Nonparametric Early Stopping Detection for c-VEP-based Brain–Computer Interfaces: A Pilot Study

Víctor Martínez-Cagigal[1,2,*], Eduardo Santamaría-Vázquez[1,2], Sergio Pérez-Velasco[1,2], Diego Marcos-Martínez[1,2], Selene Moreno-Calderón[1], and Roberto Hornero[1,2]

***Abstract*— Brain-computer interface (BCI) systems based on code-modulated visual evoked potentials (c-VEP) stand out for achieving excellent command selection accuracies with very short calibration times. One of the natural steps to democratize their use in plug-and-play environments is to develop early stopping algorithms. These methods allow real-time detection of the minimum number of code repetitions needed to provide reliable selections. However, such techniques are scarce in the current state-of-the-art for c-VEP-based BCI systems based on the classical circular shifting paradigm. Here, a novel nonparametric early stopping method is proposed, which approximates the distribution of unattended commands to a normal distribution and issues a selection when the correlation of the command is considered an outlier. The proposal has been evaluated offline with 15 healthy users, achieving an average accuracy of 97.08% and a speed of 1.37 s/command. Likewise, the algorithm has also been evaluated with an additional user in an online way, as a proof of concept to validate its technical feasibility, achieving an average accuracy of 96.88% with a speed of 1.67 s/command. These results suggest that the real time application of the proposed algorithm is feasible, significantly reducing the required selection time without compromising accuracy.**

I. INTRODUCTION

Non-invasive brain-computer interface (BCI) systems use electroencephalographic (EEG) signals to interpret the user's intentions and convert them into commands for external devices or applications [1]. Decoding such intentions from EEG is challenging, and requires the use of different control signals; i.e., strategies that generate measurable changes in the EEG when performing cognitive tasks (endogenous) or processing external stimuli (exogenous) [1]. Among exogenous signals, code-modulated visual evoked potentials (c-VEP) have recently been proposed as a new and promising control signal to offer high performance with short calibration times.

Traditional c-VEP-based BCI systems employ a pseudorandom binary code with perfect autocorrelation properties, enabling each command to be encoded using temporally shifted versions of the same sequence [2]. In this paradigm, known as circular shifting, a calibration template is calculated from the user's EEG response to the visual stimulation. The online decoding of the desired user's command is then possible by identifying the phase shift from the original template [2]. This approach easily achieves accuracy over 90% and information transfer rates (ITR) of up to 100 bpm with a calibration time of 10-30 s [3].

These c-VEP-based BCIs have shown potential to become user-friendly technologies. In this sense, the implementation of early stopping techniques would allow to optimize the selection time for each command adaptively. In other words, these methods would automatically determine the number of code repetitions needed for a reliable selection in real-time. Despite the impressive outcomes of c-VEP-based BCIs, early stopping techniques remain limited [2], frequently being incompatible with the circular shifting paradigm [4], [5], or requiring classifier dependency (wrapper) or parameter optimization (parametric) [6], [7]. Nevertheless, the use of early stopping has been proven to improve system performance, resulting in accuracy rates above 90% with command selection times of 3.26 s [7], 4.2 s [5], or 6.17 s [8].

The objective of this pilot study is to propose a novel early stopping technique for c-VEP-based BCI systems that use circular shifting. To the best of our knowledge, our method is the first one characterized by its independence from any specific classifier (i.e., it is filter-based), and its lack of necessity for parameter training (i.e., it is nonparametric). Furthermore, it does not require additional EEG recordings and can be applied in real time.

II. PARTICIPANTS

The proposed algorithm was assessed both offline using 15 healthy users (mean age: 28.80 $\pm$ 5.02 years, 10 males), and online with 1 additional healthy user (29 years old, male) as a proof of concept. All of them gave their informed consent to participate in the study. The signal was recorded by placing 16 active EEG electrodes at positions F3, Fz, F4, C3, Cz, C4, CPz, P3, Pz, P4, PO7, POz, PO8, Oz, I1 and I2, grounded at AFz and referenced to the earlobe. A g.USBamp equipment (g.Tec, Guger Technologies, Austria) with a sampling rate of 256 Hz was used. Signal acquisition, real-time signal processing, and c-VEP paradigm were programmed in MEDUSA©, a general-purpose ecosystem to develop of Python-based BCIs and neuroscience experiments (www.medusabci.com) [9]. The paradigm was displayed in

This research was supported by projects TED2021-129915B-I00, RTC2019-007350-1 and PID2020-115468RB-I00 funded by MCIN/AEI/10.13039/501100011033 and 'European Union NextGenerationEU/PRTR'; and by 'Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' through 'Instituto de Salud Carlos III' co-funded with European Regional Development Fund (ERDF) funds. ESV, SPV and DMM were in receipt of a PIF grant by the 'Consejería de Educación de la Junta de Castilla y León'.

* Corresponding author: victor.martinez.cagigal@uva.es.

[1] All authors are with the Biomedical Engineering Group, University of Valladolid, Valladolid, Spain.

[2] VMC, ESV, SPV, DMM and RH are with the 'Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)', Valladolid, Spain.

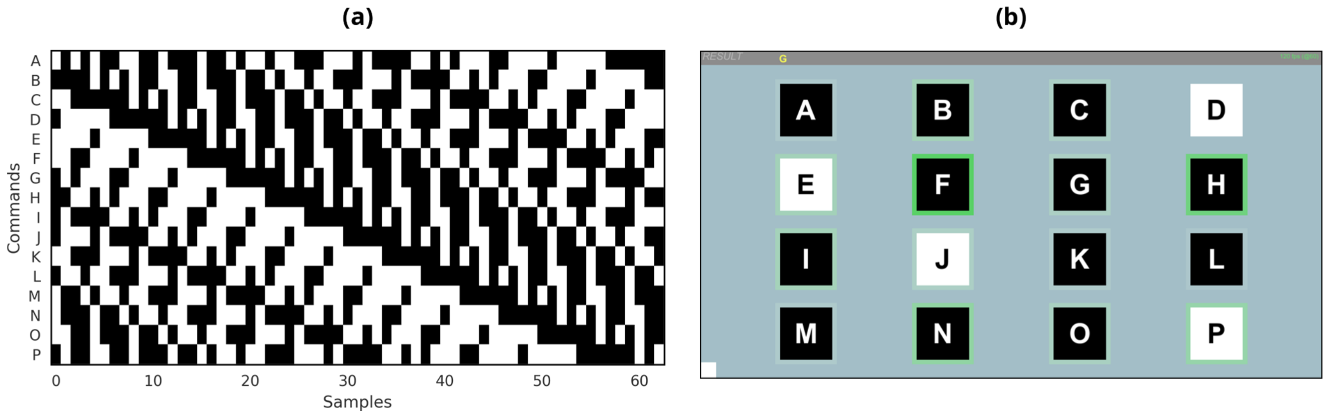

Fig. 1. (a) Encoding of each command using shifted versions of the m-sequence with a step of $\tau = 4$ samples. (b) Screenshot of the paradigm developed in MEDUSA© during a stimulation cycle. Accumulated correlation is shown in the form of a green frame surrounding the commands.

a LED FullHD @ 144 Hz monitor, attached to a PC Intel Core i9-11900KF @ 3.5 GHz, 64 GB RAM.

III. METHODS

A. Signal Processing and Paradigm

The encoding of the commands was achieved through the use of a 63-bit binary maximum length sequence (i.e., m-sequence), generated by a linear feedback shift register (LFSR) using the polynomial $x^6 + x^5 + 1$ with initial state 110000 [2]. The stimuli were displayed at a rate of 120 Hz, so a complete cycle (i.e., repetition of the m-sequence) lasted 525 ms. The c-VEP speller paradigm, publicly available as an app of MEDUSA© platform (www.medusabci.com/market/cvep_speller), consisted of 16 commands encoded with shifted versions of the original m-sequence with delays of $\theta_i = i\tau$ samples, where i indicates the command index and $\tau = 4$ was the step [2]. Command encoding and paradigm arrangement are shown in Figure 1(a).

During the calibration stage, the participant is instructed to focus on the command encoded by the original m-sequence (i.e., $\theta = 0$), which is repeated for k cycles. Following preprocessing, which includes a bank of bandpass filters (1-60Hz, 12-60Hz, and 30-60Hz) and a notch filter at 50Hz, two versions of the EEG response are generated: (1) the concatenated epochs $\boldsymbol{A} \in \mathbb{R}^{[kN_s \times N_c]}$, where N_s is the number of samples of a complete cycle and N_c is the number of channels; and (2) the epochs averaged over the k cycles $\boldsymbol{B} \in \mathbb{R}^{[N_s \times N_c]}$. Next, a canonical correlation analysis (CCA) is performed to identify the spatial filters that maximize the correlation between the projections of $\boldsymbol{A}$ and $\boldsymbol{B}$. The spatial filter $\boldsymbol{\omega_b}$ that yields the maximum correlation between the concatenated epochs and the average response is then determined. The original signal is subsequently projected using this filter to obtain the main template $\boldsymbol{x}_0 = \boldsymbol{B}\boldsymbol{\omega_b}$, while the templates for the rest of the commands are obtained by shifting the original $\theta_i = i\tau$ samples, where $i = 0 \ldots 15$ (for more information on this reference processing, see [2]). To remove noisy epochs, calibration epochs exhibiting a standard deviation three times greater than the average standard deviation of all of them were discarded.

During the command selection (i.e., test) stage: (1) individual epochs from each cycle are extracted and projected using the spatial filter $\boldsymbol{\omega_b}$; and (2) the correlation between the resulting projection and all templates is computed, i.e. $\boldsymbol{\rho}$. The selected command corresponds to the index $\arg\max_i(\boldsymbol{\rho})$ that yields the highest correlation value.

B. Proposed Early Stopping Algorithm

An early stopping algorithm must make a dichotomous decision each time a cycle is displayed: either (1) select the most probable command, or (2) repeat the stimulation for one additional cycle. With each cycle stimulation, a comparison between the EEG response from the start of the visual stimulation to the end of the current cycle and the command templates yields a vector of correlations $\rho \in \mathbb{R}^{[1\times16]}$. Upon sorting the vector in descending order, ρ_1 represents the highest correlation, which is assumed to correspond to the most likely command if the participant is indeed attending to it. The remaining correlations, $\rho_2, \rho_3, \ldots, \rho_{16}$ can then be considered as spurious correlations associated with unwanted commands.

Additionally, we can broaden the distribution of spurious correlations by determining the correlation of the EEG response with all possible shifted versions of the m-sequence, not just the 16 encoded commands (i.e., θ_j, where $j = 1, 2, \ldots, L$, and $L = 63$ is the code length). After sorting the correlations in descending order, the first correlation ρ_1 represents the most likely command, while $\boldsymbol{\rho}_{spu} = [\rho_2, \rho_3, \ldots, \rho_L]$ make up the distribution of spurious correlations. Thus, we can conclude that ρ_1 reliably corresponds to the selected command if its value stands out as an outlier from the distribution $\boldsymbol{\rho}_{spu}$.

There are several metrics to detect outliers of a distribution, such as those based on the interquartile range or on hypothesis testing. However, we propose the use of z-scores. Under the assumption that the distribution of the spurious values follows a normal distribution, i.e., $\boldsymbol{\rho}_{spu} \sim \mathcal{N}(\mu, \sigma)$, we can determine if ρ_1 is an outlier by checking if $\rho_1 - \mu >$

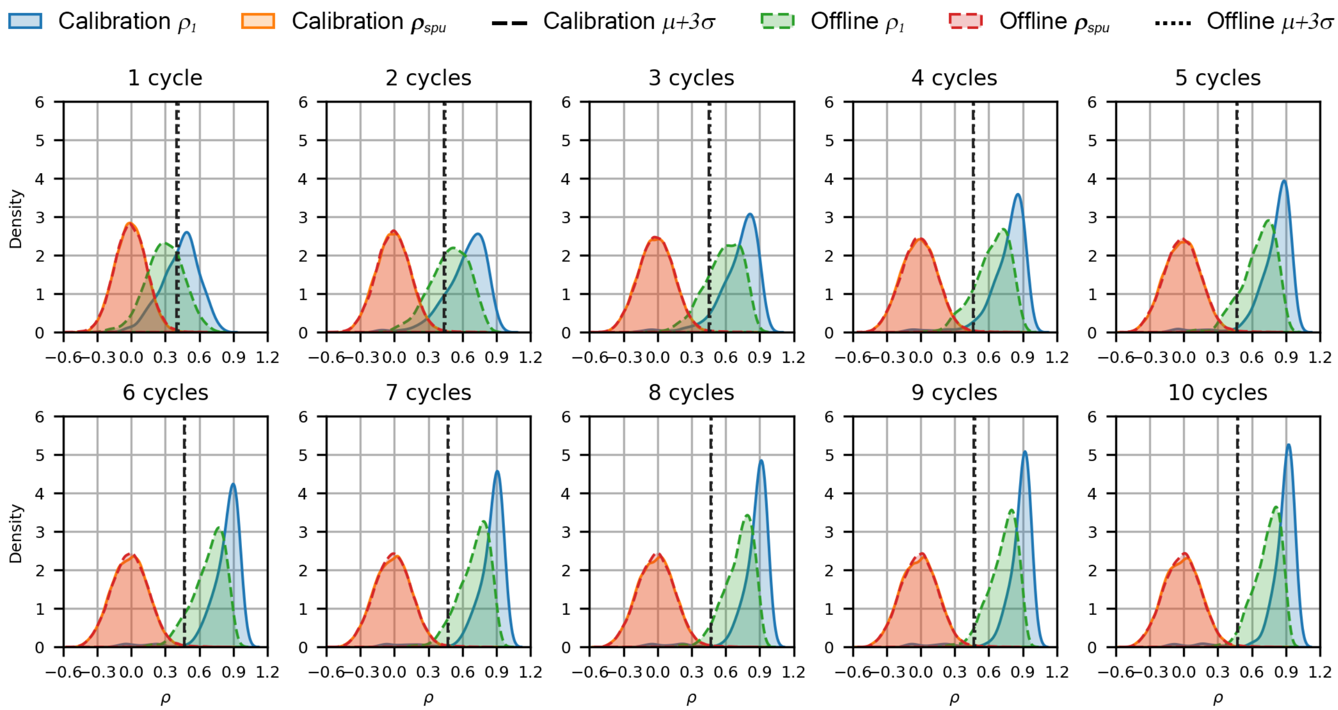

Fig. 2. Distributions of the selected correlations ρ_1 and the spurious distribution $\boldsymbol{\rho}_{spu}$ in the offline analysis for the calibration (blue and orange) and offline test (green, red) data, as well as the 99.87% percentile for each one.

$h\sigma$, where $h = 3$. This equates to determining if ρ_1 exceeds the 99.87% percentile of the distribution. If the inequality holds true, the command selection is made; otherwise, the stimulation continues with the next cycle.

C. Evaluation Protocol

To assess the method's feasibility and reliability, both an offline and online analyses were conducted. In the offline analysis, we assessed whether the distribution of unwanted command correlations follows a normal distribution and determined the average number of cycles required if the proposed early stopping algorithm is used. The analysis was based on the data collected from 6 healthy users, with 300 calibration cycles (6 runs $\times$ 5 trials $\times$ 10 cycles) and 320 test cycles (2 runs $\times$ 16 trials $\times$ 10 cycles) per user. In each run, the participants selected all the commands in lexicographic order. In the online analysis, the proposed early stopping algorithm was implemented in real time and evaluated using data from an additional subject. The data consisted of 300 calibration cycles (6 runs $\times$ 5 trials $\times$ 10 cycles) and 32 online selection trials (where the number of cycles varies).

IV. RESULTS AND DISCUSSION

A. Offline Analysis

The results of the Kolmogorov-Smirnov test indicate that the distributions of the target and spurious correlations for both calibration and offline test data are normal (p-value < 0.01), as shown in Figure 2. Therefore, the 99.87% percentile estimate when the z-score is $h = 3$ times greater than the standard deviation of the data can be considered accurate. The comparison of the distributions in the calibration and offline test data reveals significant similarity (p-value < 0.01), which suggests that the value of h could be optimized for each user without acquiring additional data. This optimization could be performed by analyzing a receiver operating characteristic (ROC) curve, by varying the value of h and observing the separation of the target and spurious distributions in the calibration.

Table I presents the accuracy results for each participant in both offline and online analyses. The theoretical maximum accuracy, calculated as the minimum number of cycles required to achieve the highest accuracy, is also included for the offline analysis. As shown, the use of the early stopping algorithm resulted in an average accuracy of 97.08% with an average of 2.61 cycles (equivalent to 1.37 s/command, ITR of 161.84 bpm) in the offline analysis. The original accuracy of 99.58% achieved with 10 cycles is higher (p-value $<$ 0.05), but our early stopping method reduced selection time per command significantly (p-value < 0.05), according to Wilcoxon signed-rank tests. Despite a slight decrease in accuracy to 97.08%, this average can be still considered an excellent result for state-of-the-art non-invasive BCIs [2].

It is important to note that the theoretical maximum suggests that a reduction to an average of 1.48 cycles is possible, which would be significantly faster than our 2.61 cycles (p-value < 0.05). The additional cycle required by our method is due to the indistinct target and spurious distributions in the first stimulation cycle, as seen in Figure 1. This phenomenon would result in an unreliable selection during the initial cycle, often requiring to wait for the next

TABLE I
OFFLINE AND ONLINE PERFORMANCE

		Early Stopping		Theoretical Maximum	
		Acc. (%)	N_c	Acc. (%)	N_c
Offline	U01	93.75	1.56	100.00	1.22
	U02	96.88	2.09	100.00	1.34
	U03	100.00	2.22	100.00	1.19
	U04	100.00	1.38	100.00	1.06
	U05	96.88	3.28	100.00	1.78
	U06	100.00	1.97	100.00	1.56
	U07	100.00	2.12	100.00	1.41
	U08	96.88	2.62	100.00	1.22
	U09	100.00	2.03	100.00	1.19
	U10	96.88	2.66	100.00	1.31
	U11	93.75	4.06	96.88	2.09
	U12	96.88	5.47	96.88	2.00
	U13	93.75	1.66	100.00	1.34
	U14	93.75	3.75	100.00	2.16
	U15	96.88	2.22	100.00	1.31
	mean	97.08	2.61	99.58	1.48
	STD	2.41	1.07	1.06	0.34
Online	U16	96.88	3.19	n.a.	n.a.

Acc.: accuracy, N_c: mean number of cycles, STD: standard deviation, n.a.: not applicable.

one. Although we consider that our results (97.08% with 1.37 s/command) demonstrate the usefulness of the method, this fact indicates that there is still room for improvement.

B. Online Analysis

Figure 1(b) displays a screenshot of the MEDUSA© app. The app utilizes the Unity graphics engine for maintaining a constant 120 Hz refresh rate and communicates with MEDUSA© through TCP/IP protocol [9]. The application provides feedback on the accumulated correlation of the early stopping method in the form of a green frame surrounding the commands. The intensity of the green color indicates the correlation level associated with each command; the higher the correlation, the more intense the green frame.

Table I also presents the results of the additional user who tested the algorithm online as proof of concept. The user achieved an accuracy of 96.88% with an average of 3.19 cycles (equivalent to 1.67 s/command, ITR of 131.76 bpm). The results demonstrate that the application of the proposed algorithm in real time is feasible and beneficial for the c-VEP-based BCI system, allowing to drastically decrease the time required to perform a selection from 5.25s (10 cycles) to 1.67 s (3.19 cycles) without compromising accuracy. Also, the performance obtained for this user is similar or exceeds the results of other studies, e.g., 3.26 s/command [7], 4.2 s/command [5], or 6.17 s/command [8].

C. Limitations and Future Lines

Despite the successful results of the proposed early stopping algorithm, there are several avenues for future research to enhance its efficacy and reliability. Firstly, it is crucial to test the method with a larger sample size of healthy users and/or motor-disabled individuals to increase the statistical power of the results. Additionally, it would be beneficial to modify the algorithm to detect the user's attention to the stimulation through an asynchronous stage. Currently, the cumulative calculation of correlations makes the control state detection challenging, as previous non-control cycles would significantly impact real-time decision-making. Instead, an asynchrony algorithm would ideally detect the attention in single-trial. Another future line of research could be the evaluation of the algorithm with a c-VEP-based BCI system using non-binary m-sequences; or even sets of codes with low cross-correlation (e.g., Gold codes, Kasami, etc.).

V. CONCLUSIONS

This study introduces a new early stopping algorithm for c-VEP-based BCI systems that boasts several advantages. The algorithm is (1) classifier-independent; (2) requires no parameter optimization; (3) does not need additional EEG signals; and (4) can be applied in real-time. The offline analysis revealed that the algorithm reduces the selection time from 5.25 s/command to 1.37 s/command while without compromising accuracy (from 99.58% to 97.08%). These findings are confirmed in the online proof of concept, achieving 96.88% accuracy with 1.67 s/command and an ITR of 131.76 bpm. It is concluded that the proposed algorithm is robust and practical, allowing for real-time detection of the number of cycles required for a selection without affecting the performance of the BCI system.

REFERENCES

[1] J. Wolpaw and E. W. Wolpaw, *Brain-computer interfaces: principles and practice*. OUP USA, 2012.

[2] V. Martínez-Cagigal, J. Thielen, E. Santamaría-Vázquez, S. Pérez-Velasco, P. Desain, and R. Hornero, "Brain-computer interfaces based on code-modulated visual evoked potentials (c-VEP): a literature review," *Journal of Neural Engineering*, vol. 18, no. 6, p. 061002, 2021.

[3] G. Bin, X. Gao, Y. Wang, B. Hong, and S. Gao, "VEP-based brain-computer interfaces: Time, frequency, and code modulations," *IEEE Computational Intelligence Magazine*, vol. 4, no. 4, pp. 22–26, 2009.

[4] S. Nagel and M. Spüler, "World's fastest brain-computer interface: Combining EEG2Code with deep learning," *PLoS ONE*, vol. 14, no. 9, pp. 1–15, 2019.

[5] J. Thielen, P. Marsman, J. Farquhar, and P. Desain, "From full calibration to zero training for a code-modulated visual evoked potentials brain computer interface," *Journal of Neural Engineering*, vol. 18, no. 5, p. 56007, 2021. [Online]. Available: https://doi.org/10.1088/1741-2552/abecef

[6] F. Gembler, P. Stawicki, A. Saboor, M. Benda, R. Grichnik, A. Rezeika, and I. Volosyak, "A Dictionary Driven Mental Typewriter Based on Code-Modulated Visual Evoked Potentials (cVEP)," in *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*. IEEE, 2018, pp. 619–624.

[7] F. Gembler and I. Volosyak, "A novel dictionary-driven mental spelling application based on code-modulated visual evoked potentials," *Computers*, vol. 8, no. 2, 2019.

[8] F. Gembler, P. Stawicki, A. Rezeika, A. Saboor, M. Benda, and I. Volosyak, *Effects of monitor refresh rates on c-VEP BCIs*. Springer International Publishing, 2018, vol. 10727 LNCS. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-91593-7_6

[9] E. Santamaría-Vázquez, V. Martínez-Cagigal, D. Marcos-Martínez, V. Rodríguez-González, S. Pérez-Velasco, S. Moreno-Calderón, and R. Hornero, "MEDUSA©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research," *Computer Methods and Programs in Biomedicine*, vol. 230, no. 107357, 2023.