

From now on the paper is organized as follows. In the next section the experimental setup and the novel paradigm is described together with EEG signals pre-processing steps. Next the analysis and optimization procedures of the ERPs at $N_{200}$ and $P_{300}$ response latencies for all experimental subjects are described. Finally classification and ITR results discussion conclude the paper together with future research directions.

II. METHODS

The EEG experiments to validate the proposed spatial auditory BCI paradigm utilizing the $N_{200}$ and $P_{300}$ latency responses have been conducted in Multimedia Laboratory in TARA Life Science Center at the University of Tsukuba, Tsukuba, Japan. All the experimental procedure details and this approach research targets have been explained to the seven human subjects who agreed voluntarily to attend. The experimental procedures are designed in accordance with ethical committee guidelines of this paper author affiliated institutions. The EEG signals are recorded by the g.USBamp EEG amplifier with the six dry g.SAHARA electrodes. The sampling frequency is set to 256Hz with a notch filter to reject the 50Hz power–line noise.

The auditory stimuli has been presented through six loudspeakers distributed with an equal radius of 1 meter around the subject’s head as depicted in Figure 1. Three speakers with equal distances are positioned at each lateral side to the head. Two short white– and pink–noise stimulus bursts are used as described in the following section. All the experiments are conducted in a silent and low reverberation room in order to limit an interference of “an environmental noise.”

A. The Offline saBCI Experimental Protocol

The experimental hypothesis is that we shall be able to distinguish from the ERP shape which direction (left or right) the subject attends based on the novel $N_{200}$ response analysis method.

To test the hypothesis we conduct a series of EEG recording experiments in the offline BCI mode [1] (no instant feedback or classification results given to the subject). The experiments are performed with the seven healthy subjects (six males and one female; age range 21 – 42 with the mean of 26.4 years old). The experimental procedure has been explained in detail to each subject and her/his consent has been obtained. The subject is seated in the center of the experimental studio and the dry EEG electrodes are attached on the scalp. The subject’s chair position is surrounded by the six loudspeakers. The elevation of the loudspeakers is fixed to the subject’s ear level. A computer display with experimental instruction is set in front of the subject. The six loudspeakers are distributed on a circle with three loudspeakers (1, 3, 5) positioned on the left side with 45–degrees angular distance. The remaining three loudspeakers (2, 4, 6) are located on the right side with the same angular distances (see Figure 1).

The sound stimulus is presented in random order one at a time from a single loudspeaker (a single trial consists of a delivery of a single target and five non-targets). We decide to use two broadband noise stimulus types in order to utilize two spatial localization mechanism 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 5mn linear attack and sustain periods has been chosen. The SOA is set to 300ms. The single session consists of the six single trials (6 targets from each direction accompanied by 30 non–targets). The target direction in each trial is presented randomly together with five non–targets. For each subject and each stimuli we perform 15 sessions (all together 90 targets and 450 non-targets are delivered). The target direction instruction is presented visually on a computer display and auditory from the same loudspeaker which subject shall latter attend to. Before each experiment the subjects are allowed for a short practice session to familiarize themselves with spatial auditory conditions.

III. THE 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 [4], [6], [8], [9]. The majority of the contemporary BCI applications aim at the $P_{300}$ response latency without consideration of the remaining ERP ranges. Only a single of recently published papers mentions the $N_{200}$ latency range as possibly useful to support classification [9], but there is no comparison made so far with $P_{300}$ only related results, what we attempt in this paper. We compare and discuss the $N_{200}$ response suitability and we show that it really improves the classification results.
Basically a concept of adding the early latency $N20c$ response is based on our previous [6] research and the recently published by other groups [5] concept of this ERP range modulation by ipsilateral vs. contralateral stimulus spatial locations. The ipsilateral $N20c$ response has higher amplitude comparing to the contralateral one. This difference confirms a feasibility to utilize the early $N200$ response latency to improve the target vs. non–target classifications outcomes.

In order to precisely analyze an impact of the early ERP repose on the saBCI paradigm classification we propose to conduct two separate analyses that shall compare how much the improvement depends only on the $N200$ response feature addition, and how much on the new feature composition based on the comparison of the ipsilateral and contralateral responses as in $N20c$ design.

A. EEG Preprocessing

The EEG signals captured by the g.MOBIIlab+ system with g.SAHARA dry electrodes are first filtered digitally with the two $5^{th}$–order Butterworth high– and low–pass filters with cutoff frequencies at 0.5Hz and 25Hz, respectively.

The high–pass filtering removes the very slow baseline drift related artifacts as well as the slow eye movements related EMG interferences. The low–pass filter limits the higher frequency EMG artifacts related to subject body muscle movements.

Next the EEG signals are segmented creating the ERP related epochs. Each epoch starts 100ms before each stimuli onset and ends 700ms after it. We use the 100ms pre–stimuli onset interval for a baseline correction procedures.

In the next step the eye movement artifacts rejection is carried out. Auditory spatial stimuli are known to cause in subjects the uncontrolled eye movements [10] which in the current approach are removed with a threshold value set at $80\mu V$ (signal amplitude level above the usual EEG activity). The rejected epochs are not further processed, since in the current approach an emphasis is focused on the spatial paradigm validation.

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

In the previously reported research on $N20c$ phenomenon [5] the anterior cluster of electrodes sites F3, F7, C3, T7, F4, F8, C4, and T8 was used, as in 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 [7]. Additionally we show that the P5 and P6 sites are also useful to differentiate the responses to lateral stimuli similarly as for left–right only comparison revealed by $N20c$. We call the new finding the $N20pc$ ($N200$–anterior–posterior–contralateral) as extension of the former one. An example in Figure 2 shows the averaged and artifact–removed classical $N20c$ responses to ipsilateral and contralateral sound stimuli as confirmed by our experiments. The presented $N200$ area responses are elucidated for ipsilateral and contralateral targets.

In order to validate statistically the differences between target and non–target responses we conduct the $t$–test analysis of the two class ERP means [12] in ipsilateral vs. contralateral experimental setting. The $t$–test method is applied to compare the differences of response distributions in single trials for each sample point of the collected ERPs. As the result we can extract discriminative information (in $N200$ and $P300$ latencies) leading to later classification optimization. The results of the above analysis are depicted in Figure 3. A color bar located on a time scale in the above figure visualizes the $t$–test’s $p$ value results, which is a probability of the null hypothesis rejections that the means from the both compared distributions are significantly different (usually $p < 0.05$ in life sciences is considered as the significant value). The color bar in the Figure 3 clearly shows that the postulated $N20pc$ differential response for lateral responses is located in the range from 100ms to 300ms, similarly to the previously published $N20c$ one. 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 contribute also to the result.
In this paper two types of binary classification problems are discussed. First we evaluate our first hypothesis that adding the early latency ERP periods as features improves the binary classification. Next we also show that the NBC can be very robust also to violations of the independence assumption.

Consider the vector $x$ with features according to the values $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

$$P(x_1|x_2) = P(x_1|x_2) = q_i,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

$$\omega_m = \arg \max \omega_i \prod_{j=1}^{l} p(x_j | \omega_i), \quad i = 1, 2, \ldots, M,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

$$P(x_1|x_2) = P(x_1|x_2) = q_i,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

$$\omega_m = \arg \max \omega_i \prod_{j=1}^{l} p(x_j | \omega_i), \quad i = 1, 2, \ldots, M,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

### C. The Offline sABCi Classification

We perform the classification steps for each subject separately in sABCi offline mode, which means that all procedures are conducted after each experiment of data collection, without any online feedback to subjects. The classification procedure is performed in a so-called binary task paradigm (we classify **target** vs. **non-target**, or **contralateral** vs. **ipsilateral** response pairs each time only).

In each classifier training and testing step we select 90 **targets** and a random subset of 90 **non-targets** (from the 450 available) to have the balanced number of the members in each class set. The resulting chance level is 50%. For the case of the **contralateral** vs. **ipsilateral** responses classification we select 30 **contralateral** and 30 **ipsilateral** events.

Based on our previous classification trials reported in [6] we decide to use a Bayesian classifier, which outperforms the linear discrimination analysis methods. The **naive-Bayses classifier** (NBC) is particularly suited for the highly dimensional features. Despite its simplicity, the NBC approach often outperforms more sophisticated classification methods [13]. The NBC application assigns an unknown sample (ERP features in our case) $x = [x_1, x_2, \ldots, x_l]^T$ based on probability maximization to the class

$$\omega_m = \arg \max \omega_i \prod_{j=1}^{l} p(x_j | \omega_i), \quad i = 1, 2, \ldots, M,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

Consider the vector $x$ with features according to the values of the ERP “hand picked” samples. The respective conditional probabilities shall be $P(x_1|\omega_1) = p_i$ and $P(x_1|\omega_2) = q_i$, in our binary classification case comparing **targets** vs. **non-targets** or **ipsilateral** vs. **contralateral** responses. In Bayesian rule, given the value of $x$ the class membership is decided according to the probabilities likelihood ratio

$$P(x_1|x_2) = P(x_1|x_2) = q_i,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

$$\omega_m = \arg \max \omega_i \prod_{j=1}^{l} p(x_j | \omega_i), \quad i = 1, 2, \ldots, M,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

$$\omega_m = \arg \max \omega_i \prod_{j=1}^{l} p(x_j | \omega_i), \quad i = 1, 2, \ldots, M,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.

$$\omega_m = \arg \max \omega_i \prod_{j=1}^{l} p(x_j | \omega_i), \quad i = 1, 2, \ldots, M,$$

with an assumption that the individual features $x_j$, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption.
based on the following substitutions

\[ w = \left[ \ln \frac{p_1(1-q_1)}{q_1(1-p_1)}, \ldots, \ln \frac{p_i(1-q_i)}{q_i(1-p_i)} \right]^T \]

\[ u_0 = \sum_{i=1}^{n} \ln \frac{1-p_i}{1-q_i} + \ln \frac{P(\omega_1)}{P(\omega_2)}. \]

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

IV. RESULTS

As the result of the presented research have obtained the results showing that for both the experimental settings of saBCI offline paradigm the classical P300 latency could be improved with the pure N200 or the more complex N2apc features identified with p–values calculated using the classical t–test for significance. We summarize below the obtained results.

A. The Classification Results from the Combined N200 and P300 ERP Latencies in the Classical target vs. non–target Setting

The first summary of classification results is presented in Table I, where classification accuracies for the features drawn from N200, P300 and the combined latencies are shown. The majority of the subjects performed already above the chance level of 50% (except subject MA for the pink noise case) for single feature latencies of N200 or P300. The proposed combination of the two “hand–picked” feature sets using the t–test significant ERP samples allowed us to boost the classification results up to 7% (only a single case of the accuracy decrease has been reported) using the leave–one–out cross validation [13] for the NBC technique.

**TABLE I**

<table>
<thead>
<tr>
<th>subject</th>
<th>noise stimulus type</th>
<th>N200 only [%]</th>
<th>P300 only [%]</th>
<th>N200/P300 combined [%]</th>
<th>N200/P300 vs. P300 [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZH</td>
<td>pink white</td>
<td>63</td>
<td>63</td>
<td>64</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TO</td>
<td>pink white</td>
<td>52</td>
<td>54</td>
<td>56</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NI</td>
<td>pink white</td>
<td>53</td>
<td>54</td>
<td>58</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MO</td>
<td>pink white</td>
<td>59</td>
<td>60</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td>pink white</td>
<td>53</td>
<td>56</td>
<td>65</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>pink white</td>
<td>64</td>
<td>66</td>
<td>66</td>
<td>2</td>
</tr>
</tbody>
</table>

B. The Classification Results from the new N2apc ERP Feature in the Ipsilateral vs. Contralateral Settings

The results of the proposed approach to compare ipsilateral and contralateral to target evoked potentials have been summarized in the Table II, based on the ERP features drawn from results of the t–test analysis as summarized in the Figures 4 and 5. The classification accuracy results have been 17% boosted in the best case, with the same method of the NBC leave–one–out cross validation.

**TABLE II**

<table>
<thead>
<tr>
<th>subject</th>
<th>noise stimulus type</th>
<th>N200 only [bit/min]</th>
<th>P300 only [bit/min]</th>
<th>N200/P300 combined [bit/min]</th>
<th>the improvement [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZH</td>
<td>pink white</td>
<td>25.84</td>
<td>25.84</td>
<td>26.88</td>
<td>5.04</td>
</tr>
<tr>
<td>TO</td>
<td>pink white</td>
<td>15.72</td>
<td>17.38</td>
<td>17.38</td>
<td>1.74</td>
</tr>
<tr>
<td>NI</td>
<td>pink white</td>
<td>16.14</td>
<td>20.02</td>
<td>20.02</td>
<td>0.00</td>
</tr>
<tr>
<td>MO</td>
<td>pink white</td>
<td>22.83</td>
<td>20.94</td>
<td>20.94</td>
<td>0.00</td>
</tr>
</tbody>
</table>

C. 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 [4]:

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

where \( C \) is the classification accuracy and \( N \) is the number of classes (\( N = 6 \) in this paper). The final bit–per–minute–rate...
In this paper we presented two approaches leading to improvements of classification accuracy and ITR in offline saBCI paradigm by introducing the novel ERP feature extraction in combined N200/P300 latencies and in the new N2apc setting which compares responses of lateral, to the head, sound sources. The first improvement analysis resulted in a comparison of classification rates for the three ERP feature sets of N200 and P300 latencies processed separately, versus the combined N200/P300. The latter combination resulted in a steady increase in classification accuracy for the majority of subjects up to 7% at maximum. Additionally the ITR improvement in this case was reported at maximum of 7bit/min. This is a very good result giving a possibility to further improve the auditory paradigm based BCI.

The second improvement step is based on the proposed extension of N2apc concept. We added a comparison of parietal electrodes responses allowing for the new feature creation from such ERP comparisons. The new ERP component was named N2apc since it combines anterior and posterior contralateral response differences. The obtained classification and ITR improvement was also very encouraging.

The two main achievements reported in the paper allowed us to improve the novel saBCI paradigm in offline mode which is a step forward in the non–vision based interfacing strategies. The obtained results reveal that not only the cortical auditory information processing centers related to the cognitive streams could be utilized to BCI purposes. Also the differences in ERPs at early latencies before 300ms are useful and they guarantee good classification results and ITRs. These results reveal that the very early spatial auditory ERPs are potentially interesting for faster BCI applications.

### V. CONCLUSIONS

The ITR results are summarized in Tables III and IV. For the both cases of the N200/P300 combination and the N2apc paradigm, there is a significant increase of ITR for the majority of subjects.

### REFERENCES

Fig. 4. ERP to pink noise stimuli grand mean averages for all subjects and the six electrodes plotted separately in each panel. The solid red lines represent the ipsilateral to target responses and the dashed blue lines to the contralateral ones. The color bars at the bottom of each panel show the t-test resulting p-values.

Fig. 5. ERP to white noise stimuli grand mean averages for all subjects and the six electrodes plotted separately in each panel. The solid red lines represent the ipsilateral to target responses and the dashed blue lines to the contralateral ones. The color bars at the bottom of each panel show the t-test resulting p-values.