Performance enhancement of wearable instrumentation for {AR}-based {SSVEP} {BCI}

Measurement (2022)

Contents lists available at ScienceDirect

Measurement

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

Performance Enhancement of Wearable Instrumentation for AR-based SSVEP BCI

Pasquale Arpaia[a,b], Egidio De Benedetto[a,b,*], Lucio De Paolis[c], Giovanni D'Errico[d], Nicola Donato[e], Luigi Duraccio[d]

[a] *Interdepartmental Research Center in Health Management and Innovation in Healthcare, University of Naples, Federico II, Naples, 80125 Italy*
[b] *Department of Electrical Engineering and Information Technology, University of Naples, Federico II*
[c] *Department of Engineering for Innovation, University of Salento, Lecce, 73100 Italy*
[d] *Department of Electronics and Telecommunications, Polytechnic University of Turin, Turin, 10129, Italy*
[e] *Department of Engineering, University of Messina, Messina, 98122, Italy*

ARTICLE INFO

Article history:
Received X Month 2021

Keywords:
Augmented Reality
Brain-Computer Interface
BCI
electroencephalography
EEG
Industry 4.0
SSVEP
Real-Time Systems
Wearable Systems
Instruments.

ABSTRACT

This work addresses an innovative processing strategy to improve the classification of Steady-State Visually Evoked Potentials (SSVEPs). This strategy resorts to the combined use of fast Fourier transform and Canonical Correlation Analysis in time domain, and manages to outperform by over 5% previous results obtained for highly wearable, single-channel Brain-Computer Interfaces. In fact, a classification accuracy of 90% is reached with only 2-s time response. Then, the proposed algorithm is employed for an experimental characterization of three different Augmented Reality (AR) devices (namely, Microsoft Hololens I, Epson Moverio BT-350, and Oculus Rift S). These devices are used to generate the flickering stimuli necessary to the SSVEP induction. Also, in the three pieces of instrumentation under test, the number of simultaneous visual stimuli was increased with respect to the state-of-art solutions. The aim of the experimental characterization was to evaluate the influence of different AR technologies on the elicitation of user's SSVEPs. Classification accuracy, time response, and information transfer rate were used as figures of merit on nine volunteers for each piece of instrumentation. Experimental results show that choosing an adequate AR headset is crucial for obtaining satisfying performance: in fact, it can be observed that the classification accuracy obtained with Microsoft Hololens is about 20% greater than Epson Moverio one.

1. Introduction

A Brain-Computer Interface (BCI) is an integration of hardware and software systems with the aim of converting the activity of the central nervous system into an artificial output. This enables communication between the brain and external devices in the form of commands, without using peripheral nerves and muscles [1]. Over the years, the adoption of BCIs has pervaded the medical field [2–5]. Recently, BCIs are also being employed in new and emerging fields, for example, for fatigue detection [6], working memory load estimation [7, 8], engagement sensing [9], gaming [10], neurofeedback [11] and neuromarketing [12]. Currently, Steady-State Visually Evoked Potential (SSVEP) represents one the most promising BCI paradigms, along with *P300* [13] and *Motor Imagery* [14]. Steady-State Visually Evoked Potentials (SSVEPs) belong to the *evoked signal* category [15]. In SSVEPs, the subject perceives visual periodic stimuli that induct a sustained cortical response [16]. In particular, the SSVEP paradigm is characterized by a specific

*Corresponding author: Tel.: +39-081-7683163;
e-mail: egidio.debenedetto@unina.it (Egidio De Benedetto)

physiological brain response to continuously flickering visual stimuli, measured at a latency lower than 100 ms. Stimulation frequency bands usually range from 6 Hz to 30 Hz. Each visual stimulus, typically a flashing light, is associated to a specific command (e.g. numbers, letters or generic control buttons): the system will allow the user to select the desired target by simply staring at the related flickering stimulus. Generally, the SSVEP shows a sinusoidal-like waveform, with a fundamental frequency equal to that of the targeted stimulus frequency and often higher harmonics [17]. Hence, through the analysis of the SSVEP features, it is possible to identify the desired target by achieving high levels of accuracy and reproducibility [18, 19].

Traditionally, SSVEP are elicited by light-emitting diodes (LEDs) [20] or by computer screens (CS) [21]. However, these rendering devices are bulky and limit the portability of the system, while also making interaction with real-world objects unintuitive. Additionally, a key requirement for the success of an SSVEP-BCI system is that the user remains focused on the stimulus: this is often difficult on a 2D rendering display, which allows for external visual noises. These constraints have so far hindered the widespread adoption of SSVEP-BCI systems.

Recently, augmented-reality head-mounted display (AR-HMD) have been used to replace traditional displays in SSVEP-BCI applications. This has led to more user friendly and portable BCIs, enhancing flexibility and mobility [22–24]. In this setup, stimuli and objects coexist in the same field of view (FOV) and this allows a more intuitive control of external devices, providing an immersive experience with a better user engagement [25–28]. However, many of the implementation and processing strategies available in the literature for traditional SSVEP-based BCI cannot be readily extended to the AR-based rendering of visual stimuli [29]; on the contrary, this requires a thorough analysis, and a comparative study between the various approaches need to be conducted. The overall performance of the combined AR-BCI system, in fact, strongly depends on the characteristics of the chosen AR-HMD. The first constraint is represented by the field of view of the AR-HMD's, which is generally limited to some tens of degrees: this affects the maximum number of flickering visual stimuli that can be accommodated in the operator's FOV. In fact, at the state of the art, only two visual stimuli have been displayed simultaneously [26]. Secondly, the non-predictability of the frame rate variations of the AR-HMD inevitably leads to a shift in the frequency values of the rendered visual stimuli: as a consequence, the SSVEP elicited on the user's EEG becomes more difficult to be detected [30]. Nevertheless, at the state of art [22, 26, 31], all the processing strategies for SSVEPs consider the rendered frequencies free of uncertainty contributions.

Starting from these considerations, in this work, an innovative classification algorithm to enhance the SSVEP classification is proposed. This algorithm is based on an adaptive strategy which resorts on the combined processing in the frequency and time domains. The aim is to mitigate the effects caused by the fps (frame per seconds) variations of the AR devices during the use and, thus, to increase the overall classification performance.

After an offline validation, using the same data set described in [22, 26], which showed significant improvement in SSVEPs classification, the proposed algorithm was also used to perform an experimental characterization of a highly wearable SSVEP-based single-channel BCI, when used with three different AR-HMDs for the generation of flickering stimuli. In particular, the visual stimuli necessary for SSVEP elicitation were provided, alternatively and comparatively, through three AR devices, namely Epson Moverio BT-350, Oculus Rift S, and Microsoft HoloLens. The aim of the experimental characterization was to evaluate the influence of different AR technologies on the elicitation of user's SSVEPs. Currently, in fact, there is no evidence of comparative tests between different AR HMD for SSVEPs elicitation. Additionally, the number of simultaneous flickering stimuli was increased up to four, with the aim to establish a starting point in the development of more immersive systems, characterized by a deeper user interaction with external devices, and a greater mobility; finally, the obtained results were compared to the two-stimuli solutions already successfully implemented over the years [22, 26].

This paper is organized as follows. Section 2 provides a background of the SSVEP paradigm, with a particular emphasis on wearable setups. In Section 3, the proposed classification algorithm is described in detail. The validation is addressed in Section 4. In Section 5, the experimental characterization of three AR HMDs is shown. In Section 6, conclusions are drawn and, finally, the future work is outlined in Section 7.

2. Related Work

One of the most historically studied training-free methods in the literature for SSVEP frequency classification is the Power Spectrum Density Analysis (PSDA). It allows to identify spectral EEG peaks in order to recognize the target stimulus [32] and evaluate the Power Spectrum Density in a certain interval. The main drawback of this method is the requirement of a minimum duration for the time windows in order to estimate the spectrum interval with an appropriate frequency resolution [22, 33]. This strongly limits the real-time performance of the SSVEP.

Another effective method is represented by the Task Related Component Analysis (TRCA) and the ensemble TRCA (eTRCA) [34]. The idea behind TRCA is to obtain spatial filters to (i) strengthen the signals task-related components, (ii) limit the noise, and (iii) measure the similarity between the filtered signal and the reference template via the correlation coefficient. One defect of TRCA is that it manages to handle only limited noise components. Then, adequate time filters must be adopted.

At the state of the art, the most effective strategy is the Canonical Correlation Analysis (CCA) in time domain [35–37]. It is a multivariate statistical method [38] which is used to calculate the correlation between EEG data and a set of sinusoidal reference patterns corresponding to the stimulus frequencies [35, 36]. Due to the effect of noise and spontaneous ongoing EEG activities, it is crucial to perform a band pass filtering step in the band of interest during the pre-processing phase. Currently, the classification performance achieved with the use of CCA is generally better than the other aforementioned approaches ([31]. However, most of SSVEP-based BCIs are focused on the traditional, monitor-based, rendering of the flick-

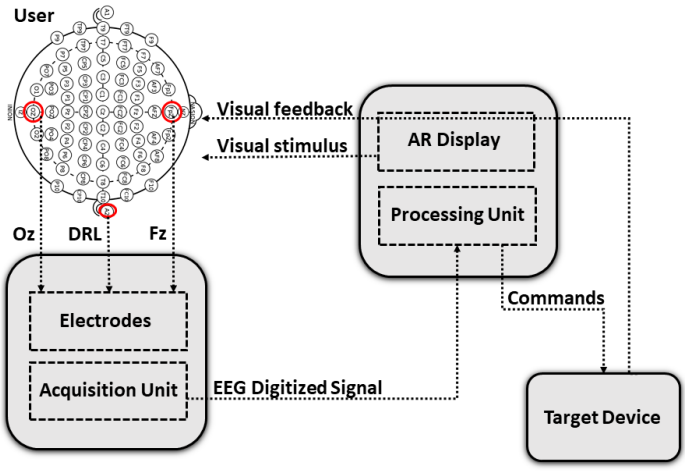


Fig. 1: General architecture of the AR-BCI SSVEP instrumentation.

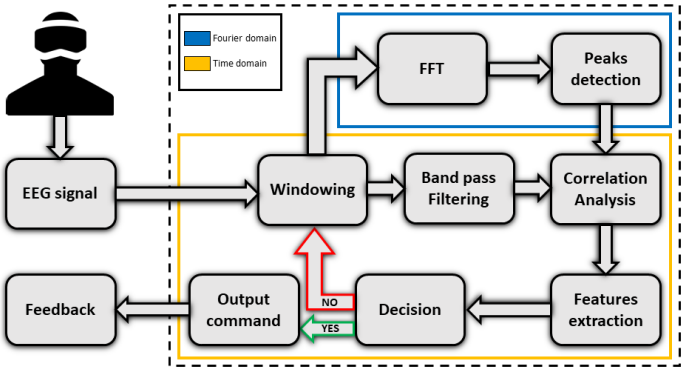


Fig. 2: Block diagram of SSVEP classification.

ering stimuli, which guarantees the best accuracy in the rendered frequencies. Instead, when using AR-based rendering, unpredictable variation of the frame rate and, consequently, of the stimuli frequencies must be handled. For this reason, an adaptive processing [30] can represent an interesting solution to achieve performance comparable with the state of the art.
Another important aspect to take into account is the EEG data acquisition: in traditional EEGs, the Ag/AgCl electrodes placed on the scalp require the use of conductive gel after a preparation of the scalp area, in order to reduce the impedance due to dead skin cells. This procedure is usually long and causes discomfort for the user. Furthermore, the use of a high number of electrodes [24] is required to obtain a better spatial resolution and better performances. *Dry* alternatives have been proposed, especially in highly wearable, AR-based BCI [22], where the aim is to make the procedure more comfortable, allowing to avoid gel by adopting reusable electrodes that are easily placed to the scalp through the hair [39]. In this context, a single-channel BCI, with only two input electrodes (CH+ and CH-) and a reference, is a highly wearable alternative to multi-channel BCI. It can strongly reduce complexity and user's discomfort, increasing the possibility of extending the adoption of BCI also in common activities. Despite the aforementioned issues related to the introduction of AR to generate the flickering stimuli, single-channel BCI has been recently employed in several studies with no degradation in overall performance [26, 40, 41].
The design and implementation of a single-channel AR-based instrumentation was addressed in [22, 26]. Fig. 1 shows the general architecture of the combined system. The *AR Display* renders the visual stimuli to elicit SSVEP activity. Three dry EEG *Electrodes* are used for a single-channel, differential acquisition to reduce common mode interference. Two active electrodes are placed on the user's scalp in *Oz* (Occipital Midline) and *Fz* (Frontal Midline) positions, according to the international 10-20 System [22], and corresponding to the positive and negative input of the acquisition unit. Furthermore, a passive electrode (*Driven Right Leg*, DRL) is placed on the earlobe (*A2*) to act as reference. The EEG signal captured by the electrodes is digitized by means of a portable *Acquisition Unit*; then, it is processed in real time by a *Processing Unit*, which can be the AR device itself or an external wearable device. After the EEG digitized signal is processed, the output command is sent to a *Target Device* (depending on the purpose of the application [26, 30]). Finally, the target device gives a visual feedback to the user, related to the selection the user performed (for example, displaying in AR the output of measuring instruments). This architecture manages to overcome the main challenges related to portability and wearability, still keeping the performance comparable to the state of art results [24].

3. Proposal

In this work, an enhanced classification algorithm for SSVEPs is designed and validated. This algorithm is optimized for highly wearable BCIs and conducts a real-time processing both in frequency and time domains, managing to outperform the results obtained in the previous applications [22, 26] based, respectively, on the use of a frequency domain [22] and a time domain [26] approach. Figure 2 shows the major blocks of the SSVEP classification. Given a time window of length T, the EEG samples are sent in real time to the *Processing Unit* (see Figure 1). Then, the corresponding signal fragment is filtered with a band-pass finite impulse response (FIR) filter between 5 Hz and 25 Hz, having 101 coefficients and a linear phase response. Such a filter conducts an effective removal of ocular and muscle artifacts.
In previous works [26], the filtered signal was correlated with a set of sine waveforms Φ_{0i} with the same nominal frequency as that of the generated flickering stimuli. In the current proposal, before this step, a fast Fourier transform (FFT) of the windowed signal is performed, with the aim to solve the issues related to the uncertainty of the stimulus frequency generation by the AR Display. In fact, a non-predictable frame rate variation of the AR application inevitably leads to a shift of the rendered frequency and, thus, of the SSVEP elicited on the user EEG. For this reason, a peaks detection in frequency domain is carried out. At this point, the time-domain filtered signal is correlated (CCA) with a set of sine waveforms Φ_i at the real generated frequencies and with variable phase ϕ. The maximum values among the Pearson correlation coefficients ρ_i are obtained, as

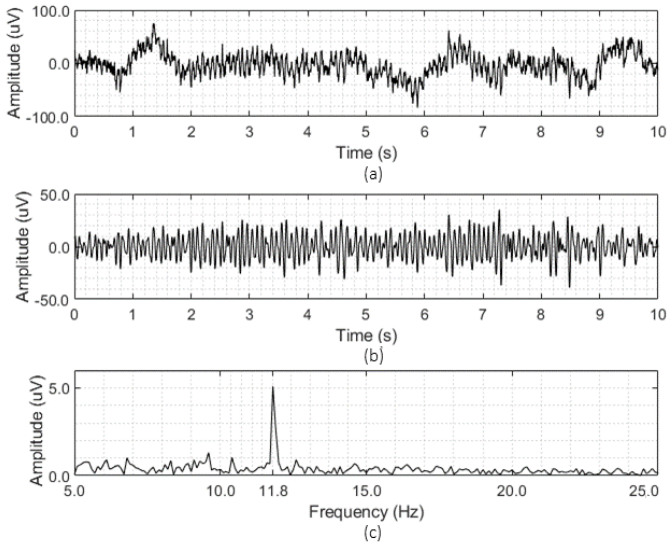


Fig. 3: EEG in time domain (a) Filtered EEG in time domain (b) EEG Magnitude in frequency domain with SSVEP peak (c).

expressed by the following equation:

$$\rho_i = \max_{\phi\in[0,2\pi]} \frac{cov(D_f, \Phi_i(\phi))}{\sigma_{D_f}\,\sigma_{\Phi_i(\phi)}} \tag{1}$$

where D_f are the filtered Data; Φ_i represents the i_{th} sinewave; ϕ is the phase; σ_D is the standard deviation of the filtered data; and σ_{Φ_i} is the standard deviation of the sinewaves. Consequently, the following features are extracted:

$$F_1 = 1^{st}\max_{i\in[1,n]}(\rho_i) \tag{2}$$

$$F_2 = 2^{nd}\max_{i\in[1,n]}(\rho_i) \tag{3}$$

$$F_3 = \frac{F_1 - F_2}{F_2} \tag{4}$$

where F_1 represents the maximum value among the correlation coefficients for all the n generated frequencies; F_2 is the second largest correlation coefficient corresponding to one of the remaining $n-1$ frequencies of stimuli; and, finally, F_3 represents the relative difference between F_1 and F_2. Given two threshold values T_1 and T_2, a signal fragment is recognized if the following condition is satisfied:

$$F_1 > T_1 \quad \cap \quad F_3 > T_2. \tag{5}$$

If (5) is not satisfied, a new signal fragment of duration T, overlapping with the previous one by $T/2$, is processed.

Fig. 3 shows a typical SSVEP when the user was observing a 12.0 Hz stimulus. In particular, Fig. 3 (a) shows the SSVEP acquired in time domain; Fig. 3 (b) shows the signal after a band-pass filtering between 5 and 25 Hz. Finally, Fig. 3 (c) shows the Fast Fourier Transform (FFT) of the original signal. As visible, the SSVEP peak is at 11.8 Hz instead of the presumed 12.0. Hence, the peaks detection in frequency domain

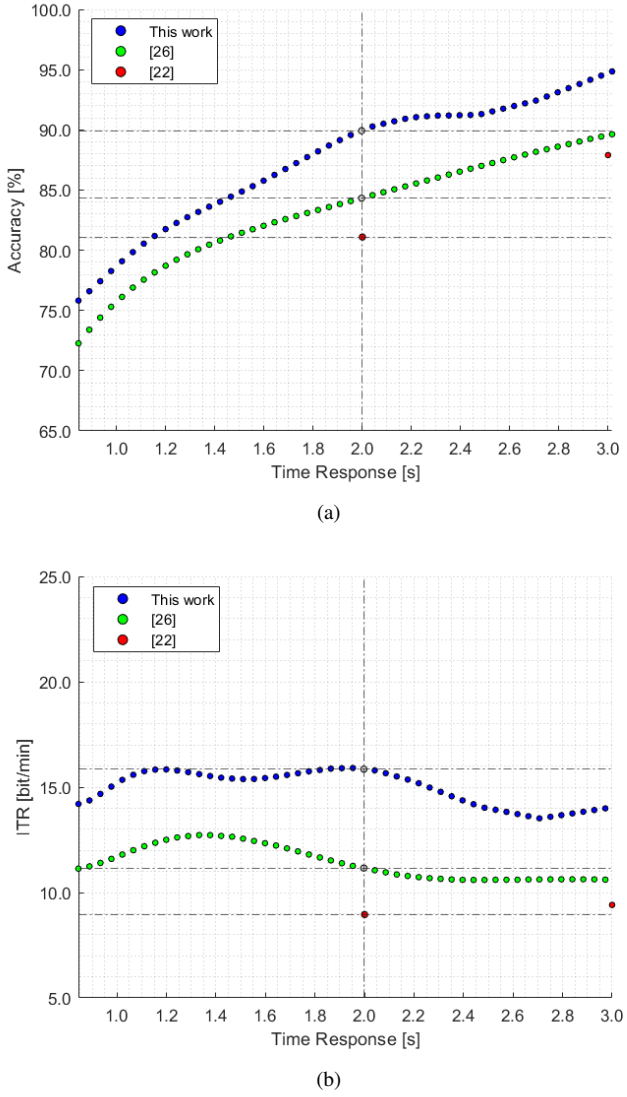


Fig. 4: Two-stimuli performances: accuracy vs time response (a) and ITR vs time response (b) for the different classification algorithms used.

represents a promising strategy to fine-tune the reference signals and optimize the Canonical Correlation analysis, improving the classification performance.

4. Numerical Validation

The proposed SSVEP classification algorithm was validated by comparing the results obtained with the data set used in [22, 26]. These works, in the knowledge of the authors, are the only ones that at the state of the art address a single-channel, AR-based setup. This data set includes the brain signals of 20 untrained and healthy volunteers. The signal acquisition is single-channel and differential: two active electrodes are placed on the users scalp in *Oz* and *Fz* positions, and connected to the Olimex EEG-SMT, a 10-bit, 256 Sa/s, Analogue-to-Digital Converter (ADC). A passive, reference electrode is placed in *A2*. The AR display used to generate the visual stimuli was the

Epson Moverio BT-200, with a 60 Hz refresh rate and a diagonal field of view of 23°. The chosen flickering frequencies were 10 Hz and 12 Hz, according to the results of the studies reported in [32]. Each subject was asked to focus on one stimulus at a time, for 10 s. A total of 24 signals per subject were acquired. Classification accuracy, time response and Information Transfer Rate (ITR) were chosen as figures of merit [31] to evaluate how the proposed algorithm performed on this data set.
The accuracy is defined as the percentage of signals correctly classified, as expressed in (6):

$$A = \frac{N - E}{N} \cdot 100 \quad (\%) \tag{6}$$

where N is the total number of signals and E is the number of misclassified signals. The time response is the average time needed to the algorithm to classify a signal. Finally, the Information Transfer Rate is defined as the amount of information conveyed by the system's output, and it is expressed by (7):

$$ITR = \left(log_2(S) + P\, log_2(P) + (1-P)\, log_2\left(\tfrac{1-P}{S-1}\right)\right) \tfrac{60}{T} \quad (bit/min) \tag{7}$$

where S is the number of visual stimuli, P the classification accuracy, and T the time response.
A grid search of the SSVEP classification parameters T and T_1 was used, while T_2 was fixed to 0.50. The values chosen for the time window T were 0.50 s, 0.60 s, 0.80 s, and 1.00 s, while those for the threshold T_1 ranged from 0.40 to 0.60 with step 0.02. Increasing values of the selected time window T and threshold T_1 leads to an increase of the overall accuracy and ITR. On the other hand, according to (5), also the time response necessary increases. Figure 4 shows the comparison of the performance of the algorithm proposed in this work with those obtained by different classification algorithms implemented over the years on the same data set. Moreover, Table 1 highlights the results achieved in function of the time response, in terms of accuracy and information transfer rate (ITR) [26]. The uncertainty is evaluated at 3-σ.
For the sake of completeness, it should be mentioned that the first version of the SSVEP classification algorithm was implemented in 2019 and was based on a power spectral density analysis in Fourier domain [22]. The second version of the algorithm was implemented in 2020, and performed a traditional CCA in time domain [26]. None of these two works considered the uncertainty contribution caused by unpredictable fps variation of the AR HMD used. Therefore, as can be noticed from Fig. 4 and Table 1, resorting to a combined processing strategy, which overcomes the limits imposed by the frame rate variations, has led to significant improvements in the results.
A focus on the results obtained at 2-s time response is provided in Table 2. As visible, the proposed algorithm reaches a classification accuracy of more than 90 %, while the traditional CCA is limited to about 85 %. A graphical interpretation of this enhancement is given in Fig. 5, where the features extracted by the two methods are compared. In particular, in Fig. 5(a) the number of misclassified signals (red dots) by the proposed algorithm is clearly less than in Fig. 5(b), where the classification was conducted by means of the traditional CCA.

Table 1: Two-stimuli Performance for the Different classification Algorithms Used on the Same Data Set

Algorithm	Time Response (s)	Accuracy (%)	ITR (bit/min)
[26]	1.04 ± 0.10	75.4 ± 11.5	11.2 ± 10.8
This work	**1.01 ± 0.07**	**79.4 ± 9.4**	**15.7 ± 10.5**
[22]	2.00 ± 0.00	81.1 ± 11.4	9.0 ± 7.2
[26]	1.94 ± 0.19	84.9 ± 8.1	12.0 ± 6.4
This work	**1.95 ± 0.17**	**90.2 ± 6.0**	**16.5 ± 6.1**
[22]	3.00 ± 0.00	87.7 ± 7.8	9.24 ± 4.4
[26]	2.99 ± 0.28	91.0 ± 6.3	11.3 ± 4.3
This work	**3.01 ± 0.30**	**94.9 ± 4.6**	**14.1 ± 4.1**

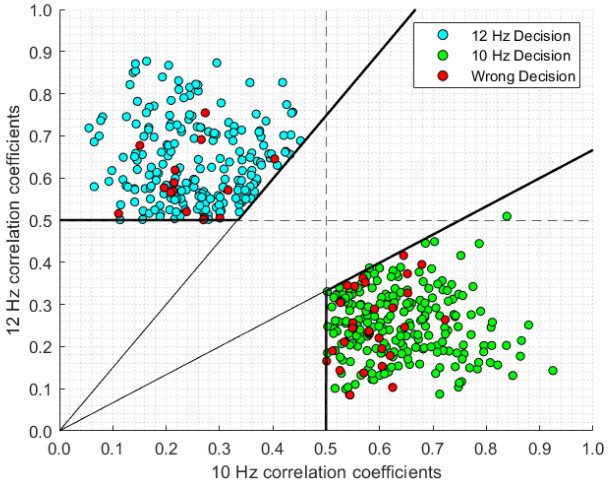
Table 2: Performance comparison at 2-s between the proposal and the traditional CCA developed in [26]

	Proposal $T = 1.0s$, $T1 = 0.50$, $T2 = 0.5$		**Traditional CCA [26]** $T = 0.8s$, $T1 = 0.44$, $T2 = 0.5$	
Volunteer	**Accuracy (%)**	**Time (s)**	**Accuracy (%)**	**Time (s)**
#1	75.0	1.83 ± 0.55	66.7	1.75 ± 0.96
#2	95.8	1.64 ± 0.59	95.8	1.76 ± 0.76
#3	70.8	2.71 ± 1.23	70.8	2.22 ± 1.05
#4	95.8	1.35 ± 0.26	100.0	1.43 ± 0.44
#5	100.0	1.67 ± 0.58	95.8	1.60 ± 0.51
#6	87.5	2.52 ± 0.76	83.3	2.62 ± 0.89
#7	100.0	1.23 ± 0.31	91.7	1.27 ± 0.44
#8	79.2	3.12 ± 1.48	78.3	3.03 ± 1.43
#9	100.0	1.23 ± 0.31	95.8	1.05 ± 0.41
#10	95.8	1.75 ± 0.72	91.7	1.38 ± 0.60
#11	100.0	1.21 ± 0.36	95.8	0.93 ± 0.16
#12	91.7	1.35 ± 0.31	91.7	1.40 ± 0.43
#13	87.5	3.37 ± 1.33	79.2	3.75 ± 1.33
#14	95.6	1.98 ± 1.15	95.6	2.17 ± 1.23
#15	100.0	1.21 ± 0.22	95.8	1.10 ± 0.26
#16	91.7	2.06 ± 0.81	91.3	2.65 ± 1.27
#17	79.2	2.15 ± 0.65	62.5	1.92 ± 0.62
#18	87.5	2.42 ± 0.81	66.7	2.55 ± 1.12
#19	87.5	2.08 ± 0.91	70.8	1.93 ± 1.06
#20	93.3	2.19 ± 0.54	79.2	2.37 ± 0.69
Results	90.2 ± 6.0	1.95 ± 0.17	84.9 ± 8.1	1.94 ± 0.19

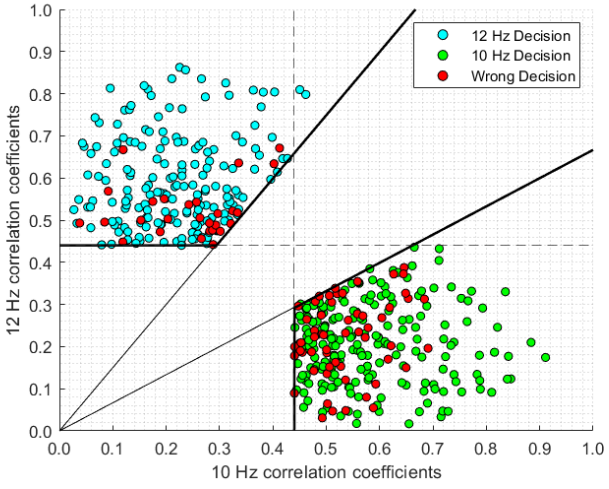
Furthermore, in Table 3, a comparison between the results achieved in this work and other single-channel systems at the state of art is provided. As visible, many of the recent single-channel BCIs [41–44], even with the use of Computer Screen (CS), are outperformed by the proposed processing strategy. To summarize, AR has the potential to provide users with a more wearable, user-friendly, and effective way to control external devices using SSVEP-BCIs. However, many recent works adopting AR are using multi-channel configurations [24, 45], and this is still a limit for the overall user experience.

5. Experimental Characterization of AR Devices

The proposed algorithm was employed for an experimental characterization of three AR devices (namely, Epson Moverio BT-350, Oculus Rift S, and Microsoft Hololens I) used to generate the flickering stimuli necessary to SSVEPs elicitation.
Currently, both Oculus Rift S and Hololens I are removed from sale. However, the goal of the experiments was to preliminarily evaluate how different technologies influence the elicitation of the user's SSVEPs and, thus, the overall performance



(a)



(b)

Fig. 5: Features scatter plot and classification for the proposed algorithm (a) and the traditional CCA [26] (b) at 2-s time response

of a single-channel AR-based BCI. Then, the resulted considerations can be applied in the choosing of the adequate HMD for future developments. In the proposed characterization, the number of simultaneous flickering stimuli was increased up to four, aiming to establish a starting point in the development of highly immersive BCIs, characterized by a deeper user interaction with external devices, and a greater mobility; finally, the obtained results were compared to the two-stimuli solutions already successfully implemented over the years [22, 26]. As mentioned in Section 4, also in this case, classification accuracy, time response and ITR were chosen as figure of merit.

5.1. Setup

Brain signals were recorded on nine healthy adult volunteers. Twenty signals per subject were acquired in a single-channel configuration (as shown in Fig. 1). The AR environment consisted of four flickering squares placed originally at the edges of the screen. For each trial, each user was asked to focus at one stimulus out of four for 10 s. At the end of each session, the users were asked to provide a feedback, aiming to express, for each HMD, their visual fatigue and comfort while observing the flickering stimuli

Table 3: Comparison Between the Obtained Accuracy and The State of Art with Single-channel Setups

Metric	**This Work**	**[41]**	**[42]**	**[43]**	**[44]**
Volunteers	**20**	8	12	20	11
Rendering	**AR**	CS	CS	CS	CS
Stimuli	**2**	5	1	4	4
Time Response (s)	**3**	3	4	4	3
Accuracy (%)	**94.9**	≈ 95	69.5	85.8	81.7

Table 4: Technical Specifications for the Considered AR Devices

Device	**FOV**	**RR (Hz)**	**See-Through**	**Cost (€)**
Moverio BT-350	23°	30	Optical	≈ 900
Oculus Rift S	110°	80	Video	≈ 350*
Hololens 1	34°	60	Optical	≈ 5000*

*removed from sale

5.1.1. Hardware

The employed devices are listed below:

- *AR Display*: The AR Devices used in the experimental campaigns were *Epson Moverio BT-350*, *Oculus Rift S*, and *Microsoft Hololens 1*. Each of these devices differs from the others both in terms of technology and market cost. The technical specifications in terms of (i) diagonal field of view (FOV), (ii) display refresh-rate (RR), (iii) see-through technology. along with the devices cost, are provided in Table 4. As visible, both Moverio and Hololens are optical-see-through devices (OST). Instead, the Oculus Rift S is a HMD mainly designed for Virtual Reality: for this reason, a Stereoscopic HD Camera (*Zed Mini*) was integrated to realize an AR video-see-through (VST) device.

- *Acquisition Unit*: The users' EEG were digitized by the *Olimex EEG-SMT*, as done in [22, 26, 30]. As mentioned in Sec. 4, it is a 10-bit, 256 Sa/s ADC. Only three dry electrodes are used: two active electrodes are placed in *Oz* and *Fz* positions, while a passive electrode is placed in *A2* and acts as a reference.

- *Processing Unit*: A single-board computer Raspberry Pi 4 was used (i) to acquire via USB the EEG digitized by Olimex and (ii) process it. A UART protocol was established with Baud Rate set to 57600 bit/s, packet size equal to 17 bytes, and no parity bit.

5.1.2. Software

Table 5 provides the details about (i) the rendered flickering frequencies for each AR device [46], and (ii) the modality of rendering.
In most cases, the chosen frequencies were sub-multiple of the

HMD Refresh Rate. This means that each square reverses between black and white after a number of frames given by the ratio between the Refresh Rate and the chosen frequency [46]. For the sake of example, Moverio BT-350 is characterized by a 30-Hz Refresh Rate: therefore, a 10-Hz frequency is realized by alternating black and white every 3 frames (i.e. two frames white and one frame black). The non-submultiples frequency values were obtained as a rounded average of a variable frequency stimulus [47].

It can be also noticed that the Moverio rendering is *fixed* to the display. On the contrary, Oculus Rift and Hololens rendering is *location-based*: this means that the four icons are not anchored to the edge of the screen, but are *associated* to a particular position in the real world. For this reason, when the users focus their attention moving their head towards one of the flickering stimuli, the remaining three squares can result less visible, thus reducing interference during the SSVEP elicitation.

Fig. 6 shows the user's view while wearing the AR equipment. For the three examples, the user's gaze is on the top-left square, and it is marked in red. As visible, the location-based rendering helps to reduce the interference coming from other concurrent light sources. In the case of Oculus Rift S (Fig. 6(d)), the surrounding environment is captured by means of the Zed Mini, on which the four flickering squares are rendered; for this reason, there is no need to use dark panels. On the other hand, these were used for Moverio (Fig. 6(b)) and HoloLens (Fig. 6(f)) to maximize the contrast between the environment and the flickering squares.

More details about the Software used for realizing the AR application are provided below:

- *Moverio BT-350*: The AR application running on the Moverio smart glasses was implemented with Android Studio. The flickering frequencies (8 Hz, 10 Hz, 12 Hz and 15 Hz) were generated with the Android library OpenGL.

- *Oculus Rift S*: The AR environment was made with Unity 3D. Unlike Moverio Glasses, Oculus Rift S can synchronize the display refresh rate and the application frame rate by means of the V-Sync technology. However, this option was disabled to avoid stuttering phenomena, which are critical for obtaining an accurate flickering. The chosen frequency values were 8.00, 10.00, 11.43, and 13.33 Hz.

- *Hololens I*: Similarly to Oculus Rift, the AR environment was realized in Unity 3D, with V-Sync kept disabled. The four frequency chosen were 8.57 Hz, 10.00 Hz, 12.00 Hz, and 15.00 Hz.

Finally, a Software running on the Raspberry and written in C was developed to (i) acquire the EEG data from the Olimex EEG via UART, (ii) process them, and (iii) send the desired command to the BCI application (i.e. [26, 30]) via TCP/IP. With regards to the UART communication, the Baud Rate was set to 57600 bit/s, the packet size was chosen equal to 17 bytes, and no parity bit was foreseen.

Table 5: Stimuli Frequencies for Each AR Device

Device	Frequency (Hz)	Rendering
Moverio BT-350	[8.00, 10.00, 12.00, 15.00]	fixed
Oculus Rift S	[8.00, 10.00, 11.43, 13.33]	location-based
Hololens 1	[8.57, 10.00, 12.00, 15.00]	location-based

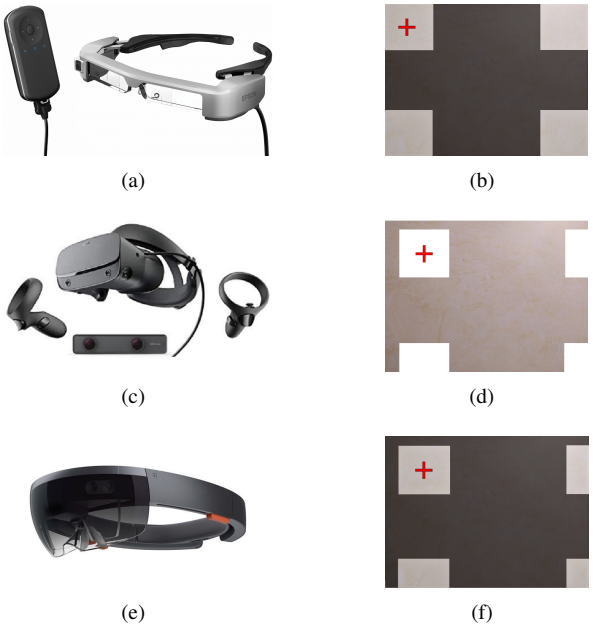
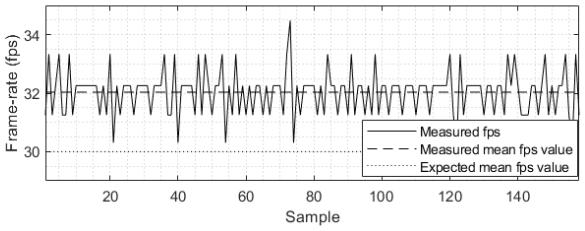


Fig. 6: AR HMD: Epson Moverio BT-350 (a), Oculus Rift S with Zed mini (c), Microsoft Hololens (e), with respective user's views (b, d, f).

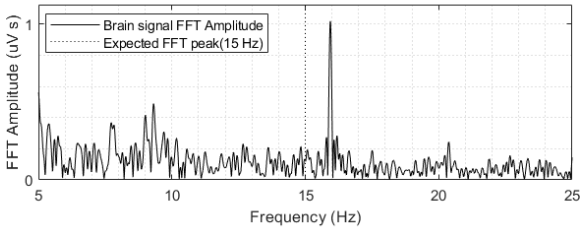
5.2. Results

The purpose of the experimental characterization was twofold: (i) to compare the performance obtained through the three AR HMD in the innovative four-stimuli setup; and (ii) to assess the overall performance of four-stimuli system with the two-stimuli one presented in [26].

With regard to the stability of the refresh rate of the AR devices, Fig. 7 shows the Epson Moverio BT-350 frame rate variation during the trials. As visible, although the declared refresh rate of the display was 30 Hz, the average fps (frames per second) obtained were about 32. This systematic error inevitably translates into a shift of the generated flickering frequencies. For instance, a white/black pixel alternation generates a 16 Hz stimulus, instead of the expected 15 Hz. This demonstrate the importance of a FFT-based approach described in Section 4 in identifying the stimulus peak and, at the same time, performing an effective Canonical Correlation Analysis with sine waves at the correct frequencies, thus mitigating both systematic errors, such as fps offset with respect to the declared values, and the unpredictability of fps variation over time. In fact, considering the set of parameters $T = 1.0$ s, $T1 = 0.50$, and $T2 = 0.50$, the results obtained by the proposed algorithm with Epson Moverio BT-350 reach an accuracy of about 54 % with a time response of about 3.9 s. Instead, a traditional CCA without a peaks detection obtained only the 32 % at 5.1 s time response.



(a)



(b)

Fig. 7: Epson Moverio BT-350: Measured and expected fps (a); Measured and expected FFT peak of the relative user brain signal (b).

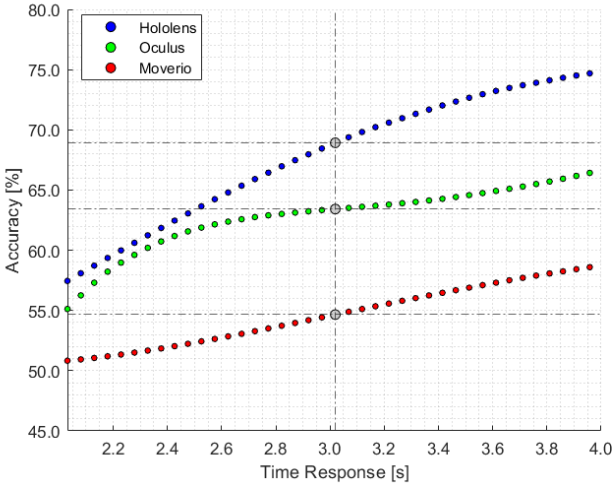
5.2.1. *Four-stimuli performance comparison between the AR HMDs*

Figure 8 shows the obtained results for the three considered HMDs. Overall, Microsoft HoloLens is the device that achieves the best performances (about 70% accuracy and 12 bit/s ITR at 3 s time response). However, Oculus Rift S reaches comparable accuracy and ITR in the range 2.0-2.5 s. These similar results are motivated by the possibility to use an object location-based rendering for these two HMDs. Instead, in the case of Epson Moverio, the presence of squares always fixed to the edges of the screen inevitably leads to undesired interferences while the user tries to look at the chosen stimulus, decreasing both the classification accuracy and the ITR. In the range 2.8-4.0 s, Oculus Rift performance slightly decreases with respect to HoloLens; in fact, during the trials, some users were affected by motion sickness effects due both to the latency introduced by the VST technology and to the device ergonomics. Table 6 summarizes the classification accuracy obtained when each volunteer used the three AR devices: results refer to a chosen set of parameters T, T_1, and T_2. The uncertainty is evaluated at 3-σ. For subject #3, #5, #7, and #9, the improvements brought by the Oculus and HoloLens location-based rendering are evident.

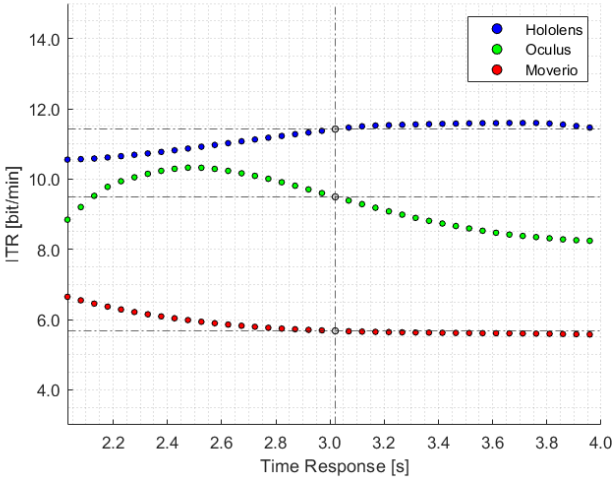
5.2.2. *Comparison between four-stimuli and two-stimuli setups*

Finally, comparing the obtained four-stimuli results (Fig. 8) with those at two stimuli (Fig. 4), a performance decrease can be observed.

In two-stimuli applications, in fact, even with a device with a small field of view (Epson Moverio BT-200, 23° FOV), the classification accuracy reached the 95% at short time responses (approximately 3 s). Considering the four-stimuli setup and a very similar device (Epson Moverio BT-350, 23° FOV), the accuracy obtained at equivalent time response (3 s) was slightly higher than 55%. Instead, with Microsoft HoloLens (34° FOV), the accuracy at 3 s increased to about 70%.



(a)



(b)

Fig. 8: Four-Stimuli performances: accuracy vs time response (a) and ITR vs time response (b) for different AR HMDs used.

The need of larger field of views becomes evident when the number of concurrent stimuli is increased. In fact, for a given maximum FOV, the flickering stimuli get inevitably too close to one another. As a result, when the user (intentionally) stares at one stimulus, another visual stimulus in close proximity may elicit an additional SSVEP, thus causing undesired interfernces. Improving the field of view for is still a technological challenge for OST AR display: this inevitably influences the possibility to interact with an arbitrary amount of flickering stimuli and, thus, to increase the immersivity of the experience and the user engagement.

6. Conclusion and Future Work

In this work, a highly wearable, SSVEP-based single-channel BCI is proposed, in which AR HMDs are chosen to

Table 6: Details of Four-Stimuli SSVEP classification accuracy for a chosen set of Time Window T and Threshold Values T_1 and T_2

	Moverio (T = 1.0 s, T_1, T_2 = 0.50)		**Oculus** (T = 1.0 s, T_1= 0.44, T_2 = 0.50)		**Hololens** (T = 1.0 s, T_1, T_2 = 0.50)	
Volunteer	**Accuracy (%)**	**Time Response (s)**	**Accuracy (%)**	**Time Response (s)**	**Accuracy (%)**	**Time Response (s)**
#1	90.0	2.65 ± 1.23	100.0	1.10 ± 0.23	84.2	2.82 ± 1.47
#2	90.0	1.87 ± 1.11	90.0	1.52 ± 0.41	80.0	2.10 ± 1.30
#3	25.0	4.90 ± 2.32	55.0	2.92 ± 1.33	95.0	2.52 ± 1.24
#4	50.0	3.05 ± 1.59	65.0	1.92 ± 0.73	42.1	3.40 ± 1.77
#5	38.9	4.42 ± 2.00	65.0	2.40 ± 0.86	65.0	2.02 ± 0.92
#6	70.0	2.65 ± 1.12	50.0	3.10 ± 1.59	75.0	2.40 ± 1.22
#7	40.0	5.47 ± 2.12	64.7	4.62 ± 2.07	76.0	4.55 ± 2.20
#8	40.0	3.05 ± 1.34	35.0	1.97 ± 0.95	40.0	2.12 ± 0.90
#9	46.2	7.00 ± 2.21	69.2	5.37 ± 2.68	76.9	6.62 ± 2.11
Results	54.5 ± 23.4	3.89 ± 0.58	65.9 ± 19.6	2.77 ± 0.47	70.5 ± 18.5	3.17 ± 0.51

generate the flickering stimuli. First, a new classification algorithm was presented and validated, in order to overcome the limits introduced by undesired variations of the AR display frame rate during the fruition of the system. The experimental validation with two stimuli confirmed a notable improvement of the performance both in terms of classification accuracy and time response.

Additionally, two main challenges, regarding (i) the performance comparisons between distinct AR devices, and (ii) the increase of the stimuli from two to four, were addressed. To this purposes, three AR HMDs with different costs were used: Epson Moverio BT-350, Oculus Rift S (integrated with Zed mini), and Microsoft HoloLens.

The experimental results with four stimuli show that the choice of effective AR devices is crucial: the overall user experience, in fact, strongly depends on the technology used to implement the flickering stimuli.

Overall, Microsoft HoloLens resulted to be the most performant, both in terms of ergonomics and SSVEP classification accuracy, thanks to a number of key features, like the location-based rendering, the Optical See-Through technology, its larger field of view (34°), and greater ergonomics and user comfort: in fact, no motion sickness effects were reported by the volunteers.

Since it was observed that increasing the number of stimuli in a AR display from 2 to 4 led to a decrease of the classification accuracy of more than 20%, further works will be dedicated to reach the state-of-art performance, in order to ensure a deeper user interaction and high wearability for a practical use in daily life. To this aim, the adoption of Machine Learning algorithms to classify the extracted features may represent an interesting solution.

Acknowledgment

This work was carried out as part of the "ICT for Health" project, which was financially supported by the Italian Ministry of Education, University and Research (MIUR), under the initiative 'Departments of Excellence' (Italian Budget Law no. 232/2016), through an excellence grant awarded to the Department of Information Technology and Electrical Engineering of the University of Naples Federico II, Italy.

References

[1] Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.. Brain–computer interfaces for communication and control. Clinical neurophysiology 2002;113(6):767–791.

[2] Yu, Y., Zhou, Z., Liu, Y., Jiang, J., Yin, E., Zhang, N., et al. Self-paced operation of a wheelchair based on a hybrid brain-computer interface combining motor imagery and p300 potential. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2017;25(12):2516–2526.

[3] Ramos-Murguialday, A., Schürholz, M., Caggiano, V., Wildgruber, M., Caria, A., Hammer, E.M., et al. Proprioceptive feedback and brain computer interface (bci) based neuroprostheses. PloS one 2012;7(10):e47048.

[4] Guy, V., Soriani, M.H., Bruno, M., Papadopoulo, T., Desnuelle, C., Clerc, M.. Brain computer interface with the p300 speller: usability for disabled people with amyotrophic lateral sclerosis. Annals of physical and rehabilitation medicine 2018;61(1):5–11.

[5] Biasiucci, A., Leeb, R., Iturrate, I., Perdikis, S., Al-Khodairy, A., Corbet, T., et al. Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke. Nature communications 2018;9(1):1–13.

[6] Zhang, X., Li, J., Liu, Y., Zhang, Z., Wang, Z., Luo, D., et al. Design of a fatigue detection system for high-speed trains based on driver vigilance using a wireless wearable eeg. Sensors 2017;17(3):486.

[7] Roy, R.N., Bonnet, S., Charbonnier, S., Campagne, A.. Mental fatigue and working memory load estimation: interaction and implications for eeg-based passive bci. In: 2013 35th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2013, p. 6607–6610.

[8] Arpaia, P., Moccaldi, N., Prevete, R., Sannino, I., Tedesco, A.. A wearable EEG instrument for real-time frontal asymmetry monitoring in worker stress analysis. IEEE Transactions on Instrumentation and Measurement 2020;69(10):8335–8343. doi:10.1109/TIM.2020.2988744.

[9] Hassib, M., Schneegass, S., Eiglsperger, P., Henze, N., Schmidt, A., Alt, F.. Engagemeter: A system for implicit audience engagement sensing using electroencephalography. In: Proceedings of the 2017 Chi conference on human factors in computing systems. 2017, p. 5114–5119.

[10] Kerous, B., Skola, F., Liarokapis, F.. Eeg-based bci and video games: a progress report. Virtual Reality 2018;22(2):119–135.

[11] Arns, M., Batail, J.M., Bioulac, S., Congedo, M., Daudet, C., Drapier, D., et al. Neurofeedback: One of today's techniques in psychiatry? L'encephale 2017;43(2):135–145.

[12] Aldayel, M., Ykhlef, M., Al-Nafjan, A.. Deep learning for eeg-based preference classification in neuromarketing. Applied Sciences 2020;10(4):1525.

[13] Yu, Y., Liu, Y., Yin, E., Jiang, J., Zhou, Z., Hu, D.. An asynchronous hybrid spelling approach based on eeg–eog signals for chinese character input. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2019;27(6):1292–1302. doi:10.1109/TNSRE.2019.2914916.

[14] Jin, J., Xiao, R., Daly, I., Miao, Y., Wang, X., Cichocki, A.. Internal feature selection method of csp based on l1-norm and dempster–shafer theory. IEEE Transactions on Neural Networks and Learning Systems 2021;32(11):4814–4825. doi:10.1109/TNNLS.2020.3015505.

[15] Ramadan, R.A., Vasilakos, A.V.. Brain computer interface: control signals review. Neurocomputing 2017;223:26–44.

[16] Vialatte, F.B., Maurice, M., Dauwels, J., Cichocki, A.. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. Progress in neurobiology 2010;90(4):418–438.

[17] Müller-Putz, G.R., Scherer, R., Brauneis, C., Pfurtscheller, G.. Steady-state visual evoked potential (ssvep)-based communication: impact of harmonic frequency components. Journal of neural engineering 2005;2(4):123.

[18] Cheng, M., Gao, X., Gao, S., Xu, D.. Design and implementation of a brain-computer interface with high transfer rates. IEEE transactions on biomedical engineering 2002;49(10):1181–1186.

[19] Volosyak, I., Gembler, F., Stawicki, P.. Age-related differences in ssvep-based bci performance. Neurocomputing 2017;250:57–64.

[20] Gao, X., Xu, D., Cheng, M., Gao, S.. A bci-based environmental controller for the motion-disabled. IEEE Transactions on neural systems and rehabilitation engineering 2003;11(2):137–140.

[21] Gergondet, P., Druon, S., Kheddar, A., Hintermüller, C., Guger, C., Slater, M.. Using brain-computer interface to steer a humanoid robot. In: 2011 IEEE International Conference on Robotics and Biomimetics. IEEE; 2011, p. 192–197.

[22] Angrisani, L., Arpaia, P., Esposito, A., Moccaldi, N.. A wearable brain–computer interface instrument for augmented reality-based inspection in industry 4.0. IEEE Transactions on Instrumentation and Measurement 2019;69(4):1530–1539.

[23] Park, S., Cha, H.S., Kwon, J., Kim, H., Im, C.H.. Development of an online home appliance control system using augmented reality and an ssvep-based brain-computer interface. In: 2020 8th International Winter Conference on Brain-Computer Interface (BCI). IEEE; 2020, p. 1–2.

[24] Ke, Y., Liu, P., An, X., Song, X., Ming, D.. An online ssvep-bci system in an optical see-through augmented reality environment. Journal of neural engineering 2020;17(1):016066.

[25] Azuma, R.T.. A survey of augmented reality. Presence: Teleoperators & Virtual Environments 1997;6(4):355–385.

[26] Arpaia, P., Duraccio, L., Moccaldi, N., Rossi, S.. Wearable brain–computer interface instrumentation for robot-based rehabilitation by augmented reality. IEEE Transactions on Instrumentation and Measurement 2020;69(9):6362–6371.

[27] Arpaia, P., De Benedetto, E., Dodaro, C.A., Duraccio, L., Servillo, G.. Metrology-based design of a wearable augmented reality system for monitoring patient's vitals in real time. IEEE Sensors Journal 2021;21(9):11176–11183. doi:10.1109/JSEN.2021.3059636.

[28] Faller, J., Allison, B.Z., Brunner, C., Scherer, R., Schmalstieg, D., Pfurtscheller, G., et al. A feasibility study on ssvep-based interaction with motivating and immersive virtual and augmented reality. arXiv preprint arXiv:170103981 2017;.

[29] Si-Mohammed, H., Petit, J., Jeunet, C., Argelaguet, F., Spindler, F., Evain, A., et al. Towards bci-based interfaces for augmented reality: feasibility, design and evaluation. IEEE transactions on visualization and computer graphics 2018;26(3):1608–1621.

[30] Arpaia, P., De Benedetto, E., Duraccio, L.. Design, implementation, and metrological characterization of a wearable, integrated ar-bci hands-free system for health 4.0 monitoring. Measurement 2021;:109280.

[31] Wang, Y., Chen, X., Gao, X., Gao, S.. A benchmark dataset for ssvep-based brain–computer interfaces. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2016;25(10):1746–1752.

[32] Wang, Y., Wang, R., Gao, X., Hong, B., Gao, S.. A practical vep-based brain-computer interface. IEEE Transactions on neural systems and rehabilitation engineering 2006;14(2):234–240.

[33] Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T.P., Gao, S.. High-speed spelling with a noninvasive brain–computer interface. Proceedings of the national academy of sciences 2015;112(44):E6058–E6067.

[34] Jin, J., Wang, Z., Xu, R., Liu, C., Wang, X., Cichocki, A.. Robust similarity measurement based on a novel time filter for ssveps detection. IEEE Transactions on Neural Networks and Learning Systems 2021;:1–10doi:10.1109/TNNLS.2021.3118468.

[35] Lin, Z., Zhang, C., Wu, W., Gao, X.. Frequency recognition based on canonical correlation analysis for ssvep-based bcis. IEEE transactions on biomedical engineering 2006;53(12):2610–2614.

[36] Bin, G., Gao, X., Yan, Z., Hong, B., Gao, S.. An online multi-channel ssvep-based brain–computer interface using a canonical correlation analysis method. Journal of neural engineering 2009;6(4):046002.

[37] Zhang, Y., Zhou, G., Jin, J., Wang, X., Cichocki, A.. Frequency recognition in ssvep-based bci using multiset canonical correlation analysis. International journal of neural systems 2014;24(04):1450013.

[38] Wang, H., Zhang, Y., Waytowich, N.R., Krusienski, D.J., Zhou, G., Jin, J., et al. Discriminative feature extraction via multivariate linear regression for ssvep-based bci. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2016;24(5):532–541.

[39] Chi, Y.M., Wang, Y.T., Wang, Y., Maier, C., Jung, T.P., Cauwenberghs, G.. Dry and noncontact eeg sensors for mobile brain–computer interfaces. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2011;20(2):228–235.

[40] Luo, A., Sullivan, T.J.. A user-friendly ssvep-based brain–computer interface using a time-domain classifier. Journal of neural engineering 2010;7(2):026010.

[41] Nguyen, T.H., Chung, W.Y.. A single-channel ssvep-based bci speller using deep learning. IEEE Access 2018;7:1752–1763.

[42] Ko, L.W., Ranga, S., Komarov, O., Chen, C.C.. Development of single-channel hybrid bci system using motor imagery and ssvep. Journal of healthcare engineering 2017;2017.

[43] Chen, S.C., Chen, Y.J., Zaeni, I.A., Wu, C.M.. A single-channel ssvep-based bci with a fuzzy feature threshold algorithm in a maze game. International Journal of Fuzzy Systems 2017;19(2):553–565.

[44] Acampora, G., Trinchese, P., Vitiello, A.. Applying logistic regression for classification in single-channel ssvep-based bcis. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). IEEE; 2019, p. 33–38.

[45] Putze, F., Weiβ, D., Vortmann, L.M., Schultz, T.. Augmented reality interface for smart home control using ssvep-bci and eye gaze. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). IEEE; 2019, p. 2812–2817.

[46] Wang, Y., Jung, T.p., et al. Visual stimulus design for high-rate ssvep bci. Electronics letters 2010;46(15):1057–1058.

[47] Wang, X., Cao, T., Wang, B., Wan, F., Mak, P.U., Mak, P.I., et al. An online ssvep-based chatting system. In: Proceedings 2011 International Conference on System Science and Engineering. IEEE; 2011, p. 536–539.