ERP and Pupillometry Synchronization Analysis on Rapid Serial Visual Presentation of Words, Numbers, Pictures

Xi Luo, Yanfei Lin, *Member, IEEE*, Rongxiao Guo, Xiaorong Gao, *Member, IEEE*, and Shangen Zhang

***Abstract*— Hybrid brain-computer interfaces (HBCI) combining eye-tracker has attracted the attentions of researchers in target recognition. However, there are still many issues to be addressed in rapid sequence visual presentation (RSVP) tasks, such as the effect of presentation rates and target types on event-related potentials (ERP) and pupillometry, synchronization analysis of electroencephalography (EEG) and eye-tracking, and so on. In this study, the RSVP experiments with three different target types of pictures, words and numbers at the presentation rates of 100 and 200 ms were conducted. EEG data and pupillometry data were synchronously collected from 20 university students. The results of ERP analysis showed that, among three different target types at the presentation rate of 100 ms, the picture P300 component had the largest amplitude and the longest latency. From the 100 ms presentation rates to 200 ms one for the three target types, the P300 amplitudes became smaller, and the P300 latencies became shorter. The results of pupillometry analysis showed that, at the presentation rates of 100 and 200 ms, the pupil dilation of pictures had the smallest amplitude and the shortest latency. At the two presentation rates, no significant differences of pupil size and latency were found for the three target types. For the early pupil dilation within 1000 ms, the picture pupil size was significantly smaller than the other ones, and the picture pupil acceleration had the largest average amplitude and the shortest latency. These pupillometry features within 1000 ms combining with the P300 features could be taken as the effective ones for target classification. Through synchronization analysis of the EEG data and pupillometry data, the effects of target type and presentation rate on ERP and pupil dilation were different. These results could contribute to developing the fusion methods between EEG and eye-tracking, and provide valuable references for the multi-target recognition of hybrid BCI based on eye-tracking.**

***Index Terms*— Rapid serial visual presentation (RSVP), hybrid brain-computer interface (HBCI), electroencephalography (EEG), P300, pupillometry.**

Manuscript received 2 November 2022; revised 20 March 2023; accepted 26 March 2023. Date of publication 31 March 2023; date of current version 5 April 2023. This work was supported in part by the Key Research and Development Program of Guangdong Province under Grant 2018B030339001; in part by the Beijing Science and Technology Program under Grant Z201100004420015; in part by the National Natural Science Foundation of China under Grant 61601028, Grant 61431007, and Grant U2241208; in part by the National Key Research and Development Program of China under Grant YFB1002505; and in part by the Fundamental Research Funds for the Central Universities of China under Grant FRF-TP-20-017A1. *(Corresponding author: Yanfei Lin.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institute Review Board of Tsinghua University under Application No. 20190021.

Xi Luo, Yanfei Lin, and Rongxiao Guo are with the School of Integrated Circuits and Electronics, Beijing Institute of Technology, Beijing 100084, China (e-mail: linyf@bit.edu.cn).

Xiaorong Gao is with the School of Medicine, Tsinghua University, Beijing 100081, China (e-mail: gxr-dea@tsinghua.edu.cn).

Shangen Zhang is with the School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: zhangsfphd@163.com).

Digital Object Identifier 10.1109/TNSRE.2023.3263502

I. Introduction

RAPID sequence visual presentation (RSVP) is to continuously present the multiple pictures to subjects in the same spatial location at high presentation rates [1]. The ERP components commonly analyzed in RSVP tasks are P300 and late potentials. P300 is a positive component that appears around 200-600 ms after the onset of target stimulus, and is mainly concentrated in the central and centro-parietal regions [2]. RSVP based brain-computer interfaces (BCI) is used to target detection with a RSVP paradigm by analyzing subjects' EEG data. RSVP-BCIs have received extensive researchers' attentions and have broad application prospects, such as Satellite image retrieval [3], [4], screening of targets by supervisors [5], RSVP speller [6], [7], and so on.

Due to the low signal-to-noise ratio of single-trial EEG data, the target recognition accuracy of single BCI systems is low, which will affect the practical applications of BCI systems. Recently, the concept of hybrid brain-computer interface (HBCI) was proposed, such as hybrid BCI based on eye-tracking. Previous studies demonstrated that pupil dilation was related to target recognition. When subjects focused on targets, the pupils dilated [8], [9], [10]. Pupil dilation is induced by the inhibition of the parasympathetic nervous system and Edinger-Westphal nucleus, and the locus coeruleus-norepinephrine system (LC-NE) controls these inhibitory processes, which dominates the regulation of attention [8], [10], [11]. Previous studies demonstrated that pupil dilation was related to cognitive processing of visual information, such as memory, attention and cognitive load [12], [13], [14]. Therefore, the fusion of EEG and eye-tracking could improve the performance of target recognition in RSVP tasks.

Recently, researchers have done many EEG studies on related parameters that could affect the performance of RSVP-BCIs, such as presentation rates [15], button press [16], target probability [17], image complexity [18], target types [19], etc. For presentation rates, Sajda et al. [15] conducted RSVP experiments at three presentation rates of 200, 100 and 50 ms. The results showed that the classification performance of the 50 ms presentation rate was significantly lower than those of the others, and there was no difference in classification performance between 100 and 200 ms presentation rate. In addition, Lees et al. [19] analyzed the effect of presentation rates on classification performance. The results indicated that the AUC of 300-400 ms presentation rate was significantly higher than that of 100-200 ms presentation rate, but not higher than that of 200-300 ms presentation rate. For button press, Gerson et al. [16] analyzed the effects of button press and motor response on the classification performance in the RSVP tasks. The results showed that there was no significant difference of classification performance between button press and no button press, and no significant difference between motor and nonmotor. For target probability, Cecotti et al. [17] studied the effect of target probability on the classification performance in the RSVP experiment. The results indicated that the percentage of target stimuli should be below 10% in RSVP paradigms, which could maximize classification accuracy. For image complexity, Lin et al. [18] analyzed the difference of P300 components under the conditions of high, middle, low image complexities. It was found that the higher the image complexity, the longer the P300 latency. For target types, Lees et al. [19] detected the classification accuracy for three different types of target image including pictures, words and numbers. The results showed that the classification performance of numbers was significantly lower than that of words and pictures, and no significant difference was found between words and pictures.

For pupillometry analysis of RSVP tasks, researchers have analyzed the effects of button press, target probability and information complexity on pupil dilation. For button press and target probability, Claudio et al. [8] found that pupil size was significantly larger with button press than without button press in RSVP paradigm. Moreover, they also found that the lower the target probability, the larger the pupil size. For information complexity, Ivory et al. [10] used pupil size to detect concealed identity information in the RSVP tasks. The results demonstrated that the pupil size of concealed identity information was significantly higher than that of control information.

In summary, previous studies showed that hybrid BCI based on eye-tracking was an effective way to improve the classification performance of RSVP tasks [20], [21]. Although there have been many ERP studies on classification performance of RSVP-BCI [19], [22], [23], pupillometry analysis about the effects of some parameters was still not explored, such as target types, presentation rates and so on. Meanwhile, it was necessary to synchronously analyze the characteristics of EEG and pupillometry to investigate the fusion methods between EEG and eye-tracking, and obtain the joint features of EEG and eye-tracking for single-trial and multi-target classification in RSVP tasks. In this study, the following issues are addressed:

(i) How are the effects of target types and presentation rates on pupil dilation in the RSVP paradigm?

(ii) What change patterns do the features of ERP and pupillometry have with different target types at different presentation rates, through synchronous analysis of EEG and pupillometry?

(iii) Which effective features of ERP and pupillometry can be combined for the target classification in RSVP tasks?

To solve these problems, EEG and pupillometry data were synchronously collected from 20 university students in the RSVP tasks with different target types at different presentation rates, and then ERP and pupillometry analysis were conducted. In the organization of this study, Section II introduces experimental procedure and data processing. Section III illustrates the results of ERP and pupillometry synchronous analysis. Section IV presents the discussion of the results and future work.

II. Methods

A. Participants

Twenty right-handed university students (10 males and 10 females, 19-24 years old) participated in the experiment. All subjects had no neurological history of disease, or mental disorders, and had normal or corrected vision. The experiment was carried out according to the Declaration of Helsinki, and obtained approval from the local ethics committee. All subjects signed the Consent Form before the experiment, and received monetary compensation for the participation.

B. Experimental

1) Materials: There were three different target types of pictures, numbers and words. All image stimuli were made into videos using Adobe Premiere software, and scaled to 560 × 360 pixels (width × height). The words and numbers stimuli had the size of 32 pixels. Pictures were from the 'morgueFile' database [24] (see Fig.1(a)). The target pictures were Dalmatians and sunflowers. Numbers were in the range of 100-999, whose images had written numbers with white font Times New Roman centered in black background (see Fig.1 (b)). The target numbers were '100' and '200'. Words consisted of common three letters, whose images had written words with white font Times New Roman centered in black background (see Fig.1 (c)). The target words were 'him' and 'her'.

For the 100 ms presentation rate, each video had 100 images with 3 target images, and each image was presented for 100 ms. The targets were placed after the 10th image and before the 80th image, and the interval between targets was at least longer than 2 s. For the 200 ms presentation rate, each video had 50 images with 1 or 2 target images, and each

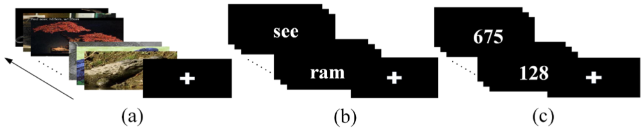

Fig. 1. Image series of RSVP. (a) Pictures. (b) Words. (c) Numbers.

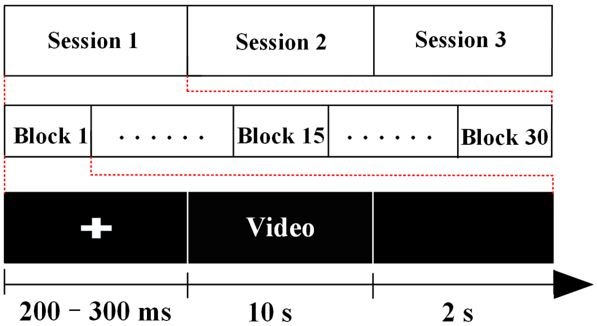

Fig. 2. Experiment procedure.

image was presented for 200 ms. The targets were placed after the 5th image and before the 40th image, and the interval between targets was at least longer than 2 s. Picture, number and word stimuli contained 15 videos for each presentation rate, respectively. The target probability was 3% for each target type in RSVP tasks. For each video, each image stimulus was strictly presented for 100 ms or 200 ms, respectively. And each image stimulus was validated by photocell in each video.

2) Experimental Procedure: Subjects sat on a comfortable chair in a soundproof room. For each block, a target prompt was first presented in the center of Virtual Reality (VR) for 2 s, and then a white cross '+' was shown for 200-300 ms. Next, a video stimulus was displayed for 10 s, and the video was played randomly. During this time, subjects were required to avoid blinking. If they detected target images, they were required to press the button as quickly as possible to maintain attention [16], [19]. After that, a black screen was kept for 2 s. And then next block entered. As shown in Fig.2, total 3 sessions were conducted in this experiment. Each session had 30 blocks. For all 90 blocks, all videos were not reused. For every 5 blocks, subjects had a rest for 3 min. For every session, subjects' rest lasted for 10 min.

C. Data Recording and Processing

1) EEG Recording and Pre-Processing: EEG data were recorded using Neuroscan Inc. Synamps2 system. All electrodes were distributed according to the international 10-20 system, and the reference electrode was located at the left mastoid. The sampling rate was 1000 Hz. During this experiment, electrode impedances were kept below 10 KΩ. EEG data were first filtered using a filter with bandwidth of 0.3-30 Hz, and a 50 Hz notch filter. And then EEG data were re-referenced offline to the average of the right mastoid. EEG epochs were extracted from 200 ms before the target onset to 1000 ms after the target onset, and the average amplitude from -200 to 0 ms was used to baseline correction. The data contaminated by ocular movements or other artifacts were eliminated manually.

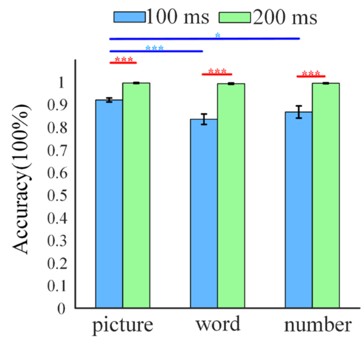

Fig. 3. Behavioral results. The error bars stand for the standard deviations.

2) Pupil Recording and Processing: Pupillometry data were recorded by an embedded infrared eye-tracking module: aGlass DKII [25]. The sampling rate was 120 Hz. According to the previous studies [8], [26], pupillometry epochs were extracted from 1000 ms before the target onset to 2500 ms after the target onset, with the time window of 1000 ms before the target onset as baseline. The time windows of pupil dilation, pupil velocity and pupil acceleration for analysis were determined by a certain period before and after their latencies. The right and left pupillometry data were averaged. Fewer than 2% of pupillometry data contaminated by head movements and excessive blinking were removed manually. The partial missing pupillometry data in a trial caused by blinks and other artifacts were filled with previous data [27], [28].

In this study, triggers were sent by stimulus program through the parallel port, and were synchronously recorded with EEG data using the amplifier. After stimulus program sent triggers, the indexes were added in the beginning of pupillometry data. Thus, the synchronous acquisition of EEG data and pupillometry data could be guaranteed.

Repeated-measures analysis of variance (ANOVA) was utilized to analyze the behavioral, ERP and pupillometry results. Pairwise comparison was tested using Least Significant Difference (LSD). The Greenhouse-Geisser correction was applied to the approximation of non-sphericity.

III. Results

A. Behavioral Results

Fig.3 shows that the recognition accuracy at two different presentation rates with three target types of pictures, numbers and words. Repeated measures ANOVA of presentation rates (2 levels) × target types (3 levels) results showed that there was a main effect on the recognition accuracy with three target types (F (2,72) = 8.593, $p = 0.000$). There was an interaction between target types and presentation rates (F (2,72) = 7.351, $p = 0.001$), and a significant difference of the recognition accuracy at presentation rates (F (1,36) = 50.26, $p = 0.000$). At the presentation rate of 100 ms, pairwise comparisons showed that the picture accuracy was significantly higher than the others ($ps < 0.05$), and no significant difference was found between words and numbers ($p > 0.05$). It was revealed that the cognitive difficulty of pictures was lower than that of words and numbers for subjects. At the 200 ms presentation rate,

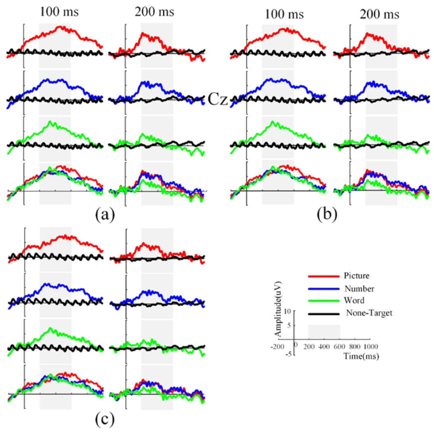

Fig. 4. Grand average ERPs of pictures, numbers and words at two different presentation rates. (a) FCz, (b) Cz, (c) CPz.

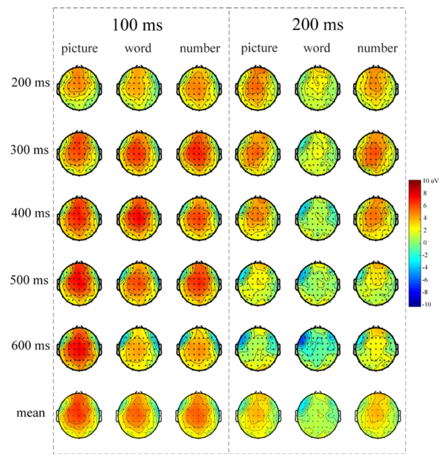

Fig. 5. Topographic maps of P300 in the time window of 200-600 ms with three target types at two presentation rates.

no significant differences were found among the three target types ($ps > 0.05$). It was indicated that there was no difference of the cognitive difficulty for three target types. In addition, the recognition accuracy at the 200 ms presentation rate was significantly higher than that at the 100 ms presentation rate for the three target types. It was indicated that it was easier to recognize targets for subjects at the lower presentation rate.

B. ERP Results

Fig.4 presents the grand average ERPs of FCz, Cz and CPz electrodes for all subjects at two presentation rates with three target types of pictures, numbers and words. It was showed that the picture P300 amplitude was higher than the other two at the 100 ms presentation rate, and the word P300 amplitude was lower than the other two at the 200 ms presentation rate. For the three target types, the P300 amplitude of the 100 ms presentation rate was higher than those of the 200 ms presentation rate, and the latencies of P300 at the 100 ms presentation rate were longer than those at the 200 ms presentation rate. Fig.5 shows topographic maps of ERPs plotted in the time window of 200-600 ms. It was illustrated that the P300 activities were mainly concentrated in the central and centro-parietal regions.

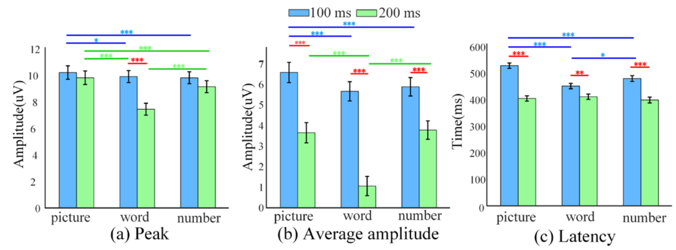

Fig. 6. The P300 peaks, average amplitudes and latencies in the time window of 200-600 ms. (a) Peak, (b) Average amplitude, (c) Latency. The error bars stand for the standard deviations.

The P300 components was analyzed in the central-parietal region (FC1, FCz, FC2, C1, Cz, C2, CP1, CPz, and CP2 electrodes) in the time window of 200-600 ms. The P300 peaks, average amplitudes and latencies with three target types at two presentation rates are shown in Fig.6. At the 100 ms presentation rate, the P300 latencies for pictures, words and numbers are 539 ms, 465 ms, 494 ms, respectively. At the 200 ms presentation rate, the P300 latencies for pictures, words and numbers are 406 ms, 375 ms, 409 ms, respectively. Repeated measures ANOVA of presentation rates (2 levels) × target types (3 levels) ×electrodes (9 levels) showed that there was a main effect on the P300 peak with three target types (F (2,648) = 40.119, $p = 0.000$). There was an interaction between target types and presentation rates (F (2,648) = 27.889, $p = 0.000$), and no significant difference of the P300 peak at presentation rates (F (1,324) = 50.26, $p = 0.068$). At the 100 ms presentation rate, pairwise comparisons revealed that the P300 peak evoked by pictures was significantly higher than that of the others ($ps < 0.05$), and no significant difference was found between numbers and words ($p > 0.05$). At the 200 ms presentation rate, the picture P300 peak was the highest, and the word P300 peak was the lowest ($ps < 0.001$).

Repeated measures ANOVA of presentation rates (2 levels) × target types (3 levels) × electrodes (9 levels) showed that there was a main effect on average amplitude with three target types (F (2,648) = 78.805, $p = 0.000$). There was an interaction between target types and presentation rates (F (2,648) = 36.071, $p = 0.000$), and a significant difference of average amplitude at presentation rates (F (1,324) = 25.075, $p = 0.000$). At the 100 ms presentation rate, pairwise comparisons revealed that the average amplitude of picture P300 was significantly higher than those of the others ($ps < 0.001$), and there was no significant difference between words and numbers ($p > 0.05$). At the 200 ms presentation rate, the average amplitude of word P300 was significantly lower than those of the others ($ps < 0.001$), and no significant difference was found between picture and number ($p > 0.05$). In addition, the average amplitudes at the 200 ms presentation rate were significantly smaller than those at the 100 ms presentation rate for the three target types ($ps < 0.001$).

Repeated measures ANOVA of presentation rates (2 levels) × target types (3 levels) × electrodes (9 levels) showed that there was a main effect on P300 latency with three

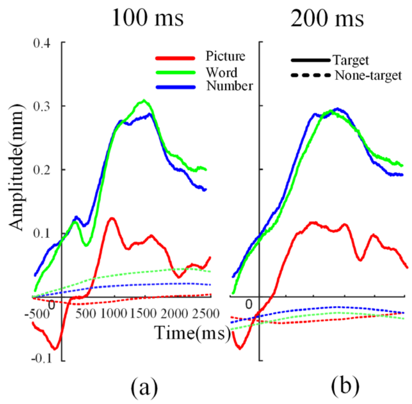

Fig. 7. Grand average pupil dilation with three target types. (a) 100 ms presentation rate, (b) 200 ms presentation rate.

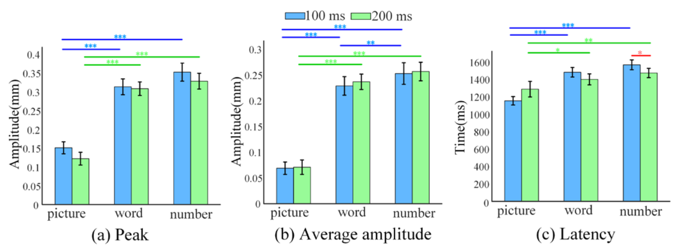

Fig. 8. The peaks, average amplitudes and latencies for pupil dilation in the time window of 500-2000 ms at two presentation rates. (a) Peak, (b) Average amplitude, (c) Latency. The error bars stand for the standard deviations.

target types (F (2,648) $= 15.96$, $p = 0.000$). An interaction was found between target types and presentation rates (F (2,648) $= 20.404$, $p = 0.000$), and there was a significant difference of latency at presentation rates (F (1,324) $=$ 42.239, $p = 0.000$). Subsequently pairwise comparisons showed that the picture latency was significantly longer than the other ones at the presentation rate of 100 ms ($ps <$ 0.001), and the word latency was significantly shorter than the other ones ($ps < 0.05$). At the 200 ms presentation rate, there were no significant differences of latency among the three target types ($ps > 0.05$). In addition, the latencies at the 200 ms presentation rate were significantly shorter than those at the 100 ms presentation rate for the three target types ($ps < 0.01$).

C. Pupillometry Results

Fig.7 shows pupil dilation for all subjects at two presentation rates with three target types of pictures, numbers and words. The pupil sizes of pupil dilation with three target types are significantly higher than those of non-target types. At two presentation rates, the amplitude of pupil dilation induced by pictures is smallest, and the picture latency of pupil dilation is shortest.

Fig.8 shows the peaks, average amplitudes and latencies for pupil dilation in the time window of 500-2000 ms with three target types at two presentation rates. At the 100 ms presentation rate, the latencies of pupil dilation for pictures, words and numbers are 1157 ms, 1485 ms, 1571 ms, respectively. At the 200 ms presentation rate, the latencies of pupil dilation for pictures, words and numbers are 1291 ms, 1403 ms, 1477 ms, respectively.

Repeated measures ANOVAs of presentation rates (2 levels) $\times$ target types (3 levels) for peak, average amplitude and latency were analyzed. The results showed that there were significant effects on the pupil dilation peak, average amplitude and latency with three target types (Peak: F (2,76) $= 79.882$, $p = 0.000$, Average amplitude: F (2,76) $= 29.311$, $p = 0.000$, Latency: F (2,76) $= 51.745$, $p = 0.000$).There were no interaction between target types and presentation rates (Peak: F (2,76) $= 0.326$, $p = 0.723$, Average amplitude: F (2,76) $= 2.058$, $p = 0.135$, Latency: F (2,76) $= 2.661$, $p = 0.076$), and no significant differences were found at presentation rates (Peak: F (1,38) $= 0.326$, $p = 0.574$, Average amplitude: F (1,38) $= 0.412$, $p =$ 0.525, Latency: F (1,38) $= 0.656$, $p = 0.423$). At the two presentation rates, pairwise comparisons revealed that the peak and average amplitude of the picture pupil dilation were smallest ($ps < 0.001$), and the picture latency was shortest ($ps < 0.05$). For two presentation rates, no significant differences of peak, average amplitude and latency were found between words and numbers ($ps > 0.05$), except a significant difference of average amplitude between word and number at the 100 ms presentation rate ($p < 0.01$).

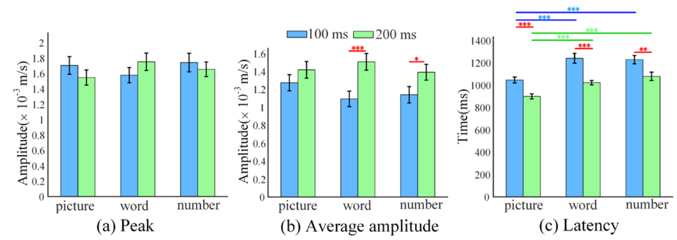

Fig. 9. The peaks, average amplitudes and latencies for pupil velocity in the time window of 500-1300 ms at two presentation rates. (a) Peak, (b) Average amplitude, (c) Latency. The error bars stand for the standard deviations.

Fig.9 shows the peaks, average amplitudes and latencies for pupil velocity in the time window of 500-1300 ms with three target types at two presentation rates. At the 100 ms presentation rate, the latencies of pupil velocity for pictures, words and numbers are 1046 ms, 1241 ms, 1228 ms, respectively. At the 200 ms presentation rate, the latencies of pupil velocity for pictures, words and numbers are 900 ms, 1022 ms, 1079 ms, respectively.

For the peak and average amplitude of pupil velocity, repeated measures ANOVAs of presentation rates (2 levels) $\times$ target types (3 levels) showed that there were no significant effects on the peak and average amplitude of pupil velocity with three target types (peak: F (2,76) $= 0.112$, $p = 0.885$, average amplitude: F (2,76) $= 1.520$, $p = 0.225$). There was no interaction between target type and presentation rate (peak: F (2,76) $= 1.956$, $p = 0.149$, average amplitude: F (2,76) $=$ 2.058, $p = 0.135$). No significant difference of peak was found at presentation rates (F (1,38) $= 0.007$, $p = 0.935$), and a significant difference of average amplitude was found at presentation rates (F (1,38) $= 8.657$, $p = 0.006$). Pairwise comparison revealed that the average amplitude of 100 ms

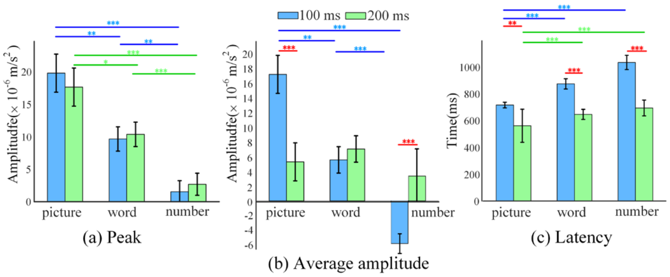

Fig. 10. The peaks, average amplitudes and latencies for pupil acceleration in the time window of 400-1300 ms at two presentation rates. (a) Peak, (b) Average amplitude, (c) Latency. The error bars stand for the standard deviations.

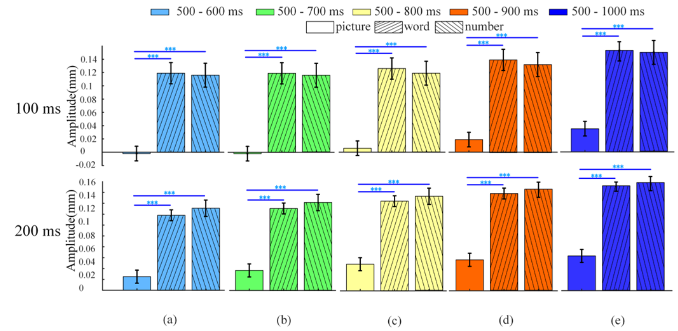

Fig. 11. The average amplitudes for pupil dilation at two presentation rates in the time windows of 500-600 ms, 500-700 ms, 500-800 ms, 500-900 ms and 500-1000 ms, respectively. The error bars stand for the standard deviations.

presentation rate was significantly smaller than that of 200 ms presentation rate ($p < 0.05$).

Repeated measures ANOVA of presentation rates (2 levels) × target types (3 levels) showed that there was a main effect on latency of pupil velocity with three target types (F (2,76) =17.184, $p = 0.000$). There was no interaction between target type and presentation rate (F (2,76) =1.988, $p = 0.144$), and a significant difference of latency at presentation rates (F (1,38) = 36.205, $p = 0.000$). Pairwise comparison revealed that, the picture latencies were significantly shorter than the others for the two presentation rates ($ps < 0.05$), and the latencies at 100 ms presentation rate were significantly longer than those at 200 ms presentation rate for the three target types ($ps < 0.01$).

Fig.10 shows the peaks, average amplitudes and latencies for pupil acceleration in the time window of 400-1000 ms with three target types at two presentation rates. At the 100 ms presentation rate, the latencies of pupil acceleration for pictures, words and numbers are 715 ms, 873 ms, 933 ms, respectively. At the 200 ms presentation rate, the latencies of pupil acceleration for pictures, words and numbers are 561 ms, 646 ms,693 ms, respectively.

For peak and average amplitude, repeated measures ANOVA of presentation rates (2 levels) × target types (3 levels) showed that there were main effects on the peak and average amplitude of pupil acceleration with three target types (peak: F (2,76) = 30.605, $p = 0.000$, average amplitude: F (2,76) = 13.522, $p = 0.000$).There was no interaction of peak between target type and presentation rate (F (2,76) = 0.355, $p = 0.702$), and an interaction of average amplitude between target type and presentation rate (F (2,76) = 9.809, $p = 0.000$). Meanwhile, no significant differences of peak and average amplitude were found between the two presentation rates (peak: F (1,38) = 0.002, $p = 0.965$, average amplitude: F (1,38) = 0.063, $p = 0.803$). For the three target types at the 100 ms presentation rate, pairwise comparisons revealed that the picture peak and average amplitude were significantly higher than the others ($ps < 0.01$), and the number peak was significantly lower than the others ($ps < 0.01$). In addition, for the three target types at the 200 ms presentation rate, pairwise comparisons revealed that the picture peak was significantly higher than the others ($ps < 0.05$), and the number peak was significantly lower than the others ($ps < 0.001$).

Repeated measures ANOVA of presentation rates (2 levels) × target types (3 levels) showed that there was a main effect on latency of pupil acceleration with three target types (F (2,76) =36.057, $p = 0.000$). There was no interaction between target type and presentation rate (F (2,76) = 1.108, $p = 0.366$), and a significant difference of latency between the two presentation rates (F (1,38) = 76.365, $p = 0.000$). Pairwise comparison revealed that, for the two presentation rates, the picture latency was significantly shorter than the others ($ps < 0.001$), and no significant differences of lantency were found between words and numbers ($ps > 0.05$). And for the three target types, the latencies at 100 ms presentation rate were significantly longer than those at 200 ms presentation rate ($ps < 0.01$).

D. Pupil Dilation Within 1000 ms

In order to utilize the respective advantages of EEG and pupillometry for fusion classification within the same time window, discriminative features of pupillometry were needed to be found out within 1000 ms. Fig.11 shows the average amplitudes of pupil dilation in the time windows including 500-600 ms, 500-700 ms, 500-800 ms, 500-900 ms, 500-1000 ms. Repeated measures ANOVAs of presentation rates (2 levels) × target types (3 levels) showed that there were main effects on average amplitude of pupil dilation with three target types in the five time windows (500-600 ms: F (2,76) = 52.593, $p = 0.000$, 500-700 ms: F (2,76) = 53.649, $p = 0.000$, 500-800 ms: F (2,76) = 55.276, $p = 0.000$, 500-900 ms: F (2,76) =59.087, $p = 0.000$, 500-1000 ms: F (2,76) = 62.026, $p = 0.000$). There were no interactions between target type and presentation rate in the five time windows (500-600 ms: F (2,76) = 0.144, $p = 0.866$, 500-700 ms: F (2,76) = 0.143, $p = 0.867$, 500-800 ms: F (2,76) = 0.131, $p = 0.877$, 500-900 ms: F (2,76) =0.084, $p = 0.919$, 500-1000 ms: F (2,76) = 0.112, $p = 0.894$), and significant differences of average amplitudes between the two presentation rates (500-600 ms: F (1,38) = 9.012, $p = 0.005$, 500-700 ms: F (1,38) = 19.044, $p = 0.000$, 500-800 ms: F (1,38) = 23.224, $p = 0.000$, 500-900 ms: F (1,38) = 21.110, $p = 0.000$, 500-1000 ms: F (1,38) = 15.409, $p = 0.000$). Pairwise comparisons revealed that, at two presentation rates, the picture average amplitude of pupil dilation were significantly smaller than the others ($ps < 0.01$), and no significant differences between words and numbers were found

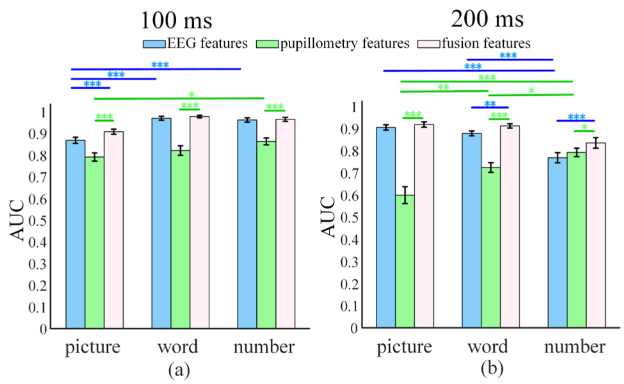

Fig. 12. The AUC values of EEG features, pupillometry features and fusion features at two presentation rates for picture, word and number target types. The error bars stand for the standard deviations.

($ps > 0.05$). And for the three target types, the average amplitudes at the 200 ms presentation rate were significantly higher than those at the 100 ms presentation rate ($ps < 0.05$).

E. Fusion Classification Performance of EEG and Pupillometry

Fusion classification performance was analyzed using EEG and pupillometry features. The average P300 amplitudes of the central-parietal regions (FC1, FCz, FC2, C1, Cz, C2, CP1, CPz and CP2 electrodes) were used as EEG features in the time window of 200-600 ms. The average amplitudes of pupil dilation in the five time windows (500-600 ms, 500-700 ms, 500-800 ms, 500-900 ms and 500-1000 ms) were used as pupillometry features. Non-target trials were randomly selected, whose trials number was the same as the target trials. 50% trials were used for training and the last trails for testing. The features were normalized using mini-max method before fusion, due to different scales between EEG and pupillometry. The features of EEG and pupillometry were cascaded and then classified using logistic regression. Fig.12 shows the classification performance of EEG features, pupillometry features and fusion features at two presentation rates.

Repeated measures ANOVA results showed a significant effect on AUC values among these features with three targets at 100 ms presentation rate (F (8,152) = 28.172, $p = 0.000$). Pairwise comparisons revealed that the picture AUC value of fusion features was significantly higher than those of EEG and pupillometry features, respectively ($ps < 0.001$). And the word and number AUC values of fusion features were significantly higher than those of pupillometry features ($ps < 0.001$). Furthermore, the word and number AUC values were significantly higher than the picture ones ($ps < 0.05$).

In addition, repeated measures ANOVA results showed a significant effect on AUC values among these features with three targets at 200 ms presentation rate (F (8,152) = 33.533, $p = 0.000$). Pairwise comparisons revealed that, the word and number AUC values of fusion features were significantly higher than those of EEG and pupillometry features, respectively ($ps < 0.05$). And the picture AUC value of fusion features was significantly higher than that of pupillometry features ($ps < 0.001$). Furthermore, the picture and word AUC values were significantly higher than the number ones ($ps < 0.05$).

IV. Discussion

A. ERP Analysis

This study showed that the picture P300 component had the largest amplitude and the longest latency at the 100 ms presentation rate. The results were consistent with those in previous studies [20], [21], [29], [30]. Tempel et al. [29] found that the P300 amplitude evoked by pictures was higher than that by words in a valence judgment task, because human brain required more cognitive capacity for the processing of pictures than for that of words. And the picture latency of P300 was slower than the word latency, because the picture stimuli were relatively unfamiliar for subjects compared with the word stimuli in this experiment, and then human brain needed relatively long time to process the picture stimuli. Lin et al. [18] demonstrated that high complexity images induced smaller P300 amplitude and longer P300 latency, because human brain required more cognitive loads and longer time to process high complexity images. In a study by Gomarus et al. [30], subjects were required to memorize 3 letters which contained all the same ('load1') or all different ('load3'). The results showed that load3 induced smaller P300 amplitude than load1, in which load3 was more difficult for subjects to memorize and more cognitive loads were required. Lees et al. [19] proposed that the P300 amplitude of numbers was significantly lower than those of pictures and words in PSVP tasks. In this experiment, the pictures and words (mother tongue) were easier to recognize for subjects compared with numbers, so the number stimuli needed higher cognitive loads to induce lower P300 amplitude.

As mentioned above, it was confirmed that a higher cognitive load could reduce P300 amplitude, and higher level cognitive processes could slow P300 latency [18], [19], [30]. In this study, behavioral results showed that the recognition accuracy of pictures was significantly higher than those of the others, which revealed that picture stimuli were easier to recognize and lower cognitive burden was required, compared with word and number stimuli. Therefore, the picture P300 amplitude was higher compared with words and numbers. Since the high complexity of pictures required higher level cognitive process, the picture latency was slower than the others.

At the 200 ms presentation rate, the P300 peak and average amplitude decreased, and the P300 latency shortened compared with 100 ms presentation rate. The behavioral results showed that the recognition accuracy of 200 ms presentation rate was significantly higher than that of 100 ms presentation rate for three target types, which indicated that human brain processed target stimuli more easily and accurately at 200 ms presentation rate. For the slow presentation rate, the brain could mobilize and integrate more cognitive resources for high-level cognitive processing, and accelerate processing speed, which would lead to smaller amplitude and shorter latency of P300 component [19], [30].

In addition, the P300 amplitude induced by words was significantly smaller than others at the 200 ms presentation rate. Since it was difficult for native Chinese speakers to maintain English words in working memory compared with the pictures and numbers, and then a higher cognitive load was required for words, and a smaller P300 amplitude was obtained. In this study, the classification performance showed that the word target achieved a high AUC value, which was consistent with Lees' study [19].

B. Pupillometry Analysis

In this study, the pupil dilation of pictures has the smallest size, and the shortest latency compared with numbers and words at the 100 ms presentation rate. The explained reasons were consistent with those in previous studies [11], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41]. Many studies showed that pupil dilation was related to cognitive load, and pupil size increased as the cognitive load increased [11], [31], [32]. For the digit span task in previous studies, pupil dilation increased with an increasing number of digits [33], [34], [35]. And Granholm et al. [36] measured pupil size using with 5 (low load), 9 (moderate load), 13 (excessive load) digits per string in a span recall task. The results further showed that pupil size increased with the increasing cognitive load, but not for more than moderate load. For the n-back task, some studies showed that pupil dilation increased with the increasing n [37], [38], [39]. For the Stroop task, Steenbergen et al. [40] proposed that pupil dilation increased for incongruent trials compared with congruent trials. For the sentence comprehension task, Just et al. [41] found that more complex sentences produced larger pupil size and longer latency.

In this study, behavioral results showed that the recognition accuracy of pictures was significantly higher than that of words and numbers at the 100 ms presentation rate. Compared with pictures, the recognition task of words and numbers were more difficult, and more cognitive resources in processing were required. Thus, the pictures induced smaller pupil size and shorter latency than the words and numbers.

In the early five time windows of 500-600 ms, 500-700 ms, 500-800 ms, 500-900 ms and 500-1000 ms, the average amplitudes of pupil dilation at 200 ms presentation rate was significantly higher than those at 100 ms presentation rate. The possible explanation was that, pupil dilation was induced by inhibition of the parasympathetic nervous system and Edinger-Westphal nucleus, and the inhibition was controlled by the LC-NE system [8], [10], [11]. At the early stage within 1000 ms, the activation intensity of the LC-NE system at 200 ms presentation rate might be greater than that at 100 ms presentation rate, thus the size of pupil dilation at 200 ms presentation rate was higher than that at 100 ms presentation rate. For the three target types, there were no significant differences in pupil size and latency between the two presentation rates in the time window of 500-2000 ms. During the whole time window, the pupil changes included dilation and contraction, and the extent of pupil change was limited, so there was no significant difference of the average pupil size for pupil dilation between the two presentation rates. According to

TABLE I

THE STATISTICAL RESULTS OF PEAKS, AVERAGE AMPLITUDES ANDLATENCIES OF P300, PUPIL DILATION, PUPIL VELOCITY ANDPUPIL ACCELERATION WITH THREE TARGET TYPES AT TWO PRESENTATION RATES.THE UP ARROW (↑) INDICATES THE LARGEST VALUE. THEDOWN (↓) ARROW INDICATES THE SMALLEST VALUE. THETILDE (∼) INDICATES NO SIGNIFICANT DIFFERENCE

Features		100 ms			200 ms			Change of presentation rate from 100 ms to 200 ms
		picture	word	number	picture	word	number	
P300 (200-600 ms)	Peak	↑			↑	↓		∼
	Average amplitude	↑				↓		↓
	Latency	↑	↓		∼	∼	∼	↓
Pupil dilation (500-2000 ms)	Peak	↓			↓			∼
	Average amplitude	↓		↑	↓			∼
	Latency	↓			↓			∼
Pupil dilation (500-1000 ms)	Average amplitude	↓			↓			↓
Pupil velocity (500-1300 ms)	Peak	∼	∼	∼	∼	∼	∼	∼
	Average amplitude	∼	∼	∼	∼	∼	∼	↑
	Latency	↓			↓			↓
Pupil acceleration (400-1000 ms)	Peak	↑		↓	↑		↓	∼
	Average amplitude	↑		↓	∼	∼	∼	∼
	Latency	↓			↓			↓

the behavioral and ERP analysis, the brain could mobilize and integrate more cognitive resources as the presentation rate changed from 100 to 200 ms. Although the increase of cognitive resources was not enough to make significant changes in the pupil size, it resulted in higher amplitudes of pupil velocity, and the shorter latency of pupil velocity and acceleration. In summary, the increase of cognitive resources from 100 ms to 200 ms presentation rate could affect the pupil velocity and acceleration, but not the pupil size.

C. Synchronous Analysis of EEG and Pupillometry

The results of synchronous analysis of EEG and pupillometry are summarized in Table I. There were different change patterns between ERP and pupil dilation with three target types at two presentation rates. At the 100 ms presentation rate, the picture P300 had the largest peak, average amplitude and the longest latency. Meanwhile, the picture pupil dilation had the smallest pupil size and the shortest latency. The picture pupil velocity had the shortest latency, and the picture pupil acceleration had the largest peak, average amplitude and the shortest latency.

As the presentation rate increased from 100 ms to 200 ms, the P300 amplitude decreased and the P300 latency became shorter. Meanwhile, no significant differences of pupil size and latency were found for the three target types. And the latencies of pupil velocity and acceleration were shortened.

As a whole, P300 and pupil dilation had opposite responses to the recognition tasks of the three target types at 100 ms presentation rate. In addition, at the 200 ms presentation rate, the P300 features had some changes, but the features of pupil dilation had no significant changes. Therefore, the fusion of EEG and eye-tracking would provide more effective information to improve the classification accuracy.

D. Effective Features of ERP and Pupillometry Fusion

Most EEG epochs were extracted within 1000 ms, but the latencies of pupil dilation and pupil velocity were more than

1000 ms. These might affect the classification performance of the synchronous fusion of EEG and eye-tracking. Therefore, it was necessary to find out the other pupil features in the early stage for the fusion classification of EEG and eye-tracking. As shown in Fig.11, the average amplitudes of pupil dilation for pictures were significantly lower than those for the others in the five time windows of 500-600 ms, 500-700 ms, 500-800 ms, 500-900 ms and 500-1000 ms. At the same time, pupil acceleration had earlier latency about 561-933 ms for the three target types. It was indicated that the pupil was in the dilation period within 1000 ms, and the pupil change was relatively sensitive due to the control of LC-NE system. Therefore, the average amplitude of pupil dilation in the time window of 500-1000 ms, the average amplitude and latency of pupil acceleration within 1000 ms, combining with the peak, average amplitude and latency of P300 in the time window of 200-600 ms, could be taken as the effective features of the fusion of EEG and eye-tracking for classification.

E. Future Work

In this study, there is still some work to be further studied for hybrid BCI based on eye-tracking in the practical application. First, more presentation rates needed to be considered to investigate whether there was a best presentation rate to improve classification performance of the hybrid BCI. Second, using the effective features of ERP and pupillometry, the best fusion methods between EEG and eye-tracking needed to be further studied for single-trial target classification. Third, the online performance of hybrid BCI based on eye-tracking needed to be tested for RSVP tasks. Finally, in this study, EEG and pupillometry features were cascaded and classified using LR to verify the effectiveness of the features. In future work, more effective classification and fusion methods will be further studied. These studies could provide valuable references for the practical applications of hybrid BCI based on eye-tracking, such as RSVP speller [6], [7], Satellite image retrieval [3], [4], screening of targets by supervisors [5], and so on.

V. Conclusion

In order to provide useful references for hybrid BCI based on eye-tracking, this study synchronously analyzed ERP and pupillometry in the RSVP tasks with three target types at two presentation rates. The results showed that, at two presentation rates, the picture P300 component had the largest amplitude and the longest latency, and the pupil dilation of pictures had the smallest size, and the shortest latency. As the presentation rate changed from 100 ms to 200 ms, the P300 amplitude became smaller, and the P300 latencies became shorter. Meanwhile, no differences of pupil size and latency were found in pupil dilation with three target types at the two presentation rates. The pupil features including the average amplitude and latency of pupil acceleration within 1000 ms, and the average amplitude of pupil dilation in the time window of 500-1000 ms, combining with the P300 features could be taken as effective ones within 1000 ms for the classification of hybrid BCI. And the synchronous analysis of EEG and pupillometry could be helpful to develop the fusion methods of EEG and eye-tracking. These findings could provide valuable references for multi-target recognition of hybrid BCI based on eye-tracking.

References

[1] S. Lees et al., "A review of rapid serial visual presentation-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 2, pp. 1–39, Apr. 2018.

[2] J. Polich, "Updating P300: An integrative theory of P3a and P3b," *Clin. Neurophysiol.*, vol. 118, pp. 2128–2148, Oct. 2007.

[3] P. Poolman, R. M. Frank, P. Luu, S. M. Pederson, and D. M. Tucker, "A single-trial analytic framework for EEG analysis and its application to target detection and classification," *NeuroImage*, vol. 42, no. 2, pp. 787–798, Aug. 2008.

[4] N. Bigdely-Shamlo, A. Vankov, R. R. Ramirez, and S. Makeig, "Brain activity-based image classification from rapid serial visual presentation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 5, pp. 432–441, Oct. 2008.

[5] A. Singh and J. Jotheeswaran, "P300 brain waves instigated semi supervised video surveillance for inclusive security systems," in *Proc. 8th. Int. Conf. Brain Inspired Cogn. Syst.*, vol. 5, Jul. 2018, pp. 184–194.

[6] B. H. Dobkin, "The evolution of neurorehabilitation and neural repair," *Neurorehabilitation Neural Repair*, vol. 28, no. 1, p. 3, Jan. 2014.

[7] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "Toward enhanced P300 speller performance," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 15–21, Jan. 2008.

[8] C. M. Privitera, L. W. Renninger, T. Carney, S. Klein, and M. Aguilar, "Pupil dilation during visual target detection," *J. Vis.*, vol. 10, no. 10, pp. 1–14, Aug. 2010.

[9] S. M. Wierda, H. Van Rijn, N. A. Taatgen, and S. Martens, "Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 22, pp. 8456–8460, 2012.

[10] I. Y. Chen, A. Karabay, S. Mathot, H. Bowman, and E. G. Akyurek, "Concealed identity information detection with pupillometry in rapid serial visual presentation," *Psychophysiology*, vol. 60, no. 1, Jul. 2022, Art. no. e14155.

[11] S. D. Goldinger and M. H. Papesh, "Pupil dilation reflects the creation and retrieval of memories," *Current Directions Psychol. Sci.*, vol. 21, no. 2, pp. 90–95, Mar. 2012.

[12] S. Sirois and J. Brisson, "Pupillometry," *Wiley Interdiscip. Rev. Cogn. Sci.*, vol. 5, no. 6, pp. 679–692, Nov. 2014.

[13] A. Bjernestedt, R. Johansson, and P. Pärnamets, "Pupil dilation reflects interference during memory retrieval," *Curr. Dir. Psychol. Sci.*, vol. 21, no. 2, pp. 90–95, 2021.

[14] N. Morris and D. M. Jones, "Memory updating in working memory: The role of the central executive," *Brit. J. Psychol.*, vol. 81, no. 2, pp. 111–121, May 1990.

[15] P. Sajda, A. Gerson, and L. Parra, "High-throughput image search via single-trial event detection in a rapid serial visual presentation task," in *Proc. 1st Int. IEEE EMBS Conf. Neural Eng.*, Mar. 2003, pp. 7–10.

[16] A. D. Gerson, L. C. Parra, and P. Sajda, "Cortically coupled computer vision for rapid image search," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 174–179, Jun. 2006.

[17] H. Cecotti, J. Sato-Reinhold, J. L. Sy, J. C. Elliott, M. P. Eckstein, and B. Giesbrecht, "Impact of target probability on single-trial EEG target detection in a difficult rapid serial visual presentation task," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2011, pp. 6381–6384.

[18] Z. Lin, Y. Zeng, L. Tong, H. Zhang, C. Zhang, and B. Yan, "Method for enhancing single-trial P300 detection by introducing the complexity degree of image information in rapid serial visual presentation tasks," *PLoS ONE*, vol. 12, no. 12, pp. 1–14, Dec. 2017.

[19] S. Lees, P. McCullagh, P. Payne, L. Maguire, F. Lotte, and D. Coyle, "Speed of rapid serial visual presentation of pictures, numbers and words affects event-related potential-based detection accuracy," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 113–122, Jan. 2020.

[20] M. Qian et al., "Decision-level fusion of EEG and pupil features for single-trial visual detection analysis," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 7, pp. 1929–1937, Jul. 2009.

[21] S. Cheng, J. Wang, L. Zhang, and Q. Wei, "Motion imagery-BCI based on EEG and eye movement data fusion," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2783–2793, Dec. 2020.

[22] B. Li, Y. Lin, X. Gao, and Z. Liu, "Enhancing the EEG classification in RSVP task by combining interval model of ERPs with spatial and temporal regions of interest," *J. Neural Eng.*, vol. 18, no. 1, Feb. 2021, Art. no. 016008.

[23] H. Cecotti, M. P. Eckstein, and B. Giesbrecht, "Single-trial classification of event-related potentials in rapid serial visual presentation tasks using supervised spatial filtering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 11, pp. 2030–2042, Nov. 2014.

[24] *Morguefile*. Accessed: Dec. 16, 2018. [Online]. Available: http://www.morguefile.com

[25] *aGlass DKII, 7invensun*. Accessed: Apr. 10, 2018. [Online]. Available: https://www.7invensun.com

[26] G. Porter, T. Troscianko, and I. D. Gilchrist, "Effort during visual search and counting: Insights from pupillometry," *Quart. J. Exp. Psychol.*, vol. 60, no. 2, pp. 211–229, 2007.

[27] C. Strauch, I. Koniakowsky, and A. Huckauf, "Decision making and oddball effects on pupil size: Evidence for a sequential process," *J. Cognition*, vol. 3, no. 1, pp. 1–17, Mar. 2020.

[28] N. Cohen, N. Moyal, and A. Henik, "Executive control suppresses pupillary responses to aversive stimuli," *Biol. Psychol.*, vol. 112, pp. 1–11, Dec. 2015.

[29] K. Tempel, L. Kuchinke, K. Urton, L. H. Schlochtermeier, H. Kappelhoff, and A. M. Jacobs, "Effects of positive pictograms and words: An emotional word superiority effect?" *J. Neurolinguistics*, vol. 26, no. 6, pp. 637–648, Nov. 2013.

[30] H. K. Gomarus, M. Althaus, A. A. Wijers, and R. B. Minderaa, "The effects of memory load and stimulus relevance on the EEG during a visual selective memory search task: An ERP and ERD/ERS study," *Clin. Neurophysiol.*, vol. 117, no. 4, pp. 871–884, Apr. 2006.

[31] J. Beatty and B. L. Wagoner, "Pupillometric signs of brain activation vary with level of cognitive processing," *Science*, vol. 199, no. 4334, pp. 1216–1218, Mar. 1978.

[32] P. Van Der Wel and H. van Steenbergen, "Pupil dilation as an index of effort in cognitive control tasks: A review," *Psychonomic Bull. Rev.*, vol. 25, no. 6, pp. 2005–2015, Feb. 2018.

[33] T. Piquado, D. Isaacowitz, and A. Wingfield, "Pupillometry as a measure of cognitive effort in younger and older adults," *Psychophysiology*, vol. 47, no. 3, pp. 560–569, May 2010.

[34] W. S. Peavler, "Pupil size, information overload, and performance differences," *Psychophysiology*, vol. 11, no. 5, pp. 559–566, Sep. 1974.

[35] E. L. Johnson, A. T. M. Singley, A. D. Peckham, S. L. Johnson, and S. A. Bunge, "Task-evoked pupillometry provides a window into the development of short-term memory capacity," *Frontiers Psychol.*, vol. 5, p. 218, Mar. 2014.

[36] E. Granholm, R. F. Asarnow, A. J. Sarkin, and K. L. Dykes, "Pupillary responses index cognitive resource limitations," *Psychophysiology*, vol. 33, no. 4, pp. 457–461, Jul. 1996.

[37] D. Pehlivanoglu, S. Jain, R. Ariel, and P. Verhaeghen, "The ties to unbind: Age-related differences in feature (un)binding in working memory for emotional faces," *Frontiers Psychol.*, vol. 5, p. 253, Apr. 2014.

[38] A.-M. Brouwer, M. A. Hogervorst, M. Holewijn, and J. B. F. Van Erp, "Evidence for effects of task difficulty but not learning on neurophysiological variables associated with effort," *Int. J. Psychophysiol.*, vol. 93, no. 2, pp. 242–252, Aug. 2014.

[39] J. F. Hopstaken, D. Van Der Linden, A. B. Bakker, and M. A. J. Kompier, "A multifaceted investigation of the link between mental fatigue and task disengagement," *Psychophysiology*, vol. 52, no. 3, pp. 305–315, Mar. 2015.

[40] H. van Steenbergen, G. P. H. Band, and B. Hommel, "Does conflict help or hurt cognitive control? Initial evidence for an inverted U-shape relationship between perceived task difficulty and conflict adaptation," *Frontiers Psychol.*, vol. 6, p. 974, Jul. 2015.

[41] M. A. Just and P. A. Carpenter, "The intensity dimension of thought: Pupillometric indices of sentence processing," *Can. J. Exp. Psychol.*, vol. 47, no. 2, pp. 310–339, Jun. 1993.