

A novel brain-computer interface based on the rapid serial visual presentation paradigm

Laura Acqualagna, Matthias Sebastian Treder, Martijn Schreuder, and Benjamin Blankertz

Abstract—Most present-day visual brain computer interfaces (BCIs) suffer from the fact that they rely on eye movements, are slow-paced, or feature a small vocabulary. As a potential remedy, we explored a novel BCI paradigm consisting of a central rapid serial visual presentation (RSVP) of the stimuli. It has a large vocabulary and realizes a BCI system based on covert non-spatial selective visual attention. In an offline study, eight participants were presented sequences of rapid bursts of symbols. Two different speeds and two different color conditions were investigated. Robust early visual and P300 components were elicited time-locked to the presentation of the target. Offline classification revealed a mean accuracy of up to 90% for selecting the correct symbol out of 30 possibilities. The results suggest that RSVP-BCI is a promising new paradigm, also for patients with oculomotor impairments.

I. INTRODUCTION

Research in brain-computer interfacing aims to establish a direct link between mind and machine. The main target audience of a brain-computer interface (BCI) is patients deprived of other means of communication, for instance patients suffering from amyotrophic lateral sclerosis. Since a BCI circumvents motor activity, it can serve as a surrogate information channel. One of the most widely used type of control signals is the event-related potential (ERP). An ERP is a phasic brain response time-locked to an external or internal event.

Much BCI research on selective attention and information processing focused on the P300 ERP component, a positive potential which can be observed at central and parietal electrode sites approximately 300-400 ms after stimulus onset. However, the timing of this component depends on task and stimulus parameters and may range from 250 ms to 900 ms [1]. It is most frequently elicited in an oddball paradigm, wherein rare target events are interspersed with frequent non-target events.

Farwell and Donchin [2] were the first to realize an ERP speller, that is, a communication device based on the detection of ERP components. Symbols were arranged in a 6x6 matrix, where the rows and the columns were intensified in a random order. The participant had to focus the attention on the target symbol, and when row and the column containing it were intensified, the P300 component was

typically enhanced. Recently, it was shown that the matrix spellers relies on eye movements [3]. For patients with unimpaired oculomotor control, eyetrackers are a faster and more accurate communication tool. For patients with impaired oculomotor control, it is crucial to develop paradigms that do not require eye movements.

One approach is to rely on other modalities such as audition [4] and touch [5]. Another approach is to develop visual BCIs that can be operated using covert attention. One promising approach is the ERP Hex-o-Spell, wherein items are arranged in six groups on the vertices of an invisible hexagon [3][6]. In the Hex-o-Spell, symbol selection is a two-stages process. First, the participant had to focus on the circle containing the group of letters wherein the target was; then, after selected the symbol group, the individual target symbol could be selected in the second level. The Hex-o-Spell outperformed the matrix speller, but performance was still not feasible.

In the present study, we introduce a novel visual speller that is fast-paced, has a large vocabulary (30 symbols), and does not depend on eye movements. The speller features rapid serial visual presentation (RSVP), that is, fast bursts of symbols presented successively at a single central location. Symbols are chosen by selectively attending to the desired symbol, which leads to an enhancement of ERP components elicited by the visual presentation.

In an offline study, we investigated four different conditions. Since stimulus presentation time was shown to be a crucial factor in previous studies [7], we analyzed how the performances and the accuracy changed increasing the speed of the RSVP. We used two stimulus onset asynchronies (SOAs), namely 83 ms and 133 ms.

Furthermore, in order to determine whether the color would help distinguish the target stimuli, two different color-modes were used: either all the letters black or divided into three colored groups: red,green,blue.

[□]This work was partly supported by the European Information and Communication Technologies (ICT) Programme Project FP7-224631 and 216886, by grants of the Deutsche Forschungsgemeinschaft (DFG) (MU 987/3-1) and Bundesministerium für Bildung und Forschung (BMBF) (FKZ 01IB001A, 01GQ0850) and by the FP7-ICT Programme of the European Community, under the PASCAL2 Network of Excellence, ICT-216886. This publication only reflects the authors' views. Funding agencies are not liable for any use that may be made of the information contained herein.

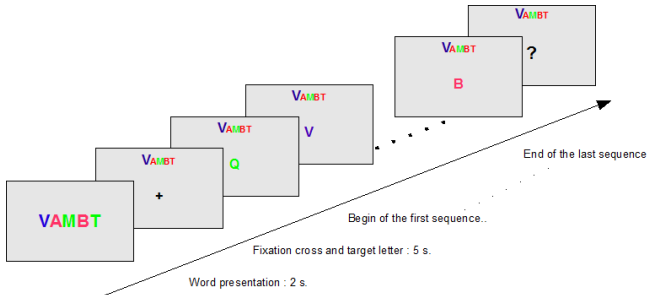


Fig. 1. RSVP paradigm (color mode). At the beginning the word is presented, then the target letter in highlighted and the fixation cross is displayed. The first sequence starts: all the sequences consist of three bursts of 10 items each with an inter-burst interval of 0.3 s. When all the sequences are presented, a question mark is displayed: the participant is asked to enter the number of target letters that he counted.

II. METHODS

A. Participants

Nine participants (8 males and 1 female, aged 24-31) participated in the experiment. All had normal or corrected-to-normal visual acuity. Normal color vision was confirmed by a Nishihara color blindness test. None of them had a history of a neurological disease or injury. Written informed consent was obtained from each participant prior to the beginning of the experiment.

B. Apparatus

EEG was recorded using a 64-channels actiCAP system (Brain Products, Munich, Germany). The electrodes Fp1, AF7, FT7, AF8 and FT8 were replaced with the additional occipital electrode PO9, I1, I2, PO10. All electrodes were placed according to the international 10-20 system and referenced against a nose reference. Eye movements were monitored using an EOG electrode placed below the right eye. All impedances were kept below 15 k. EEG data were sampled at a rate of 1000 Hz and subjected to off-line analysis. The bandpass of the hardware filter was 0.016-250 Hz. Stimuli were presented on a 24" TFT screen with a refresh rate of 60 Hz and a resolution of 1920 x 1200 px². Participants entered their response via a standard computer keyboard.

C. Design and procedure

We used a 2 x 2 design, so that we had 4 conditions : 1)no-color-83ms; 2)no-color-133ms; 3)color-83ms; 4)color-133ms.

The modes about the color of the items were: no-color : all the letters were black; color : the 30 items were divided into 3 groups : red (ABCDEFGHIJ-), green (KMNWEQRST+), blue (UVOXYZLP!).

The different conditions about the time of presentation of the letters were : 83ms and 133 ms.

Participants sat at a distance of approximately 80 cm from a computer screen. They engaged in a copy-spelling task whereby they had to copy 5-6 letters words. There was a set of six words, chosen such that each letter in the English

alphabet was covered at least 1 time. Each participant was tested in eight blocks (two per experimental condition) with a total duration of approximately 3 hours. Each block consisted of 3 words with 5-6 letters per word.

When a new word was introduced, it was shown on the screen prior to the start of the trial. Then, the current word was shown on the top of the display with the current letter being highlighted. The participants had 3 seconds to identify the current target letter. Subsequently, a fixation cross appeared for 2 s and the RSVP started.

Stimuli were presented in the center of the screen on a gray background. A total of 30 symbols was used, the 26 letters of the English alphabet and 4 punctuation marks. They had an height of 3.5 cm (1° visual angle).

Symbols were randomly shuffled and divided into 3 bursts of 10 items each. These three bursts formed a sequence. Each symbol was repeated ten times, there was a total of 10 x 3 = 30 bursts, with an interval of 0.3 s between successive bursts. Figure 1 depicts the course of one trial in the color mode.

In each sequence, the order of the presentation of the items was pseudo-randomized in order to avoid that the same symbols appeared together close in time. Furthermore, to obtain meaningful behavioral data, we varied the number of occurrences of the target symbol. To this end, each trial featured an additional prequel and a sequel (both not used in the analysis), adding up to a total of 12 sequences per trial. In both the prequel and the sequel, the target was randomly presented 0-2 times.

In the color mode, the order of the presentation of the colors was fixed to red, green, blue.

Participants were instructed to silently count the number of the intensifications of the target symbol and, at the end of each trial, to enter their count via the computer keyboard.

The RSVP paradigm was implemented in the open-source BCI framework Pyff [8] with Vision Egg [9], and remote-controlled via Matlab.

D. Data analysis

For ERP analysis, EEG data were down-sampled to 200 Hz and divided into epochs ranging from -200ms to 1200ms relative to the onset of each symbol. Baseline correction was performed on the pre-stimulus period of either 133 ms (for the 133ms mode) or 166 ms (for the 83 ms mode). Epochs containing artifacts were rejected based on a variance criterion. Epochs containing eye movements were detected and rejected using a min-max criterion (75μV) on the EOG channels.

Due to the fast pace of visual stimulation, ERP components of successive epochs were overlapping. For ERP analysis, contamination of the non-target epochs by target presentations was reduced by considering only those target epochs wherein the three preceding and the three following symbols were also non-targets. For classification, all epochs were used. For the grand average, the ERP curves were averaged across all trials and participants. To compare the ERP curves of two conditions or classes, $\text{sgn } r^2$ -values based on the point-biserial correlation coefficient were

calculated as a measure of how much information one feature x carries about the class labels y . The r^2 -value was then multiplied by its sign. Sgn r^2 -values were also averaged across all trials and participants.

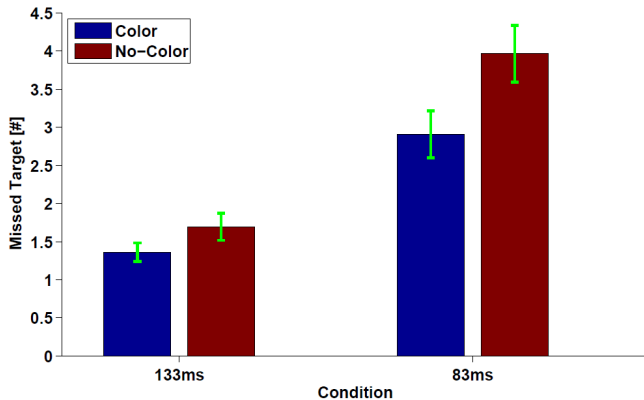


Fig. 2. Behavioral analysis. Subjects matched a greater number of targets in the 133 ms condition than in the 83 ms. They also performed better in the color mode than in the no-color.

III. RESULTS

A. Behavioral data

Figure 2 shows the mean counting error and the standard error of the mean. Performance is significantly better in the 133 ms condition than in the 83ms condition, which is consistent with the classification results shown below. However, participants reached a higher counting performance with colored letters when compared to the no-color mode, which contradicts the poorer classification score for both colored modes.

B. Event-related potentials (ERPs)

Figure 3 shows the time course of the ERPs at channels CPz and PO7 for all four sub-conditions and the scalp topographies for the time intervals that are shaded in the ERP plot for the corresponding condition.

Two main components can be distinguished. First, a N2 component, a pronounced negative visual component in the range of 280-360 ms with a parieto-occipital focus. Second, a positive P300 component with central distribution in the range of 400-600 ms. It shows a delay in comparison to the typical P300 observed in an oddball paradigm.

As revealed by an analysis of variance (ANOVA), there is a significant modulation of these components in all four conditions, with significantly larger amplitudes for targets than for non targets for both the N2 ($F = 37.71, p < .001$) and the P300 ($F = 25.9, p < .001$) components. The N2 component has a larger amplitude in the 133 ms mode than in 83 ms mode for both color modes ($F = 4.28, p < .05$), whereas no significant differences in amplitude are found between the no-color and the color modes for both speeds ($p = .94$). For the P300 component, we did not find significant effects of color ($p = .9$) or speed ($p = .23$). Interactions between the factors were not significant.

C. Classification

The ERP analysis in the previous section showed that the N2 and the P300 components are strongly modulated by attention. The task of a BCI is to detect these modulations in order to deduce whether a target or a non-target was presented.

For off-line classification, we used linear discriminant analysis (LDA) with shrinkage of the covariance matrix. The first block of each condition was taken as training set, the second one as test set. The feature vector consisted of 55 spatial and 7 temporal features.

Figure 4 depicts the results of the classification as a function of the number of sequences. The different colored lines represent the mean of all participants for each condition. Mean accuracy increases significantly with the number of sequences ($F = 13.23; p < .001$) and approaches high levels for the 133 ms conditions: after ten sequences, about 90% (6 participants 100%) in the no-color mode and about 85% (3 participants 100%) in the color mode.

Higher accuracy is obtained in the 133 ms condition than in the 83 ms condition ($F = 51.29; p < .001$), but there is no significant effect of color ($p = .28$).

IV. DISCUSSION

A mean classification performance of up to 90% was found, which is substantial given a chance level of 3.33%. The highest accuracy was obtained for the no-color 133 ms condition, but in contrast to speed, color did not have a significant effect on the results.

We also found a discrepancy between the behavioral data and classification performance. Counting results were better in the color mode than in the no-color mode, but this difference was not found in the classification data, where the color condition even tended to be worse. A possible explanation could be that since a subset of non-target letters has the same color as the target in the color condition, they can act as distractors, influencing the top-down attentional capture and affecting the classification performance. The effects of colored distractors on ERPs and their psychological consequences on target selection in a RSVP sequence has been investigated in [10].

Our results show that the RSVP-speller has a considerably higher classification performance than both the Matrix and the Hex-o-Spell in the covert condition. A higher accuracy can also be observed with respect to the classification accuracy of another BCI system in which covert attention is used to modulate a steady-state visual evoked potential (SSVEP) [11]. Furthermore, in contrast to most other BCIs based on SSVEP, the accuracy of our setup refers to a one-in-thirty multi-class decision and not to a binary one. So far, our paradigm has only been tested offline, while Zhang et al. [11] reported an averaged classification online accuracy, achieved during the last 3 days of a training program.

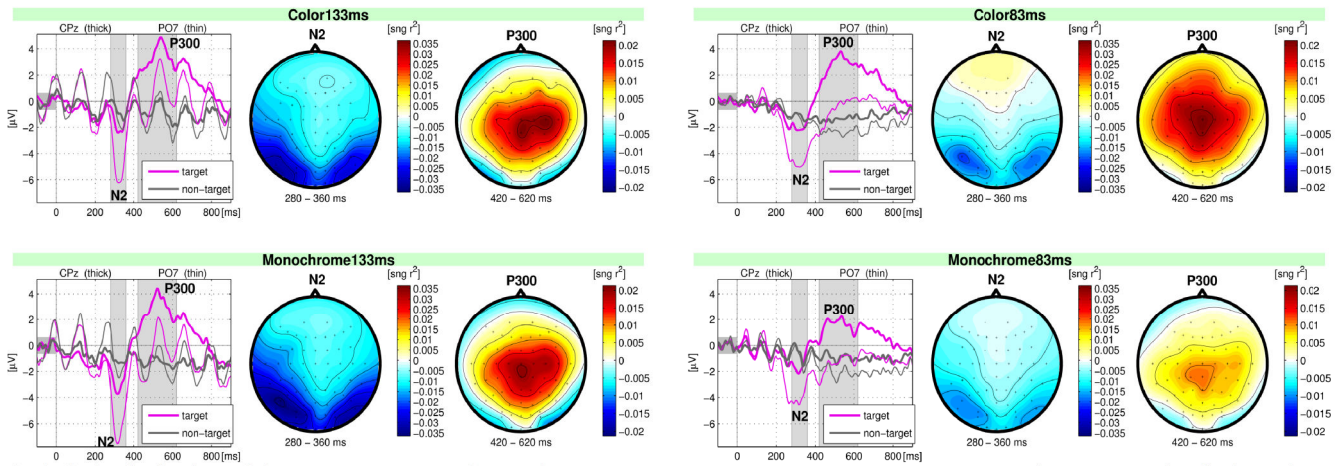


Fig. 3. Scalp distributions of the ERP components and ERP plots target vs. non-target. Dark shaded areas in the ERP plots correspond to the intervals for which topographies are shown adjacent. N2 and P300 components have high amplitudes for target intensifications especially for the 133 ms conditions.

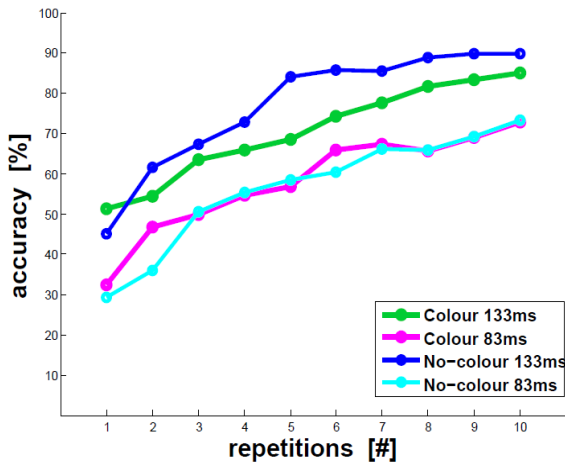


Fig. 4. Results of the classification as a function of number of sequences. The colored lines represent the mean accuracy over all the participants in the different conditions. After 10 sequences, accuracy approaches 85-90% in the 133 ms condition, while in the 83 ms is around 70%.

Currently, the bottleneck of the RSVP-speller seems to be the relatively high number of sequences required for reliable performance.

Both classification and stimulus presentation have to be optimized to reduce this number and thereby increase information throughput.

To conclude, the present study demonstrated that a visual speller based on RSVP constitutes a promising new paradigm for EEG-based BCIs. In comparison to other approaches that exploit covert visual attention, classification performance is competitive and the vocabulary is large. The RSVP-speller requires no eye movements and is therefore applicable to patients with impaired oculomotor control.

ACKNOWLEDGMENT

We are grateful to Torsten Schmits for programming the stimulus presentation.

REFERENCES

- [1] Salil H. Patel and Pierre N. Azzam, "Characterization of N200 and P300: Selected Studies of the Event-Related Potential", *Int J Med Sci.* 2005; 2(4): 147-154.
- [2] 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.
- [3] M. Tredor and B. Blankertz, "(c)over attention and speller design in visual attention based brain-computer interfaces," *Behavioral & Brain Functions*, 6:28, May 2010.
- [4] 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, 2010.
- [5] A.-M. Brouwer and J. van Erp, "A tactile P300 BCI and the optimal number of factors: effects of target probability and discriminability," in *Proceedings of the 4th International BCI Workshop and Training Course*, 2008, pp. 280-285.
- [6] J. Williamson, R. Murray-Smith, B. Blankertz, M. Krauledat, and K.-R. Müller, "Designing for uncertain, asymmetric control: Interaction design for brain-computer interfaces," *International Journal of Human-computer Studies*, Volume 67, Issue 10, October 2009, Pages 827-841.
- [7] Brendan Z. Allison, Jaime A. Pineda, "Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: Implications for a BCI system", *International Journal of Psychophysiology* 59 (2006) 127 - 140.
- [8] B. Venthur and B. Blankertz, "A platform-independent open-source feedback framework for BCI systems," in *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course 2008*. Verlag der Technischen Universitaet Graz, 2008, pp. 385-389.
- [9] Straw, Andrew D. (2008) Vision Egg: An Open-Source Library for Realtime Visual Stimulus Generation. *Frontiers in Neuroinformatics*. doi: 10.3389/neuro.11.004.2008.
- [10] Esther Vierck, Jeff Miller, "Distraction by color and its electrophysiological correlates", *Psychophysiology*, Volume 46, Issue 3, Pages 593-606, Mar 2009.
- [11] Dan Zhang, Alexander Maye, Xiaorong, Gao, Bo Hong, Andreas K Enge and Shangkai Gao, "An independent brain-computer interface using covert non-spatial visual selective attention", *J. Neural Eng.* 7 (2010) 016010 (11pp)