Proceedings of the
8th Graz Brain-Computer Interface Conference 2019 DOI: 10.3217/978-3-85125-682-6-53

AN EXPECTATION-BASED EEG MARKER FOR THE SELECTION OF MOVING OBJECTS WITH GAZE

D.G. Zhao[1], A.N. Vasilyev[1], B.L. Kozyrskiy[1], A.V. Isachenko[1,2], E.V. Melnichuk[1], B.M. Velichkovsky[1,2], S.L. Shishkin[1]

[1] Lab. for Neurocognitive Technologies, NRC "Kurchatov Institute", Moscow, Russia
[2] Institute of NBCIS S&T, Moscow Institute of Physics and Technology, Moscow, Russia

E-mail: sergshishkin@mail.ru

ABSTRACT: The use of an EEG expectation-related component, the expectancy wave (E-wave), in brain-machine interaction was proposed more than 50 years ago, but active exploration of this possibility has started only recently, in the context of developing passive brain-computer interfaces for the enhancement of gaze interaction. We report, for the first time, the results of a systematic experimental study that revealed an EEG marker for selecting intentionally an object among other moving objects using smooth pursuit eye movements. This marker appeared to have the same nature as the E-wave previously observed in the EEG accompanying the selection of static objects with gaze fixations. A convolutional neural network classified the intentional and spontaneous smooth pursuit eye movements with average ROC AUC 0.69±0.13 (M±SD). These results suggest that the E-wave might be robust enough to serve, after further improvement of the methodology, as the basis of hybrid eye-brain-computer interfaces applied for selection in dynamically changing visual environments.

INTRODUCTION

One of the oldest proposals for the design of a brain-computer interface (BCI), put forward by Grey Walter more than 50 years ago, was the use of the expectancy wave (E-wave), the electroencephalogram (EEG) slow negative wave that arises when a person is waiting for a stimulus. In this paradoxical BCI design expectation of the interface triggering was proposed to be used to trigger the interface [1]. Although in the subsequent decades the E-wave (later known also under the name of stimulus preceding negativity, SPN) and the more complex phenomenon, the contingent negative variation (CNV), were among the most studied EEG phenomena, they still not received much attention in the BCI world. One important exception was the proposal to use an expectation based passive BCI to support gaze interaction [2, 3] – more specifically, to solve the Midas touch problem, one of the most serious obstacles for using gaze as a tool to control computers and other devices. The Midas touch problem is the inability of gaze based human-machine interfaces to differentiate, in many cases, eye movements and fixations intentionally used for control from spontaneous ones, e.g. those used for visual scene exploration [4]. Later, an EEG component with the time course and topography similar to a typical SPN/E-wave and enabling classification of the gaze fixation intentionally used to make moves in a game vs. spontaneous gaze fixations was described [5] and first online tests of a hybrid eye-brain-computer interface (EBCI) based on this design were run [6]. In these studies, static visual layout was used, while the Midas touch problem may be more severe when the scene is changing dynamically.

Recently, the selection of moving objects with smooth pursuit eye movement was found to be very effective, in addition to the long-used gaze fixations and saccades [7]. In the presence of multiple moving objects, however, any of such objects can easily attract attention and become pursued with gaze automatically. The use of an EBCI instead of a merely gaze-based interface for moving object selection thus seems reasonable, however, only experimental studies can show if this is feasible, as multiple factors may change EEG components in such dynamical settings (see discussion on such factors for the case of selection of moving objects with the P300 BCI in [8] and for the case of using the fixation-related EEG P300 wave in the context of visual search in [9]). The EEG under smooth pursuit eye movements was studied very little, and we found no prior studies of EEG accompanying intentional object selection with the smooth pursuit eye movements.

In our preliminary studies [10, 11] we observed an EEG component resembling the SPN/E-wave prior to moving object selection by gaze, but a similar wave also accompanied smooth pursuit eye movements in control conditions, where the intention to select an object could not appear. We concluded that certain form expectation or at least estimation of time to execute certain operation (a related cognitive activity) easily appears under various conditions when a pursuit is intentionally used to follow instruction. In the current study, we designed a control condition that was free from this shortcoming and corrected some other details of the study design. With the new design, we collected data

 DOI: 10.3217/978-3-85125-682-6-53

from a larger group of participants, which allowed us to analyze the difference between the EEG accompanying spontaneous smooth pursuit and intentional selection of moving objects by gaze, and also to model their classification in the framework of an EBCI using a convolutional neural network.

METHODS

Participants: 14 healthy volunteers (8 male, 6 female), aged 22–52 (M±SD 28±8) years participated in our study (participant #1 participated in two sessions on different days). Except for four participants, others had prior experience with a gaze-based interface. Ten participants had normal vision, others had corrected to normal vision (3 with glasses and 1 with contact lenses).

Data acquisition: A consumer grade eye tracker *Tobii 4C* (*Tobii*, Sweden) was attached to the lower edge of the monitor. EEG was recorded from 19 locations (F3, Fz, F4, C3, Cz, C4, P1, P3, Pz, P2, P4, PO7, PO3, POz, PO4, PO8, O1, Oz, O2) at 500 Hz with *NVX52* amplifier (*MCS*, Russia). Digitally connected electrodes at earlobes served as reference. The vertical electrooculogram (EOG) was recorded using electrodes about 2 cm below the right eye, and the horizontal EOG with electrodes above the eyebrow and near the outer canthi, both with the same amplifier. Impedance was kept below 20 KOhm. In the last four experiments, a signal from a microphone was also recorded with the same amplifier to enable detection of time periods when the participants reported their answers.

Task: The participants were presented with 15 "balls" (circles), each 80 px in diameter (2.8°), displayed on an 18.5" monitor, at a distance of about 64 cm from their eyes (Fig. 1). The balls were numbered 1 to 15 and contained 5 to 8 dots around the number (Fig. 1, right). In the "dynamic" conditions (SMB, FA, CT) balls moved linearly on the screen at ~230 px/s (6.8°/s) speed, changing their movement direction in a natural way when hitting each other or the screen edges.

Balls with randomly varying numbers of dots were generated as 15 different images, separately for each block of four runs. Each image was assigned to each ball number at the start of a run.

In a run, the participants were asked to complete one of the following tasks:

Select moving balls (SMB) – select 15 balls one by one according to their numbers in ascending order and then (only for subjects 7-14) in descending order (see details about selection below).

Select static balls (SSB) – select 15 balls one by one as in the SMB condition, but in this case, the balls were fixed at their initial positions.

Find accelerated (FA) – find the ball that was moving faster than the other balls. This target ball, randomly chosen, accelerated once after 15 to 20 s from the block start by 1.3˚/s (19% of the basic speed). Participants were advised to announce the number of the ball that appeared to be the fast one only when they made sure that it was actually faster than the others.

Counting task (CT) – find, one by one, five objects with a specified number of dots and summarize the objects' numbers. In the CT condition, participants were asked, five times in a row, to find a ball with the number of dots specified by the experimenter (this number could be from 5 to 8, chosen randomly each time). After they had heard the number, they had to find an appropriate ball and summarize its labeling number with previous the number(s) of the previous ball(s).

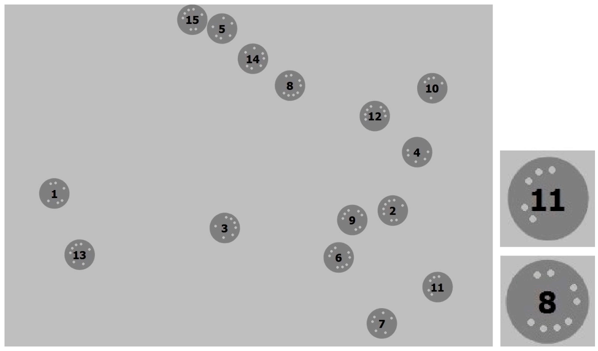

Figure 1: A screenshot of a typical experimental layout (**left**), and examples of "balls" with different amount and positioning of dots (**right**).

The SMB and SSB tasks modeled intentional gaze selection, while FA and CT tasks served to collect data on smooth pursuit eye movements without intention to make a selection. The SMB task was similar to the task used in [12] where gaze based selection was compared to a mouse-based one. However, in the current study the number of balls and their speed were set at lower levels in all tasks, otherwise, FA and CT tasks appeared to be too difficult.

At the beginning of the run, all the balls were dark grey. In the SMB and SSB conditions, a selected ball was highlighted with a green color. Ball color returned to dark grey when another ball was selected. Participants were asked to select the 15 balls in ascending order (all participants) and then in descending order (only participants 7-14) as soon as possible. In the FA and CT conditions, the balls that were considered selected by the online selection algorithm were not highlighted.

Online Gaze-Based Selection: In all conditions, the median distance from the gaze position to each ball's center was calculated in a moving window of ~867 ms length. If this measure computed for a ball did not exceed 55 pixels (in other words, the distance from the gaze position to this ball's center did not exceed 55 px longer than half of the length of this window) and was the smallest among all balls, this ball was selected. Some samples from the eye tracker (sampling rate 90 Hz) were discarded from time to time to maintain synchrony with the monitor (its refresh rate was 75 Hz). Using a high-speed video camera (240 fps) we found out that a delay between a saccade start and the start of a gaze-controlled cursor moving was approximately 140 ms. To improve synchronization between the gaze and

 DOI: 10.3217/978-3-85125-682-6-53

balls coordinate streams, the latter was intentionally delayed by the same value. The eye trackers' built-in alignment filter was not applied. With a few exceptions, the selection algorithm followed one described in [12].

Algorithm triggering time was recorded using a photosensor attached to the lower left edge of the screen, where a circle not visible to a participant changed its brightness as selection algorithm triggered (in the SMB and SSB conditions, it was in the same video frame when the visual feedback was given). The signal from the sensor was recorded synchronously with the EEG by the amplifier. In addition, the ID of the selected ball was recorded regardless of its highlighting.

The visual task and the selection algorithm were implemented in C++ and QML with the use of the Qt framework.

Procedure: Participants were seated in an adjustable chair in front of an office table with the monitor. To ensure head position stabilization a chin rest was used.

The experiment consisted of six blocks, each containing all four tasks (conditions). The order of conditions was the same over the blocks for a single participant but counterbalanced in the group of participants. Conditions where balls were intentionally selected with gaze (SMB, SSB) were not allowed to immediately follow each other. Each condition in a block included one run of ~2 min duration.

Before the start of the first block, the experimenter showed the moving balls and in detail explained the tasks. The first block was considered as a practice, and the participants were allowed to ask any questions during the tasks. The data from the subsequent five blocks were used for analysis. Before each block, native 7-points calibration of the eye tracker was used. Also, calibration was monitored by the experimenter during SMB and SSB conditions. If the participant felt discomfort at those conditions, re-calibration was run and the condition was overwritten. Within a single block, participants were asked to avoid movements to prevent calibration distortion. After completing one block, the participants were allowed to rest for a few minutes. 10 minutes break came after the third block.

In the end of an experiment, participants were asked to rate how difficult was performing a task by putting a mark on four visual analog scales (VAS), one per condition. Left end of the scale was labeled “very easy” and the right end was labeled "very difficult".

Data Analysis: In SMB and SSB conditions we considered only selections in right order (balls #1, #2, … #15, #14, … #1).

Averaging of pursuit-related EEG was triggered using a sensor (photodiode) placed in the bottom left screen corner, which was activated (not visible to the participant) as the algorithm triggered. In SMB and SSB conditions, this was followed with a ball highlighting. In FA and CT conditions, highlighting didn't happen, but the data about fixations still recorded.

Preprocessing: All EEG analysis was made offline using the data collected during the experiments. For individual trial analysis, EEG data were filtered with lowpass Butterworth 4th order filter with cutoff at 30 Hz and then cut into epochs around the time of interest, including the time of online gaze-based selection algorithm triggering and the time of gaze pursuit initial fixation onset. The latter was found using gaze velocity estimated from the EOG prior to the algorithm triggering (our *Tobii* eye tracker license did not allow for gaze coordinate offline analysis). Epochs (trials) were discarded from analysis if either of the following conditions were met:

1) Pursuit duration was shorter than 450 ms or exceeded 1200 ms
2) The trial was not in the correct selection sequence (for SMB and SSB)
3) The trial resulted in the selection of the accelerated ball (for FA)
4) Speech sounds were detected between fixation onset and algorithm triggering (for participants 11 and 14)
5) Amplitude range in any EEG or EOG signal exceeded 150 µV during gaze pursuit

The resulting dataset was used for the EEG data analysis and classification.

Signal preprocessing, segmentation and visualization was performed using MATLAB (*MathWorks*, USA). Function *topoplot* from the EEGLAB package [13] was used to create topographical amplitude maps.

Classification: For training, the SMB and SSB data were joined together and the second class consisted of the CT data, while during testing the SMB and SSB data were considered separately, and the second class again was CT. Features were amplitudes from the time interval starting -400..0 ms relative to online gaze-based selection triggering. Average over -100..0 ms interval was subtracted from the same channel data in each trial, i.e., was used as a baseline. Note that a saccade typically preceded the start of pursuit, thus we could not use the pre-pursuit EEG as the baseline in single-trial processing without special measures to remove the EOG, while such measures could introduce data distortions that are difficult to control.

We used Keras implementation of the convolutional neural network EEGNet [14] (https://github.com/vlawhern/arl-eegmodels) under Python 3.6. The procedure consisted of 4-fold cross-validation (80% of the whole dataset) and testing on the holdout set (20% of the whole dataset). Target to non-target class ratio was preserved when forming the test subset. For each fold, we saved the best model after the optimal number of training iterations with respect to performance on the validation set of each fold, and for each trial from the testing set, we averaged predicted probabilities from these four models. We used hyperparameters optimized for the case of classifying the EEG accompanying static object selection by gaze in our study [15] (see it also for other details of the classification procedure). Since the EBCI design makes possible the use of classifiers with relatively low sensitivity while maintaining high specificity, but the balance between them might be dependent of the use

 DOI: 10.3217/978-3-85125-682-6-53

case [3, 5, 6], we found more relevant to present the test results using the ROC AUC measure which is not related to the choice of the classifier threshold.

RESULTS

In the responses to the questionnaire, the FA condition was ranked as the most "difficult" (M±SD 71.8±17.6), while for the CT condition the difficulty score was 46.3±22.7 (the score could be between 0 and 100). Only two participants perceived CT as more difficult than FA. The intentional selection conditions, SMB and SSB, were scored as less difficult: 19.2±21.3 and 10.2±9.1, respectively (the difference from the CT condition was significant, with p=0.33 in both cases, according to Wilcoxon matched pairs test).

According to visual inspection of the averaged EEGs with and without rejection of epochs containing speech signal, they did not substantially differ. For further EEG analysis and classification, we excluded data obtained from two participants: #12, who often did not follow instructions correctly enough, and #13, who had issues with several EEG electrodes. The number of EEG epochs obtained in each condition for the remaining 12 participants is shown in Table 1.

Fig. 2 shows the grand average EEG amplitude at the electrode POz. In all conditions, a similar EEG pattern was observed about 0.5-0.6 s prior to the selection algorithm triggering, i.e., near the beginning of the pursuit (the largest positive peak could be the lambda wave, well known in the EEG accompanying static gaze fixations). In SMB and SSB, a slow negative wave was developing up to the time of the online gaze-based selection algorithm triggering; soon after it, the amplitude abruptly went in the opposite direction (evidently, in response to the visual feedback). The time course of EEG amplitude in the FA condition was quite similar to the SMB and SSB conditions along most time before the selection algorithm triggering, while in the CT condition only very light deviation to negative direction was observed.

The EEG amplitude head maps (Fig. 2, bottom) demonstrated a pattern similar to previously reported for the intentional static object selection with gaze [5] in the SMB and SSB conditions, and its weaker variant in the FA condition, but nothing similar in the CT conditions.

Mean EEG amplitude at POz over the last 100 ms before selection triggering varied significantly across conditions, as determined by one-way ANOVA ($F(3,43) = 6.97$, $p = 0.0006$). A Tukey post hoc test revealed that POz was significantly more negative in SMB and SSB compared to CT (p=0.003 and p=0.001, respectively). Amplitude in FA showed no significant difference when compared to any other condition.

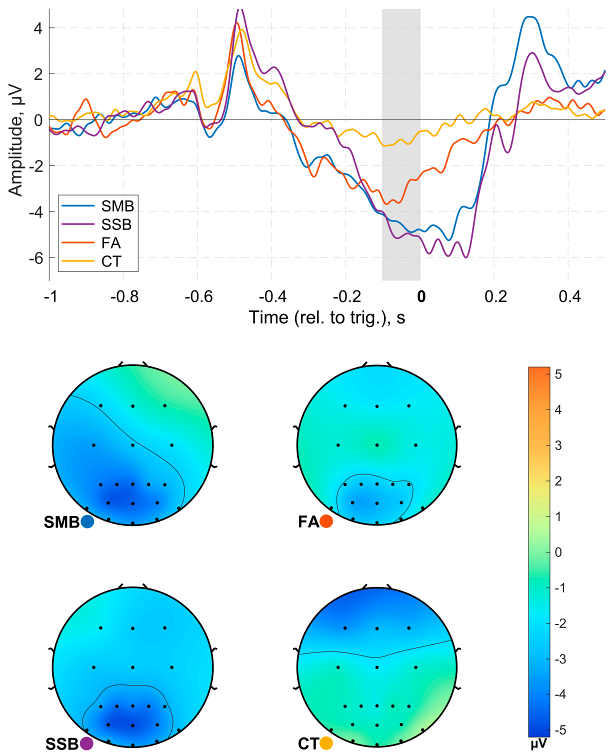

Figure 2: (**Top**) The grand average EEG amplitude at electrode POz relative to gaze-based selection algorithm triggering (N=12, baseline -1200...-700 ms). Feedback was provided to the participants at zero time in SMB and SSB but not in CT and FA conditions. (**Bottom**) Topographical maps of the EEG amplitude averaged over the last 100 ms before triggering.

Fig. 3 demonstrates the variability of the individual EEG data and shows (in its right part) the EEG dynamics with a more precise synchronization with the pursuit start. As in the studies of static object selection with gaze [5, 6], the E-wave was evident in the intentional selection conditions in all participants. Improvement of synchronization with the pursuit start tended to improve tracking the E-wave, although the difference between the two types of averaging was not large. In some participants (here, S09 and S10) the EEG dynamics in the pre-triggering time interval in the CT condition was similar to the SMB and SSB. Visual comparison of the averaged posterior EEG and EOG waveforms revealed no correspondence between them.

Classification results are presented in Table 1. We did not obtain enough data in the FA condition, so the classifier was trained and tested on both intentional conditions only against the CT condition (for training, the SMB and SSB subsets were collapsed). Note that intentional moving object selection with smooth pursuit eye movements (SMB) was classified almost as good as the selection of static objects with gaze fixations (SSB).

DOI: 10.3217/978-3-85125-682-6-53

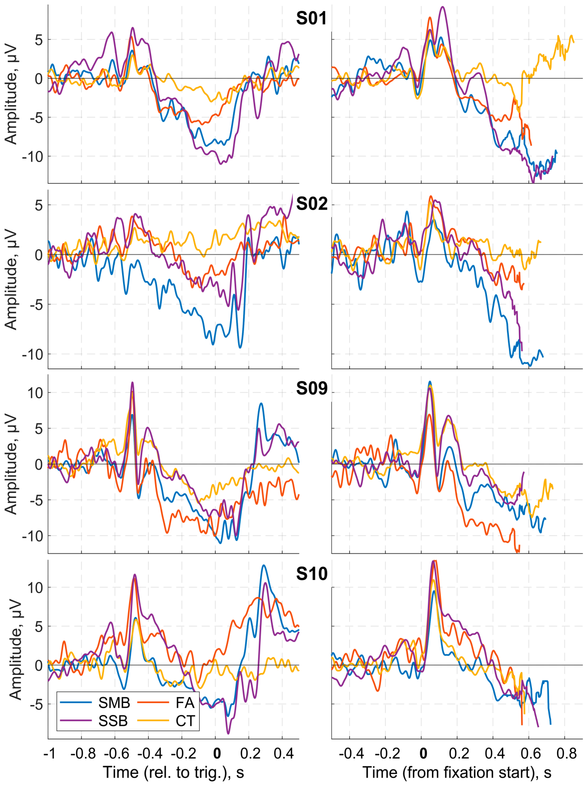

Figure 3: Averaged EEG amplitude at POz for two participants with good (S01, S02) and weak (S09, S10) separation of the SMB and SSB conditions from the CT condition: (**Left**) Synchronized (zero time) with gaze-based selection algorithm triggering (baseline -1200...-700 ms). (**Right**) Synchronized (zero time) with smooth pursuit start and truncated at online algorithm triggering (averaged curves truncated where n of trials was < 25; baseline -700 to -200 ms).

DISCUSSION

This study, with the preceding preliminary studies [10, 11], demonstrated for the first time that intentional selection of a moving object with smooth pursuit eye movements can be accompanied by a posterior EEG negativity, similar to the phenomenon observed earlier in the selection of static objects with gaze fixations.

In our preliminary studies in small groups of participants who selected moving objects with their gaze [10, 11] we found it difficult to find appropriate control condition in which moving objects are often pursued but without expecting any upcoming event. Moreover, when the participants were instructed just to pursue a moving object for a certain time without any other task, a similar negative wave also developed, likely due to mental counting time. In the current study, pursuit under the instruction to look for a faster ball also evoked a similar EEG dynamics, which could be related to waiting for a good opportunity to compare the speed of the pursued ball with the speed of some other ball(s) (a common strategy to solve this task). Surprisingly, even the less expectation-loaded condition, the CT (counting the dots on the moving balls and adding up the number of the ball to the previous sum), was accompanied with a similar deviation of the averaged EEG in some participants. Thus, cautions are due concerning the specific relation of the marker to intentional selection. Nevertheless, in the majority of participants and in the grand average (Fig. 2) the difference between the intentional selection and this control condition was clear.

Table 1: Number of epochs per condition used for EEG analysis and classification, and the classification results.

Subject	N of epochs			ROC AUC	
	SMB	SSB	CT	SMB vs CT	SSB vs CT
1	124	148	149	0.59	0.69
2	47	62	99	0.50	0.46
3	62	62	112	0.70	0.65
4	40	34	49	0.89	0.86
5	58	62	90	0.45	0.53
6	60	64	121	0.74	0.74
7	184	165	140	0.76	0.70
8	155	161	139	0.74	0.77
9	123	123	155	0.68	0.76
10	122	128	66	0.69	0.85
11	123	130	61	0.89	0.89
14	118	128	74	0.71	0.67
M	101.3	105.6	104.6	0.69	0.71
SD	46.4	45.6	36.8	0.13	0.13

Another concern could be inflation of the classification results due to the EEG contamination by the EOG. However, we used only features from the pursuit time interval. While it was accompanied by certain EOG dynamics, the longest possible distance during the time interval used for feature extraction could be 3° and in the majority of trials the distance was much lower, so the EOG contribution to EEG could not be strong. Removing the frontal electrodes from the analysis did not lead to a substantial decrease in ROC AUC, demonstrating that the EOG contribution could not strongly support classification. Nevertheless, additional study will be needed to quantify possible effects from the EOG or other sources not specific to the intentional selection that may inflate the classification results. Moreover, cleaning data from EOG might even help to improve classification.

Unlike in the studies of static object selection with gaze fixations [5, 6], the task in the current study was not free from visual search, and the pursuit could often start just after identification of the target object. Thus, the P300 wave that provides the most useful features for classification in the case of visual search [9] could also contribute to classification in our case. However, we did not observe a strong P300 wave in the averaged waveforms (Fig. 2, 3; as the first large positive peak was likely the lambda wave, the P300 peak could be expected 200 ms after it or later), so its contribution to classification could not be high.

We used an advanced classifier that demonstrated impressive performance in several BCI paradigms [14]

 DOI: 10.3217/978-3-85125-682-6-53

and also in the EBCI paradigm, i.e. in the classification of intentional vs. spontaneous gaze fixations [15]. The hyperparameters used in this study, however, were tuned for a different setup. It is possible that their more specifical tuning would provide better classification results. It is also possible that other improvements can enhance the interface design and the data processing pipeline, making the EBCI performance sufficiently good at least for certain applications requiring selection in a dynamically changing visual environment.

CONCLUSION

The results obtained in this study suggest that the EEG negative wave that is developing when a gaze-based input user expects feedback from the interface might be robust enough to serve, after further improvement of the methodology, as the basis of hybrid eye-brain-computer interfaces applied for selection in dynamically changing visual environment.

ACKNOWLEDGMENT

This study was supported in part by the Russian Science Foundation, grant 18-19-00593 (development, adaptation and application of techniques for data preprocessing and classification).

REFERENCES

[1] Walter WG. Expectancy waves and intention waves in the human brain and their application to the direct cerebral control of machines. Electroenceph. Clin. Neurophysiol. 1966; 21: 616
[2] Ihme K, Zander TO. What you expect is what you get? Potential use of contingent negative variation for passive BCI systems in gaze-based HCI, in Proc. ACII'11, 2011, 447-456
[3] Protzak J, Ihme K, Zander TO. A passive brain-computer interface for supporting gaze-based human-machine interaction, in Proc. UAHCI'13, 2013, 662-671
[4] Jacob RJK. The use of eye movements in human-computer interaction techniques: what you look at is what you get. ACM Trans. Inf. Syst. 1991; 9(2): 152-169
[5] Shishkin SL, Nuzhdin YO, Svirin EP, Trofimov AG, Fedorova AA, Kozyrskiy BL, Velichkovsky BM. EEG negativity in fixations used for gaze-based control: Toward converting intentions into actions with an Eye-Brain-Computer Interface. Front. Neurosci. 2016; 10: 528
[6] Nuzhdin YO, Shishkin SL, Fedorova AA, Kozyrskiy BL, Medyntsev AA, Svirin EP, Korsun OV, Dubynin IA, Trofimov AG, Velichkovsky BM. Passive detection of feedback expectation: Towards fluent hybrid eye-brain-computer interfaces, in Proc. 7th Graz BCI Conf., 2017, 361-366
[7] Vidal M, Bulling A, Gellersen H. Pursuits: spontaneous interaction with displays based on smooth pursuit eye movement and moving targets, in ACM UbiComp'13, 2013, 439-448
[8] Ganin IP, Shishkin SL, Kaplan AY. A P300-based brain-computer interface with stimuli on moving objects: Four-session single-trial and triple-trial tests with a game-like task design. PloS ONE. 2013; 8(10): e77755
[9] Uščumlić M, Blankertz B. Active visual search in non-stationary scenes: coping with temporal variability and uncertainty. J. Neural Eng. 2016; 13(1): 016015
[11] Zhao DG, Isachenko AV, Melnichuk EV, Kozyrskiy BL, Shishkin SL. EEG potentials related to moving object selection with gaze: a possible basis for more flexible eye-brain-computer interfaces. Opera Medica et Physiologica. 2018; 4(S1): 109-110
[12] Zhao D, Isachenko I, Melnichuk E, Shishkin S. Eye-brain-computer interfacing with smooth pursuit eye movements, in Proc. 4th Int. Conf. BCI: Science and Practice, Samara, 2018, 21
[13] Isachenko AV, Zhao DG, Melnichuk EV, Dubynin IA, Velichkovsky BM, Shishkin SL. The pursuing gaze beats mouse in non-pop-out target selection, in 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, 3518-3523
[14] Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods. 2004; 134(1): 9-21
[15] Lawhern VJ, Solon AJ, Waytowich NR, Gordon SM, Hung CP, Lance BJ. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. J. Neural Eng. 2018; 15: 056013
[16] Kozyrskiy BL, Ovchinnikova AO, Moskalenko AD, Velichkovsky BM, Shishkin SL. Classification of the gaze fixations in the eye-brain-computer interface paradigm with a compact convolutional neural network. Procedia Computer Science. 2018; 145: 293-299