Comparison of Adaptive Symbol Presentation Methods for RSVP Keyboard

U. Orhan[1], M. Akcakaya[1], D. Erdogmus[1], B. Roark[2], M. Moghadamfalahi[1], M. Fried-Oken[2]
[1]CSL, Northeastern University Boston, MA; [2]Oregon Health and Science University, Portland, OR

Correspondence: U. Orhan, Northeastern University, Boston, MA. E-mail: rsvpkeyboard@coe.neu.edu

Abstract. RSVP Keyboard™ is an EEG based letter-by-letter typing system specifically designed for people with locked-in-syndrome (LIS). It uses rapid serial visual presentation of symbols and classification of event relatedpotentials with the aid of a language model. We designed various adaptive symbol presentation methods for each sequence of visual stimuli and compared the efficacy of these methods using estimated task completion accuracy and speed with rigorous modeling and Monte Carlo simulations, based on EEG data collected for calibration purposes.

Keywords: EEG, ERP, BCI, Language Model, Kernel Density Estimate, Symbol Presentation design

1. Introduction

Brain computer interfaces (BCI) have been designed to assist people with severe motor disabilities or locked-in syndrome with communication and control when motor control is not possible. Practicality of noninvasive portable acquisition systems, like electroencephalography (EEG), further attracted the researchers to BCI systems. Existing noninvasive BCIs for typing use many repetitions of stimuli to increase accuracy at the cost of speed [Pfurtscheller et al., 2000; Krusienski et al., 2008]. However, speed is also crucial aspect of peer-to-peer communication.

To develop a system that achieves high accuracy and speed simultaneously, in our previous work we have demonstrated rapid serial visual presentation (RSVP) in conjunction with language models (RSVP Keyboard™) in order to assist letter selection during the brain-typing process. RSVP relies on temporal rather than spatial separation of stimuli, and EEG responses for the visual stimuli are classified using regularized discriminant analysis applied to stimulus-onset-locked temporal features from all channels. Fusion of language and EEG evidence is achieved using a probabilistic framework, assuming that these are conditionally independent given class labels [Orhan et al., 2012].

Currently, RSVP Keyboard™ presents random permutations of the 26 letters in English alphabet, a space symbol and a backspace symbol (we call this set a sequence). If repetition is needed, all symbols are repeated multiple times (maximum number of repetitions is bounded) to improve classification accuracy until a desired confidence level is reached [Orhan et al., 2012]. However, using all 28 symbols in a sequence might not be necessary for a specific target symbol to be selected successfully. Depending on the context, some symbols might be highly unlikely to be chosen. Therefore, an adaptive symbol presentation method based on the evidence provided by EEG and language model (LM) might decrease the number of symbols in a sequence and the total symbol selection time. In this study, to further increase the typing accuracy and communication speed, we investigate different sequence selection methods, demonstrating their typing performances using Monte Carlo simulations on prerecorded EEG data.

2. Methods

In RSVP Keyboard™, the user is assumed to show positive intent exactly for all occurrences of one symbol per epoch (section in which user attempts to select a target symbol for typing). Each epoch contains a set of sequences, currently containing all 28 symbols. In this study, using Monte Carlo simulations on multiple pre-recorded calibration data with different performance (different area under the receiver operating characteristic (ROC) curve (AUC) values), and changing the number of symbols in a sequence (N_T) and maximum number of sequences in an epoch (N_{MAX}), we propose to compare the typing performance of three different subset selection methods. We select ten different sentences and aim to spell a phrase in each sentence (called the copy phrase task). Task difficulty is determined by requiring each letter of the target phrase to have a likelihood ratio against the highest competing non-target letter within a specified interval: (1) Hard: (0.3, 0.5], (2) Very hard: (0, 0.3].

We employ a 6-gram character-based LM that is trained using a one-million sentence sample of the New York Times portion of the English Gigaword corpus. We apply our simulation model to the copy phrase task and report the estimated performance in terms of average time to complete the whole task (T_{est}), and probability of successful phrase completion (P_{est}). In other words, P_{est} represents the typing accuracy and T_{est} represents the typing duration. We use the target and non-target EEG responses from the calibration data and perform kernel density estimation with cross validation on the features extracted from these raw responses. During simulation, we draw samples from these densities to obtain simulated EEG evidence for target and non-target symbols. The EEG evidence is fused with the LM to compute the posterior probabilities used for decision making. Simulations, utilizing the EEG responses to the RSVP paradigm from participants, are a close representation of the typing performance.

Three subset selection methods we compare here are as follows: Method 1 (M1) uses the full set (28 symbols); Method 2 (M2) displays a k-element subset of the alphabet having the highest posterior probabilities after fusion with EEG evidence at the end of each sequence; Method 3 (M3) displays a k-element subset with the highest probabilities, while giving priority to symbols with fewer number of repetitions, for instance, for $N_T = 7$ after 4 sequences every symbol will be shown exactly one time.

3. Results

Using 100 Monte Carlo simulations and calibration data with AUC = 0.7662 and 0.8330, respectively, we obtain the results summarized for analysis in Fig. 1.

		AUC =0.7662					AUC = 0.8330				
		M 1	M2		M3		M 1	M2		M3	
		N_T=28	N_T=14	N_T=7	N_T=14	N_T=7	N_T=28	N_T=14	N_T=7	N_T=14	N_T=7
T_{est} (min)	N_{MAX} = 2	26.44	15.06	9.68	13.770	8.6324	22.02	12.55	8.79	14.821	9.509
	N_{MAX} = 4	38.44	22.66	15.02	24.867	14.223	21.40	13.31	10.27	19.372	15.140
	N_{MAX} = 8	40.81	24.48	18.03	39.086	27.707	20.91	12.41	9.52	19.277	21.703
	N_{MAX} = 16	39.25	23.32	17.22	41.616	45.334	20.31	12.30	9.13	18.258	21.308
P_{est}	N_{MAX} = 2	0.093	0.117	0.096	0.023	0.018	0.644	0.647	0.464	0.247	0.219
	N_{MAX} = 4	0.512	0.468	0.370	0.117	0.012	0.984	0.973	0.861	0.721	0.261
	N_{MAX} = 8	0.962	0.959	0.879	0.522	0.123	1.000	1.000	0.999	0.991	0.697
	N_{MAX} = 16	0.982	1.000	0.998	0.930	0.559	0.999	1.000	1.000	1.000	0.991

Figure 1. *Performance analysis results for three different symbol selection methods for two different accuracy levels.*

4. Discussion

The current version of RSVP Keyboard^{TM} operates using method 1 with $N_{MAX} = 4$. We first observe that a simple change in our current system by changing the N_{MAX} from 4 to 8 or 16 does not increase the estimated total task duration, but increases the accuracy drastically. Secondly, we demonstrate that the second method has the best typing performance among the three. To see this from Fig. 1, if we compare the (M2, N_{MAX} = 8 or 16) with (M1, N_{MAX} = 4 or 8), we observe that M2 improves both speed and accuracy. A similar comparison between M1 and M3 illustrates that M3 performs similar to M1 and may not be the best symbol presentation option. In our future work, we will develop symbol presentation methods relying on sequential dynamic state space models and provide a more quantitative analysis of adaptive symbol selection methods to increase speed without sacrificing accuracy.

5. Acknowledgements

This work is supported by NSF (CNS-1136027, IIS-0914808, IIS-1149570) and NIH (1RO1DC009834-01).

References

Pfurtscheller G, Neuper C, Guger C, Harkam W, Ramoser H, Schlogl A, Obermaier B, Pregenzer M. Current trends in Graz brain-computer interface (BCI) research. *IEEE Trans Rehabil Eng*, 8 (2):216-219, 2000.

Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. *J Neurosci Meth*, 167(1):15-21, 2008.

Orhan U, Hild KE, Erdogmus D, Roark B, Oken B, Fried-Oken M. RSVP Keyboard: An EEG based typing interface. *In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 645-648, 2012.