ARTICLE IN PRESS
JID: COMM [m5G;January 21, 2020;23:40]
Computer Methods and Programs in Biomedicine xxx (xxxx) xxx
Contents lists available at ScienceDirect
Computer Methods and Programs in Biomedicine
journal homepage: www.elsevier.com/locate/cmpb

A novel hybrid BCI speller based on RSVP and SSVEP paradigm

Shayan Jalilpour, Sepideh Hajipour Sardouie*, Amirmohammad Mijani

Human-Machine Interfaces Laboratory (HMIL), Sharif University of Technology, Tehran, Iran

A R T I C L E I N F O

Article history:
Received 24 May 2019
Revised 30 November 2019
Accepted 8 January 2020
Available online xxx

Keywords:
Brain computer interface (BCI): Hybrid paradigm: Rapid serial visual presentation (RSVP): Steady-state visual evoked potential (SSVEP): P300

A B S T R A C T

Background and objective: Steady-state visual evoked potential (SSVEP) and rapid serial visual presentation (RSVP) are useful methods in the brain-computer interface (BCI) systems. Hybrid BCI systems that combine these two approaches can enhance the proficiency of the P300 spellers.

Methods: In this study, a new hybrid RSVP/SSVEP BCI is proposed to increase the classification accuracy and information transfer rate (ITR) as compared with the other RSVP speller paradigms. In this paradigm, RSVP (eliciting a P300 response) and SSVEP stimulations are presented in such a way that the target group of characters is identified by RSVP stimuli, and the target character is recognized by SSVEP stimuli.

Results: The proposed paradigm achieved accuracy of 93.06%, and ITR of 23.41 bit/min averaged across six subjects.

Conclusions: The new hybrid system demonstrates that by using SSVEP stimulation in Triple RSVP speller paradigm, we could enhance the performance of the system as compared with the traditional Triple RSVP paradigm. Our work is the first hybrid paradigm in RSVP spellers that could obtain the higher classification accuracy and information transfer rate in comparison with the previous RSVP spellers.

1. Introduction

Brain-computer interface (BCI) systems enable users to communicate with the environment by translating brain electrical activities into specific commands. These systems can help patients with motor disorders in several real-life situations, such as controlling wheelchairs or typing characters [1–3].

Recording surface electrical activity from the scalp, namely Electroencephalography (EEG), is usually used in most BCI studies because of its high temporal resolution, low cost, and ease of recording [4]. Two of the popular applications of EEG signals in this domain are the BCI systems based on event-related potentials (ERP) [5–7] and steady-state visual evoked potential (SSVEP) [8–10].

SSVEP refers to stable oscillation in brain signals when subject focuses attention on a visual stimulus flickering at a constant frequency [11]. The effect of stimulation frequency and its harmonics is more evident in parietal and occipital regions than the other areas of the scalp/brain [12]. Due to the high SNR of SSVEP, this potential can be easily detected by using frequency analysis methods.

In common/traditional SSVEP BCI systems, the subject should pay attention to one of the stimuli that are flickering in different frequencies. The attention to the target stimulus increases the effect of the corresponding frequency in brain activity. Therefore, it is possible to determine which stimulus is given attention by the user, by identifying the main frequency of the power spectral density of the brain activity [13].

In [14,15], unlike conventional work done in the SSVEP BCIs, a single flicker stimulus was used to perform multi-class classification. The idea was using the spatial distribution of SSVEP power in scalp topography. To this end, instead of defining distinct flicker stimulus for each class, a single flicker was employed, and classes were defined as non-flickering targets around the SSVEP stimulus. When the subject looked at each of the targets at different positions surrounding the flicker stimulus, the spatial properties of SSVEP frequency changed in different regions of the head. Then, the classifier used these spatial features to determine the direction of the subject gaze in order to identify the target.

The P300 is an event-related potential that is elicited as a positive voltage approximately 300 ms after the onset of the target stimulus [16,17]. This component is evoked when a rare stimulus (target stimulus) is presented between several relevant stimuli (non-target stimuli) in an oddball paradigm. The task of the participants during the experiment is to focus on the target stimulus and count the number of its occurrence [18].

The P300 component is known by its amplitude and latency. Various factors can affect the amplitude and latency of P300 where

* Corresponding author.
E-mail address: hajipour@sharif.edu (S. Hajipour Sardouie).

https://doi.org/10.1016/j.cmpb.2020.105326
0169-2607/© 2020 Published by Elsevier B.V.

Please cite this article as: S. Jalilpour, S. Hajipour Sardouie and A. Mijani, A novel hybrid BCI speller based on RSVP and SSVEP paradigm, Computer Methods and Programs in Biomedicine, https://doi.org/10.1016/j.cmpb.2020.105326

the most important factors are the target probability occurrence and the target-to-target interval. The probability of appearance of target stimuli versus non-target stimuli in an oddball paradigm should be at least 25% [19].

One of the most widely used P300-based BCI systems is the P300 speller, in which the user can type the letters using his/her brain signals. The first P300 speller system (matrix speller) was presented by Farwell et al. [20]. In this system, the protocol was constructed by 36 characters which are embedded in a 6 × 6 matrix, and the rows and columns of the matrix blink randomly. To select the target, the subject focused attention on target character in the matrix, and the corresponding row and column of that target induced P300 component.

A lot of research has been done to improve the accuracy and speed of the spellers, which can refer to [21]. In this work, the accuracy and bit rate were improved by replacing the flashing rows and columns with the familiar face. By adding SSVEP to the P300 speller systems, hybrid BCI systems were introduced to enhance the performance of the system compared with the conventional P300 speller. SSVEP stimulation could differentiate between characters in the same row or column [22].

Despite the advantages of matrix spellers, these paradigms are gaze-dependent, and the subject should move his/her eyes during the experiment. This condition can deteriorate the classification accuracy. It was shown later that the accuracy is reduced as the size of the letters in the matrix speller becomes smaller, and these systems are space dependent [23]. Therefore, the matrix speller system is not suitable for patients suffering from visual disorders.

To solve these problems, researchers have proposed two general solutions: (1) the type of stimulation changes from the visual to auditory or sensory-tactile [24,25] and (2) using another visual protocol instead of matrix speller. Region-based (RB) speller paradigms are utilized among the second group of solutions [26–28]. The stimulation paradigm in these systems is such that the letters were grouped into different categories. Character selection is done in a two-step procedure. In the first step, the category containing the target character is selected; then the selected category is split into its constitutive letters. In the second step, these characters are presented as separated stimuli, and the target is chosen. Although these patterns almost resolved the problems, they were still not efficient.

After that, rapid serial visual presentation (RSVP) speller paradigms were presented. In these protocols, single characters are randomly displayed at the same location on the screen, so they are gaze-independent, and it is not necessary to look at different locations of the screen during the experiment [29–33]. Despite the benefits of this paradigm, it takes considerable time to perform the experiment because each character should be displayed separately, and this process should be repeated several times. Therefore, the information transfer rate of these protocols is low. To solve this problem, Multi RSVP paradigms were proposed [33]. In the triple paradigm, three different characters appears simultaneously on the screen, and the subject should look at the three characters at the same time. When the subject observes the target character, the P300 component appears for those letters. This process is repeated three times. There are two important points in the design of the Triple RSVP protocol: (1) each set of letters can appear with each other only once and (2) The position of the each character must be unique in all iterations of the trial. It means that, if a character is on the left, right, or bottom, it should be displayed in the same position in all repetitions. To detect the target character, the EEG signals are averaged over each character. Each stimulus contains one target and two non-target characters. Since non-target characters are different in the repetitions, the averaging reduces the effect of the P300 component for non-target characters and increases the amplitude of the P300 for the target character.

By using the Triple RSVP paradigm [33], the experiment time was reduced to approximately one-third of the single RSVP paradigm, and the information transfer rate was enhanced. The average ITR in the experiment was reported 19.87 and 20.25 bit/min in offline and online environments, respectively. Although they could achieve good results in their work, the great disadvantage of this protocol was the existence of the P300 component in some non-target characters, which deteriorated the classification accuracy.

In this paper, we will present a new BCI speller protocol that addresses the problems described above and improves the Triple RSVP paradigm. As mentioned in [14,15], by using a single flicker stimulus, classes are coded by their spatial position. In this paper, we intend to solve the P300 occurrence problem for non-target characters by adding an SSVEP stimulation to the Triple RSVP protocol, which we will discuss below.

2. Method and experiment

2.1. SSVEP/RSVP-hybrid stimuli

The protocol presented in this article is a combination of the Triple RSVP and SSVEP paradigms. Using the SSVEP properties, the target direction is detected and by using the Triple RSVP protocol (eliciting a P300 response), the target stimulus is recognized.

For the design of the Triple RSVP protocol, 27 alphabetical letters are divided into 9 symbol groups (each of them contains 3 letters). Each symbol group is shown as a stimulus, and each trial consists of five-time repetitions of the 9 randomized stimuli. The symbol group and position of each character are unique/same in all repetitions of the trial and just the order of the stimulation display changes in repetitions so that the stimuli are randomized and presented in an oddball paradigm. In terms of appearance, the font of letters is adjusted on Times New Roman with the size of 80 pt. The distance between every two letters is approximately 255 pixels. Also, we use a 14-inch laptop screen with a resolution of 1366 × 768 pixels.

In addition, for the SSVEP design, a flickering white square with the size of 240×240 pixels is placed in the center of the screen which is surrounded by 3 characters on the three sides of it. The flickering frequency is 15 Hz. By analyzing the spatial properties of the brain response to this flickering frequency, we can recognize which of the characters in the symbol group is attended.

In this protocol, two distinct classifications are performed. In the first classification, the target symbol group is identified between 9 classes (9 symbol groups). In the second classification procedure, by using the SSVEP stimulation, a three-class classification is performed to specify the direction of the target character.

The experiment begins with the fixation cross appearing in the center of the screen for 2 s in order to prepare the user. Then at the top of the screen, the three characters that will be spelled in a run are displayed for 2 s, and eventually, the target character of the trial is presented for 2 s (Fig. 1). After that, 45 stimuli (i.e., 5 time repetitions of the 9 randomized stimuli) are presented one after another and simultaneously a white square located in the center of the screen flickers at 15 Hz. We tried to select 9 symbol groups in such a way that similar characters (such as O and Q, or M and N) do not appear in the same location. These groups are listed in Table 1. The color of the characters and the background are black and gray, respectively.

2.2. Participants

Six healthy volunteers (male, aged 22–27, mean 24.8) participated in our experiment. All subjects provided written informed

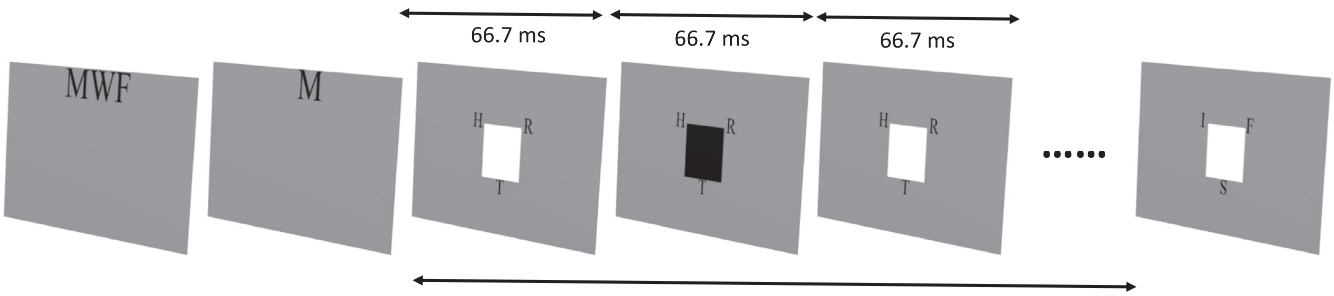

Fig. 1. The order of display in one run of the experiment: First, the three characters that will be spelled in a run are displayed for 2 s. Then, the target character of the trial is presented for 2 s. After that, 45 stimuli are presented one after another. The interval between two sequential stimuli lasted for 14 cycles of the 60 Hz refresh rate period (233 ms). Simultaneously, a white square located in the center of the screen flickers at 15 Hz (periodicity of 66.7 ms). The top arrows indicate the period of the SSVEP stimulation created by the flickering of the central square, and the bottom arrow represent one period of the RSVP stimulation.

Table 1
Symbol groups of characters corresponding to 9 stimuli.

Stimulus Number	1	2	3	4	5	6	7	8	9
Left direction	**H**	**I**	**C**	**B**	**O**	**M**	**X**	**A**	**E**
Right direction	**R**	**F**	**N**	**Q**	**J**	**U**	**K**	**L**	**G**
Bottom direction	**T**	**S**	**.**	**V**	**Z**	**Y**	**P**	**W**	**D**

consents. They do not have eye problems or history of neurological diseases, and three subjects did not have any experience with BCI systems. Iran's Medical Sciences ethics committee approved the protocol of our experiments and data recording was conducted in the National Brain Mapping Lab (NBML). Before the beginning of the experiment, the participants were instructed to minimize eye movements and seat comfortably in the chair facing the screen.

2.3. Experiment setup

Each subject performed 24 offline runs in the experiment. Each run comprised spelling of 3 characters (3 trials), and in each trial, 45 stimuli were presented for the subject. Subjects must focus on the symbol group, which includes the target character and count the number of target stimulus occurrence silently. During the experiment, the target character appears five times in the same symbol group and position in all repetitions of the trial. Since the position of characters on the screen didn't change, the subject had to gaze at the target location whenever he/she could find the target character and ignore the other two locations. Table 2 shows all details about these definitions.

There were 7 and 60 s rest time between every two trials and every two runs, respectively. These rest times were selected almost similar to previous works [36]. Each P300 stimulus lasted for 230 ms without any inter-stimulus interval (ISI), resulting in the total length of 10.5 s for selecting a character. Subjects completed one additional pre-run to make sure that the experimental procedure was really understood.

2.4. Data acquisition

This experiment was recorded using a 32-channel g.Hlamp (G.Tech Company) device with 32 active electrodes in accordance with the international 10–20 system (Fig. 2). The sampling rate of the signal was 512 Hz, and all channels were referenced to the right earlobe with a forehead ground (GND). All channels were used for P300 analysis and only 9 channels: including P7, P3, Pz, P4, P8, PO3, PO4, O1, and O2 which cover parietal and occipital regions were used in the SSVEP analysis. The stimuli were shown in a 19.5-inch monitor with a refresh rate of 60 Hz. Psychtoolbox was used for designing the stimulation protocol [34]. The processes and analyses were performed in Matlab 2017b by using the BBCI (Berlin Brain-Computer Interface) toolbox [35]. After data acquisition, EEG data were separately segmented into epochs for the P300 and SSVEP analysis with different lengths. For the P300 analysis, EEG data was segmented into 1 s epochs relative to stimulus onset. In the SSVEP analysis, the entire duration of the trial (10.5 s) was considered as an epoch. Note that during the 10.5 s of a trial, the subject looked at a fixed direction.

2.5. Feature extraction and feature reduction

As mentioned above, the target group is identified by RSVP paradigm, and the target location is determined by SSVEP stimulation. So the P300 and SSVEP analysis are done separately, and different processing methods are used to extract the feature vectors.

In the SSVEP analysis, the signal is first filtered with a 1 to 47 Hz bandpass FIR filter. Then, according to [14,15], the canonical correlation analysis (CCA) method is applied to extract the features. The CCA method is used in many SSVEP-based BCI works [37] and it determines spatial filters **A** and **B** for two signals **X** and **Y** in such a way that, the correlation between the corresponding rows of the two filtered signals (**AX** and **BY**) is maximized. For our work, **X** and **Y** are EEG and reference signals, respectively. Reference signals are considered as a sinusoidal-cosine signal of 15 Hz

Table 2
All details about the definitions: Experiment, Run, Trial, and Repetition. The description, the duration, and the number of stimuli of each experiment, run, trial, and repetition are presented in the first, second, and third rows of the table.

	Experiment	Run	Trial	Repetition
Description	24 runs	3 trials	5 repetitions	–
Duration (seconds)	2472	45.5	10.5	2.1
Number of shown stimuli	3240	135	45	9

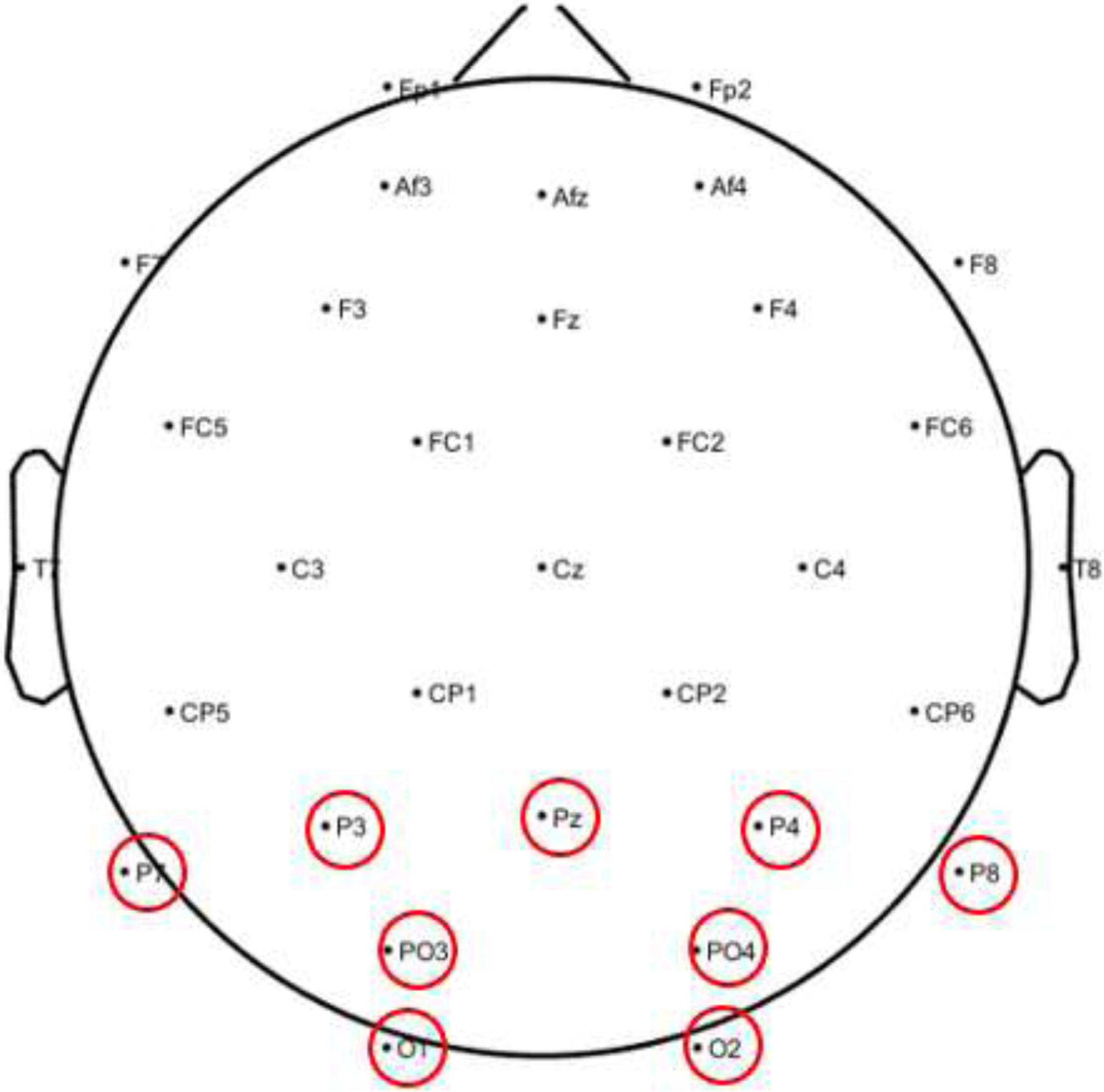

Fig. 2. Electrode placement according to the international 10–20 system. All channels are used in P300 analyzes and 9 channels (including P7, P3, Pz, P4, P8, PO3, PO4, O1, and O2) are used in SSVEP analyzes which are specified by red circles.

and its harmonics as described below:

$$\mathbf{Y} = \begin{bmatrix} \sin(2\pi ft) \\ \cos(2\pi ft) \\ \sin(4\pi ft) \\ \cos(4\pi ft) \\ \sin(6\pi ft) \\ \cos(6\pi ft) \end{bmatrix} \tag{1}$$

where f is the stimulation frequency of the SSVEP paradigm (i.e., 15 Hz here) and t is the data point. As described above, there are 3 classes (3 directions) for classification in the SSVEP paradigm. The matrices $\mathbf{A}_c$and $\mathbf{B}_c$ will be obtained to maximize the correlation between the corresponding rows of the two $\mathbf{A}_c\mathbf{X}$ and $\mathbf{B}_c\mathbf{Y}$matrices for each of the three locations (or classes) c = 1,2 and 3. In the training phase, for constructing spatial matrixes $\mathbf{A}_c$and $\mathbf{B}_c$, c = 1, 2, 3, the CCA method is applied on the concatenated trials in which the subject gazed to the (direction) class c. Then, for each training trial, the M highest correlation coefficients are calculated between each filtered trial ($\mathbf{A}_c\mathbf{X}$) and the reference signal ($\mathbf{B}_c\mathbf{Y}$) for all c filters. These correlation coefficients are then concatenated to construct the feature vector. Finally, a feature vector with a length of $C \times M$ is achieved for each training trial so that C is the number of classes ($C = 3$) and $M$is the minimum of the number of channels and the number of the sine and cosine harmonics ($M = 6$). Using the obtained features, we train a multi-class (3-class) classifier.

By looking at each direction, the SSVEP power distribution varies in the individual's scalp. By applying spatial filters ($\mathbf{A}_c$and$\mathbf{B}_c$,$c = 1, 2, 3$) on the brain signal in the test phase, the feature vector is obtained similar as for the training data. The output of the classifier determines that the subject gazed to which direction.

In the P300 analysis, the EEG data are filtered using a 0.5–25 Hz bandpass FIR filter. Then, the data are segmented into 1000 ms epochs which are started at the onset of each stimulus. Baseline correction is performed using the 150 ms pre-stimulus interval. We use two types of features, namely the amplitude of the time samples and the wavelet coefficients, to classify these P300/non-P300 epochs.

More specifically, to extract the amplitude of the time samples, each epoch is decimated with a factor of 10. Therefore, the length of each epoch is reduced to 51 data points. Then, by concatenating the features of all channels, 1632 features are achieved for each trial.

On the other hand, wavelet has been used as a powerful feature extraction method in the BCI studies and especially for classifying the P300 component [38,39]. To take advantage of this type of features, we use the discrete wavelet transform (DWT) with Daubechies mother wavelet which has been widely used in many EEG researches [40,41]. The wavelet decomposition of the EEG signal is performed at 7 levels. Since the sampling rate is 512 Hz, we have the coefficients of the four frequency bands of the EEG (delta, theta, alpha, and beta). As a result, 1792 wavelet features are obtained. To apply this decomposition on our data, we use the Matlab wavelet toolbox [42].

Finally, using both time samples and wavelet coefficients, a total of 3224 features are obtained for the P300 analysis. Since this number of features is large compared to the number of training trials, we use least absolute shrinkage and selection operator (Lasso) algorithm to reduce the number of features to avoid overfitting and eliminate redundant information.

2.6. Classification

To evaluate the proposed protocol, the data set was sub-divided into three partitions using the three-fold cross-validation. In each repetition of the cross-validation, a single fold (partition) was used for validation and two folds were used as the training data. The accuracy was achieved by averaging the results of the three folds. Each fold of the cross-validation contained 8 runs; therefore, for each fold, 24 characters were reserved as the validation data.

Two different classifiers were used for the P300 and SSVEP analysis. The regularized linear discriminant analysis (RLDA) is one of the best classifiers which is applied for the P300 detection. This classifier is one type of the Fisher linear discriminant analysis (FLDA) that can solve the problem of overfitting by regularizing the features [43]. So we used RLDA to perform the classification for the P300 analysis. Also, multi-class kernel support vector machine (SVM) is the well-known classifier for pattern recognition, especially in the SSVEP based BCI researches in which we should do multi-class detection [44]. This classifier maps the features into a new space in which the transformed features are more distinct than to each other. Radial basis function (RBF) kernel is one of the popular nonlinear kernels to map the features. In this article, RBF-SVM classification was performed in three classes to recognize the correct directions in the SSVEP mode. This was done using the Libsvm toolbox [45].

2.7. Evaluation

In addition to the usual criterion, namely the classification accuracy, the performance of the proposed system was evaluated by the information transfer rate (ITR). The ITR is an important criterion to determine the system performance in BCI studies, especially P300 speller protocols. It shows the amount of information that can be transferred in a minute [46] and is defined as:

$$ITR = \left\{ \log_2 N + P\log_2 P + (1-P)\log_2 \frac{1-P}{N-1} \right\} / T \quad (2)$$

Table 3
The achieved accuracy and ITR for each subject.

	Acc (%)	ITR (bits/min)
Sub1	90.28	21.99
Sub2	91.67	22.63
Sub3	100	27.25
Sub4	94.44	23.98
Sub5	88.89	21.37
Sub6	93.06	23.29
Avg	93.06	23.41

where, N denotes the number of classes, P is the accuracy of classification and T indicates the time interval that character selection is performed in minute.

3. Results

3.1. The SSVEP and P300 ERP analysis

For the SSVEP analysis, we first show the power topography of the EEG signal for each target direction, averaged on all subjects. The power was obtained in the range of 14.5 to 15.5 Hz over occipital and parietal areas where the maximum power of SSVEP can be seen. As shown in Fig. 3, the effect of the SSVEP frequency in the left occipital area is more evident than the other regions for all three directions. When subjects gazed at the left-hand corner, the flickering square is placed in their right visual field so that the maximum power can be observed over the left hemisphere. On the other hand, when they attend to the right, the flickering square is placed in their left visual field, and the activity of the left areas is reduced and slightly shifted to the right. When the subjects change the gaze to bottom, the flickering square is placed in their upper visual field, and the maximal SSVEP power propagates towards the centro-parietal region. Furthermore, by comparing our achieved results with those of [15], we can see the same changes in power, when subject gaze to different directions.

The grand average ERPs with the corresponding topographies for two-time intervals are plotted over the Cz and Pz channels in Fig. 4. In ERP plots, the red and blue colors represent the grand average signals for the target and non-target trials, respectively.

In addition, the scalp maps corresponding to the occurrence of the P300 component are shown in Fig. 4. As shown in these maps, the center and parietal regions show the most considerable difference between the two classes. This component occurs between 400 and 530 ms and reaches to its maximum value (8.5 microvolts) in the central region, especially in Cz channel.

3.2. Cross-validation accuracy

The character detection accuracy and ITR of cross-validation for all subjects are given in Table 3. In our system, the average accuracy and ITR of 93.06 and 23.41 bits/min were obtained, respectively. One of the six subjects could reach the accuracy of 100% with ITR of 27.25 bits/min, and 5 subjects could achieve accuracy above 90% using all 5 repetitions. In addition, all of the subjects have ITRs higher than 21 bits/min. With regard to the fact that ITR is a function of both the experiment time and the accuracy, we can claim that with the same experiment time as reported in [33], we could enhance the accuracy significantly. To illustrate the possibility of using our hybrid system, the accuracy and ITR obtained for each subject in terms of the number of repetitions are shown in Fig. 5 and Fig. 6, respectively. Obviously,

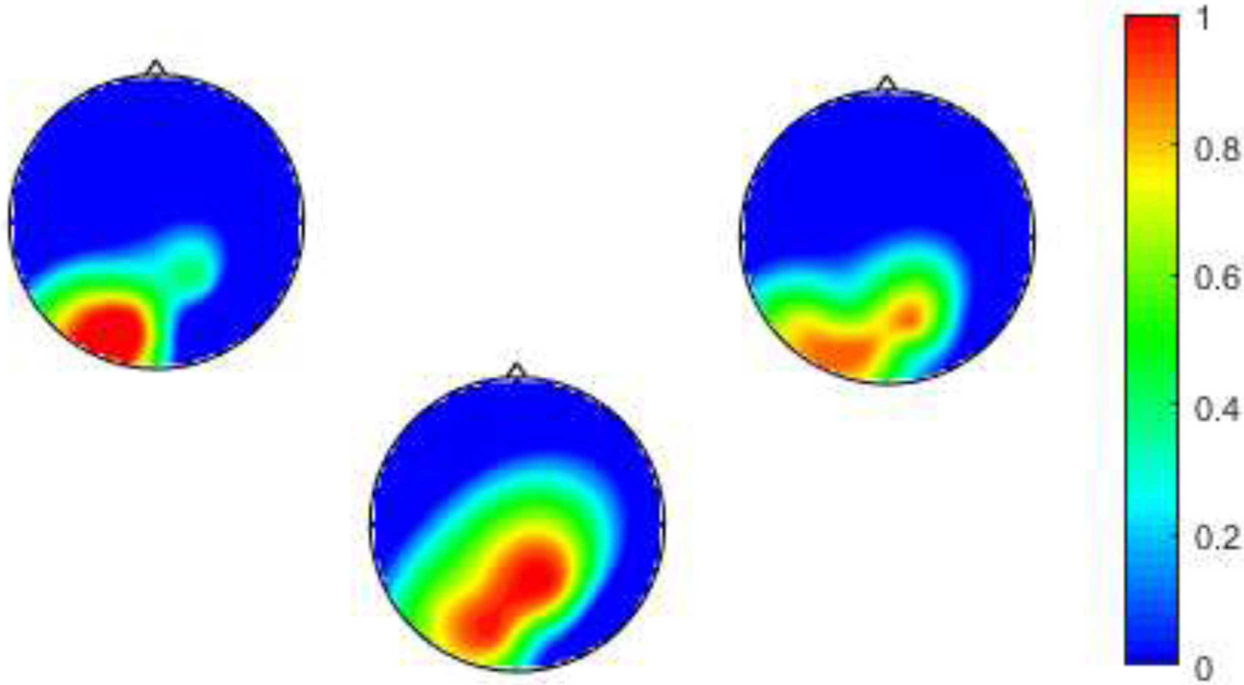

Fig. 3. The averaged power topographies for three classes. (Left, right and bottom topographies indicate three locations of the target stimuli). The maximum power is shifted from left to right occipital areas when the subjects gaze from left to right direction. Furthermore, when the subjects gaze to bottom direction, maximum power can be observed in centro-parietal areas.

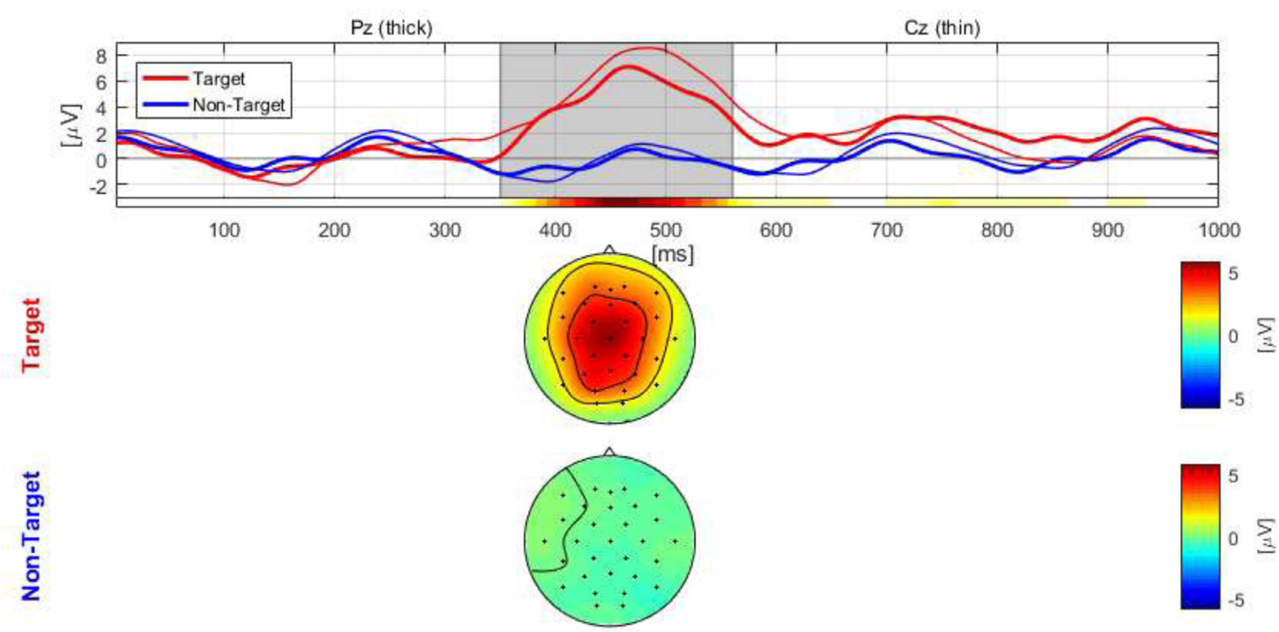

Fig. 4. The grand average signals over all subjects at Pz (thick lines) and Cz (thin lines) channels. Red and blue colors represent the target and non-target trials, respectively. The shadowed area shows specific interval that the P300 component occurs. Also, scalp topographies of the P300 component are plotted for the target and non-target stimuli.

the accuracy increases as more iterations of stimuli are used in a trial.

Although the classification accuracy increases by increasing the number of repetitions, but this causes the system to execute a command in a longer time. The averaged accuracy and ITR for the different number of repetitions are shown in Table 4. The maximum ITR (34.38 bits/min) is achieved when just two repetitions of stimuli were used in a trial; in this condition, the averaged accuracy is 68.98%.

Note that the chance level is 3.7% since classification is done between 27 classes. By performing one repetition of a trial, a higher accuracy (44.21%) than the chance level can be obtained.

Furthermore, we investigated the impact of the RSVP and SSVEP classification accuracy on the overall accuracy. In this way, the ac-

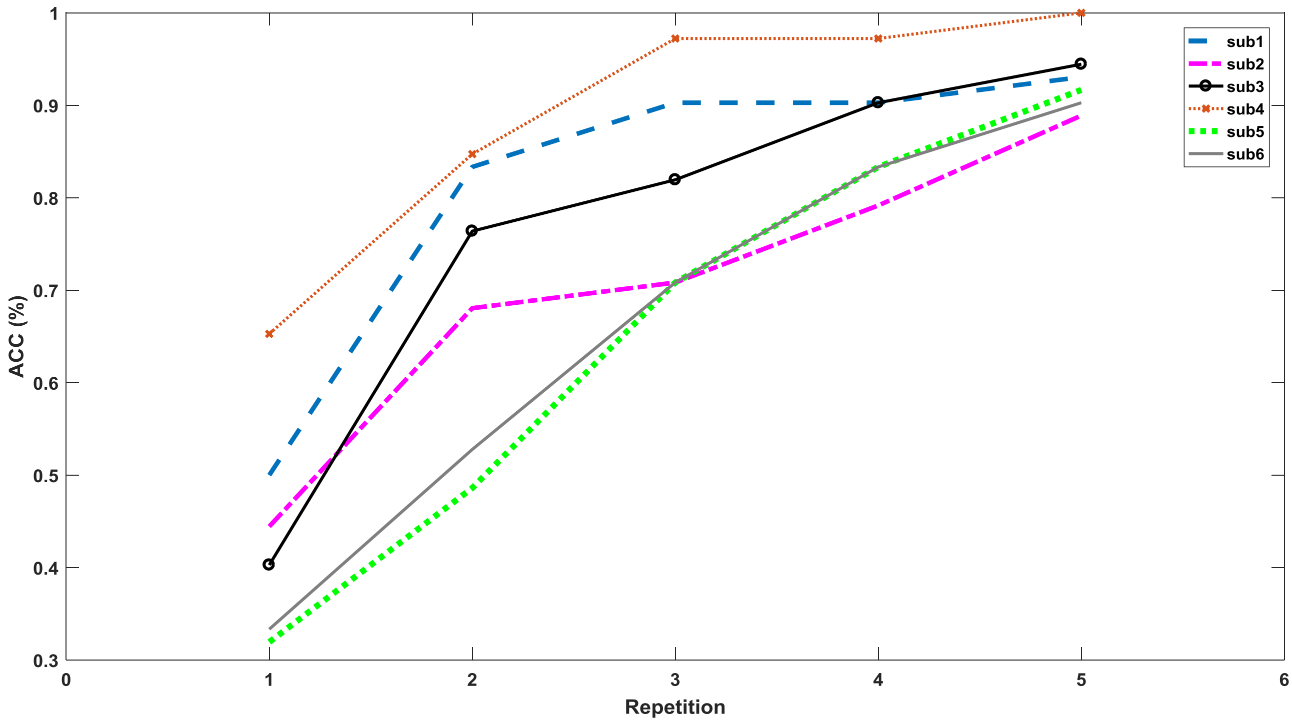

Fig. 5. The accuracy (%) in terms of the number of repetitions for each subject.

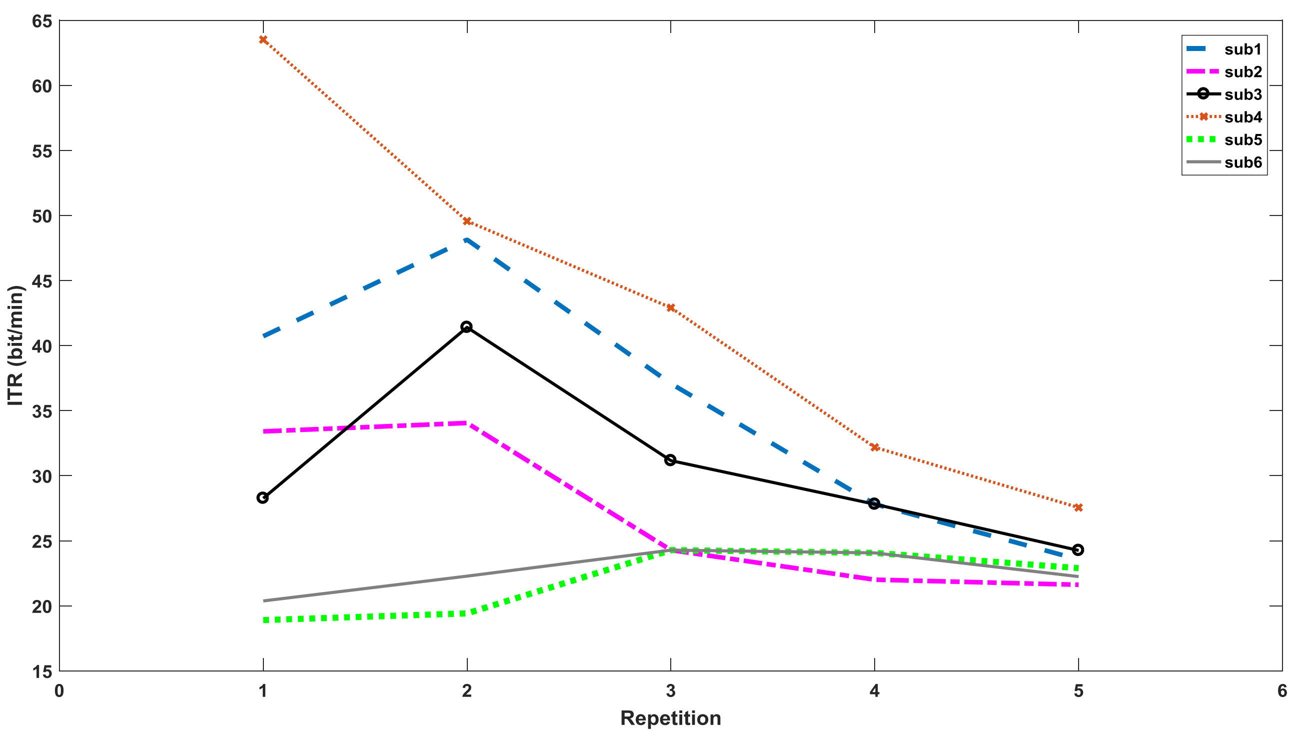

Fig. 6. The ITR (bits/min) in terms of the number of repetitions for each subject.

curacy of the target group detection was examined just by using the P300 component. We also studied the accuracy of the target location detection by exploiting the SSVEP. The averaged accuracies for the different number of repetitions are shown separately for RSVP and SSVEP in Table 5.

As can be seen, both RSVP and SSVEP accuracies are higher than overall accuracies in all repetitions of the experiment. It can be reasoned that the overall accuracy is the intersection of RSVP and SSVEP accuracies. It means that for character selection, it is essential that both the group and location are identified correctly. It is

Table 4
The averaged accuracy (%) and ITR (bits/min) in terms of the number of repetitions.

Number of repetitions	Acc (%)	ITR (bits/min)
1	44.21	32.68
2	68.98	34.38
3	80.32	29.71
4	87.27	25.80
5	93.06	23.41

Table 5
The averaged RSVP and SSEVP accuracy (%) in terms of the number of repetitions.

Number of repetitions	RSVP Acc. (%)	SSVEP Acc. (%)
1	59.95	47.88
2	77.08	73.61
3	86.11	87.5
4	90.51	93.06
5	96.29	96.75

Table 6
the averaged accuracy and ITR of all protocols. For our proposed protocol, the same accuracy as the single RSVP protocol [32] is obtained. Additionally, we achieve the higher ITR as compared with triple RSVP protocol [33].

	Accuracy (%)	ITR (bit/min)
Single RSVP [32]	93.6	7
Triple RSVP [33]	78	19.876
Proposed Protocol	93.06	21.41

worth noting that chance levels are 33.3% and 11.1% for SSVEP and RSVP paradigms, respectively.

4. Conclusion and discussion

In this paper, we proposed a new protocol that can select the character more accurately and quickly as compared with the other RSVP speller paradigms. By studying the history of the spellers, we found that most of the work done in this field was presented in a matrix structure [20,21]. In these protocols, the subject should look at different points on the screen during the experiment. Therefore, patients with visual impairment confront problem using this type of spellers. RSVP systems [29,32] were proposed to solve this problem, but the defect of these protocols was the long duration of the experiment time that deteriorate the ITR. After these systems, Triple RSVPs were introduced in [33] that were able to reduce the experiment time to a large extent. In these paradigms, symbol groups (containing three characters) were shown as stimuli instead of displaying a single character. But on the other hand, the occurrence of the P300 component in non-target characters could affect the accuracy.

As presented in another group of studies, the use of hybrid RSVP-SSVEP systems can increase the accuracy and speed in these systems [22]. It has recently been shown in [14] that by using a single-flicker stimulus, the user can perform multi-class selections based on different angle views. In these systems, non-flicker targets are placed around the SSVEP stimulation with a fixed flickering frequency. Therefore, there is no need to use different SSVEP stimulation for each command. In work presented in [14], it was shown that with a single flicker stimulus, up to 9 classes could be detected. These 9 classes are 9 different directions that the user should gaze to them.

By combining ideas from the two groups of studies mentioned above, we proposed our hybrid protocol by adding the SSVEP stimulation to the triple RSVP protocol. In the proposed paradigm, each symbol group (containing three characters) is considered as a single stimulus. In all iterations of the trial, the group and position of the characters do not change. In this case, the classification is done for 9 symbol groups instead of 27 characters. To select the target character from the characters in a group, we use SSVEP stimulation. To this end, SSVEP stimulation is considered as a flashing square in the center of the screen, and three characters of a group are placed in the three different locations of the square. When the subject looks at each of these three targets, different spatial maps related to the SSVEP frequency are generated in his/her head. These spatial features are used to determine the target character.

By comparing the proposed protocol with the Triple RSVP one [33], the following points are noteworthy. For the triple RSVP paradigm [33], in each repetition of the trial containing the target character, two non-target characters also appear with the target character. By considering three repetitions in this experiment, there are three P300 components for the target character and one P300 component for the six non-target characters. But in our proposed paradigm, symbol groups appear together in all repetitions. As explained above, in the P300 analysis, we should only identify the target group. Therefore, the target group is identified between 9 classes (9 groups of characters). In this case, with any number of iterations, the P300 component appears only for the target group, and we do not find a P300 component for the non-target groups.

The functionality of the BCI systems, especially speller systems, is evaluated by two criteria: accuracy and information transfer rate. There is always a trade-off between these two criteria. In other words, by increasing the time of trial, the accuracy increases and the ITR drops, and vice versa.

About the previously proposed spellers, we can say that the Triple RSVP paradigm [33] improved the ITR significantly in the RSVP-based speller systems, but this protocol has significantly decreased the accuracy of the character recognition. On the other hand, in the protocol that we call "single RSVP" and was provided by [32], the character recognition accuracy was very high, but the ITR was low. The accuracies and ITRs of articles [32,33] are compared with those of our proposed protocol in Table 6. We reach the same accuracy as the single RSVP [32] while the ITR is enhanced 23.41 bits/min in our protocol. Also, the achieved ITR in our protocol is approximately equal to the reported ITR in [33], while the accuracy increases by 15% as compared with triple RSVP.

By combining the two stimulation protocol of the RSVP and the SSVEP, we not only improved the accuracy achieved in the Triple RSVP protocol, but also increased the ITR of the character recognition in the single RSVP paradigm.

References

[1] J. Wolpaw, EW. Wolpaw, Brain-computer interfaces: Principles and Practice, OUP USA, 2012.
[2] J.R. Wolpaw, N Birbaumer, D.J. McFarland, G Pfurtscheller, T.M. Vaughan, "Brain–computer interfaces for communication and control, Clinical Neurophysiol. 113 (6) (2002) 767–791.
[3] Li Yuanqing, J Pan, F Wang, Z Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control, IEEE Trans. Biomed. Eng. 60 (11) (2013) 3156–3166.
[4] M Teplan, "Fundamentals of EEG measurement, Measurement Sci. Rev. 2 (2) (2002) 1–11.
[5] J Jin, BZ. Allison, X Wang, C Neuper, "A combined brain–computer interface based on P300 potentials and motion-onset visual evoked potentials, J. Neurosci. Methods 205 (2) (2012) 265–276.
[6] J Jin, BZ. Allison, Yu Zhang, X Wang, A Cichocki, "An ERP-based BCI using an oddball paradigm with different faces and reduced errors in critical functions, Int. J. Neural. Syst. 24 (08) (2014) 1450027.
[7] Yu Zhang, Q Zhao, J Jin, X Wang, A Cichocki, "A novel BCI based on ERP components sensitive to configural processing of human faces, J. Neural Eng. 9 (2) (2012) 026018.

[8] R Zerafa, T Camilleri, O Falzon, KP Camilleri, To train or not to train? a survey on training of feature extraction methods for SSVEP-based BCIs, J. Neural Eng. 15 (5) (2018) 051001.
[9] H Lim, J Ku, "Multiple-command single-frequency SSVEP-based BCI system using flickering action video, J. Neurosci. Methods 314 (2019) 21–27.
[10] M Wang, I Daly, Z. Brendan, JJ Allison, Yu Zhang, Lanlan Chen, X Wang, A new hybrid BCI paradigm based on P300 and SSVEP, J. Neurosci. Methods 244 (2015) 16–25.
[11] N.R. Galloway, Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine, Br. J. Ophthalmol. 74 (4) (1990) 255.
[12] GR. Müller-Putz, R Scherer, C Brauneis, G Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components, J. Neural Eng. 2 (4) (2005) 123.
[13] C-H Wu, H-C Chang, Po-L Lee, K-S Li, J-J Sie, C-W Sun, C-Y Yang, Po-H Li, H-T Deng, K-K Shyu, "Frequency recognition in an SSVEP-based brain computer interface using empirical mode decomposition and refined generalized zero-crossing, J. Neurosci. Methods 196 (1) (2011) 170–181.
[14] A Maye, D Zhang, AK Engel, Utilizing retinotopic mapping for a multi-target SSVEP BCI with a single flicker frequency, IEEE Trans. Neural Syst. Rehabilitat. Eng. 25 (7) (2017) 1026–1036.
[15] J Chen, D Zhang, K. Andreas, QG Engel, A Maye, Application of a single-flicker online SSVEP BCI for spatial navigation, PLoS ONE 12 (5) (2017) e0178385.
[16] TW. Picton, The P300 wave of the human event-related potential, J. Clinical Neurophysiol. 9 (4) (1992) 456–479.
[17] J Polich, in: "Neuropsychology of P300, 159, Oxford Handbook of Event-Related Potential Components, 2012, p. 88.
[18] S Sur, V.K. Sinha, Event-related potential: an overview, Ind. Psychiatry J. 18 (1) (2009) 70.
[19] CJ. Gonsalvez, J Polich, "P300 amplitude is determined by target-to-target interval, Psychophysiology 39 (3) (2002) 388–396.
[20] LA Farwell, E Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials, Electroencephalogr. Clin. Neurophysiol. 70 (6) (1988) 510–523.
[21] T Kaufmann, S.M. Schulz, C Grünzinger, A Kübler, "Flashing characters with famous faces improves ERP-based brain–computer interface performance, J. Neural Eng. 8 (5) (2011) 056016.
[22] E Yin, Z Zhou, J Jiang, F Chen, Y Liu, D Hu, "A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm, J. Neural Eng. 10 (2) (2013) 026012.
[23] M Salvaris, F Sepulveda, "Visual modifications on the P300 speller BCI paradigm, J. Neural Eng. 6 (4) (2009) 046011.
[24] I Käthner, CA. Ruf, E Pasqualotto, C Braun, N Birbaumer, S Halder, "A portable auditory P300 brain–computer interface with directional cues, Clinical Neurophysiol. 124 (2) (2013) 327–338.
[25] R Ortner, Z Lugo, R Prückl, C Hintermüller, Q Noirhomme, C Guger, "Performance of a tactile P300 speller for healthy people and severely disabled patients, in: Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2013, pp. 2259–2262.
[26] J Pan, Y Li, Z Gu, Z Yu, "A comparison study of two P300 speller paradigms for brain–computer interface, Cogn. Neurodyn. 7 (6) (2013) 523–529.
[27] MH Chang, JSu Lee, J Heo, KS Park, "Eliciting dual-frequency SSVEP using a hybrid SSVEP-P300 BCI, J. Neurosci. Methods 258 (2016) 104–113.
[28] MS. Treder, B Blankertz, "(C) overt attention and visual speller design in an ERP-based brain-computer interface, Behav. Brain Funct. 6 (1) (2010) 28.
[29] U Orhan, KE. Hild, D Erdogmus, B Roark, B Oken, M Fried-Oken, "RSVP keyboard: an EEG based typing interface, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2012, pp. 645–648.
[30] L Acqualagna, MS Treder, M Schreuder, B Blankertz, "A novel brain-computer interface based on the rapid serial visual presentation paradigm, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology, IEEE, 2010, pp. 2686–2689.
[31] L Acqualagna, B Blankertz, "A gaze independent spelling based on rapid serial visual presentation, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2011, pp. 4560–4563.
[32] L Acqualagna, B Blankertz, "Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP), Clinical Neurophysiol. 124 (5) (2013) 901–908.
[33] Z Lin, C Zhang, Y Zeng, Li Tong, B Yan, "A novel P300 BCI speller based on the triple RSVP paradigm, Sci. Rep. 8 (1) (2018) 3350.
[34] M Kleiner, D Brainard, D Pelli, A Ingling, R Murray, C Broussard, "What's new in psychtoolbox-3, Perception 36 (14) (2007) 1.
[35] B Blankertz, L Acqualagna, S Dähne, S Haufe, M Schultze-Kraft, I Sturm, M Uščumlic, A. Markus, GC Wenzel, M Klaus-Robert, The berlin brain-computer interface: progress beyond communication and control, Front. Neurosci. 10 (2016) 530.
[36] N. Kaongoen, S. Jo, A novel hybrid auditory bci paradigm combining assr and P300, J. Neurosci. Methods 279 (2017) 44–51.
[37] Y.U. Zhang, G Zhou, J Jin, X Wang, A Cichocki, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis, Int. J. Neural Syst. 24 (04) (2014) 1450013.
[38] B Perseh, AR Sharafat, An efficient P300-based bci using wavelet features and IBPSO-based channel selection, J. Med. Signals Sens. 2 (3) (2012) 128.
[39] M Salvaris, F Sepulveda, "Wavelets and ensemble of FLDs for P300 classification, in: Proceedings of the 4th International IEEE/EMBS Conference on Neural Engineering, IEEE, 2009, pp. 339–342.
[40] SA Markazi, S. Qazi, LS. Stergioulas, A Ramchurn, D Bunce, "Wavelet filtering of the P300 component in event-related potentials, in: Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2006, pp. 1719–1722.
[41] SA Markazi, K. Lampros, AR Stergioulas, B David, Latency corrected wavelet filtering of the P300 event-related potential in young and old adults, in: Proceedings of the 3rd International IEEE/EMBS Conference on Neural Engineering, IEEE, 2007, pp. 582–586.
[42] “Matlab Wavelet toolbox, ” Jul. 2008. [Online]. Available: http://www.mathworks.com
[43] J Höhne, M Schreuder, B Blankertz, M Tangermann, "A novel 9-class auditory erp paradigm driving a predictive text entry system, Front. Neurosci. 5 (2011) 99.
[44] R Singla, A Khosla, R Jha, "Influence of stimuli color on steady-state visual evoked potentials based BCI wheelchair control, J. Biomed. Sci. Eng. 6 (11) (2013) 1050.
[45] Chang, C-C. "" LIBSVM: a library for support vector machines," ACM transactions on intelligent systems and technology, 2: 27: 1–27: 27, 2011." http://www. csie. ntu. edu. tw/~ cjlin/libsvm 2 (2011).
[46] JR. Wolpaw, N Birbaumer, DJ. McFarland, G Pfurtscheller, TM. Vaughan, "Brain–computer interfaces for communication and control, Clinical Neurophysiol. 113 (6) (2002) 767–791.

Please cite this article as: S. Jalilpour, S. Hajipour Sardouie and A. Mijani, A novel hybrid BCI speller based on RSVP and SSVEP paradigm, Computer Methods and Programs in Biomedicine, https://doi.org/10.1016/j.cmpb.2020.105326