Bit-wise reconstruction of non-binary visual stimulation patterns from EEG using deep learning: a promising alternative for user-friendly high-speed c-VEP-based BCIs

Eduardo Santamaría-Vázquez[1,2][0000−0002−7688−4258], Víctor Martínez-Cagigal[1,2][0000−0002−3822−1787], and Roberto Hornero[1,2][0000−0001−9915−2570]

[1] Biomedical Engineering Group (GIB), E.T.S Ingenieros de Telecomunicación, University of Valladolid, Paseo de Belén 15, Valladolid, 47011, Spain
[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), Spain
eduardo.santamaria.vazquez@uva.es

Abstract. Brain-computer interfaces (BCI) based on code-modulated visual evoked potentials (c-VEP) have shown great potential for communication and device control. These systems encode each command using different sequences of visual stimuli. Normally, the stimulation pattern is binary (i.e., black and white), but non-binary stimuli sequences with different shades of gray could reduce eyestrain and improve user-friendliness. This study introduces a novel approach to decode non-binary visual stimuli patterns from electroencephalography (EEG) signals using deep learning. The proposed method uses, for the first time, a bit-wise reconstruction strategy for stimulation patterns encoded with 2, 3, 5, 7 and 11 levels of gray. The performance of the proposed approach was evaluated on a dataset of 16 subjects, reaching an average command decoding accuracy over 95% for all stimulation sequences. The high accuracy and speed of the proposed method make it a promising alternative for user-friendly, high-speed c-VEP-based BCIs.

Keywords: Brain-computer interfaces · c-VEP · EEG · Deep learning

1 Introduction

Brain-Computer Interfaces (BCI) enable direct communication between the brain and an external device, bypassing traditional motor pathways such as muscles and nerves [14]. BCIs have the potential to transform the way we interact with technology and the world around us, providing new possibilities for people with disabilities to control their environment and enhancing the performance of able-bodied individuals. In this regard, the field of non-invasive BCI research is rapidly advancing, with new technologies and applications being developed and tested in both clinical and non-clinical settings [14].

The electroencephalography (EEG) is the most used technique to measure the brain activity in BCIs due to its non-invasiveness and high temporal resolution, providing a cost-effective option for real-time control of external devices in a variety of settings [14]. On the other hand, this signal has low-spatial resolution and is greatly affected by noisy artifacts, such as power line frequency, muscular activity or eye blinks, among others. As a consequence of its low signal-to-noise ratio (SNR), EEG cannot be used to decode fine neural activity, such as inner speech or activation of individual muscles [14].

In order to increase the SNR of EEG, BCIs use different paradigms and control signals to decode user's intentions from EEG. Non-invasive BCIs are typically categorized as exogenous and endogenous [1]. Exogenous BCIs rely on external stimuli, such as flashing lights, sounds, or tactile sensations, to evoke neural responses that encode user's intentions in the EEG. These BCIs allow to select commands among a predefined set of options, being suitable for communication and control applications due to their greater accuracy and robustness [1]. In contrast, endogenous BCIs rely on self-generated brain activity to control the BCI system, being more adequate for clinical applications, such as neurorehabilitation [1].

The most widespread exogenous paradigms are P300 evoked potentials and steady-state visual evoked potentials (SSVEP). The P300 paradigm detects differences in the neural response to target and non-target stimuli, showing to be effective in a variety of BCI applications [11]. However, this paradigm has important limitations in terms of accuracy and speed, needing long calibration sessions to achieve peak performance [10]. The SSVEP paradigm solved some of these drawbacks by encoding each command with visual stimuli that flicker at different frequencies, increasing the overall performance [5]. However, its susceptibility to visual fatigue, and the low number of selectable commands that can be displayed at the same time without affecting systems's accuracy, limit the use of SSVEP BCIs for practical applications [5].

Code-modulated visual evoked potentials (c-VEPs) represent a promising alternative to these paradigms, as they offer several advantages, including higher accuracy, selection speed and robustness to environmental factors [7]. C-VEP-based BCIs present specific stimulation patterns for each command following a predefined sequence. The most widely used technique is the circular shifting paradigm [7]. This method uses shifted versions of maximal length sequences (m-sequences), which present very low correlation between them, to encode the commands [7]. Therefore, the different options are encoded with the same stimulation sequence, which is circularly shifted a certain number of bits. In the calibration phase, the EEG response to this sequence is recorded, calculating a subject-specific template. Then, the target command is decoded by correlating the EEG signal of each trial with the different circularly-shifted versions of the template. The main advantages of this paradigm are: (1) it reduces the calibration time, as it only needs to record the EEG response to 1 stimulation sequence; and (2) it ensures that the shifted versions of the m-sequence have very low correlation, maximizing the performance of the system. However, as only 1

stimulation sequence is used, the number of commands that can be encoded with the circular shifting paradigm is limited. To solve this limitation, Nagel *et al.* [8] introduced EEG2Code, a method that models the brain response to individual stimulation events. Its main advantage is that it allows to decode commands by predicting, bit by bit, the stimulation sequence. Thus, it allows the use of random stimulation patterns to encode more commands than m-sequences. EEG2Code showed high accuracy, which was later improved in more recent studies using deep learning [9]. Despite these advances, which show the potential of c-VEPs to significantly enhance the usability and effectiveness of exogenous BCIs, there is still room for improvement. For instance, as SSVEPs, c-VEPs also provoke visual fatigue and eyestrain to the user, which reduces the usability of the system for practical applications. To alleviate this effect, Gembler *et al.* [4] recently proposed low-contrast stimuli based on different shades of gray as a mean to reduce user discomfort. They applied the m-sequences with 5 levels of gray and canonical correlation analysis (CCA) to classify between 8 different commands with accuracies above 95% [4]. Despite of the novelty of Gembler's study, their method can only decode a limited number of commands due to the use of the circular shifting paradigm. On the other hand, bit-wise reconstruction approaches, such as EEG2Code, could cope with an unlimited number of commands [9]. However, these techniques have not been applied to non-binary stimulation patterns yet.

In this study, we propose, for the first time, a novel classification paradigm based on bit-wise code reconstruction for non-binary stimulation patterns encoded with 2, 3, 5, 7 and 11 shades of gray. In our approach, the stimulation code is reconstructed bit by bit, rather than use the whole trial to calculate correlations, as in the circular shifting paradigm. Our method uses a flavored version of EEG-Inception, a convolutional neural network (CNN) for EEG classification tasks, which has been specifically tailored to better fit the needs of the bit-wise reconstruction approach [10].

2 Materials & Methods

2.1 Subjects and signals

In this study, 16 healthy participants (11 males, 5 females; 28.8±5.0 years old) took part in the experiments [6]. EEG signals were recorded using a g.USBamp amplifier with 16 channels at F3, Fz, F4, C3, Cz, C4, CPz, P3, Pz, P4, PO7, PO7 and Oz, according to the International System 10-10. The reference was placed at the right earlobe, whereas ground was placed at FPz. The sampling rate was 256 Hz. MEDUSA©, a novel Python-based BCI platform, was used to present and process stimuli in real-time [12]. Concretely, we used the app "P-ary c-VEP Speller", which is publicly available at www.medusabci.com [6]. The experiment was conducted on a computer with an Intel Core i7-7700 processor, 32 GB RAM, and a LED FullHD screen with a refresh rate at 144 Hz [6].

Table 1. P-ary m-sequences used to generate the stimulation sequences

	Base	Order	Length	Cycle duration
$GF(2^6)$	2	6	63 bits	0.525 s
$GF(3^4)$	3	4	80 bits	0.667 s
$GF(5^3)$	5	3	124 bits	1.033 s
$GF(7^2)$	7	2	48 bits	0.408 s
$GF(11^2)$	11	2	120 bits	1.000 s

GF: Galois Field; Cycle duration: time of each stimulation cycle with an stimulus presentation rate of 120 Hz.

2.2 System design

Stimulation patterns. The stimulation patterns of each command were designed using m-sequences, which are pseudorandom periodic time series that exhibit almost orthogonal behavior to circularly shifted versions of themselves [7]. These sequences, which are widely used in c-VEP-based BCIs, are generated through linear-feedback shift registers (LFSRs) that utilize a linear function of the immediate previous state to compute new values. The base p, order r, and arrangement of taps determine the LFSR, which is expressed as a polynomial whose coefficients are bounded on a Galois Field (GF). More information about the generation of m-sequences is available in the study of Buračas *et al.* [2]. In our study, 5 different p-ary m-sequences were used, as displayed in Table 1. Importantly, an m-sequence cannot be of arbitrary length and instead consists of exactly $N = p^r - 1$ bits. The 16 commands of our BCI system are encoded with shifted versions of the each m-sequence, which were selected to ensure near-to-zero correlation to maximize the performance of the c-VEP-based BCI. See the study of Martínez-Cagigal *et al.* [6] for more information about the design of the m-sequences. The stimuli were presented with a refresh rate of 120 Hz.

Pre-processing. The preprocessing stage of the algorithm involves two main steps: the application of a bandpass Infinite Impulse Response (IIR) Butterworth filter between 1 and 60 Hz, and a notch IIR Butterworth filter between 49 and 51 Hz. Both filters had order 7. The purpose of the band-pass filter is to remove waveforms outside of the desired range, which can include noise and artifacts that can interfere with the analysis of the signal. The notch filter is designed to eliminate the powerline interference at 50 Hz.

Feature extraction. After pre-processing, the feature extraction stage decimates the EEG signals to a sampling frequency of 200 Hz. Then, the system extracts an EEG epoch for each stimulus, with a temporal window of 0 to 500 ms after the stimulus onset. This temporal window is designed to capture the entire VEP response of the brain to each stimulus. Z-score baseline normalization is applied by taking the 250 ms of signal before the stimulus onset. After this process, each observation has a feature vector of 100 samples $\times$ 16 channels.

Classification. The classification of EEG epochs is performed with a modified version of EEG-Inception [10]. This CNN was specifically designed for P300 detection. Therefore, it has been adapted for its use in c-VEP-based BCIs. However, most of the architectural advantages, which proved to be effective for P300 detection, can be applied in this context as well. These include [10]: (1) efficient integration of depthwise convolutions, dropout regularization, batch normalization, and average pooling; (2) inclusion of Inception modules specifically designed for EEG processing to allow a multi-scale analysis in the temporal domain; (3) special design to avoid overfitting, including an output block that synthesizes the information extracted by Inception modules in very few, high-level features; and (4) optimized hyperparameters (i.e., dropout rate, activation functions and learning rate) for EEG classification tasks. For this study, we the input layer has been modified to allow the new shape of input feature vectors, the temporal scales of the 3 Inception branches, the number of filters of each branch and the dropout rate. Additionally, it must be noticed that the number of classes for each m-sequence is different, being equal to p, in accordance to the different shades of gray. Table 2 provides a detailed overview of the architecture and hyperparameters of the model used in this study. The dropout rate is set to 0.15. The model is trained using the calibration trials from the evaluation experiment (see section 2.3). In this study, we use subject-specific models, i.e., models that are trained and tested with data from the same subject. Specifically, the training dataset consisted of $30 \times L_s \times 10$ observations, where L_s is the length of each m-sequence with base p and order r, as described in Table 1. Then, the test trials are used to evaluate the model. To decode the test trials, the epochs corresponding to individual stimulation events are fed to each instance of EEG-Inception. Finally, the model predicts the class (i.e., shade of gray) of each stimulus based on the EEG data to reconstruct the m-sequence bit by bit.

Command decoding. Once the stimulation pattern has been reconstructed with EEG-Inception, the system correlates the predicted stimulation sequence with each one of the sequences of the 16 commands. The command with maximum Pearson correlation coefficient is selected.

2.3 Experimental procedure

Participants performed a single evaluation session with the c-VEP speller. The session had 5 blocks, one for each m-sequence, with calibration and test recordings. The graphical user interface of the speller, with the different modes, is displayed in Fig. 1. The stimulus presentation rate was 120 Hz. For each m-sequence, participants performed 30 calibration trials focusing on the unshifted version of the corresponding m-sequence using the calibration matrix. Then, the test task consisted of selecting the 16 commands, 2 times each. Therefore, we recorded 32 test trials with the test matrix for each m-sequence and participant. In summary, the experiment had 5 blocks of 30 calibration trials and 32 test trials per participant. Importantly, each trial consisted of 10 stimulation cycles

Table 2. Details of the modified version of EEG-Inception for c-VEP detection

Block	Type	Filt.	Depth	Kernel	Padding	Output	Conn. to
IN	Input	–	–	–	–	$100 \times 16 \times 1$	C1, C2, C3
C1	Conv2D	12	–	50×1	Same	$100 \times 16 \times 12$	CO1
C2	Conv2D	12	–	25×1	Same	$100 \times 16 \times 12$	CO1
C3	Conv2D	12	–	12×1	Same	$100 \times 16 \times 12$	CO1
CO1	Concatenate	–	–	–	–	$100 \times 1 \times 36$	A1
A1	AveragePooling2D	–	–	2×1	–	$50 \times 1 \times 36$	D1
D1	DepthwiseConv2D	–	2	1×16	Valid	$50 \times 1 \times 72$	A1
A2	AveragePooling2D	–	–	2×1	–	$25 \times 1 \times 72$	C4, C5, C6
C4	Conv2D	12	–	12×1	Same	$25 \times 1 \times 12$	CO2
C5	Conv2D	12	–	6×1	Same	$25 \times 1 \times 12$	CO2
C6	Conv2D	12	–	3×1	Same	$25 \times 1 \times 12$	CO2
CO2	Concatenate	–	–	–	–	$25 \times 1 \times 36$	A3
A3	AveragePooling2D	–	–	2×1	–	$12 \times 1 \times 36$	C7
C7	Conv2D	18	–	6×1	Same	$12 \times 1 \times 18$	A4
A4	AveragePooling2D	–	–	2×1	–	$6 \times 1 \times 18$	C8
C8	Conv2D	9	–	3×1	Same	$6 \times 1 \times 9$	A5
A5	AveragePooling2D	–	–	2×1	–	$3 \times 1 \times 9$	OUT
OUT	Dense	–	–	–	–	p	–

The type specifies the class of each block in Keras framework [3]. All convolutional blocks (i.e., Conv2D and DethpwiseConv2D) include batch normalization, activation and dropout regularization (dropout rate of 0.15). The model has 26948 parameters, of which 26606 are fitted during training. The base of the m-sequence p is the number of output classes of the model, one for each shade of gray.

(i.e., repetitions of the stimulation sequence). This allows to analyze the performance of the system as a function of this parameter. More stimulation cycles will increase the decoding accuracy of the system at the expense of reducing the selection speed, as more stimulation time is needed. The order of the blocks was randomized across participants to avoid bias. Users were not aware of which specific p-ary m-sequence was used to avoid unintentional biases.

3 Results

The results of the experiments are presented in Tables 3, 4, 5, 6 and 7. Each of these tables show the command decoding accuracy per participant and number of cycles considered in the analysis for each m-sequence. The mean accuracy across cycles for each subject is detailed at the left side of the tables, whereas the mean accuracy across subjects for each number of cycles is reported at the bottom. Moreover, the grand-average accuracy provided in the bottom-right corner gives an overall picture of the system's performance in a single metric.

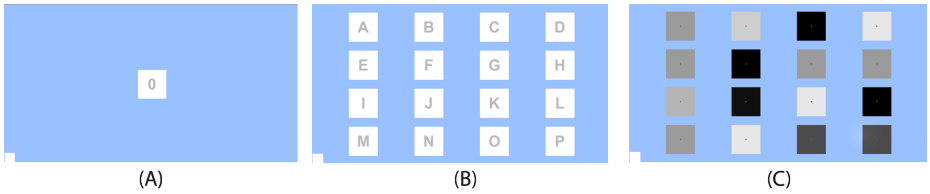

Fig. 1. Screenshots of the "P-ary c-VEP Speller" app in MEDUSA©. (A) Calibration matrix with 1 command that is highlighted following the original p-ary m-sequence. (B) Test matrix with 16 commands that are highlighted following shifted versions of the p-ary m-sequence. (C) Screenshot during the stimulation period for $GF(11^2)$.

Table 3. Command decoding accuracy (%) for base $p = 2$.

Subj.	Cycles										**Mean**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
1	87.5	96.9	100	100	100	100	100	100	100	100	**98.4**
2	96.9	100	100	100	100	100	100	100	100	100	**99.7**
3	93.8	100	100	100	100	100	100	100	100	100	**99.4**
4	96.9	100	100	100	100	100	100	100	100	100	**99.7**
5	75	93.8	100	100	100	100	100	100	100	100	**96.9**
6	100	100	100	100	100	100	100	100	100	100	**100**
7	96.9	100	100	100	100	100	100	100	100	100	**99.7**
8	87.5	96.9	100	100	100	100	100	100	100	100	**98.4**
9	93.8	100	100	100	100	100	100	100	100	100	**99.4**
10	93.8	100	100	100	100	100	100	100	100	100	**99.4**
11	75	87.5	90.6	93.8	93.8	93.8	93.8	93.8	96.9	100	**91.9**
12	53.1	84.4	93.8	96.9	96.9	96.9	96.9	96.9	96.9	96.9	**90.9**
13	90.6	96.9	96.9	100	100	100	100	100	100	100	**98.4**
14	50	71.9	84.4	84.4	96.9	93.8	96.9	100	100	100	**87.8**
15	93.8	100	100	100	100	100	100	100	100	100	**99.4**
16	81.2	96.9	100	100	100	100	100	100	100	100	**97.8**
Mean	**85.4**	**95.3**	**97.9**	**98.4**	**99.2**	**99.0**	**99.2**	**99.4**	**99.6**	**99.8**	**97.3**

The results show that the proposed approach achieved high command decoding accuracy for all m-sequences. Unsurprisingly, the number of cycles has a positive impact in the systems' performance: more cycles, which imply longer stimulation times, increase the command decoding accuracy. The results indicate that binary m-sequences with high-contrast black and white stimulus achieved the highest overall performance, reaching a grand-average accuracy of 97.3%. Non-binary m-sequences achieved slightly lower performances in the range between 90.9% for base 3 and 93.7% for base 5.

4 Discussion

In this study, the feasibility of bit-wise reconstruction of non-binary stimulation patterns from EEG data using deep learning has been tested. To this end, we

Table 4. Command decoding accuracy (%) for base $p = 3$.

Subj.	Cycles										Mean
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
1	75	96.9	96.9	100	100	100	100	100	100	100	**96.9**
2	90.6	93.8	93.8	93.8	93.8	93.8	96.9	100	100	100	**95.6**
3	96.9	96.9	100	100	96.9	96.9	100	100	100	100	**98.8**
4	87.5	96.9	100	100	100	100	100	100	100	100	**98.4**
5	37.5	78.1	96.9	93.8	96.9	100	100	100	100	100	**90.3**
6	87.5	90.6	93.8	93.8	93.8	93.8	93.8	96.9	96.9	96.9	**93.8**
7	96.9	100	100	100	100	100	100	100	100	100	**99.7**
8	96.9	100	100	100	100	100	100	100	100	100	**99.7**
9	59.4	71.9	81.2	84.4	90.6	90.6	93.8	93.8	90.6	90.6	**84.7**
10	96.9	100	100	100	100	100	100	100	100	100	**99.7**
11	46.9	53.1	68.8	71.9	71.9	68.8	78.1	78.1	81.2	81.2	**70**
12	40.6	59.4	78.1	90.6	87.5	93.8	96.9	96.9	96.9	96.9	**83.8**
13	96.9	100	100	100	96.9	100	96.9	100	100	100	**99.1**
14	15.6	28.1	31.2	43.8	50	50	50	59.4	59.4	75	**46.2**
15	87.5	93.8	100	100	100	100	100	100	100	100	**98.1**
16	96.9	100	100	100	100	100	100	100	100	100	**99.7**
Mean	**75.6**	**85.0**	**90.0**	**92.0**	**92.4**	**93.0**	**94.1**	**95.3**	**95.3**	**96.3**	**90.9**

Table 5. Command decoding accuracy (%) for base $p = 5$.

Subj.	Cycles										Mean
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
1	87.5	100	100	100	100	100	100	100	100	100	**98.8**
2	90.6	96.9	96.9	100	100	100	100	100	100	100	**98.4**
3	100	100	100	100	100	100	100	100	100	100	**100**
4	90.6	93.8	93.8	93.8	93.8	93.8	93.8	96.9	96.9	96.9	**94.4**
5	31.2	75	90.6	84.4	93.8	96.9	96.9	100	100	100	**86.9**
6	100	96.9	100	100	100	100	100	100	100	100	**99.7**
7	93.8	93.8	100	100	100	100	100	100	100	100	**98.8**
8	93.8	96.9	100	100	100	100	100	100	100	100	**99.1**
9	87.5	96.9	96.9	96.9	96.9	100	100	100	100	100	**97.5**
10	84.4	100	100	100	100	100	100	100	100	100	**98.4**
11	56.2	68.8	75	78.1	78.1	75	75	75	78.1	78.1	**73.8**
12	34.4	59.4	68.8	81.2	84.4	90.6	96.9	100	100	100	**81.6**
13	87.5	100	100	100	100	100	100	100	100	100	**98.8**
14	53.1	65.6	71.9	75	81.2	81.2	87.5	87.5	87.5	90.6	**78.1**
15	78.1	96.9	100	100	100	100	100	100	100	100	**97.5**
16	90.6	96.9	96.9	96.9	100	100	100	100	100	100	**98.1**
Mean	**78.7**	**89.8**	**93.2**	**94.1**	**95.5**	**96.1**	**96.9**	**97.5**	**97.7**	**97.9**	**93.7**

adapted EEG-Inception for multi-class classification of VEPs elicited by stimuli with different shades of gray. This approach was evaluated in a c-VEP-based BCI, reaching high command decoding accuracy.

Table 6. Command decoding accuracy (%) for base $p = 7$.

Subj.	Cycles										Mean
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
1	71.9	90.6	96.9	100	100	100	100	100	100	100	**95.9**
2	65.6	87.5	100	100	100	100	100	100	100	100	**95.3**
3	96.9	100	100	100	100	100	100	100	100	100	**99.7**
4	71.9	90.6	100	100	100	100	100	100	100	100	**96.2**
5	68.8	84.4	93.8	100	100	96.9	96.9	100	100	100	**94.1**
6	78.1	87.5	87.5	90.6	90.6	90.6	90.6	93.8	93.8	96.9	**90**
7	96.9	90.6	100	100	100	100	100	100	100	100	**98.8**
8	75	90.6	96.9	96.9	96.9	100	100	100	100	100	**95.6**
9	84.4	90.6	96.9	96.9	96.9	96.9	96.9	96.9	96.9	96.9	**95**
10	81.2	100	100	100	100	100	100	100	100	100	**98.1**
11	43.8	46.9	56.2	68.8	75	71.9	71.9	75	75	75	**65.9**
12	59.4	62.5	93.8	96.9	100	100	100	100	100	100	**91.2**
13	81.2	90.6	87.5	93.8	96.9	100	100	100	100	100	**95**
14	46.9	59.4	71.9	78.1	75	81.2	84.4	81.2	87.5	84.4	**75**
15	96.9	100	96.9	100	100	100	100	100	100	100	**99.4**
16	90.6	100	100	100	100	100	100	100	100	100	**99.1**
Mean	**75.6**	**85.7**	**92.4**	**95.1**	**95.7**	**96.1**	**96.3**	**96.7**	**97.1**	**97.1**	**92.8**

Table 7. Command decoding accuracy (%) for base $p = 11$.

Subj.	Cycles										Mean
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
1	96.9	100	100	100	100	100	100	100	100	100	**99.7**
2	87.5	93.8	100	100	100	100	100	100	100	100	**98.1**
3	96.9	100	100	100	100	100	100	100	100	100	**99.7**
4	100	100	100	100	100	100	100	100	100	100	**100**
5	75	96.9	96.9	96.9	96.9	100	100	100	100	100	**96.2**
6	100	100	100	100	100	100	100	100	100	100	**100**
7	96.9	100	100	100	100	100	100	100	100	100	**99.7**
8	96.9	100	100	100	100	100	100	100	100	100	**99.7**
9	46.9	53.1	50	53.1	56.2	56.2	53.1	53.1	53.1	53.1	**52.8**
10	93.8	93.8	100	100	100	100	100	100	100	100	**98.8**
11	50	59.4	68.8	84.4	75	81.2	78.1	84.4	81.2	81.2	**74.4**
12	75	90.6	100	100	100	100	100	100	100	100	**96.6**
13	100	100	100	100	100	100	100	100	100	100	**100**
14	37.5	50	65.6	81.2	81.2	87.5	87.5	90.6	96.9	93.8	**77.2**
15	90.6	96.9	100	100	100	100	100	100	100	100	**98.8**
16	78.1	93.8	96.9	100	100	100	100	100	100	100	**96.9**
Mean	**82.6**	**89.3**	**92.4**	**94.7**	**94.3**	**95.3**	**94.9**	**95.5**	**95.7**	**95.5**	**93.0**

The stimulation paradigm that has been used in this work has 2 key features designed to improve user-friendliness of c-VEP-based BCIs. First, we used a presentation rate of 120 Hz, in contrast to 60 Hz as in the majority of c-VEP studies [7]. In this respect, there is a consensus in the literature stating that higher pre-

sentation rates result in less visual fatigue while reducing the selection time [4]. In addition, non-binary stimulation patterns, which present stimulus with lower contrast due to the utilization of intermediate levels of gray, are presumably less prone to provoke visual fatigue and eyestrain [9]. In this study, we have tested stimulation patterns with 3, 5, 7 and 11 shades of gray, showing the feasibility of these non-binary sequences [6]. Given the advantages of our system, it could significantly enhance the user-friendliness of applications that demand continuous control, such as assistive systems designed for severely disabled people.

The proposed method for bit-wise reconstruction of non-binary stimulation patterns has the potential to enhance the development of c-VEP-based BCIs by providing a more general classification framework that can be used in a wide range of applications. In contrast to more extended methods based on CCA, which require EEG recordings for all the stimulation sequences used in the system, bit-wise reconstruction methods showed that they are able to cope with arbitrary binary stimulation patterns once calibrated [9]. The successful bit-wise decoding of non-binary sequences achieved in this work, with up to 11 levels of gray, is an indication of the potential for our strategy to handle more complex and diverse patterns of neural activity. This can help to improve the usability of c-VEP-based BCIs, making them more user-friendly while maintaining their reliability and efficiency. Additionally, this approach can provide new insights into the neural mechanisms underlying visual perception, which can be further explored in future studies. Overall, this study highlights the promising potential of bit-wise reconstruction methods for advancing c-VEP-based BCIs and other neurotechnologies.

Regarding the command decoding results of our evaluation experiments, there are several points that are worth discussing. All the tested m-sequences reached accuracies above 95% with high selection speed. However, there are significant differences between the command decoding accuracy achieved with the binary m-sequence and the rest ($p - value < 0.05$, Wilcoxon Signed Rank Test). In this regard, it would be desirable to reduce these differences as much as possible. Further improvements in the architecture of EEG-Inception could enhance multi-class classification. Interestingly, the system's performance did not decrease more for the sequences with more classes (e.g., $p = 7$ or $p = 11$), despite that these patterns provide less training examples per-class. Therefore, stimulation sequences with more levels of gray, which could be more comfortable for the user, would improve the system's usability without affecting performance. Similarly, the length of the m-sequences (see Table 1) did not have as much influence as it could be expected in advance, given that longer sequences have more stimuli per cycle. Thus, the length of the stimulation sequences could be reduced in future designs to increase selection speed. In this case, random stimulation patterns, instead of m-sequences with predefined length, could be more practical [2].

Despite the promising results of this work, we acknowledge several limitations. We focused on the adaptation of our bit-wise reconstruction strategy for non-binary sequences. Thus, we used well-tested stimulation patterns, such as

the m-sequences, which show very low correlation between shifted versions [7]. Although this is a desirable characteristic to avoid selection errors, the use of m-sequences limits the total number of commands that can be encoded in the system. In future studies, our strategy should be tested with other types of stimulation patterns to enable an arbitrary number of commands, making an in-depth comparison with the circular shifting paradigm. The proposed deep learning framework could also be enhanced by applying transfer learning to increase the performance, robustness and calibration time of the system [13]. In this regard, a pretraining stage with data from other subjects could be beneficial and should be tested in the future. Finally, the proposed system has been tested with healthy subjects. Although this is an important step to show the feasibility of our approach, further validation in more practical settings with severely disabled subjects must be addressed in future works.

5 Conclusion

This study evaluated for the first time the feasibility of a bit-wise reconstruction approach of non-binary stimulation patterns for c-VEP-based BCIs using deep learning. Non-binary sequences, which use low-contrast stimuli, could reduce visual fatigue and eyestrain in these systems. The results showed that the proposed approach based on EEG-Inception achieved high command decoding accuracy regardless of the number of levels of gray used to encode the different commands. Therefore, the stimulation paradigm used in this work, which combines a high stimulus presentation rate at 120 Hz with non-binary sequences, together with our bit-wise reconstruction method, increase the usability of c-VEP-based BCIs while maintaining their characteristic reliability and efficiency.

Acknowledgements This study was partially funded by "Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación" and European Regional Development Fund (ERDF) [TED2021-129915B-I00, RTC2019-007350-1, and PID2020-115468RB-I00]; and "CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)" through "Instituto de Salud Carlos III". E. Santamaría-Vázquez, was in receipt of a PIF grant by the "Consejería de Educación de la Junta de Castilla y León".

References

1. Abiri, R., Borhani, S., Sellers, E.W., Jiang, Y., Zhao, X.: A comprehensive review of eeg-based brain–computer interface paradigms. Journal of Neural Engineering **16**, 011001 (2 2019). https://doi.org/10.1088/1741-2552/aaf12e, https://iopscience.iop.org/article/10.1088/1741-2552/aaf12e
2. Buračas, G.T., Boynton, G.M.: Efficient design of event-related fmri experiments using m-sequences. NeuroImage **16**, 801–813 (7 2002). https://doi.org/10.1006/nimg.2002.1116
3. Chollet, F.: Keras (2015), https://keras.io

4. Gembler, F.W., Rezeika, A., Benda, M., Volosyak, I.: Five shades of grey: Exploring quintary m -sequences for more user-friendly c-vep-based bcis. Computational Intelligence and Neuroscience **2020**, 1–11 (3 2020). https://doi.org/10.1155/2020/7985010
5. Lin, Z., Zhang, C., Wu, W., Gao, X.: Frequency recognition based on canonical correlation analysis for ssvep-based bcis. IEEE Transactions on Biomedical Engineering **53**, 2610–2614 (2006)
6. Martínez-Cagigal, V., Santamaría-Vázquez, E., Pérez-Velasco, S., Marcos-Martínez, D., Moreno-Calderón, S., Hornero, R.: Non-binary m-sequences for more comfortable brain-computer interfaces based on c-veps. Expert Systems With Applications (under review) (2023)
7. Martínez-Cagigal, V., Thielen, J., Santamaría-Vázquez, E., Pérez-Velasco, S., Desain, P., Hornero, R.: Brain-computer interfaces based on code-modulated visual evoked potentials (c-vep): A literature review. Journal of Neural Engineering **18**, 1–21 (2021)
8. Nagel, S., Spüler, M.: Modelling the brain response to arbitrary visual stimulation patterns for a flexible high-speed brain-computer interface. PLOS ONE **13** (10 2018). https://doi.org/10.1371/journal.pone.0206107
9. Nagel, S., Spüler, M.: World's fastest brain-computer interface: Combining eeg2code with deep learning. bioRxiv (2019)
10. Santamaría-Vázquez, E., Martinez-Cagigal, V., Vaquerizo-Villar, F., Hornero, R.: Eeg-inception: A novel deep convolutional neural network for assistive erp-based brain-computer interfaces. IEEE Transactions on Neural Systems and Rehabilitation Engineering **28**, 2773–2782 (2020)
11. Santamaría-Vázquez, E., Martínez-Cagigal, V., Gomez-Pilar, J., Hornero, R.: Asynchronous control of erp-based bci spellers using steady-state visual evoked potentials elicited by peripheral stimuli. IEEE transactions on neural systems and rehabilitation engineering **27**, 1883–1892 (2019)
12. Santamaría-Vázquez, E., Martínez-Cagigal, V., Marcos-Martínez, D., Rodríguez-González, V., Pérez-Velasco, S., Moreno-Calderón, S., Hornero, R.: Medusa©: A novel python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research. Computer Methods and Programs in Biomedicine **230**, 107357 (2023). https://doi.org/10.1016/j.cmpb.2023.107357
13. Santamaría-Vázquez, E., Martínez-Cagigal, V., Pérez-Velasco, S., Marcos-Martínez, D., Hornero, R.: Robust asynchronous control of erp-based brain-computer interfaces using deep learning. Computer Methods and Programs in Biomedicine **215**, 1–10 (2022)
14. Wolpaw, J., Wolpaw, E.W.: Brain-computer interfaces: principles and practice. OUP USA (2012)