SCIENTIFIC REPORTS

OPEN

A novel P300 BCI speller based on the Triple RSVP paradigm

Zhimin Lin[1], Chi Zhang [1], Ying Zeng[2,1], Li Tong[1] & Bin Yan[1]

A brain–computer interface (BCI) is an advanced human–machine interaction technology. The BCI speller is a typical application that detects the stimulated source-induced EEG signal to identify the expected characters of the subjects. The current mainstream matrix-based BCI speller involves two problems that remain unsolved, namely, gaze-dependent and space-dependent problems. Some scholars have designed gaze-independent and space-independent spelling systems. However, this system still cannot achieve a satisfactory information transfer rate (ITR). In this paper, we propose a novel triple RSVP speller with gaze-independent and space-independent characteristics and higher ITR. The triple RSVP speller uses rapid serial visual presentation (RSVP) paradigm, each time presents three different characters, and each character is presented three times to increase the ITR. The results of the experiments show the triple RSVP speller online average accuracy of 0.790 and average online ITR of 20.259 bit/min, where the system spelled at a speed of 10 s per character, and the stimulus presentation interface is a 90×195 pixel rectangle. Thus, the triple RSVP speller can be integrated into mobile smart devices (such as smartphones, smart watches, and others).

A brain-computer interface (BCI) system based on electroencephalography (EEG) is a popular research direction in the field of human-computer interaction. The system can provide a direct communication channel to connect the human brain and computer. An EEG speller system is a typical brain-computer interface system. The speller system uses a clever paradigm to induce specific event-related potential (ERP) components (for example, P300 component). Then, according to the ERP components, a symbol of the expected subjects can be determined. The EEG speller system can re-establish the disabled communication, and normal people can use the technology to obtain convenient interactive methods[1–6]. At present, most EEG speller systems are based on modified P300 speller. The P300 component is a common ERP component, which shows a peak when small probability events are observed after approximately 300–500 ms[7–10]. The P300 component also exhibits significant waveform characteristics in the time domain[11]. A P300 detection algorithm is essential because it determines the accuracy and reliability of BCI systems. Thus, some scholars have attempted and proposed many P300 detection algorithms such as independent component analysis (ICA)[12], common spatial pattern (CSP)[13], xDawn[14,15], hierarchical discriminant component analysis (HDCA)[16–19], sliding HDCA (sHDCA)[20,21], and convolutional neural network (CNN)[22]. On this basis, many BCI system performances were improved significantly.

Farwell and Donchin first proposed P300 alphabet speller system (FD speller)[23]. The system is a 6×6 matrix, and each element is a specified character. The system has a total of 36 characters (26 letters and other control characters) and uses a stepwise linear discriminant analysis (SWLDA) algorithm to detect P300 components. Krusienski et al.[24] compared the performances of various P300 detection algorithms and concluded that SWLDA and Fisher's linear discriminant (FDA) are suitable for the P300 Speller system. In the FD speller, the matrix rows or columns blink randomly, and when the locked symbol is hit by rows or columns (probability is 1/6), it induces a P300 component. The FD speller determines the expected symbol position using the detected P300 component. Cuntai Guan et al.[25] believed that if the probability of a character being hit is small, then the P300 component induced is evident and detection is easier. Thus, Guan et al. proposed a single-character display random flashing speller system (SC speller) based on the FD speller (the probability that the character was hit was 1/36). Furthermore, Townsend et al.[26] proposed a checkerboard paradigm speller system, that is, a random flash hits multiple characters, and then a special coding is employed to confirm the anticipant character of the subject. This method can achieve higher accuracy and mean bit rate. Brian Roark et al.[27,28] proposed Huffman scanning, which uses Huffman coding to select the symbols to highlight, based on the FD speller. Qi LI et al.[29] proposed

[1]China National Digital Switching System Engineering and Technological Research Center, Zhengzhou, China. [2]Key Laboratory for NeuroInformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China. Correspondence and requests for materials should be addressed to B.Y. (email: ybspace@hotmail.com)

1

a familiar face (FF) spelling paradigm, which introduces a familiar face to improve recognition accuracy based on the FD speller. They further proposed a green familiar face (GFF) paradigm[30]. Dewen Hu et al.[31] proposed a hybrid BCI speller that fuses P300 with steady-state visual evoked potential (SSVEP) and introduced an SSVEP feature on P300-row and P300-column feature. Chen et al.[32,33] used the SSVEP feature and introduced the phase feature to build a high-speed speller to achieve the highest information transfer rate (approximately one character per second). These speller systems are matrix-based spellers.

In the last two decades, research on the EEG speller has shown exciting achievements, but many problems remain unsolved. Brunner et al.[34] showed that FD speller-based recognition accuracy is more dependent on the gaze of the subject; it requires the user to not move his eyes, and then often cannot achieve the desired accuracy. Therefore, the matrix-based speller has a gaze-dependent feature. This condition limits the use of the matrix-based speller in some patients with severe neuromuscular disability. Brendan et al.[35] showed that the larger the matrix, the better the performance. Salvaris et al.[36] demonstrated that when the character is small, the performance is worse. Therefore, the matrix-based speller has a space-dependent feature. The matrix-based speller cannot easily produce sufficiently small results that people between the mild disabled and normal cannot easily spread.

To solve these problems, many scholars have proposed non-matrix structure speller paradigm. Gabriel et al.[37] proposed a GIBS block speller that effectively overcomes the problem of the matrix-based speller. Treder et al.[38] proposed an ERP Hex-o-Spell paradigm, a two-level speller consisting of six discs arranged on an invisible hexagon. The method can achieve higher accuracy than the FD speller when eye movements are not permitted. Fabio Aloise et al.[39] proposed a Geometric Speller to optimize the design of Hex-o-Spell. Treder et al.[40] compared the Hex-o-Spell, Cake Speller, and Center Speller. The three different variants of a two-stage visual speller are based on covert spatial attention and non-spatial feature attention. The results show that they can achieve higher accuracy and independent of eye gaze. However, this method is still not sufficiently convenient. Orhan et al.[41–43] proposed RSVP keyboard, using rapid serial visual presentation (RSVP) paradigm for spelling. In the RSVP paradigm, each candidate letter is shown at the same place on the screen in a temporally ordered sequence at a comfortably high presentation rate. Chennu et al.[44], Moghadamfalahi et al.[45], and Acqualagna et al.[46,47] showed that the EEG classifiability with the RSVP speller was as good as that with the FD speller, and the RSVP speller has gaze-independent and space-independent characteristics. Therefore, the RSVP speller can be used by persons with sight disabilities, and the design of the interface can be extremely small. The RSVP speller has a tempting potential to be integrated into small intelligent devices (such as smartphones, augmented reality devices, and virtual reality devices). However, relative to most matrix-based spellers, the RSVP Speller information transfer rate (ITR) is not satisfactory. Literature[44,45] indicates that the RSVP speller (which contains 28 characters) has only one-third of the ITR of the Matrix speller (which contains 36 characters). Therefore, improving the ITR of the RSVP speller system under the condition of ensuring RSVP speller recognition accuracy is an extremely meaningful subject of study.

The typical RSVP speller presents a scene (a character) at a time. If we present multiple characters at a time, then we can effectively shorten the time to display all characters, and then enhance the ITR. We propose a triple-RSVP speller. The system uses the RSVP paradigm and shows three different characters, with each character appearing three times. We asked the subjects to observe the three characters at the same time. Finally, we group average the corresponding EEG signal for each character, and determine the P300 component and expected characters of the subject.

Results

Visual stimuli and procedure. The triple-RSVP speller presentation interface is shown in Fig. 1(A). The three symbol presentation areas are rectangles with size of 60.00×65.00 pixels, parallel arranged, and the middle symbol presentation area is offset down by 30 pixels to ensure that subjects can view three characters at the same time, in a short time. We use a 17.3-inch mobile workstation screen with a screen resolution of 1920×1080 pixels. Every 250 ms (4Hz), the three symbols in presentation areas refresh once at the same time. The new symbol is presented in three presentation areas. The symbol group presentation order is determined by the symbol group sequence, such as block 1, block 2, and block 3. Each block contains 36 characters, each three characters make up a symbol group, and each symbol group appears in three symbol presentation areas. Thus, each block contains 12 symbol groups (36 characters make up 12 symbol groups and each symbol group contains 3 characters). The three blocks are different, and the symbol group and positions of the different blocks are specially designed to ensure that the position of each character appears unique in different blocks. The subjects must watch three blocks to identify an expected character. During this period, the expected character appears three times in different symbol groups of different blocks. We average the EEG to correspond to the each character, determine the P300 component, and identify the expected symbol, as shown in Fig. 2.

As shown in Fig. 2, the subject looks at the symbol group sequence that consists of three blocks, each block consists of 12 symbol groups (12 trials), each trial shows three characters, and each character appears three times. For example, when the target character is "B," the subjects notice three symbol groups: "BDF" in block 1, "BEH" in block 2, and "BNZ" in block 3. Each symbol group has a corresponding EEG signal; thus, the character of the same symbol group corresponds to the same EEG signal. Therefore, we designed unique symbol groups in three blocks. When we group average the EEG signals of a character, the P300 component of the target character can be easily detected. For example, in Fig. 2, we group average the EEG signal of symbol group "BDF", "BEH", and "BNZ" as the group average signal of character "B". Similarly, the group average signal of non-target character is composed of the EEG of all symbol groups that contain the non-target character, and the target adjoint character is also true. The symbol groups of the target adjoint character contain the target character; thus, the group average EEG of the target adjoint character contains some of the P300 component. For example, the group average EEG of character "D" consists of "BDF," "ADG", and "DP2", where "BDF" contains the target character B. Thus, a trial

A.

B.
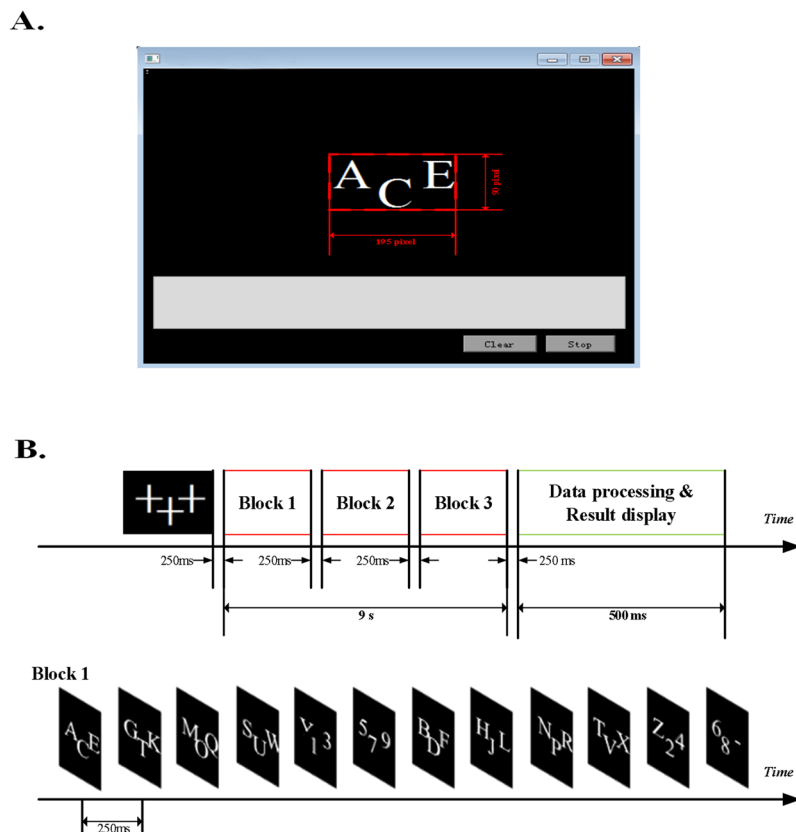
Figure 1. Stimulus presentation interface of triple RSVP speller.
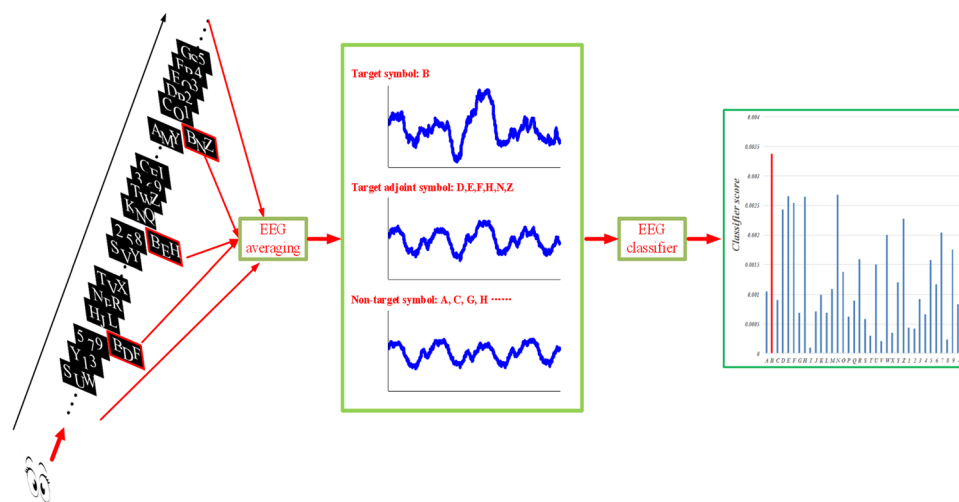
Figure 2. Triple RSVP speller system overview. The subject looks at the symbol group sequence consisting of three blocks, and focus on the specific character, for example "B". In the triple RSVP speller, each trial shows three characters, and each character appears three times. The EEG signal that corresponds to each character is group averaged, and the score is calculated by a trained P300 classifier. Finally, the highest score of all symbols is identified as the target character for the subjects.

P300 component is introduced during the mean processing of the letter "D" to allow the classifier score of the target adjoint character to be closer to the target character score. The design details of the symbol group sequence are in the Methods section.

Event-related responses. We calculate the mean of the signal corresponding to the target character, target adjoint character, and non-target character, as shown in Fig. 3(A,B and C), respectively. An evident peak is
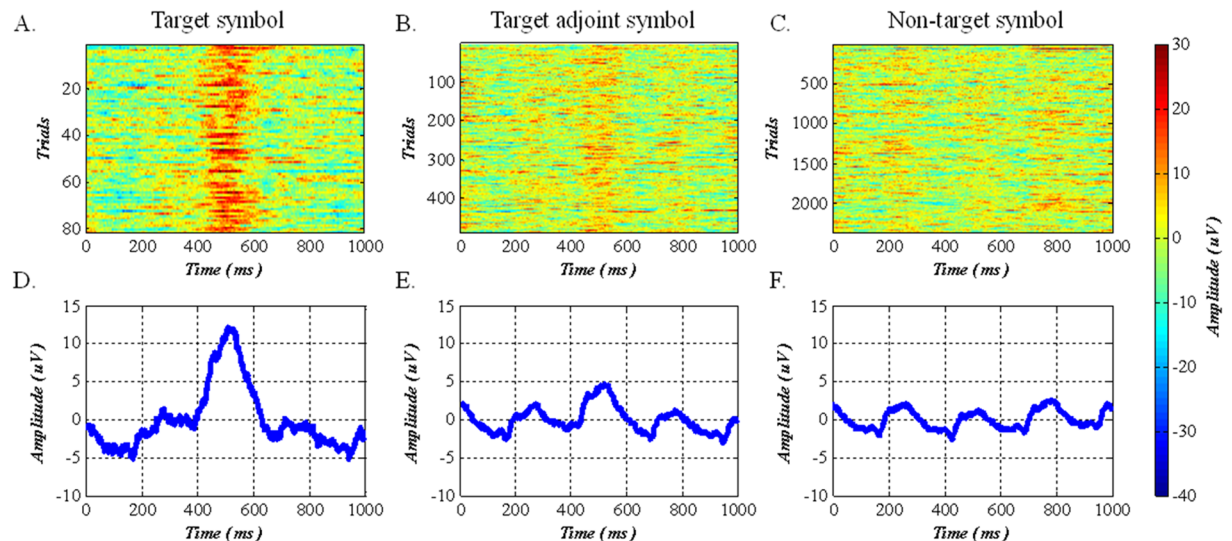
Figure 3. EEG signals of different categories (target, target adjoint, and non-target symbol), measured by electrode Pz. (**A**) The target character corresponds to the mean of the EEG signal. Each line represents the group average of the EEG signals of the target character (three symbol groups, each symbol group has a trial EEG, and the signal of the target character is the mean of the three symbol group), and the color indicates the signal amplitude value. Similarly, (**B**) and (**C**) indicate the cases of target adjoint symbol and non-target symbol, respectively. (**D**) indicates the averages across all signals of (**A**). (**E**) indicates the averages across all signals of (**B**). (**F**) indicates the averages across all signals of (**C**).

Subject	Accuracy, %		ITR, bit/min	
	Training (offline)	Testing (online)	Training (offline)	Testing (online)
sub1	0.742	0.732	18.137	17.740
sub2	0.812	0.831	21.050	21.886
sub3	0.663	0.702	15.117	16.575
sub4	0.663	0.732	15.117	17.740
sub5	0.756	0.783	18.701	19.813
sub6	0.951	0.922	27.819	26.248
sub7	0.733	0.716	17.780	17.114
sub8	0.821	0.883	21.443	24.295
sub9	0.779	0.785	19.646	19.897
sub10	0.828	0.814	21.753	21.137
sub11	0.776	0.753	19.521	18.579
sub12	0.715	0.741	17.075	18.097
sub13	0.902	0.882	25.228	24.247
Mean	0.780	0.790	19.876	20.259
SD	0.084	0.071	3.665	3.101

Table 1. Classification accuracy (%) and ITR (bit/min) in triple RSVP speller.

observed in the mean signal corresponding to the target character in Fig. 3(A), between 400 and 600 ms. The target adjoint character has an undetectable peak between 400 and 600 ms because the three symbol groups that correspond to the target adjoint character contain a target character. Therefore, the P300 component is included in the corresponding EEG signal of a symbol group, resulting in the target adjoint character mean EEG signal containing some of the P300 components. In addition, the P300 component is not included in the signal of a non-target character. All categories of the signals contain 4 Hz harmonic components because of the symbol group presentation speed of 4 per second.

Triple-RSVP speller performance. This study tested the triple RSVP speller using a copy spelling task. Table 1 lists the selected accuracy and ITR for all subjects. The triple RSVP speller offline average accuracy is 0.780, and the average ITR is 19.876 bit/min. The online average accuracy is 0.790, and the average ITR is 20.259 bit/min. The system spelled at a speed of 10 s per character, and the stimulus presentation interface is a 90 × 195 pixel rectangle.
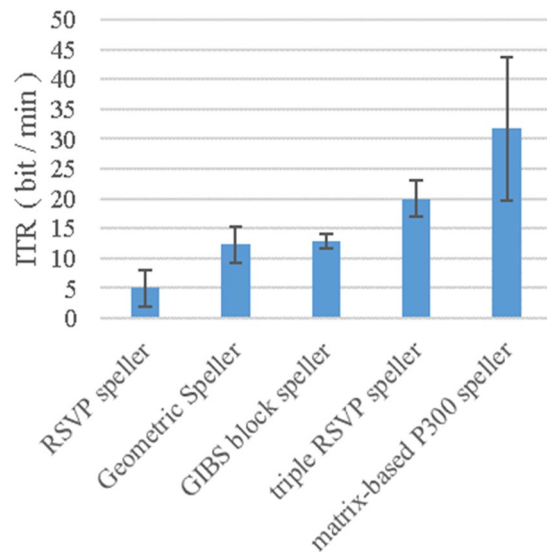
Figure 4. Information transfer rates for different speller systems. The RSVP speller, Geometric speller, GIBS block speller, and triple RSVP speller are gaze-independent spelling system. The matrix-based P300 speller refers to the current typical P300 spelling system to achieve the ITR level.

Discussion

We propose a novel triple RSVP speller with two notable features as follows: gaze-independent and space-independent features. The current mainstream BCI speller system is often based on the matrix structure. This structure of the speller system can achieve a higher information transfer rate. However, two major problems are not resolved, namely, gaze-dependent and space-dependent problems. The gaze-dependent problem refers to the subjects that must look at specific clues when spelling a character. Thus, when spelling a character in a covert attention approach, satisfactory select accuracy and information transfer rate are often not achieved. The space-dependent problem refers to the stimulation presentation interface that requires a large screen. Some studies have shown that in the matrix-based speller, when the matrix achieved is large, the performance is better; when the character is small, the performance is worse. Thus, the matrix-based speller cannot use a smaller screen (for example, a mobile phone screen) to ensure a certain accuracy and information transfer rate conditions; it is detrimental to BCI speller spread among normal people.

In the face of gaze dependent issue, many scholars have proposed covert attention speller systems such as the GIBS block speller[37], Hex-o-spell[38], Geometric Speller[39], and RSVP speller[42,43,48]. However, these methods cannot achieve a satisfactory information transfer rate. The GIBS block speller, Hex-o-spell, and Geometric Speller in the design cannot solve the problem of space dependency. Compared with these speller systems, we designed the triple RSVP speller system that can achieve a higher information transmission rate.

As shown in Fig. 4, we compared the triple RSVP speller with the mainstream gaze-independent spell system (the RSVP speller, Geometric speller, and GIBS block speller, and the data are derived from references[44,39], and[37], respectively) and the matrix-based P300 speller (data are derived from literature[31]) information transfer rate.

The RSVP keyboard was proposed by Orhan et al.[41–43], using rapid serial visual presentation (RSVP) paradigm for spelling. In the RSVP paradigm, each candidate letter is shown at the same place on the screen in a temporally ordered sequence at a high presentation rate, however, the ITR of this method is not satisfactory. The Geometric Speller was proposed by Aloise et al.[39], a total of N^2 characters are grouped into 2N sets of N characters (analogous to rows and columns of the FD speller). In this arrangement, each character belongs to exactly two sets. In the visual layout of each set, characters are displayed at the vertices of a regular geometric figure. During the presentation, each set of characters is displayed transiently on the screen. The GIBS block speller was proposed by Gabriel et al.[37], the symbols are grouped into four blocks following an alphabetical order. This layout is composed by a group of 9 symbols in the center and 3 lateral small blocks with the remaining symbols. To select a symbol, the user has first to select the small block where it belongs. When the respective block is selected, the symbols of that block move to the center and the respective small block disappears. The Geometric speller and GIBS block speller have higher ITRs, but require larger screen space. The matrix-based P300 speller is a classic BCI system, but the gaze-dependent and space-dependent problems cause it cannot be widely used in severe neuromuscular disability and normal people.

In this paper, the proposed triple RSVP speller proposed has space-independent features. Compared to the matrix-based BCI speller, the triple RSVP speller can be integrated in a small screen. In this experiment, we reduced the stimulus presentation interface to a size of 195 * 90 pixels. Compared to that of the matrix-based spelling system, the RSVP sequence is less demanding to stimulate size and spatial location requirements. Thus, the proposed triple RSVP speller can be embedded in the vast majority of portable smart devices (such as smartphones, smart watches, and others). The speller is useful for the BCI technology spread in daily life, and disabled persons can conveniently use it.

Index	1	2	3	4	5	6	7	8	9	10	11	12
block1	ACE	GIK	MOQ	SUW	Y13	579	BDF	HJL	NPR	TVX	Z24	68-
Index	13	14	15	16	17	18	19	20	21	22	23	24
block2	ADG	JMP	SVY	258	BEH	KNQ	TWZ	369	CFI	LOR	UX1	47-
Index	25	26	27	28	29	30	31	32	33	34	35	36
block3	AMY	BNZ	CO1	DP2	EQ3	FR4	GS5	HT6	IU7	JV8	KW9	LX-

Table 2. Symbol group sequence.
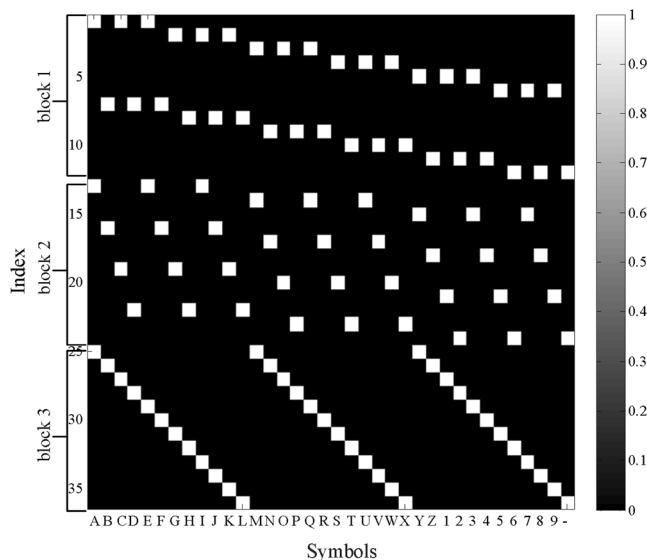
Figure 5. Rule of symbol group sequence. The abscissa represents 36 characters, and the ordinate represents the order in which the characters are presented, where 1 to 12 are block 1, 13 to 24 are block 2, and 25 to 36 are block 3. The symbol group presented in different trials is composed of white-marked characters.

We propose the triple RSVP speller with gaze-independent and space-independent characteristics. Compared to the traditional gaze-independent spelling system, the triple RSVP speller can achieve higher letter spelling speed on a smaller screen. However, compared to the matrix-based spelling system, the triple RSVP speller ITR is still relatively low. Our goal is to improve the triple RSVP speller ITR.

Methods

Symbol group sequence. The design of the symbol group sequence in the proposed triple RSVP speller is crucial. Our design criteria are as follows. First, each symbol must be presented again at a longer time interval. Thus, we designed three blocks, and each block contains the same 36 characters. Second, the symbol group of each character is different in different blocks to avoid confusion with the target character and the target adjoint character when distinguishing the classifier score. We designed the symbol group sequence as shown in Table 2 and Fig. 5.

Participants. A total of 13 subjects (11 males and 2 females, average age of 22.1 ± 1.5 years, and right-handed) participated in the experiment. All subjects were students of Zhengzhou University and did not have any previous training in the task. The subjects exhibited normal or corrected-to-normal vision with no neurological problems and were financially compensated for their participation. This study was conducted after we obtained informed consent and Ethics Committee approval of China National Digital Switching System Engineering and Technological Research Center, and was carried out in accordance with the approved guidelines. All of the participants provided their written informed consent to participate in this study.

EEG acquisition. EEG data were acquired by a g.USBamp system (G.Tec), using 16 electrodes distributed in accordance with the international 10–20 system. The electrooculographic (EOG) activity were recorded by two electrodes positioned above and below the left eye. We collected a group of EOC samples before the experiment and implemented the EOC artifact removal by using the method proposed by Chi Zhang[49,50].

The EEG data were sampled at 2400 Hz using 200 Hz low-pass and 50 Hz notch filters. Prior to scoring the images, we pre-processed the EEG data through the following steps: downsampling to 600 Hz, band-pass filtering (0.1–60 Hz), and baseline correction. Zero-delay filtering was implemented using the filtfilt() function in MATLAB. Then, the EEG data were divided into epochs. Each epoch consisted of 1000 ms of EEG data after the stimulus onset.

EEG data analysis.   Analysis of the ERP using HDCA algorithm was performed as described by Parra *et al.*[16,17,51]. The HDCA algorithm can be divided into two layers. First, the HDCA algorithm was employed to obtain the average data and divide the original EEG data by time window size. The weight of each channel was then calculated in each time window to maximize the differences between the target and non-target classes. The time window size cannot be determined in advance. Thus, we selected 25 ms as the time window size after numerous experimental repetitions. The weight of each channel in each time window was calculated by Fisher linear discriminant (FLD). In each time window, the EEG signal was reduced to one dimension, such as in Equation (1), as follows:

$$y_k = \left(\frac{1}{N}\right)\sum_n \sum_i w_{ki} x_{i[(k-1)N+n]},$$

(1)

where $x_{i[(k-1)N+n]}$ represents the *k*th separate time-window value from the single-trial data. The variable corresponds to the EEG activity at the data sample point *n* measured by electrode *i*. *w* is a set of spatial weights. Weight vector w_{ki} is identified for the *k*th window and *i* electrode following each image presentation (*T* is the temporal resolution of the time window, which is 0.025 in this paper, *N* is the sampling time point of the time window, F_S is the sampling rate, *K* is the number of time windows, and $n = 1, 2, \cdots, N,\quad N = T \times F_S, 0 \le k \le K$). y_k is the signal after reduced dimension in *k*th separate time-window. The time window size cannot be determined in advance. Thus, we selected 25 ms as the time window size after numerous experimental repetitions. The weight of each channel in each time window was calculated by Fisher linear discriminant (FLD).

$$y_{IS} = \sum_k v_k y_k,$$

(2)

The results for the separate time windows (y_k) were then combined in a weighted y_k average to provide a final interest score (y_{IS}) for each image. FLD analysis was used to calculate the spatial coefficient w_{ki}, and logistic regression was adopted to calculate the temporal coefficient v_k, such as in (2).

The time window size is 25 ms, $k = 40$.

Evaluation of triple RSVP speller performance.   To evaluate the performance of our triple RSVP speller, we computed the classification accuracy and ITR, which is widely used in the BCI speller community[52]. The ITR is given by

$$ITR = \frac{\left(\log_2 N + P\log_2 P + (1 - P)\log_2 \frac{1 - P}{N - 1}\right)}{Time},$$

(3)

where *Time* is the time interval per selection. In this paper, we designed the triple RSVP speller to select a character that takes 10 seconds. *N* is the number of possible choices, and *P* is the probability that the desired choice will be selected in a process called target identification accuracy or classifier accuracy.

In the performance evaluation stage, we calculated the precision accuracy and information transfer rate of all subjects in the offline and online environments, respectively. In the training phase, we asked the subjects to focus on specific characters and collected EEG data for offline analysis and P300 classifier training. Then, we asked the subjects to spell a series of specific characters in the copy mode to calculate the online accuracy and ITR.

References
1. Gao, X., Xu, D., Cheng, M. & Gao, S. A BCI-based environmental controller for the motion-disabled. *Neural Systems & Rehabilitation Engineering IEEE Transactions on* **11**, 137–140 (2003).
2. Wolpaw, J. R., Birbaumer, N., Mcfarland, D. J., Pfurtscheller, G. & Vaughan, T. M. Brain–computer interfaces for communication and control. *Supplements to Clinical Neurophysiology* **57**, 607–613 (2004).
3. Birbaumer, N. Breaking the silence: brain-computer interfaces (BCI) for communication and motor control. *Psychophysiology* **43**, 517–532 (2006).
4. Fabiani, G. E., Mcfarland, D. J., Wolpaw, J. R. & Pfurtscheller, G. Conversion of EEG activity into cursor movement by a brain-computer interface (BCI). *IEEE Transactions on Neural Systems & Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine & Biology Society* **12**, 331–338 (2004).
5. Guger, C. *et al.* How many people are able to control a P300-based brain–computer interface (BCI)? *Neuroscience Letters* **462**, 94–98 (2009).
6. Buch, E. *et al.* Think to move: a neuromagnetic brain-computer interface (BCI) system for chronic stroke. *Stroke* **39**, 910–917 (2008).
7. Polich, J. Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* **118**, 2128–2148, https://doi.org/10.1016/j.clinph.2007.04.019 (2007).
8. Gonsalvez, C. J. & John, P. P300 amplitude is determined by target-to-target interval. *Psychophysiology* **39**, 388–396 (2002).
9. Polich, J. P300 Probability, and Interstimulus Interval. *Psychophysiology* **27**, 396–403 (1990).
10. Linden, D. E. The P300: where in the brain is it produced and what does it tell us? *The Neuroscientist* **11**, 563–576 (2005).
11. Sutton, S., Braren, M., Zubin, J. & John, E. R. Evoked-potential correlates of stimulus uncertainty. *Science* **150**, 1187–1188 (1965).
12. Bigdely-Shamlo, N., Vankov, A., Ramirez, R. R. & Makeig, S. Brain activity-based image classification from rapid serial visual presentation. *IEEE Trans Neural Syst Rehabil Eng* **16**, 432–441, https://doi.org/10.1109/TNSRE.2008.2003381 (2008).
13. Ramoser, H., Muller-Gerking, J. & Pfurtscheller, G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans Rehabil Eng* **8**, 441–446 (2000).
14. Rivet, B., Cecotti, H., Souloumiac, A., Maby, E. & Mattout, J. Theoretical analysis of xDAWN algorithm: application to an efficient sensor selection in a P300 BCI. *19th Signal Processing Conference, European, IEEE* (2011).
15. Rivet, B., Souloumiac, A., Attina, V. & Gibert, G. xDAWN algorithm to enhance evoked potentials: application to brain-computer interface. *IEEE Trans Biomed Eng* **56**, 2035–2043, https://doi.org/10.1109/TBME.2009.2012869 (2009).
16. Gerson, A. D., Parra, L. C. & Sajda, P. Cortically coupled computer vision for rapid image search. *IEEE Trans Neural Syst Rehabil Eng* **14**, 174–179, https://doi.org/10.1109/TNSRE.2006.875550 (2006).
17. Parra, L. C. *et al.* Spatiotemporal linear decoding of brain state. *Signal Processing Magazine, IEEE* **25**, 107–115 (2008).

18. Sajda, P. *et al*. In a blink of an eye and a switch of a transistor: cortically coupled computer vision. *Proceedings of the IEEE* **98**, 462–478 (2010).
19. Pohlmeyer, E. A. *et al*. Closing the loop in cortically-coupled computer vision: a brain-computer interface for searching image databases. *J Neural Eng* **8**, 036025, https://doi.org/10.1088/1741-2560/8/3/036025 (2011).
20. Marathe, A. R., Ries, A. J. & McDowell, K. In *Foundations of Augmented Cognition* 345–352 (Springer, 2013).
21. Marathe, A. R., Ries, A. J. & McDowell, K. Sliding HDCA: single-trial EEG classification to overcome and quantify temporal variability. *IEEE Trans Neural Syst Rehabil Eng* **22**, 201–211, https://doi.org/10.1109/TNSRE.2014.2304884 (2014).
22. Cecotti, H., Eckstein, M. P. & Giesbrecht, B. Single-trial classification of event-related potentials in rapid serial visual presentation tasks using supervised spatial filtering. *IEEE Transactions on Neural Networks & Learning Systems* **25**, 2030–2042 (2014).
23. Farwell, L. A. & Donchin, E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography & Clinical Neurophysiology* **70**, 510–523 (1988).
24. Krusienski, D. J. *et al*. A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering* **3**, 299–305 (2006).
25. Guan, C., Thulasidas, M. & Wu, J. High performance P300 speller for brain-computer interface, *IEEE International Workshop on Biomedical Circuits and Systems*. IEEE (2004).
26. Townsend, G. *et al*. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clinical Neurophysiology Official Journal of the International Federation of Clinical Neurophysiology* **121**, 1109–1120 (2010).
27. Roark, B., Beckley, R., Gibbons, C. & Fried-Oken, M. Huffman scanning: using language models within fixed-grid keyboard emulation. *Computer Speech & Language* **27**, 1212–1234 (2013).
28. Roark, B., Villiers, J. D., Gibbons, C. & Fried-Oken, M. Scanning methods and language modeling for binary switch typing. *Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies* 28–36 (2010).
29. Li, Q., Liu, S. & Li, J. Neural Mechanism of P300-Speller Brain-Computer Interface Using Familiar Face Paradigm. *International Conference on Network and Information Systems for Computers*, IEEE (2015).
30. Li, Q., Liu, S., Li, J. & Bai, O. Use of a Green Familiar Faces Paradigm Improves P300-Speller Brain-Computer Interface Performance. *Plos One* **10**, e0130325 (2015).
31. Yin, E. *et al*. A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm. *Journal of Neural Engineering* **10**, 026012 (2013).
32. Chen, X., Chen, Z., Gao, S. & Gao, X. A high-ITR SSVEP-based BCI speller. *Brain-Computer Interfaces* **1**, 181–191 (2014).
33. Chen, X. *et al*. High-speed spelling with a noninvasive brain-computer interface. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E6058 (2015).
34. Brunner, P. *et al*. Does the 'P300' speller depend on eye gaze? *Journal of Neural Engineering* **7**, 056013–056019 (2010).
35. Allison, B. Z. & Pineda, J. A. ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system. *Neural Systems & Rehabilitation Engineering IEEE Transactions on* **11**, 110–113 (2003).
36. Salvaris, M. & Sepulveda, F. Visual modifications on the P300 speller BCI paradigm. *Journal of Neural Engineering* **6**, 046011 (2009).
37. Pires, G., Nunes, U. & Castelo-Branco, M. GIBS block speller: Toward a gaze-independent P300-based BCI, *International Conference of the IEEE Engineering in Medicine & Biology Society*. IEEE (2011).
38. Treder, M. S. & Blankertz, B. (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions* **6**, 28 (2010).
39. Aloise, F. *et al*. A covert attention P300-based brain-computer interface: Geospell. *Ergonomics* **55**, 538–551 (2012).
40. Treder, M. S., Schmidt, N. M. & Blankertz, B. Gaze-independent brain-computer interfaces based on covert attention and feature attention. *Journal of Neural Engineering* **8**, 066003 (2011).
41. Higger, M., Akcakaya, M., Orhan, U. & Erdogmus, D. Robust Classification in RSVP Keyboard. **8027**, 443–449 (2013).
42. Orhan, U. RSVP Keyboard (TM) : An EEG Based BCI Typing System with Context Information Fusion. *Dissertations & Theses - Gradworks* (2013).
43. Orhan, U. *et al* RSVP Keyboard: An EEG Based Typing Interface, *IEEE International Conference on Acoustics*, IEEE (2012).
44. Chennu, S., Alsufyani, A., Filetti, M., Owen, A. M. & Bowman, H. The cost of space independence in P300-BCI spellers. *Journal of Neuroengineering & Rehabilitation* **10**, 82 (2013).
45. Moghadamfalahi, M. *et al*. Language-Model Assisted Brain Computer Interface for Typing: A Comparison of Matrix and Rapid Serial Visual Presentation. *IEEE Transactions on Neural Systems & Rehabilitation Engineering A Publication of the IEEE Engineering in Medicine & Biology Society* **23**, 910–920 (2015).
46. Acqualagna, L. & Blankertz, B. A gaze independent spelling based on rapid serial visual presentation, *International Conference of the IEEE Engineering in Medicine & Biology Society*, IEEE (2011).
47. Acqualagna, L., Treder, M. S., Schreuder, M. & Blankertz, B. A novel brain-computer interface based on the rapid serial visual presentation paradigm. *Conf Proc IEEE Eng Med Biol Soc* **2010**, 2686–2689 (2010).
48. Friedoken, M., Mooney, A., Peters, B. & Oken, B. A clinical screening protocol for the RSVP Keyboard™ brain-computer interface. *Disability and Rehabilitation: Assistive Technology* **10**, 11–18 (2015).
49. Zhang, C. *et al*. Prior artifact information based automatic artifact removal from EEG data, *Neural Engineering* (*NER*) *International IEEE/EMBS Conference on*. IEEE (2015).
50. Zhang, C. *et al*. Automatic artifact removal from electroencephalogram data based on a priori artifact information. *BioMed research international* **2015** (2015).
51. Sajda, P., Gerson, A. & Parra, L. High-throughput image search via single-trial event detection in a rapid serial visual presentation task, *First International IEEE EMBS Conference on*. IEEE (2003).
52. Yuan, P. *et al*. A study of the existing problems of estimating the information transfer rate in online brain-computer interfaces. *Journal of Neural Engineering* **10**, 26014 (2013).

Acknowledgements

Author Contributions

Zhimin Lin, Chi Zhang, Ying Zeng, Li Tong and Bin Yan designed research; Zhimin Lin proposed ideal and the preparation of the corresponding procedures and algorithms and draws all the pictures; Chi Zhang assist to start experimenting and analyzing the data; Ying Zeng and Li Tong guided the writing of the paper; and Bin Yan supervised the project.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.