NeuroImage 284 (2023) 120446

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Burst c-VEP Based BCI: Optimizing stimulus design for enhanced classification with minimal calibration data and improved user experience

Kalou Cabrera Castillos [a,*], Simon Ladouce [a], Ludovic Darmet [a], Frédéric Dehais [a,b]

[a] Human Factors and Neuroergonomics, Institut Supérieur de l'Aéronautique et de l'Espace, 10 Av. Edouard Belin, Toulouse, 31400, France
[b] Biomedical Engineering, Drexel University, Philadelphia, 19104, PA, United States

ARTICLE INFO

Keywords:
Code-VEP
Reactive BCI
CNN
Amplitude depth reduction
Visual comfort

ABSTRACT

The utilization of aperiodic flickering visual stimuli under the form of code-modulated Visual Evoked Potentials (c-VEP) represents a pivotal advancement in the field of reactive Brain–Computer Interface (rBCI). A major advantage of the c-VEP approach is that the training of the model is independent of the number and complexity of targets, which helps reduce calibration time. Nevertheless, the existing designs of c-VEP stimuli can be further improved in terms of visual user experience but also to achieve a higher signal-to-noise ratio, while shortening the selection time and calibration process. In this study, we introduce an innovative variant of code-VEP, referred to as "Burst c-VEP". This original approach involves the presentation of short bursts of aperiodic visual flashes at a deliberately slow rate, typically ranging from two to four flashes per second. The rationale behind this design is to leverage the sensitivity of the primary visual cortex to transient changes in low-level stimuli features to reliably elicit distinctive series of visual evoked potentials. In comparison to other types of faster-paced code sequences, *burst* c-VEP exhibit favorable properties to achieve high bitwise decoding performance using convolutional neural networks (CNN), which yields potential to attain faster selection time with the need for less calibration data. Furthermore, our investigation focuses on reducing the perceptual saliency of c-VEP through the attenuation of visual stimuli contrast and intensity to significantly improve users' visual comfort. The proposed solutions were tested through an offline 4-classes c-VEP protocol involving 12 participants. Following a factorial design, participants were instructed to focus on c-VEP targets whose pattern (burst and maximum-length sequences) and amplitude (100% or 40% amplitude depth modulations) were manipulated across experimental conditions. Firstly, the full amplitude *burst* c-VEP sequences exhibited higher accuracy, ranging from 90.5% (with 17.6 s of calibration data) to 95.6% (with 52.8 s of calibration data), compared to its m-sequence counterpart (71.4% to 85.0%). The mean selection time for both types of codes (1.5 s) compared favorably to reports from previous studies. Secondly, our findings revealed that lowering the intensity of the stimuli only slightly decreased the accuracy of the *burst* code sequences to 94.2% while leading to substantial improvements in terms of user experience. Taken together, these results demonstrate the high potential of the proposed *burst* codes to advance reactive BCI both in terms of performance and usability. The collected dataset, along with the proposed CNN architecture implementation, are shared through open-access repositories.

1. Introduction

Steady-State Visually Evoked Potentials (SSVEPs) refer to the sustained rhythmic brain activity elicited by repetitive visual stimuli and are detected over the occipital cortex through surface electroencephalography (EEG) (Kritzman et al., 2022; Zhang et al., 2021). While modern template-based approaches have established SSVEPs as a leading paradigm for designing reactive Brain–Computer Interfaces (BCIs) and achieving unparalleled Information Transfer Rates (ITR) (Chevallier et al., 2021; Zerafa et al., 2018; Nakanishi et al., 2018), this method's reliance on templates presents end-user issues, notably an extended calibration period (e.g., 40 min in Chen et al., 2022). Recent studies (Thielen et al., 2021; Nagel and Spüler, 2019b; Martínez-Cagigal et al., 2021, 2023) have highlighted potential solutions to these limitations by utilizing code-modulated Visual Evoked Potentials (c-VEP) that substitutes the periodic flickers with aperiodic, binary (zeroes and ones corresponding to black/off and white/on states of the flicker), random code sequences — for a detailed review, see Martínez-Cagigal et al. (2021).

* Corresponding author.
E-mail address: kalou.cabrera-castillos@isae-supaero.fr (K. Cabrera Castillos).

https://doi.org/10.1016/j.neuroimage.2023.120446
Received 18 August 2023; Received in revised form 31 October 2023; Accepted 6 November 2023
Available online 8 November 2023

While brain activity associated with distinct c-VEP stimuli can be discriminated using template-based approaches (Martínez-Cagigal et al., 2023; Liu et al., 2018; Wei et al., 2018; Bin et al., 2011), alternative decoding methods have been explored. These approaches (Thielen et al., 2015; Nagel and Spüler, 2019a,b) aim to establish a direct relationship between EEG signals extracted through short sliding windows and the elementary bits of the c-VEP code sequence, represented as either *off* (zero) or *on* (one) states. By sequentially decoding each bit, a reconstructed sequence is generated, which can then be compared to the stimulation patterns associated with the different targets. This approach allows the model to be trained using a limited number of trials (i.e., presentation of the entire code), as many instances of the *on* and *off* states are presented within each trial. Moreover, it offers flexibility in decoding time, as this approach does not require the complete reconstruction of the c-VEP sequence in its entirety, as opposed to classical template-based approaches that rely on the matching of fixed epoch lengths. Nagel and Spüler (2019a) have referred to this approach as "EEG2Code", while Thielen et al. (2015) named it "Re(con)volution".

Bitwise decoding utilizes short sliding windows with less information than template-based approaches employing full epoch stimulation data. Such an approach requires handling significant data volumes as information aggregated across several windows is required to decode the whole stimulation pattern. Moreover, this approach only represents an improvement over previous template-based approaches if the target stimulation can be effectively decoded based on the accumulation of a limited number of windows (shorter than the whole duration of the stimulation pattern), hence requiring higher discriminative power. To achieve this, Nagel and Spüler (2019b) improved the EEG2code approach by introducing convolutional neural networks (CNNs) to the approach. The CNN implementation presents the advantage of handling nonlinear aspects of neural responses. This updated approach has exhibited remarkable performance, boasting exceptional Information Transfer Rates (ITR) along with unprecedentedly swift selection times in less than 2 s (Nagel and Spüler, 2019b; Li and Huang, 2021). Moreover, the CNN implementation to the approach offers interesting prospects for transfer learning (Huang et al., 2023).

While CNNs offer unique advantages that have the potential to further enhance BCI experience from an end-user perspective (i.e., high responsiveness, shorter calibration), several challenges need to be considered. Notably, the training of a CNN has a substantial cost both in terms of computational power and time (e.g., 3 min in Darmet et al., 2023). Ideally, the development of BCI should thrive toward solutions that can be readily run on embedded hardware systems with minimal setup time (Fairclough and Lotte, 2020). In the present study, we present a solution for addressing this issue by optimizing the design of c-VEP stimuli within the framework of a CNN decoding approach. Currently, the maximum-length sequence (*m-sequence*, or *m-seq* for short) is one of the most popular techniques used for c-VEP implementation, which relies on pseudo-random and aperiodic binary sequences alternating short plateaus of "one" (visual stimuli on) and "zero" (visual stimuli off). Our contribution is to propose an alternative to the m-sequences through the use of less frequent and more brief bursts of visual stimuli (i.e., ones), offering two significant advantages. First, it leverages the sensitivity of the primary visual cortex to short-lasting, aperiodic light stimulations, which are known to elicit robust visual event-related responses (Luo and Ding, 2020; Regan, 1989). The rationale is to ease the decoding and classification processes by enhancing the amplitude of the VEP. Secondly, this approach simplifies the classification task into a two-class problem, distinguishing between the presence of the *burst* stimulus (*on*) and its absence (*off*), rather than a complex template-matching problem. This reduction in complexity is expected to reduce computational costs and further shorten CNN training times.

Another significant improvement pursued in this study involves enhancing visual comfort in c-VEP BCIs. Flickering stimuli can lead to user discomfort, ranging from mild visual irritation (Volosyak et al., 2011) to sustained eye strain (Zhu et al., 2010), causing mental fatigue or even resulting in drowsiness (Cao et al., 2014; Ortner et al., 2011; Patterson Gentile and Aguirre, 2020). A potential solution to enhance visual comfort involves increasing the flickering frequency to render the visual stimuli less perceptible to the user. For instance, recent c-VEP studies (Martínez-Cagigal et al., 2023; Başaklar et al., 2019) disclosed that increasing the presentation rate of the aperiodic stimuli considerably improves visual user experience. However, these stimuli still remain visible and this approach has detrimental effects on the discriminability of the codes due to the reduced orthogonality among templates as they begin to overlap more extensively (Başaklar et al., 2019). An alternative solution to improve user experience and visual comfort is to reduce the contrast and intensity of visual stimuli by attenuating their amplitude modulation depth. Stimulus amplitude depth refers to the contrast difference between the two antagonist states of a c-VEP (or SSVEP) stimulus. In a recent study conducted by our research group, we demonstrated that reducing the amplitude depth of periodic flickers significantly enhances the user experience while maintaining a high classification accuracy, as evidenced in Ladouce et al. (2022). Adopting a similar approach for c-VEP stimuli holds the potential to elevate visual comfort. This approach aligns with the findings of Gembler et al. (2020) and Martínez-Cagigal et al. (2023), where they utilized a "p-ary m-sequences" comprising various shades from white to grey to assure a more comfortable visual experience than binary black-and-white m-sequences.

The primary focus of this study is to address end-user-related challenges in the c-VEP BCI paradigm. Our efforts are directed toward minimizing calibration time, optimizing performance, and enhancing overall visual comfort for users. To achieve these objectives, we are introducing an innovative *burst* c-VEP sequence design. This design aims to leverage the capabilities of Convolutional Neural Networks (CNNs) to achieve greater accuracy and reduced selection times, all while minimizing the demands on training and computational resources. In addition to the burst-type design, our approach is to prioritize the comfort of users during c-VEP interactions. Building on our previous research on SSVEP (Ladouce et al., 2022, 2021), we opted for a reduction of amplitude modulation depth down to 40% of the maximal amplitude between two states, as such a reduction achieves the best compromise between retaining high classification performance and visual comfort improvement. We anticipate that this reduction, combined with the robust responses elicited by the burst-type stimuli, will not only enhance visual comfort but also maintain classification accuracy, further optimizing the c-VEP BCI end-user experience. Using a factorial experimental design, burst and m-sequence c-VEP will be presented at either 100% or 40% of the maximal amplitude modulation depth. Our primary objective is to evaluate the efficacy of the *burst* c-VEP compared to the m-sequence c-VEP, while also examining the effects of amplitude reduction on classification performance and user experience. For this purpose, we will report various metrics of BCI performance, including CNN training time, classification accuracy, and selection time. Additionally, we will investigate the relationship between the number of data points used for calibration and these metrics. In order to provide a comprehensive analysis, we will also examine evoked responses and inter-trial coherence (ITC) metrics. These analyses can reveal the distinct brain responses elicited by the *burst* c-VEP in comparison to the m-sequence c-VEP. Finally, we conduct subjective assessments to evaluate several dimensions of user experience such as visual comfort, visual tiredness, and intrusiveness associated with each type of code and amplitude depth. These assessments aim to provide valuable insights into the subjective experiences of the participants.

2. Material and methods

2.1. Participants

Twelve healthy volunteers (4 women, mean age: 30.6 years, standard deviation: 7.1), all students and staff at ISAE-SUPAERO with

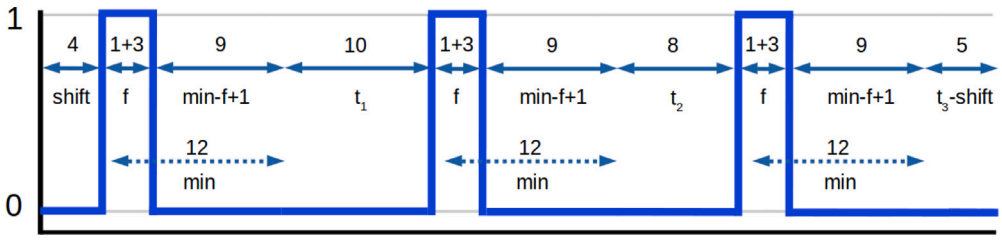

Fig. 1. Burst code from compact representation (4, 12, [10, 8, 9], 4)$_{60}$. To create the unfolded sequence, we iterate through all the additional times in the sequence, adding the *burst* frames ($f = 4$), the minimal duration ($min = 12$, minus the burst frames, thus adding only $min - f + 1$) and the current additional duration ($t_1 = 10$, $t_2 = 8$ and $t_3 = 9$). The last step is to circularly shift the code to the right, by 4 frames, giving the final *burst* code. The total number of frames for this code amounts to $3 \times 4 + 3 \times 9 + 10 + 8 + 9 = 76$ frames and since the code is made for a 60 Hz screen refresh rate, the code duration is then meant to be 1.26 s.

normal or corrected-to-normal vision participated in this original study conducted in our research department. None of the participants reported any of the exclusion criteria (neurological antecedents and being under psychoactive medication at the time of the study). The study was approved by the ethics committee of the University of Toulouse (CER approval number 2020-334) and was carried in accordance with the declaration of Helsinki. Participants gave informed written consent prior to the experiment. The anonymized data collected within the frame of this study are available at https://zenodo.org/record/8255618.

2.2. Stimuli design

We utilized two types of code design to generate the visual stimulation patterns in our study: maximum-length sequences (m-sequence), which are widely recognized as one of the most popular techniques for implementing c-VEP (Martínez-Cagigal et al., 2021), and the proposed *burst* codes.

The *bursts* codes consist of brief flashes (typically around 50 ms), referred to as *bursts*, with a minimum interval of 200 ms between the onset of one *burst* and the onset of the next. Additionally, we incorporated a variable duration between consecutive bursts to create a code rather than a periodic signal. The maximum duration between consecutive flashes is limited to a fixed bound, set at 500 ms. More formally, we can represent a *burst* sequence using the compact 4-tuple notation $(f, min, seq, shift)_R$, where f is the duration of a *burst*, *min* is the minimal duration between the onset of two *bursts*, $seq = [t_1, \ldots, t_n]$ is the sequence of the variable part between each *burst*, including the minimal duration, *shift* is the phase of the code (circular right-shift), and R is the screen refresh rate the code is designed for (as the duration is expressed in frames rather than milliseconds). This notation contains all the necessary information to generate the code data array. Refer to Fig. 1 for an illustration of how they relate.

In assessing the correlation across distinct *burst* sequences, we capitalize on the concept of a burst's "neighborhood" – instances where bursts from divergent codes are in closer proximity, thus exhibiting heightened correlation. By examining the neighborhood surrounding bursts within one code concerning another, we ascertain their relative closeness. In the context of a code c_1, each individual *burst* undergoes scrutiny for its distance in relation to *bursts* within a separate c_2 code. For this analysis, we establish a defined vicinity around the commencement of a burst (set at 100 ms in our instance, as we are looking for clean P100 responses). We then quantify the linear gap between the two bursts, assigning a value of 1 if both bursts initiate exactly simultaneously, and 0 if a burst lies beyond the predetermined neighborhood range. Subsequently, a mean value is computed from the accumulated correlation scores. Should this mean value exceed a predetermined threshold, it signifies an excessive proximity between the two codes in their present phase, prompting their classification as overly convergent.

One may raise the question of potential duplication in counting a burst within c_2, should it find itself in the vicinity of two c_1 bursts. This scenario, however, does not impart symmetry to the correlation under consideration. Our specific circumstances preclude such an occurrence: with an inter-burst interval of 200 ms and a neighborhood zone spanning 100 ms, even under the most unfavorable conditions (wherein no temporal gap exists between the two bursts), there remains an absence of overlap within the adjoining zones of consecutive bursts within the same code. Thus, it becomes evident that a burst within c_2 cannot simultaneously inhabit the neighboring zones of both bursts, thereby confirming the adequacy of this straightforward correlation definition in our context.

With our correlation function in place, we embark on the generation of our *burst* codes, initiating a systematic assessment of correlations through phased adjustments. Within each code pair, our focus narrows to the paramount correlation score derived from the array of scores originating from these shifted codes. A pivotal facet, this peak correlation score, is then juxtaposed against an established threshold. When the resulting score falls below this predefined threshold, it signifies a diminished likelihood of misinterpretation, an outcome resilient against variations in relative phases. As a result, the crux of our code selection process revolves around upholding minimal pairwise correlation—a fundamental criterion guiding our final code choices. The four *burst* codes designed for the purpose of this study and their visual representation can be found in the Supplementary Material Section 1.

Concerning the m-sequences, we drew upon codes originally formulated in a preceding study, achieving a 98% accuracy within an 11-class problem (Darmet et al., 2023). Our approach involved the utilization of a Fibonacci-type linear feedback shift register (LFSR), employing two distinct polynomials: $x^9 + 1$ and $x^9 + x^6 + x^3 + 1$. The initial states chosen were 11111110011, 01111010011, 10110110011, 10111110011, and 01111110011. This configuration enabled the construction of a codebook containing m-sequence-like codes. The codes derived from our codebook underwent segmentation into subsequences, each spanning 132 frames. Given the four-class nature of our experiment, we ultimately used the initial four codes as the most optimal candidates, after making sure that the correlation between them and across all their phased versions is not higher than 0.5.

Finally, in addition to the variation of the type of code, we presented both *m-sequences* and *burst* c-VEP at two different levels of amplitude depth of stimulation. The background color and the "off" state of the visual stimuli are set to a medium grey, i.e. 50% of the maximum screen luminance. This medium grey background and off state had a luminance of 124 lux (measured using a digital light meter from Extech Instruments). For the stimuli *on* states, we chose two conditions: 100% amplitude depth which is the modulation depth between the base medium grey luminance to the maximal screen luminosity (brightest white, 168 lux), and a relative 40% amplitude depth (142 lux) between the two aforementioned luminance levels. As such, the 40% amplitude modulation depth stimuli alternate between the base medium grey luminance to 40% of the medium grey — brightest white range screen luminance (as depicted in the bottom left part of Fig. 2. The rightmost part of Fig. 2 illustrates these four conditions.

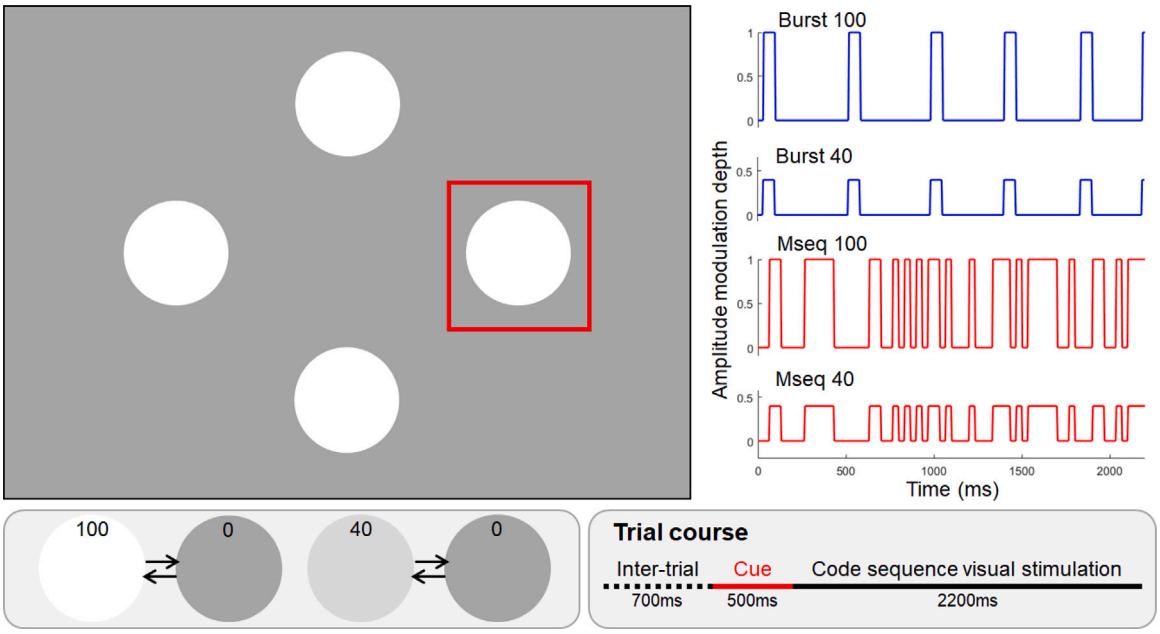

Fig. 2. Left: the presentation of *burst* and m-sequence based c-VEP varied in amplitude depth, with options for either maximum (100%) or reduced to 40%, aiming to enhance visual comfort. Top-right (blue): the brief aperiodic *burst* based c-VEP presented at a rate of 2 to 3 flashes per second. Bottom-right (red): The m-sequence is an aperiodic binary sequence that alternates short-lasting plateaus of visual stimuli *on* and *off* at a rate of approximately 10 flashes per second. The x-axis is time (in milliseconds). Bottom-left: Representation of the two alternating states of visual stimuli for both amplitude modulation depth (100 and 40%). Bottom-right: time course of a trial, an inter-trial interval of 700 ms separates the current trial from the previous one. A red-bordered square cue then appears for 500 ms around the stimulus to be attended. The cue is followed by the onset of visual stimulation where all the different codes corresponding to the current conditions (Burst 100, Burst 40, Mseq 100, Mseq 40) are presented for 2200 ms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.3. Experimental protocol

Participants were comfortably seated and instructed to read and sign the informed consent. EEG data were recorded using a BrainProduct LiveAmp 32 active electrodes wet-EEG setup with a sample rate of 500 Hz to record the surface brain activity. The 32 electrodes were placed following the 10–20 international system on a BrainProduct Acticap. The ground electrode was placed at the FPz electrode location and all electrodes were referenced to the FCz electrode. The impedance of all electrodes was brought below 25kΩ prior to recording onset.

Once equipped with the EEG system, volunteers were asked to focus on four targets that were cued sequentially in a random order for 0.5 s, followed by a 2.2 s stimulation phase, before a 0.7 s inter-trial period. The cue sequence for each trial was pseudo-random and different for each block. After each block, a pause was observed and subjects had to press the space bar to continue. The participants were presented with fifteen blocks of four trials for each of the four conditions (burst or m-sequence × 40% or 100%), see Fig. 2 - left. The task was implemented in Python using the Psychopy toolbox.[1] The four discs were all 150 pixels, without borders, and were presented on the following LCD monitor: Dell P2419HC, 1920 × 1080 pixels, 265 cd/m^2, and 60 Hz refresh rate. After completing the experiment and removing the EEG equipment, the participants were asked to provide subjective ratings for the different stimuli conditions. These stimuli included *burst* c-VEP with 100% amplitude, *burst* c-VEP with 40% amplitude, m-sequences with 100% amplitude, and m-sequences with 40% amplitude. Each stimulus was presented three times in a pseudo-random order. Following the presentation of each stimulus, participants were presented with three 11-points scales and were asked to rate the visual comfort, visual tiredness, and intrusiveness using a mouse. In total, participants completed 12 ratings (3 repetitions x 4 types of stimuli) for each of the three scales.

2.4. EEG pre-processing

EEG data and markers were synchronized during recording using Lab Streaming Layer.[2] EEG analyses were conducted only on a subset of specific electrodes: O1, O2, Oz, Pz, P3, P4, P8, and P9. The raw continuous EEG data were average re-referenced. An IIR cut-band filter between 49.9 and 50.1 Hz of order 16 was then applied to remove line noise. The raw continuous data were then epoched from 0 to 2.2s around timestamps of flickers onset. Finally, a baseline removal was performed on the epochs to remove eventual slow drifts.

2.5. Convolutional neural network and pattern decoding

The CNN architecture[3] used in this study is the same as in our previous work (Darmet et al., 2023). It is an improved version of EEG2Code (Nagel and Spüler, 2019b), which demonstrated superior performance in terms of selection time and accuracy compared to the original architecture. This CNN architecture comprises three blocks of convolutional layers and ends with dense layers block. Each convolutional block is defined as a sequence of the following layers: (1) a 2D convolutional layer, with the number of filters, kernel size, stride and dilation parameters dependent on the block's depth, (2) a batch normalization layer, scaled but not centered (3) a Leaky Rectified Linear Unit (LeakyReLU) activation function layer, with a small negative slope of 0.3 for negative values to address saturation issues during training (4) a 2D Max pooling with a pool size of 2, and (5) a dropout layer with a dropout rate of 0.5. The first block performs a 1D spatial filtering with 16 filters and an 8 × 1 kernel (8 being the number of electrodes) without dilation. These kernels act as spatial filters and as 16 different convolutional kernels are used, this corresponds to 16 spatial filters in total. This block notably lacks the activation function so as not to lose too much information on the feature maps at this step. The second

[1] https://www.psychopy.org/

[2] https://github.com/sccn/labstreaminglayer

[3] https://github.com/neuroergoISAE/burst_codes

block performs a 1D temporal filtering with 8 filters and a 1 × 32 kernel with a dilation of 2. As with the first block, the kernels act as temporal filters, and with the 8 filters of this convolutional block, we thus have 8 different temporal filters. The third block performs a 2D convolution with 4 filters and a 5 × 5 kernel size, that aims to integrate information across the channels of the feature maps. Finally, the dense block flattens the output of the previous block and connects to a 256-long dense layer, followed by a Leaky ReLU and finally ends in a 2-class dense layer, the output. An illustration depicting the architecture of the CNN can be found in the Supplementary Material Section 2 dedicated to the CNN.

To predict the target the subject intends to select, the decoding process consists of two phases. Firstly, similarly to Nagel and Spüler (2019b), short windows of 250 ms with 2 ms steps of EEG data are input to the CNN to decode the electrophysiological response, thus creating a sequence of ones (a VEP has been detected in the window) and zeroes. Secondly, this decoded sequence is compared to the stimulation patterns of the targets using Pearson correlation. However, the actual state switch rate of the simulation is slower due to the screen refresh rate limitation of 60 Hz. Thus, downsampling of the decoded sequence to match the screen refresh rate is performed using majority voting. In the second phase, when the decoded sequence reaches a minimal length of 0.7 s, a Pearson correlation coefficient is computed between the downsampled binarized decoded sequence and the different target templates. The label of the highest correlating template is temporarily considered as the prediction. The decoded sequence is then expanded with the next stimulation decoding using the following EEG data (i.e., the downsampled next 8 windows), and a new temporary prediction is made. If the system consecutively outputs the same temporary prediction 60 times, a closing prediction for the trial is triggered. Computation is stopped, and the classification starts for the next trial. In cases where no closing prediction could be made after 2 s, the system outputs −1, indicating no prediction, and therefore does not issue any command.

2.6. Classification analysis

In this study, we evaluated the number of needed calibration data with an iterative approach: starting with one block of calibration (2.2 s × 4 stimuli × 1 block = 8.8 s of data) and gradually increasing it to six blocks of calibration (2.2 s × 4 × 6 = 52.8 s of data). This approach allowed us to compare the effects of two stimuli (burst versus m-sequence) and amplitude depth reduction on CNN training time, 4-class selection time, and 4-class decoding accuracy. To simulate an online setting, we performed sequential train/test splits in each of these computations. The epochs were segmented into windows of 250 ms with a stride of 2 ms (one data point of EEG data sampled at 500 Hz). We computed the standard deviation within each time window of the calibration dataset enabling standard deviation normalization. This normalization value was then applied to the testing data as well. Additionally, we allocated 10% of the calibration data for validation purposes: this subset was used to monitor the loss and accuracy during the CNN training process.

2.7. VEP analysis

The effect of experimental manipulations on Visual Evoked Potentials (VEP) responses was investigated to better understand the neural underpinnings that may explain differences in terms of classification performance. All the EEG analyses were ran using Matlab (R2021b) (The MathWorks Inc., 2021) and EEGlab (2019.1) (Delorme and Makeig, 2004). For this purpose, the continuous EEG data was bandpass filtered between 0.1 and 40 Hz (FIR, 16,501 filter order, cut-off frequencies at 0.05 Hz and 40.05 Hz) and then average rereferenced. The onset of singular visual stimuli (within the whole sequence) was defined based on the timings at which the code sequences alternated from the dark to the bright states. The continuous EEG data were then segmented into epochs around the onset of visual stimulation onset (from −0.5 s to 1 s). The VEP responses were averaged across trials for each electrode and participant. The mean amplitude of the VEP (computed over a 60 ms to 110 ms time window) and its peak latency were extracted from the resulting waveforms at electrode Oz over which the VEP response is most prominent.

2.8. Statistical analyses

Statistical analyses were carried out with JASP 0.18 software (JASP Team, 2023). 2 × 2 repeated measure ANOVA, with factors of c-VEP type (burst, m-sequence) and amplitude depth (low, high), were computed for each subjective metrics (visual comfort, mental tiredness and intrusiveness) with types of codes and luminance as within factor. Similar 2 × 2 repeated measure ANOVA were computed for each objective metrics (classification accuracy, selection time, CNN training time, information transfert rate) when considering 6 blocks of calibration data (4 flickers × 2.2 s × 6 blocks = 52.8 s). We also ran some descriptive analyses to report the effect of the number of blocks of calibration trials (from one block to 6 blocks) on classification accuracy and CNN training time. Eventually, we ran the same repeated measures ANOVA to assess the effect of types of codes and amplitude depth on our visual evoked potentials metrics (amplitude, latency and inter-trial coherence). The Greenhouse–Geisser correction was applied when the assumption of sphericity was violated. The Bonferroni test was used for all post-hoc comparisons. Significance level was set at $p < 0.05$ for all analyses.

3. Results

3.1. Subjective results

The following subsections present the results of statistical analyses performed on the subjective ratings in terms of Visual Comfort, Mental Fatigue, and Intrusiveness for the different code sequences (Burst and Mseq) and amplitude (100 and 40) stimuli. These subjective results are visually presented in Fig. 3. A detailed view on the answers of the participants regarding their subjective experience can be found in the Supplementary Material Section 3.

3.1.1. Visual comfort

The repeated measure ANOVA disclosed a main effect of the type of code on visual comfort ($F(1,11) = 7.18$, $p = 0.021$, $\eta_p^2 = 0.395$) and a main effect of the amplitude depth on visual comfort ($F(1,11) = 28.03$, $p = 0.001$, $\eta_p^2 = 0.718$) but no significant interactions ($F(1,11) = 0.407$, $p = 0.53$, $\eta_p^2 = 0.036$). Post-hoc analyses revealed that the *burst* c-VEP led to better visual comfort than their m-sequence counterpart ($p = 0.021$, Cohen's d = 0.522). Post-hoc analyses also revealed that lowering the amplitude depth led to increased visual comfort ($p < 0.001$, Cohen's d = 1.418).

3.1.2. Mental tiredness

The repeated measure ANOVA highlighted a main effect of the type of code on mental tiredness ($F(1,11) = 10.043$, $p = 0.009$, $\eta_p^2 = 0.477$) and a main effect of the amplitude depth on mental tiredness ($F(1,11) = 16.584$, $p = 0.002$, $\eta_p^2 = 0.601$) but no significant interactions ($F(1,11) = 0.340$, $p = 0.571$, $\eta_p^2 = 0.030$). Post-hoc analyses revealed that the *burst* c-VEP led to lower mental tiredness than the m-sequence c-VEP ($p = 0.009$, Cohen's d = 0.437) and that reducing the amplitude decreased mental tiredness ($p = 0.002$, Cohen's d = 0.854).

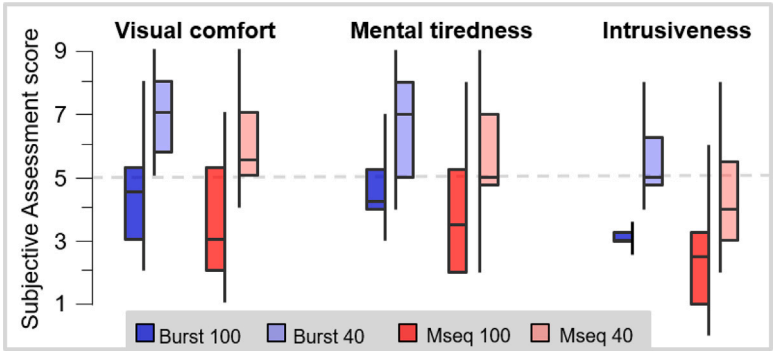

Fig. 3. Distribution (N = 12) of subjective assessment of the different visual stimulation code sequences (*Burst 100*, *Burst 40*, *Mseq 100*, *Mseq 40*) across three dimensions: visual comfort, mental tiredness, and presence.

3.1.3. Intrusiveness

The repeated measure ANOVA revealed a main effect of the type of code on intrusiveness ($F(1,11) = 7.18$, $p = 0.02$, $\eta_p^2 = 0.395$) and a main effect of the amplitude depth on intrusiveness ($F(1,11) = 31.93$, $p < 0.001$, $\eta_p^2 = 0.744$) but no significant interactions ($F(1,11) = 0.256$, $p = 0.62$, $\eta_p^2 = 0.023$). Post-hoc analyses revealed that the *burst* c-VEP were less intrusive than the m-sequence c-VEP ($p = 0.02$, Cohen's d = 0.515) and that reducing the amplitude also led to decreased intrusiveness ($p < 0.001$, Cohen's d = 1.201).

3.2. BCI performances

The following subsections present analyses performed on the main metrics commonly used to assess BCI performance such as classification accuracy, selection time, CNN training time, and Information Transfer Rate using neural data evoked by the different code sequences (Burst and Mseq) and amplitude (100 and 40) visual stimuli. The last subsection documents the impact of calibration data reduction on classification performance.

3.2.1. Classification accuracy

The repeated measure ANOVA revealed a significant main effect of the type of code on classification accuracy ($F(1,11) = 16.62$, $p = 0.002$, $\eta_p^2 = 0.602$) and a significant main effect of the amplitude depth on classification accuracy ($F(1,11) = 9.86$, $p = 0.009$, $\eta_p^2 = 0.473$). However, no significant interaction of these variables was observed on classification accuracy ($F(1,11) = 2.49$, $p = 0.14$, $\eta_p^2 = 0.185$). Post-hoc analyses indicated that the *burst* c-VEP resulted in higher accuracy compared to their *m-sequence* counterparts ($p = 0.002$, Cohen's $d = 1.177$). Similarly, post-hoc analyses revealed that reducing the amplitude depth led to a decrease in classification accuracy ($p = 0.009$, Cohen's d = 0.907) - see Fig. 4 (upper left).

3.2.2. Selection time

The repeated measure ANOVA did not reveal a significant main effect of the type of code on selection time ($F(1,11) = 1.29$, $p < 0.28$, $\eta_p^2 = 0.105$) or a significant main effect of the amplitude depth on selection time ($F(1,11) = 4.26$, $p = 0.06$, $\eta_p^2 = 0.094$). Additionally, no significant interaction was observed on selection time ($F(1,11) = 0.66$, $p = 0.43$, $\eta_p^2 = 0.016$ - see Fig. 4 (upper right)).

3.2.3. CNN training time

The repeated measure ANOVA revealed a significant main effect of the type of code on CNN training time ($F(1,11) = 36472.86$, $p < 0.001$, $\eta_p^2 = 1.0$). However, no main effect of the amplitude depth on CNN training time was observed ($F(1,11) = 15804.50$, $p < 0.001$, $\eta_p^2 = 0.999$). Additionally, no significant interactions were found on CNN training time ($F(1,11) = 0.83$, $p = 0.04$, $\eta_p^2 = 0.0001$). Furthermore, a significant interaction was observed on selection time ($F(1,11) = 7.92$, $p = 0.01$, $\eta_p^2 = 0.419$). Post-hoc analyses indicated that the *burst* c-VEP with full amplitude and 40% amplitude resulted in shorter training time compared to their *m-sequence* counterparts ($p < 0.001$). Additionally, the full amplitude *m-sequence* exhibited shorter training time compared to the 40% amplitude *m-sequence* ($p = 0.03$).

3.2.4. Information transfer rate

The repeated measure ANOVA revealed a significant main effect of the type of code on Information Transfer Rate (ITR) ($F(1,11) = 30.144$, $p < 0.001$, $\eta_p^2 = 0.733$) and a significant main effect of the amplitude depth on ITR ($F(1,11) = 8.892$, $p = 0.012$, $\eta_p^2 = 0.447$). However, no significant interactions were found on ITR ($F(1,11) = 0.529$, $p = 0.482$, $\eta_p^2 = 0.046$). Post-hoc analyses revealed that the *burst* c-VEP resulted in higher ITR compared to their *m-sequence* counterparts ($p < 0.001$, Cohen's d = 1.585). Similarly, post-hoc analyses revealed that reducing the amplitude depth led to a decrease in ITR ($p = 0.012$, Cohen's d = 0.861). The mean ITR values for each condition were as follows: *burst* 100% (mean = 67.49, SD = 11.75), *burst* 40% (mean = 62.713, SD = 11.58), *m-sequence* 100% (mean = 48.7, SD = 19.08), *m-sequence* 40% (mean = 38.83, SD = 23.9).

3.2.5. Impact of calibration trial blocks on classification performance

The results we reported in the previous subsections on the effects of code type and amplitude were all given with 6 calibration blocks as calibration data. We further explored the impact of the amount of calibration data fed to the model for its training. We thus incrementally reduced the number of calibration blocks dedicated to training, from 6 (equivalent to 52.8 s) to 1 (equivalent to 8.8 s of data), the others being used for testing. The descriptive results of our analysis reveal compelling patterns. Notably, the accuracy of the *burst* codes consistently maintains a level above 85%, surpassing the accuracy achieved by the m-sequence codes. This observation is visually represented in Fig. 4, located in the lower left quadrant. Moreover, in terms of training time, the *burst* codes demonstrate a significant advantage over the m-sequence codes. The CNN training time associated with the *burst* codes is consistently at least four times faster than that of the m-sequence codes. This disparity is graphically depicted in Fig. 4, displayed in the lower right quadrant. Our descriptive findings suggest that the utilization of *burst* codes not only ensures robust classification accuracy, consistently surpassing 85%, but also offers a substantial reduction in CNN training time compared to the usage of m-sequence codes.

3.3. Visual evoked potentials results

The following subsections present statistical analyses performed on ERP features such as area under curve amplitude, peak latency, and inter-trial coherence elicited by the different code sequences (Burst and Mseq) and amplitude (100 and 40) visual stimuli (see Fig. 5).

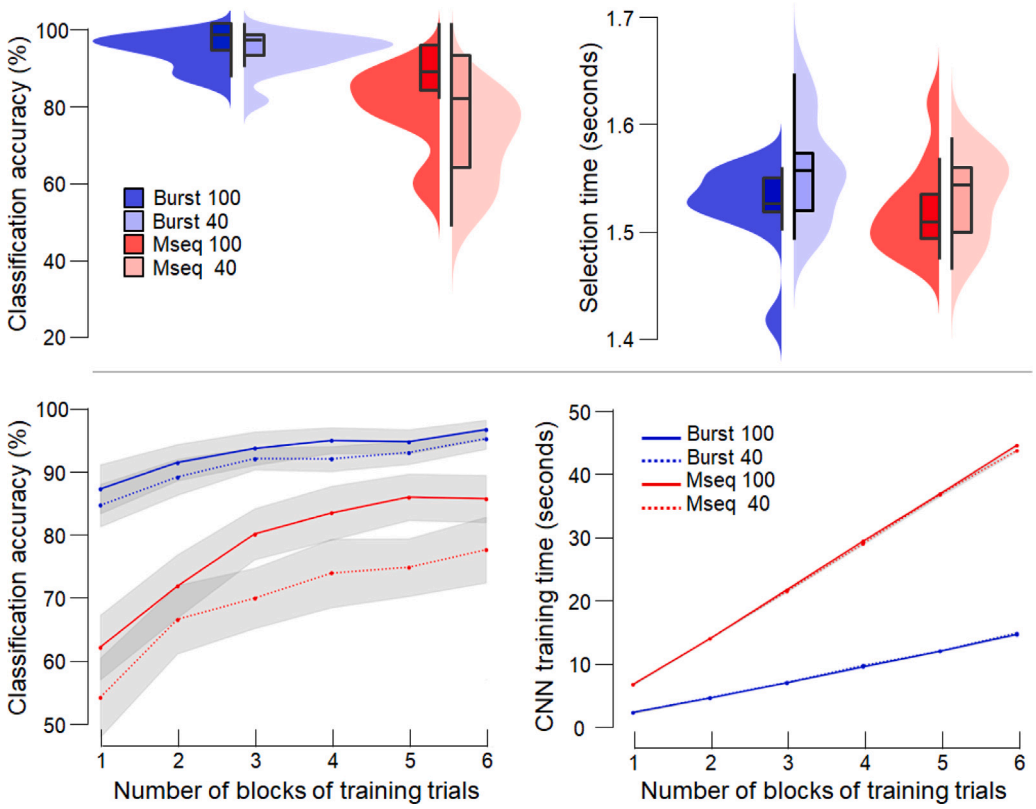

Fig. 4. Performance assessment of BCI on the 4-class problem across code sequences (Burst 100, Burst 40, Mseq 100 and Mseq 40) through different metrics. Top-left: Classification accuracy with 6 trials of calibration data across visual stimulation. Top-right: Comparison of output commands selection time achieved by the CNN with 6 trials of calibration data per class. Bottom-left: Classification accuracy achieved by the CNN as a function of the number of calibration trials per class for each code sequence. Bottom-right: CNN training time as a function of the number of calibration trials. In the present experiment, each additional calibration trial collected lengthened the calibration time by 8.8 s (4 classes x (2.2 s stimuli duration)).

3.3.1. Amplitude

The repeated measure ANOVA revealed a main effect of code sequence type on VEP amplitude ($F(1,11) = 28.25$, $p < 0.001$, $\eta_p^2 = 0.739$) but no main effect of the amplitude depth ($F(1,11) = 2.616$, $p = 0.123$, $\eta_p^2 = 0.207$) and no interaction between these two factors ($F(1,11) = 0.004$, $p = 0.949$, $\eta_p^2 = 0.0004$). The *burst* code sequences evoked significantly larger VEPs than the m-sequence code sequences ($p < 0.001$, Cohen's d = 1.637).

3.3.2. Latency

The repeated measure ANOVA revealed a main effect of code sequence type on VEP latency ($F(1,11) = 46.485$, $p < 0.001$, $\eta_p^2 = 0.823$) and a main effect of amplitude depth ($F(1,11) = 7.888$, $p = 0.019$, $\eta_p^2 = 0.441$). There was no interaction effect on VEP latency between these two factors ($F(1,11) = 4.047$, $p = 0.072$, $\eta_p^2 = 0.288$). The latency of VEPs responses was found to be significantly earlier in response to *burst* than m-sequence codes ($p < 0.001$, Cohen's d = 1.794). Moreover, VEP responses appeared earlier following the onset of full amplitude visual stimulation than in the case of reduced amplitude stimuli ($t(11) = 2.808$, $p = 0.019$, Cohen's d = 0.988).

3.3.3. Inter-trial coherence

The repeated measure ANOVA revealed a main effect of code sequence type on theta-alpha ITC ($F(1,11) = 62.036$, $p < 0.001$, $\eta_p^2 = 0.861$) and a main effect of amplitude depth ($F(1,11) = 6.752$, $p = 0.027$, $\eta_p^2 = 0.403$). There was an interaction effect on VEP ITC between these two factors ($F(1,11) = 12.030$, $p = 0.006$, $\eta_p^2 = 0.546$). Post-hoc analyses revealed that the *burst* c-VEP induced higher ITC than their m-sequence counterpart ($p < 0.001$, Cohen's d = 1.790). The ITC coefficient of VEPs responses was higher in response to full amplitude visual stimulation than in the case of reduced amplitude stimuli ($p = 0.027$, Cohen's d = 0.288). There was no significant difference in ITC between the burst 100% and 40% conditions ($p = 1.00$, Cohen's d = 0.094). In contrast, a significant decrease in ITC was observed when the amplitude depth of m-sequence codes was decreased from 100% to 40% ($p = 0.009$, Cohen's d = 0.481).

4. Discussion

The present study focused on the evaluation of novel solutions aimed at enhancing the performance and user experience of c-VEP BCI paradigms. More specifically, we introduced a novel c-VEP pattern design, referred to as *burst*, with the goal of improving classification accuracy, reducing decoding time, and decreasing the data and computational requirements for training CNNs. Our burst patterns comprise brief, irregular flashes occurring at a rate of 2 to 4 flashes per second. In light of this, we designed an experiment to compare the proposed burst codes against the commonly used m-sequence c-VEP patterns. Another significant contribution of this study was the enhancement of user comfort during code presentation through a reduction in amplitude modulation depth. Consequently, we presented both code patterns at two amplitude levels: 100% and 40% of the maximal modulation amplitude. This allowed us to examine how reducing amplitude depth influences both overall BCI performance and the user experience.

4.1. Classification performance

In order to compare the two types of codes (m-sequence and *burst*), let us first consider the classic case in which stimuli are presented at maximal amplitude modulation depth. Our findings indicated that the *burst* c-VEP outperformed the m-sequence code on most of the

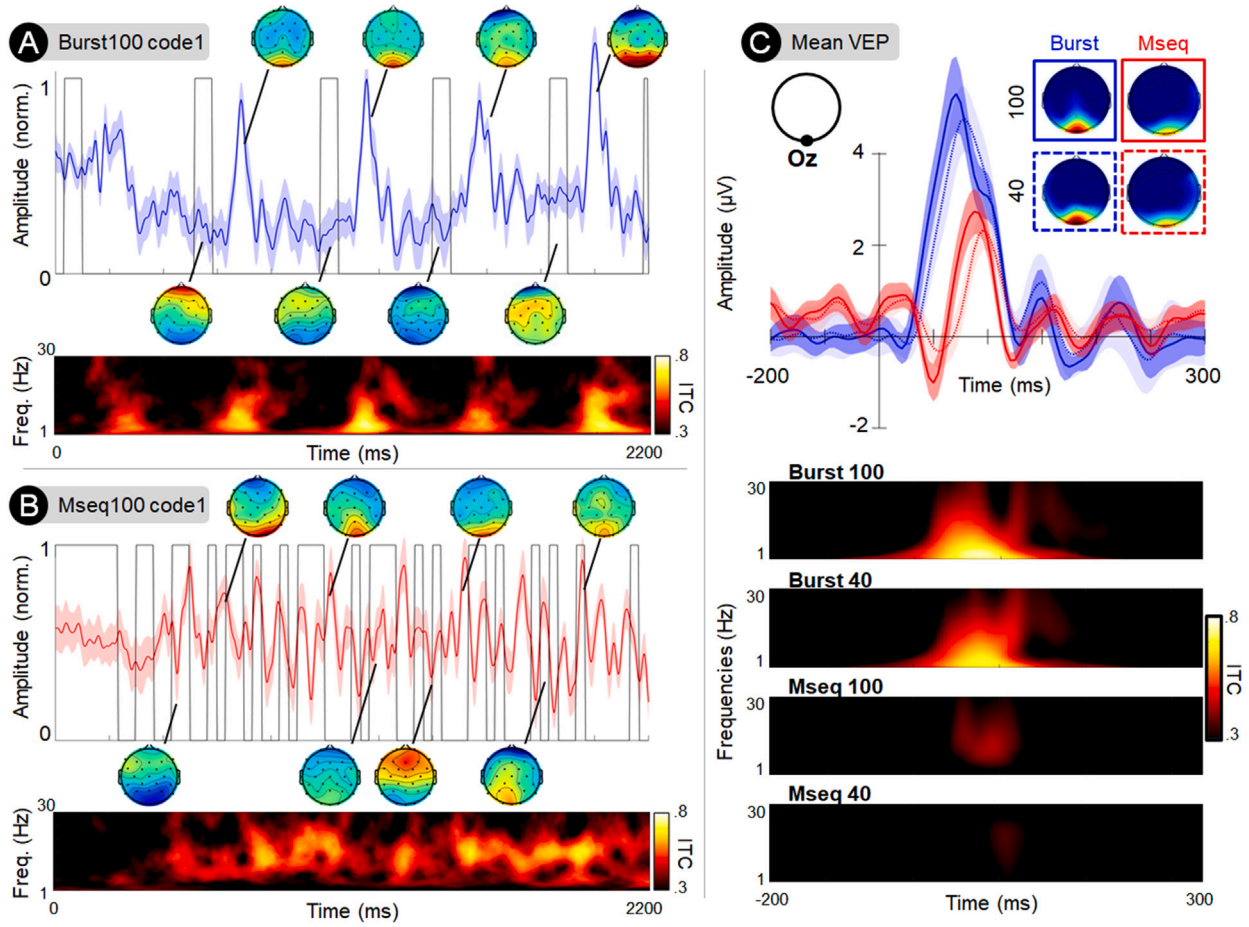

Fig. 5. Grand average (N = 12) EEG responses recorded at occipital electrode Oz elicited by the presentation *burst* and m-sequence code sequences. (A) The blue line presents the grand average of the Visually Evoked Potentials (VEP) response to the first *burst* code-VEP sequence (black line) lasting for 2.2 s. The VEP amplitude was normalized and the shaded area represents the VEP amplitude standard error across individuals computed over time samples. The scalp maps present the instantaneous topographical distribution of the VEP responses at different phases of the stimuli input function. The bottom subplot presents the Inter-Trial Coherence (ITC) of VEPs computed across frequencies (y-axis) and over time (x-axis), where brighter shades denote higher ITC. (B) Similarly to subplot A, the red line presents the grand average VEP response to the first m-sequence code-VEP (black line) lasting for 2.2 s. C) Top: The line plots present the grand average VEP responses locked to the onset of visual stimulation for both types of code sequences (burst in blue and m-sequence in red) and amplitude modulation depth (100% amplitude in solid line and lower and 40% amplitude in dotted line). The scalp maps present the spatial distribution of average VEP response over the VEP time window (150 to 210 ms after stimulus onset). The bottom plots highlight the ITC across all four experimental conditions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

classification performance metrics. When using 6 blocks of calibration data (amounting to 52.8s of calibration data), the full amplitude *burst* c-VEP exhibits a significantly higher classification accuracy (95.6%) compared to the m-sequence codes (85.0%). Moreover, the proposed *burst* codes demonstrated greater robustness to reductions of calibration data compared to the m-sequence. With no more than two blocks of calibration (corresponding to only 17.6 s of data acquisition), the classification accuracy of neural responses to *burst* c-VEP sequences remained high, reaching 90.5%, while classification accuracy related to the m-sequence codes dropped substantially to 71.4%. This finding highlights the relevance of the *burst* code design to reduce calibration time without sacrificing classification accuracy. To provide context, the original EEG2Code study, which first utilized CNN to decode the m-sequence, required 384 s of calibration data for training (Nagel and Spüler, 2019b). Our study shows favorable comparisons to previous research, such as (Spüler et al., 2012)'s study, which achieved an average accuracy of 95% in 1min08s, and Gembler et al. (2018)'s study, which achieved 92% accuracy in 1min24s. Furthermore, a recent study (Thielen et al., 2021) showcased that c-VEP-based BCI can be calibrated using only a few trials at the start of a session, or even without any calibration data altogether. On a related issue, the training time for the CNN monotonically increases as a function of calibration data. However, the steepness of the slope characterizing this relationship is significantly higher for the m-sequence than the *burst* codes, amounting to 40 against 15 s when using 6 blocks of calibration trials (see bottom-right quadrant of Fig. 4). Both of these measures clearly demonstrate that the *burst* c-VEP stimulus design requires the acquisition of less training data which substantially reduces calibration and CNN training time as compared to m-sequence. This is an important consideration within the frame of non-invasive BCI development as a new calibration phase is required at the start of each session.

4.2. Selection time performance

Another critical aspect of BCI performance closely tied to user experience is how long it takes for the system to identify the command selected by the user and perform the associated action. In this study, the mean selection time of the two types of codes was found to be similar (1.5 s). This fast (yet accurate) decoding was achieved by taking advantage of the reactivity and flexibility of the CNN approach. The proposed CNN implementation outperformed the CNN implementation of the original EEG2Code study (1.7 s) (Nagel and Spüler, 2019b), and showed faster selection times than our previous study by Dehais et al. (2022) in which we reported selection duration varying between 1.6 s to 2.6 s depending on workload condition. It should also be noted that selection times achieved through the CNN classification approach are

significantly shorter than those reportedly achieved with the canonical correlation analysis approach, where the best performances range between 3.1 s and 3.8 s (Başaklar et al., 2019; Gembler et al., 2019; Gembler and Volosyak, 2019; Thielen et al., 2021; Martínez-Cagigal et al., 2023). The short selection time in our study contributes to a user-friendly experience and further enhances the applicability of BCI in operational settings.

4.3. End-user assessment

In this study, our aim was also to enhance the end-user experience of BCI by modifying the features of c-VEP stimuli. The subjective reports indicated that the slow *burst* codes were comparably more comfortable visually than their faster m-sequence counterparts. This finding is surprising in light of previous research demonstrating that higher flashing rates are associated with higher subjective ratings of user experience (Martínez-Cagigal et al., 2021; Ladouce et al., 2021, 2022; Martínez-Cagigal et al., 2023; Başaklar et al., 2019). This divergence could be attributed to the fact that our *burst* codes involved fewer flashes of shorter duration and therefore emitted less overall luminance than the m-sequences. Indeed burst c-VEP stimulation features a significantly lower ratio, with only a few frames in the on-phase per cycle. This lower duty cycle ratio implies that the stimuli light up for a shorter duration within each aperiodic cycle of stimulation. In contrast, m-sequence codes exhibit a variable duty cycle but tend to have an distribution, indicating a balanced proportion of *on* and *off* phases throughout the sequence. We hypothesize that the reduced *on* phase time in burst codes compared to m-sequence codes may explain the improved user experience reported by participants. It is also worth noting that our subjective findings underscored the significant contribution of amplitude reduction to enhance the user experience, as evidenced by the more pronounced size effect observed in this variable, consistent with previous studies (Ladouce et al., 2021, 2022). It is essential to highlight that the combined effects of flash scarcity and amplitude reduction established the 40% reduced *burst* code as the preferred choice for user experience. Although amplitude reduction had a negative impact on classification accuracy for both code patterns, *burst* showed a minimal 1.4% loss in accuracy with 6 blocks of training data. Even with 40% luminance and only three blocks of calibration (26.4s of training data), *burst* achieved an accuracy of 90.7% (whereas it decreased to 66% for the m-sequence). These results underscore the practicality and robustness of *burst* code for improved user experience with high classification performance, especially when compared to m-sequence codes.

4.4. Factors influencing performance variations

The CNN architecture we employed was designed for the bitwise decoding of m-sequences, demonstrating robust performance in that context (Nagel and Spüler, 2019b). Consequently, any performance disparities observed between the *burst* and m-sequence can be primarily attributed to two key factors. Firstly, the *burst* code effectively elicited consistent and coherent event-related responses over time thereby enhancing the CNN's learning and training process. This stemmed from the provision of more reliable and predictable inputs to the network. Electrophysiological analyses revealed that the *burst* codes are associated with prototypical visual P1 event-related responses that are prominent over occipital recording sites (corresponding to activation of primary visual cortex areas recorded at surface level). The amplitude elicited by the *burst* codes (approximately 4.4μV) was found to be twice as high as that of the m-sequence (approximately 2.2μV). These latter amplitudes for the m-sequence align with those reported by Thielen et al. (2015). Furthermore, the ITC evoked by the *burst* exhibited significantly stronger and more distinguishable responses when compared to those elicited by the m-sequence. This difference may be attributed to the higher rate of visual stimulation provided by the m-sequences. The rapid succession of visual stimulus onsets can induce overlapping responses, resulting in non-linear interactions (Nguyen et al., 2019). Another hypothesis is that the scarcity of the *burst* allows neural populations to complete their refractory period and therefore be more sensitive to following stimuli (Regan, 1989). Secondly, the utilization of *burst* codes reduces the computation cost and training time for the CNN. The *burst* code streamlines the classification task into a 2-class problem, involving the detection of zeros and ones in the electrophysiological response, which are also inherently more distinguishable, as previously explained. Despite having more instances of zeros than ones in our burst codes, balancing the classes yields a smaller but higher-quality training dataset, consequently accelerating CNN training.

4.5. Conclusion and future work

Collectively, our findings underscore the significant potential of the burst code sequence design in advancing BCI technology. Implementing burst c-VEP yielded superior classification performance with minimal calibration data requirements. The reported variations in user experience across code types underscore the paramount importance of end-user satisfaction and emphasize the advantages of employing burst codes with reduced amplitude depth in developing BCI solutions. This approach not only ensures effective adoption but also promotes long-term user retention, offering a balanced combination of performance and user experience.

While the proposed approach shows promising prospects, several limitations of the present work should be considered. Firstly, our experimental design primarily focused on comparing the *burst* and m-sequence code design. For both types of c-VEP stimulation, a CNN-based decoding method was applied. This CNN-based decoding approach was demonstrated to yield superior performance in terms of classification accuracy and selection time for the decoding of m-sequence c-VEP than traditional template-matching methods (Nagel and Spüler, 2019b; Darmet et al., 2023; Huang et al., 2023; Li and Huang, 2021). Nonetheless, it is critical to consider alternative classification and decoding methods to explore solutions that will further enhance BCI performance. Future work should therefore focus on a more systematic comparison across state-of-the-art methods traditionally used for the decoding c-VEP signals (for a review see Martínez-Cagigal et al., 2021). By making the dataset publicly available, we encourage fellow researchers to apply their own classification and decoding methods to both types of c-VEP stimulation patterns. Secondly, while our study was primarily dedicated to demonstrating the advantages of burst c-VEP over the conventional m-sequence, we did not explicitly showcase its application within a reactive BCI framework, such as a speller interface. This approach would allow us to illustrate how burst c-VEP can facilitate hands-free interactions by harnessing user responses to external stimuli, aligning seamlessly with the traditional utilization of c-VEP in reactive BCI settings. Moreover, it is important to recognize that c-VEP, including the *burst* c-VEP variant, may have significant relevance in passive BCI applications (Nagel and Spüler, 2019b). This specific neuroadaptive technology enables implicit interaction by tracking degraded mental states and it adjusts human–machine interactions to overcome cognitive limitations (Zander et al., 2010; Ewing et al., 2016). By embedding c-VEP-based flickers into various regions of interest, one can gauge the intensity of the brain's response and infer the level of vigilance or attention allocated to these particular areas (Dehais et al., 2022). It is also important to acknowledge that this study was conducted in a controlled laboratory environment using a wet EEG system, which optimized signal quality, facilitating the detection of visually evoked potentials (P100) by the CNN. However, it is essential for future research to prioritize the seamless implementation of dry EEG, especially in real-world settings. In such scenarios, detecting single-trial P100 Event-Related Potentials (ERPs) can present significantly greater challenges compared to highly controlled laboratory conditions (Dehais et al., 2019).

CRediT authorship contribution statement

Kalou Cabrera Castillos: Conceptualization, Data collection, Data analysis, Writing. **Simon Ladouce:** Conceptualization, Data collection, Data analysis, Writing. **Ludovic Darmet:** Conceptualization, CNN implementation. **Frédéric Dehais:** Conceptualization, Data collection, Data analysis, Writing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that may be considered as potential competing interests or could have appeared to influence the work reported in this paper.

Data availability

The links to access the EEG data (hosted on a Zenodo repository) and the stimulus presentation software (hosted on GitHub) included in the paper are shared in the Methods section.

Acknowledgments

This work was funded by AID (Powerbrain project), France, the AXA Research Fund Chair for Neuroergonomics, France and Chair for Neuroadaptive Technology, Artificial and Natural Intelligence Toulouse Institute (ANITI), France.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.neuroimage.2023.120446.

References

Başaklar, T., Tuncel, Y., Ider, Y.Z., 2019. Effects of high stimulus presentation rate on EEG template characteristics and performance of c-VEP based BCIs. Biomed. Phys. Eng. Express 5 (3), http://dx.doi.org/10.1088/2057-1976/ab0cee.

Bin, G., Gao, X., Wang, Y., Li, Y., Hong, B., Gao, S., 2011. A high-speed BCI based on code modulation VEP. J. Neural Eng. 8 (2), http://dx.doi.org/10.1088/1741-2560/8/2/025015.

Cao, T., Wan, F., Wong, C.M., da Cruz, J.N., Hu, Y., 2014. Objective evaluation of fatigue by EEG spectral analysis in steady-state visual evoked potential-based brain-computer interfaces. BioMed. Eng. Online 13 (1), 1–13. http://dx.doi.org/10.1186/1475-925X-13-28.

Chen, X., Liu, B., Wang, Y., Gao, X., 2022. A spectrally-dense encoding method for designing a high-speed SSVEP-BCI with 120 stimuli. IEEE Trans. Neural Syst. Rehabil. Eng. 30, 2764–2772. http://dx.doi.org/10.1109/TNSRE.2022.3208717.

Chevallier, S., Kalunga, E.K., Barthélemy, Q., Monacelli, E., 2021. Review of Riemannian distances and divergences, applied to SSVEP-based BCI. Neuroinformatics 19 (1), 93–106. http://dx.doi.org/10.1007/s12021-020-09473-9.

Darmet, L., Ladouce, S., Dehais, F., 2023. Shortened calibration of code-VEP based BCI by improved deep learning architecture and golden subjects pre-training.

Dehais, F., Duprès, A., Blum, S., Drougard, N., Scannella, S., Roy, R.N., Lotte, F., 2019. Monitoring pilot's mental workload using ERPs and spectral power with a six-dry-electrode EEG system in real flight conditions. Sensors 19 (6), http://dx.doi.org/10.3390/s19061324.

Dehais, F., Ladouce, S., Darmet, L., Nong, T.-V., Ferraro, G., Torre Tresols, J.J., Velut, S., Labedan, P., 2022. Dual passive reactive brain-computer interface: A novel approach to human-machine symbiosis. Front. Neuroergonomics 3, http://dx.doi.org/10.3389/fnrgo.2022.824780.

Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. http://dx.doi.org/10.1016/j.jneumeth.2003.10.009.

Ewing, K.C., Fairclough, S.H., Gilleade, K., 2016. Evaluation of an adaptive game that uses EEG measures validated during the design process as inputs to a biocybernetic loop. Front. Human Neurosci. 10, http://dx.doi.org/10.3389/fnhum.2016.00223.

Fairclough, S.H., Lotte, F., 2020. Grand challenges in neurotechnology and system neuroergonomics. Front. Neuroergonomics 1, http://dx.doi.org/10.3389/fnrgo.2020.602504.

Gembler, F.W., Rezeika, A., Benda, M., Volosyak, I., 2020. Five shades of grey: Exploring quintary m-sequences for more user-friendly c-VEP-based BCIs. Comput. Intell. Neurosci. 2020, http://dx.doi.org/10.1155/2020/7985010.

Gembler, F., Stawicki, P., Saboor, A., Benda, M., Grichnik, R., Rezeika, A., Volosyak, I., 2018. A dictionary driven mental typewriter based on code-modulated visual evoked potentials (cVEP). In: 2018 IEEE International Conference on Systems, Man, and Cybernetics. SMC, IEEE, pp. 619–624. http://dx.doi.org/10.1109/SMC.2018.00114.

Gembler, F., Stawicki, P., Saboor, A., Volosyak, I., 2019. Dynamic time window mechanism for time synchronous VEP-based BCIs—Performance evaluation with a dictionary-supported BCI speller employing SSVEP and c-VEP. PLoS One 14, http://dx.doi.org/10.1371/journal.pone.0218177.

Gembler, F., Volosyak, I., 2019. A novel dictionary-driven mental spelling application based on code-modulated visual evoked potentials. Computers 8 (2), http://dx.doi.org/10.3390/computers8020033.

Huang, Z., Liao, Z., Ou, G., Chen, L., Zhang, Y., 2023. Authentication using c-VEP evoked in a mild-burdened cognitive task. Front. Hum. Neurosci. http://dx.doi.org/10.3389/fnhum.2023.1240451.

JASP Team, 2023. JASP (Version 0.18.0)[Computer software]. URL https://jasp-stats.org/.

Kritzman, L., Eidelman-Rothman, M., Keil, A., Freche, D., Sheppes, G., Levit-Binnun, N., 2022. Steady-state visual evoked potentials differentiate between internally and externally directed attention. NeuroImage 254, 119–133. http://dx.doi.org/10.1016/j.neuroimage.2022.119133.

Ladouce, S., Darmet, L., Torre Tresols, J.J., Velut, S., Ferraro, G., Dehais, F., 2022. Improving user experience of SSVEP BCI through low amplitude depth and high frequency stimuli design. Sci. Rep. 12 (8865), http://dx.doi.org/10.1038/s41598-022-12733-0.

Ladouce, S., Torre Tresols, J.J., Darmet, L., Ferraro, G., Dehais, F., 2021. Improving user experience of SSVEP-BCI through reduction of stimuli amplitude depth. In: 2021 IEEE International Conference on Systems, Man, and Cybernetics. IEEE, pp. 2936–2941. http://dx.doi.org/10.1109/SMC52423.2021.9659135.

Li, J., Huang, Z., 2021. A lightweight convolutional neural network for personal identification based on code-modulated visual-evoked potentials. In: 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics. CISP-BMEI, IEEE, pp. 1–5. http://dx.doi.org/10.1109/CISP-BMEI53629.2021.9624212.

Liu, Y., Wei, Q., Lu, Z., 2018. A multi-target brain-computer interface based on code modulated visual evoked potentials. PLoS One 13 (8), http://dx.doi.org/10.1371/journal.pone.0202478.

Luo, C., Ding, N., 2020. Visual target detection in a distracting background relies on neural encoding of both visual targets and background. NeuroImage 216, http://dx.doi.org/10.1016/j.neuroimage.2020.116870.

Martínez-Cagigal, V., Santamaría-Vázquez, E., Pérez-Velasco, S., Marcos-Martínez, D., Moreno-Calderón, S., Hornero, R., 2023. Non-binary m-sequences for more comfortable brain–computer interfaces based on c-VEPs. Expert Syst. Appl. http://dx.doi.org/10.1016/j.eswa.2023.120815.

Martínez-Cagigal, V., Thielen, J., Santamaria-Vazquez, E., Pérez-Velasco, S., Desain, P., Hornero, R., 2021. Brain–computer interfaces based on code-modulated visual evoked potentials (c-VEP): A literature review. J. Neural Eng. 18 (6), http://dx.doi.org/10.1088/1741-2552/ac38cf.

Nagel, S., Spüler, M., 2019a. Asynchronous non-invasive high-speed BCI speller with robust non-control state detection. Sci. Rep. 9 (1), http://dx.doi.org/10.1038/s41598-019-44645-x.

Nagel, S., Spüler, M., 2019b. World's fastest brain-computer interface: Combining EEG2Code with deep learning. PLoS One 14 (9), 1–15. http://dx.doi.org/10.1371/journal.pone.0221909.

Nakanishi, M., Wang, Y., Chen, X., Wang, Y.T., Gao, X., Jung, T.P., 2018. Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis. IEEE Trans. Biomed. Eng. 65 (1), 104–112. http://dx.doi.org/10.1109/TBME.2017.2694818.

Nguyen, K.T., Liang, W.K., Lee, V., Chang, W.S., Muggleton, N.G., Yeh, J.R., Huang, N.E., Juan, C.H., 2019. Unraveling nonlinear electrophysiologic processes in the human visual system with full dimension spectral analysis. Sci. Rep. 9 (1), 1–13. http://dx.doi.org/10.1038/s41598-019-53286-z.

Ortner, R., Allison, B.Z., Korisek, G., Gaggl, H., Pfurtscheller, G., 2011. An SSVEP BCI to control a hand orthosis for persons with tetraplegia. IEEE Trans. Neural Syst. Rehabil. Eng. 19 (1), 1–5. http://dx.doi.org/10.1109/TNSRE.2010.2076364.

Patterson Gentile, C., Aguirre, G.K., 2020. A neural correlate of visual discomfort from flicker. J. Vis. 20, 1–10. http://dx.doi.org/10.1167/jov.20.7.11.

Regan, D., 1989. Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. Elsevier.

Spüler, M., Rosenstiel, W., Bogdan, M., 2012. Online adaptation of a c-VEP brain-computer interface (BCI) based on error-related potentials and unsupervised learning. PLoS One 7 (12), http://dx.doi.org/10.1371/journal.pone.0051077.

The MathWorks Inc., 2021. MATLAB version: 9.10.0 (r2021b). URL https://www.mathworks.com.

Thielen, J., van den Broek, P., Farquhar, J., Desain, P., 2015. Broad-band visually evoked potentials: Re(con)volution in brain-computer interfacing. PLOS ONE 10, http://dx.doi.org/10.1371/journal.pone.0133797.

Thielen, J., Marsman, P., Farquhar, J., Desain, P., 2021. From full calibration to zero training for a code-modulated visual evoked potentials brain computer interface. J. Neural Eng. 18, http://dx.doi.org/10.1088/1741-2552/abecef.

Volosyak, I., Valbuena, D., Lüth, T., Malechka, T., Gräser, A., 2011. BCI demographics II: How many (and what kinds of) people can use a high-frequency SSVEP BCI? IEEE Trans. Neural Syst. Rehabil. Eng. 19 (3), 232–239. http://dx.doi.org/10.1109/TNSRE.2011.2121919.

Wei, Q., Liu, Y., Gao, X., Wang, Y., Yang, C., Lu, Z., Gong, H., 2018. A novel c-VEP BCI paradigm for increasing the number of stimulus targets based on grouping modulation with different codes. IEEE Trans. Neural Syst. Rehabil. Eng. 26 (6), 1178–1187. http://dx.doi.org/10.1109/TNSRE.2018.2837501.

Zander, T.O., Kothe, C., Jatzev, S., Gaertner, M., 2010. Enhancing human-computer interaction with input from active and passive brain-computer interfaces. pp. 181–199. http://dx.doi.org/10.1007/978-1-84996-272-8_11.

Zerafa, R., Camilleri, T., Falzon, O., Camilleri, K.P., 2018. To train or not to train? A survey on training of feature extraction methods for SSVEP-based BCIs. J. Neural Eng. 15 (5), http://dx.doi.org/10.1088/1741-2552/aaca6e.

Zhang, G., Cui, Y., Zhang, Y., Cao, H., Zhou, G., Shu, H., Yao, D., Xia, Y., Chen, K., Guo, D., 2021. Computational exploration of dynamic mechanisms of steady state visual evoked potentials at the whole brain level. NeuroImage 237, 118–166. http://dx.doi.org/10.1016/j.neuroimage.2021.118166.

Zhu, D., Bieger, J., Molina, G.G., Aarts, R.M., 2010. A survey of stimulation methods used in SSVEP-based BCIs. Comput. Intell. Neurosci. 2010, 1–12. http://dx.doi.org/10.1155/2010/702357.