34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

How stimulation speed affects Event-Related Potentials and BCI performance

Johannes Höhne[1] and Michael Tangermann[2]

Abstract— **In most paradigms for Brain-Computer Interfaces (BCIs) that are based on Event-Related Potentials (ERPs), stimuli are presented with a pre-defined and constant speed. In order to boost BCI performance by optimizing the parameters of stimulation, this offline study investigates the impact of the stimulus onset asynchrony (SOA) on ERPs and the resulting classification accuracy. The SOA is defined as the time between the onsets of two consecutive stimuli, which represents a measure for stimulation speed. A simple auditory oddball paradigm was tested in 14 SOA conditions with a SOA between 50 ms and 1000 ms. Based on an offline ERP analysis, the BCI performance (quantified by the Information Transfer Rate, ITR in bits/min) was simulated. A great variability in the simulated BCI performance was observed within subjects (N=11). This indicates a potential increase in BCI performance ($\geq$ 1.6 bits/min) for ERP-based paradigms, if the stimulation speed is specified for each user individually.**

I. INTRODUCTION

Using a Brain-Computer Interface (BCI), users can send control signals to an application, even if they are unable to control any muscle. Recent research aims to develop novel BCI paradigms with a high rate of communication. Most of these approaches are based on Event-Related Potentials (ERPs), which are the EEG responses triggered by a perceived event or stimulus. Various paradigms were proposed using the visual [1] or auditory [2], [3], [4] modality of stimulation. Most of those ERP paradigms follow the oddball principle of rare target and frequent non-target events. But they differ in the choice and presentation mode of stimuli. Thus, it is reasonable to boost the classification accuracy and BCI performance by optimizing the stimulus characteristics. For the visual and auditory modality, this can be achieved by finding stimulation procedures that elicit the strongest possible class-discriminative components [5], [6], [7], [8].

Another parameter that can be modified is the stimulation speed, which is often described by the stimulus onset asynchrony (SOA) or inter stimulus intervals (ISI). The SOA specifies the time between the onsets of two consecutive stimuli. Most BCI paradigms are applied with a SOA value between 83 ms [1] and 500 ms [9]. Comparing the visual BCI performance of two SOA levels (175 ms and 350 ms), Sellers et al. [10] already stated in 2006 that the choice of SOA highly affects the BCI performance, concluding that "it appears to be worthwhile to test multiple ISI values and thereby determine the optimal value for each user". Nevertheless, the exact choice of stimulation speed has not yet been considered to be crucial, thus it was not optimized by any means.

In the present study, the parameter SOA was investigated with respect to the impact on classification accuracy and BCI performance in a simple auditory oddball paradigm. Classical ERP literature [11] describes decreasing amplitudes of class-discriminative ERP components such as P300 for decreasing SOA values and target-to-target intervals (TTI). Consequently, it is expected that the binary classification accuracy (target vs. non-target) correlates with the SOA, such that fast SOA conditions result in a lower accuracies than slow SOA conditions. But although speeding up the stimulation might lead to a reduced classifiability per stimulus, the rate of stimulation is increased. Thus, there may be more stimuli, with each stimulus carrying less discriminative information, which could result in an increased BCI performance. Accordingly, finding the best SOA for a BCI user corresponds to finding the optimal trade-off between the rate of stimuli and the classifiability per stimulus.

II. METHODS

A. Experimental design

Within a single session of about 3 hours, a simple auditory oddball paradigm was tested in 14 SOA conditions. The same type of experiment was performed with varying stimulation speed: a SOA between 50 ms and 1000 ms. The exact SOA conditions are shown at the bottom of Fig. 3c. The experiment was divided into four parts, each part consisting of eight blocks with randomized order of conditions. Within each block, there were four consecutive trials of the same condition. In each trial, participants had to concentrate on a rare target tone while neglecting the frequent (83.4 %) non-target tone. Both types of stimuli were sinusoidal with a duration of 40 ms. The target tone had a high pitch (1000 Hz) and the non-target tone had a low pitch (500 Hz). Each trial consisted of 72–90 stimulus presentations (16.6 % targets), and the participant had the task to mentally count the occurrences of the target stimulus. In total, this leads to 1296 events (216 targets and 1080 non-targets) in each condition. Within one trial, the sequence of targets and non-targets was randomized, while it was assured that there were at least three non-targets between two consecutive target stimuli. While attending to the auditory stimuli, the participants

*This work is supported by the European ICT Programme Project FP7-224631 and by GRK 1589/1. This publication only reflects the authors' views. Funding agencies are not liable for any use that may be made of the information contained herein.

[1]J. Höhne is with the Machine Learning Department, Berlin Institute of Technology, Berlin, Germany j.hoehne@tu-berlin.de

[2]M. Tangermann is with the Machine Learning Department, Berlin Institute of Technology, Berlin, Germany michael.tangermann@tu-berlin.de

were asked to fixate a fixation cross and to not use any muscles. After the first block, the subjects were asked which stimulation speed they preferred.

B. EEG acquisition

EEG signals were recorded using a Fast'n Easy Cap (EasyCap GmbH) with 61 wet monopolar Ag/AgCl electrodes placed at symmetrical positions. Channels were referenced to the nose. Additionally, Electrooculogram (EOG) was acquired under the right eye. Signals were amplified using two 32-channel amplifiers (Brain Products), sampled at 1 kHz and band-pass filtered between 0.4 and 40 Hz. The data was epoched between -150 ms and 1000 ms relative to each stimulus onset.

C. Data analyses

All ERP analyses were performed in Matlab and the EEG data was downsampled to 200 Hz. In total, 216 target epochs and 1080 non-target epochs were obtained for each participant and each condition. To remove artifacts, epochs were excluded if their peak-to-peak voltage difference in any EEG or EOG channel exceeded 100μV. For classification, the mean potentials in 12 globally selected intervals at each channel were taken as features, leading to a 732-dimensional (12×61) feature vector for each epoch. The intervals were chosen between 100 ms and 700 ms after stimulus onset with shorter intervals for early responses. A binary classifier that separates between target and non-target epochs was trained for each participant and condition. To account for the the ill-posed ratio of number of data points vs. dimensionality of the feature vector, (linear) Fisher Discriminant Analysis (FDA) was applied with shrinkage regularization [12]. The classification accuracy was estimated by a cross validation with 5 folds and 5 shuffles. To account for the imbalance between non-targets and targets, the classwise balanced classification accuracy was calculated, which is the average decision accuracy across classes (target vs. non-target, chance level 50 %).

D. Simulating the ITR

Based on the empirically obtained binary classification accuracy for each SOA condition, the corresponding BCI performance (in bits/minute) was assessed by simulation. A BCI experiment with a 6-class ERP paradigm was simulated for each subject and SOA condition. Therefore, classifier outputs for target and non-target events were generated according to the binary accuracy, which was determined for the two-class oddball data. Thus, it is assumed that the binary classification accuracy (targets vs. non-targets) of the 6-class paradigm corresponds to the classification accuracy of the 2-class paradigm with equal stimulation speed. Based on the generated classifier outputs, trials were simulated and a multiclass decision was made as soon as an early-stopping criterion was fulfilled, at the latest after 15 presentations of each stimulus [13]. The duration of a trial and the selection accuracy of the corresponding one-out-of-six decision thus

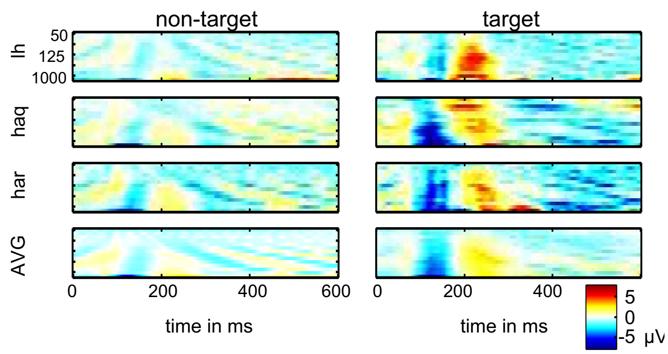

Fig. 1. Target and non-target ERPs maps for three subjects and the grand average over all subjects at electrode *Fz*. Each image depicts the course of an ERP over time and each row corresponds to one SOA condition. All color legends are equal, with red colors coding for positive amplitudes and blue colors coding for negative amplitudes.

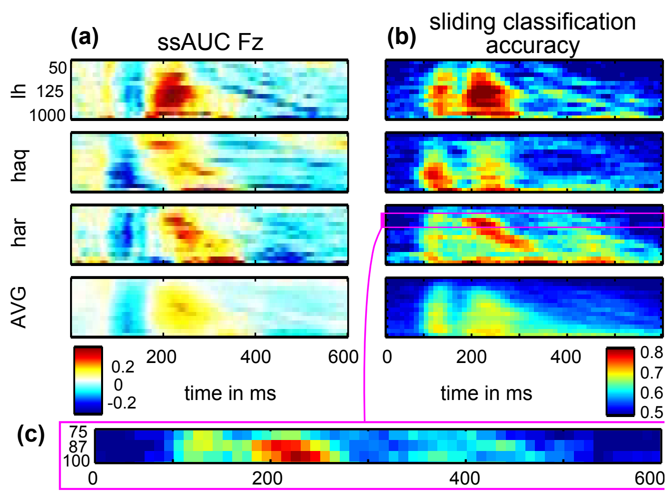

Fig. 2. Class discrimination maps over time for each SOA condition: ssAUC values at electrode *Fz* over time (a) and binary classification accuracy based on the mean amplitude of a sliding 50 ms EEG epoch with all electrodes (b). A close-up of the binary classification accuracy *har* for the SOA conditions 75, 87, 100 is shown in (c).

depended on the SOA and the binary classification accuracy. To account for pauses in between trials, a fixed time of 7 seconds was added after each selection. The ITR (as defined in [14]) was then computed based on the number of correct and incorrect decisions after the simulated BCI session, which lasted 60 minutes.

III. RESULTS

An analysis of the EEG data revealed that the stimulation time strongly impacts the shape of ERP components for non-target and target epochs. Fig. 1 depicts ERP responses to target and non-target stimuli for three subjects and the grand average. The ERP response is color-coded with blue (red) colors coding negative (positive) amplitudes. Each of the 14 rows in the image corresponds to one SOA condition where the top row shows the fastest stimulation (SOA = 50 ms) and the bottom row reflects the slowest stimulation (SOA = 1000 ms). As a general trend, the amplitudes of the ERPs increase with slower stimulation speed, which is in line with

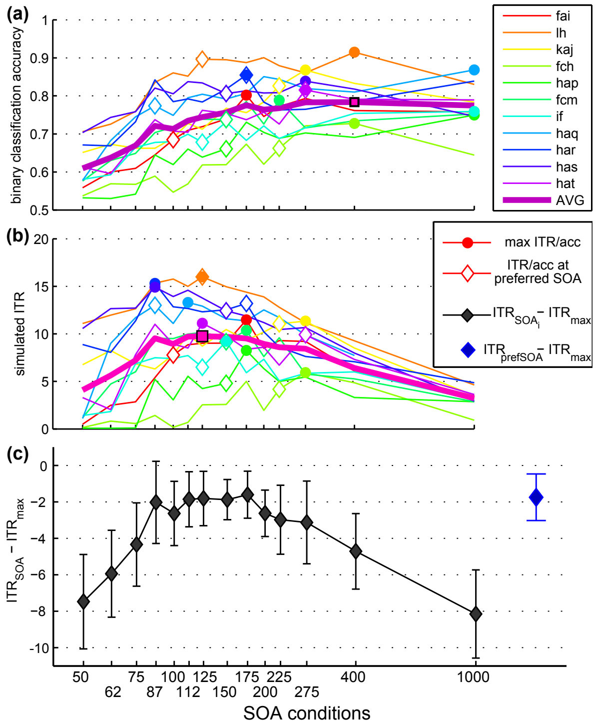

Fig. 3. Classwise balanced binary classification accuracy (a) for each subject and SOA condition. Simulated ITR (b) for each subject and SOA condition. Individual maximum values are marked with colored circle, individually preferred conditions are marked with a diamond. Average absolute difference (c) between the ITR in individually preferred SOA and individual optimal SOA. The whiskers show the standard deviation across subjects.

classical ERP literature [11]. This holds particularly for non-target ERPs.

For target and non-target responses, one can observe a negative deflection 150 ms after stimulus onset. This leads to a vertical blue pattern in the images. For the target events, this N150 component is considerably stronger which is often referred to as Mismatch Negativity (MMN) in neurophysiology literature [15]. Target responses show a positive deflection that starts 200 ms after stimulus onset. Amplitude and duration of this P200 component increases with increasing SOA (and decreasing stimulus speed, respectively).

For non-targets, one can additionally find a diagonal pattern between 200 ms and 400 ms after stimulus onset. This pattern reflects the shift in the steady state response, caused by consecutive stimuli. Thus, those responses are directly affected by stimulation speed.

Fig. 2 depicts the class discrimination between targets and non-targets over time. Fig. 2a shows the course of class discrimination for electrode *Fz*, while Fig. 2b displays a measure of class discrimination that incorporates all 61 EEG channels. To quantify class discrimination for one channel over time, the area under the ROC-curve (AUC) was computed and slightly modified (signed and linearly scaled to the range range of [0, 1]). The resulting measure (called ssAUC, see also [3]) provides information about the strengths and the direction of an effect. In Fig. 2a, an early negative class-discriminative component (MMN) and a later positive discriminative component (P2) can be observed at *Fz*.

To obtain a measure for class discrimination that considers all 61 EEG channels, classification accuracy was estimated with a sliding window as features: mean amplitudes of a 50 ms interval were computed for all electrodes, resulting in a 61-dimensional feature vector for each stimulus. Based on those features, the classification accuracy (targets vs. non-targets) was computed for the given interval. The averaging interval was sliding between 0 ms and 600 ms after stimulus onset. Fig. 2b depicts the obtained classification accuracy, with red (blue) coding for high (low) classification accuracy. Fig. 2a-b reveal, that the latency of the class discriminative N150 component is the same for all conditions. Thus, stimulation speed does not affect the latency of the N150. In contrast, the latency of the class discriminative P200 component is affected by the stimulation speed, in particular for subject *har* and *haq*. Moreover, one can observe the general trend of increasing amplitudes and class discrimination with increasing SOA for both the N150 and the P200 component, which is known from classic ERP literature [11].

This correlation of class discrimination and SOA is also reflected in Fig. 3a, where classification accuracy is plotted for each subject and each condition. On average, the binary classification accuracy is highest for a SOA of 1000 ms (SOA_{1000}). Although this observation is in line with classic ERP literature, classification accuracy is not decreasing monotonously with faster stimulation. For example, Fig. 3a shows clear peaks for subject *har* at SOA_{87} and SOA_{175}, which means that those stimulation conditions induce evoked potentials that can be classified more accurately than other (even slower) stimulation speeds. For *har*, the classification accuracy at SOA_{87} (0.84) is considerably higher than the accuracy for SOA_{75} (0.73) and also higher than SOA_{100} (0.78). The reason for that increase is explained in Fig. 3c, showing that for SOA_{75}, there is only early discriminative information centered at 120 ms after stimulus onset. For SOA_{87}, a strong P200 component is observed additionally, which explains the increase in classification accuracy from 73 % to 84 %. Reducing the stimulation speed from 87 ms to 100 ms (SOA_{100}), the P200 latency increases, but more importantly, the early component at 120 ms diminishes, which results in a reduction of overall class discrimination and classification accuracy (84 % to 78 %). This is only one example for individual variability in ERP components and classification accuracies for slightly different stimulation speeds.

Fig. 3b shows the ITR that was simulated for each subject and condition as described above. One can observe that the optimal stimulation speed (with respect to ITR) is between 87 ms and 200 ms for most subjects. The maximum ITR value for each subject is marked with a circle. Due to considerable variability in the binary classification accuracy, the ITR is also varying for single subjects, leading to peaks in the curve, such as SOA_{87} for *har*. Fig. 3c quantifies how much BCI performance is lost by a globally defined stimulation speed that is used for all subjects: the individual maximum ITR (ITR_{max}) is subtracted from the individual ITR (ITR_{SOA_i}) for each SOA condition *i*. Thus, the curve in Fig. 3c can only reach the value 0 if all subjects have their maximum ITR at the same stimulation speed. The graph shows that if the stimulation speed is globally chosen between 87 ms and 200 ms, the average BCI performance is $\sim$2 bits/min lower than the individually optimized ITR. Across all conditions, SOA_{175} performs best with with a loss of 1.6 bits/min. Thus, if the individually optimal SOA was used as stimulation speed, the average increase in ITR would be at least 1.6 bits/min, even if the globally optimum was known.

Moreover, it was found that using the individually preferred stimulation speed leads to a very good performance as well (loss of $SOA_{prefSOA}$ = 1.74).

IV. DISCUSSION

In typical BCI paradigms based on ERPs, such as [1], [6], [3], the stimulation speed (here SOA) is pre-defined and thus equal for each subject. Changing the stimulation speed, one observes varying ERPs as shown in Fig. 1. In the study presented here, it is demonstrated, that even in one of the simplest types of ERP paradigms (2-class auditory oddball), a slight change in stimulation speeds may result in non-linear variations of class-discriminative ERP components and the resulting classification accuracy. Discriminative ERP component are suppressed or enhanced for specific stimulation speeds, as it is shown for one subject in Fig. 2.

As a result, this study points out that an individual choice of the stimulus onset asynchrony is highly beneficial with respect to BCI performance. The analyses of a simulated online BCI experiment with 14 SOA conditions reveal that BCI performance (assessed by ITR) is increased by $\sim$2 bits/min, if the SOA is defined for each subject individually. The work by Sellers [10] already showed that the choice of SOA highly impacts the BCI performance. The presented study underlines these findings and quantifies the systematic error which is made due to the global selection of the SOA. Moreover, we show that the personally preferred stimulation speed also leads to a very good BCI performance, being almost as good as the (mostly unknown) global optimum.

As the next steps, multiclass ERP experiments should be performed with varying stimulation speed to validate the presented findings. Moreover, machine learning methods will be elaborated to find the individually optimal -or suboptimal- stimulation speed within a short time.

REFERENCES

[1] L. Acqualagna, M. S. Treder, M. Schreuder, and B. Blankertz, A novel brain-computer interface based on the rapid serial visual presentation paradigm, in Conf Proc IEEE Eng Med Biol Soc, vol. 2010, 2010, pp. 2686–2689.

[2] J. Höhne, M. Schreuder, B. Blankertz, and M. Tangermann, Two-dimensional auditory P300 speller with predictive text system, In Conf Proc IEEE Eng Med Biol Soc, Volume 1, p. 4185–4188, 2010.

[3] J. Höhne, M. Schreuder, B. Blankertz, and M. Tangermann, A novel 9-class auditory ERP paradigm driving a predictive text entry system, Front Neuroscience, vol. 5, p. 99, 2011.

[4] M. Schreuder, T. Rost, and M. Tangermann, Listen, you are writing! speeding up online spelling with a dynamic auditory BCI, Front Neuroscience, vol. 5, p. 112, 2011.

[5] M. Tangermann, M. Schreuder, S. Dähne, J. Höhne, S. Regler, A. Ramsay, M. Quek, J. Williamson, and R. Murray-Smith, Optimized stimulation events for a visual ERP BCI, Int J Bioelectromagnetism, vol. 13, no. 3, pp. 119–120, 2011.

[6] T. Kaufmann, S. M. Schulz, C. Grünzinger, and A. Kübler, "Flashing characters with famous faces improves ERP-based braincomputer interface performance, Journal of Neural Engineering, vol. 8, no. 5, p. 056016, 2011.

[7] J. Hill, J. Farquhar, S. Martens, F. Bießmann, and B. Schölkopf, Effects of Stimulus Type and of Error-Correcting Code Design on BCI Speller Performance, Advances in Neural Information Processing Systems, vol. 21, pp. 665–672, 2009.

[8] J. Höhne, K. Krenzlin, S. Dähne, and M. Tangermann, Natural stimuli improve auditory BCIs with respect to ergonomics and performance, Journal of Neural Engineering, 2012, in press.

[9] L. Farwell and E. Donchin, Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials, Electroencephalogr Clin Neurophysiol, vol. 70, pp. 510–523, 1988.

[10] E. Sellers, D. Krusienski, D. McFarland, T. Vaughan, and J. Wolpaw, A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance, Biol Psychol, vol. 73, pp. 242–252, 2006.

[11] C. Gonsalvez and J. Polich, P300 amplitude is determined by target-to-target interval, Psychophysiology, vol. 39, no. 03, pp. 388–396, 2002.

[12] B. Blankertz, S. Lemm, M. S. Treder, S. Haufe, and K.-R. Müller, Single-trial analysis and classification of ERP components – a tutorial, Neuroimage, vol. 56, pp. 814–825, 2011.

[13] M. Schreuder, J. Höhne, M. S. Treder, B. Blankertz, and M. Tangermann, Performance optimization of ERP-based BCIs using dynamic stopping, in Conf Proc IEEE Eng Med Biol Soc, 2011, pp. 4580–4583.

[14] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, Brain-computer interfaces for communication and control, Clin Neurophysiol, vol. 113, pp. 767–791, 2002.

[15] R. Näätänen, P. Paavilainen, T. Rinne, and K. Alho, The mismatch negativity (MMN) in basic research of central auditory processing: A review, Clin Neurophysiol, vol. 118, no. 12, pp. 2544 – 2590, 2007.