Latest Advancements in SSVEPs Classification for Single-Channel, Extended Reality-based Brain-Computer Interfaces

25th IMEKO TC4 International Symposium
23rd International Workshop on ADC and DAC Modelling and Testing
IMEKO TC-4 2022
Brescia, Italy / September 12-14, 2022

Latest Advancements in SSVEPs Classification for Single-Channel, Extended Reality-based Brain-Computer Interfaces

Leopoldo Angrisani[1], Pasquale Arpaia[1,2], Egidio De Benedetto[1], Nicola Donato[3], Luigi Duraccio[4]

[1]*University of Naples Federico II - Department of Information Technology and Electrical Engineering, Naples, Italy*
[2]*University of Naples Federico II - Interdepartmental Research Center in Health Management and Innovation in Healthcare (CIRMIS), Naples, Italy; pasquale.arpaia@unina.it*
[3]*University of Messina - Department of Engineering, Messina, Italy*
[4]*Polytechnic University of Turin - Department of Electronics and Telecommunications, Turin, Italy*

***Abstract* – This work details the latest advancements on a single-channel, reactive Brain-Computer Interfaces developed at the Interdepartmental Research Center in Health Management and Innovation in Healthcare (CIRMIS) of the University of Naples Federico II. The proposed instrumentation is based on Extended Reality (XR) and exploits the acquisition and classification of the Steady-State Visually Evoked Potentials (SSVEPs). In particular, an XR headset is employed for generating the flickering stimuli necessary to the SSVEP elicitation. The users brain signals are captured by means of a highly wearable and portable electroencephalografic acquisition unit, which is connected to a portable processing unit in charge of processing in real time the incoming data. In this way, a deeper interaction between users and external devices with respect to traditional architectures is guaranteed. The classification capability of the proposed instrument has been significantly improved over the years. Currently, in fact, a classification accuracy up to 90 % is obtained with at least 2 s of acquisition time.**

I. INTRODUCTION

A Brain-Computer Interface (BCI) is a technology employed in order to provide a direct communication path between the human brain and external devices [1, 2, 3]. Since a huge amount of information can be extracted from the user brain signals, a distinction is typically made between active, reactive, and passive BCIs [4]. Reactive BCIs rely on the acquisition and processing of brain waves produced in response to external stimuli [5]. Among all the reactive-BCI paradigms, Steady-State Visually Evoked Potentials (SSVEPs) have gained momentum in the development of applications regarding healthcare [6], entertainment [7], and industry [8]. Typically, SSVEPs are a sinusoidal-like waveform with a fundamental frequency equal to that of the observed flickering stimulus. Often, higher harmonics can also be detected [9].

In SSVEP-based applications, N flickering stimuli at different frequencies are associated to specific commands, so that the user can select the desired target by simply looking at the corresponding flickering stimulus. These stimuli are traditionally displayed on LCD monitors. Moreover, multi-channel EEG acquisition units are often adopted [10]. However, this setup is often very bulky and limits the portability of these systems.

For this reason, the Interdepartmental Research Center in Health Management and Innovation in Healthcare (CIRMIS) of the University of Naples Federico II has developed an innovative solution which can facilitate the adoption of BCI-SSVEP in daily-life activities. In particular, the proposed BCI instrumentation is based on a single-channel electroencephalographic (EEG) acquisitions [11] and on the use of Extended Reality (XR) devices to render the flickering stimuli [12] for the SSVEPs elicitation. The number of the stimuli accommodated on the XR display was set to two [8], so that the user was able to take a binary decision by gazing one stimulus out of two. The data coming from the wearable EEG acquisition unit are processed in real time by a portable processing unit, which is in charge of classify the received brain signals and send the related command to external devices.

The classification capability of the proposed system has been significantly improved over the years. The first processing method employed was based on the traditional Power Spectral Density Analysis (PSDA) in frequency domain. It managed to classify one frequency out of two with an accuracy greater than 80 % and an acquisition time of 2 s [8]. More recently, a time-domain approach based on the Canonical Correlation Analysis (CCA) was employed by improving the obtained accuracy of about 5 % [6]. However, the main drawback of these methods was related to the impossibility to detect undesired frame per second (fps) variation of the XR headset during the generation of the flickering stimuli. For the sake of the example, given a refresh rate of 60 Hz, a variation of about 5 % inevitably

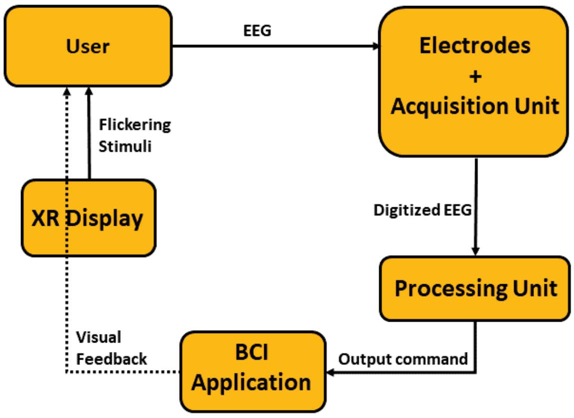

Fig. 1. Major blocks of the system architecture.

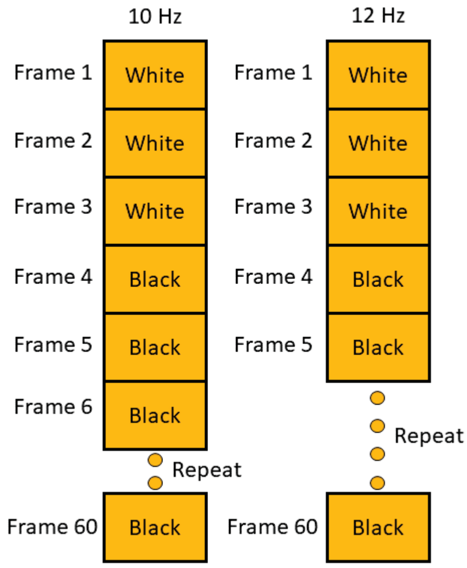

Fig. 2. Rendering of the two flickering stimuli in a 1-s time interval

leads to a shift of the rendered frequencies from 10.0 Hz up to 9.5 Hz or 10.5 Hz. This means that the elicited SSVEPs are shifted accordingly. Therefore, the classification performance of the proposed algorithms decreases since the acquired signals and the reference ones are no longer consistent.

For this reason, an adaptive strategy to find the actual SSVEP peaks has been employed in recent months [5]. Moreover, the adoption of Machine Learning (ML) classifers such as K-Nearest Neighbor (K-NN), Support Vector Machine (SVM), and Feed-Forward Neural Networks (NN) has also been performed to further improve the classification of the features which are extracted from the original samples. This new technique, defined *Features Reduction (FR)*, has achieved an accuracy greater than 90 % at 2-s acquisition time.

The paper is organized as follows. Section ii. provides a description of the system architecture, along with the modality of rendering of the flickering stimuli, and the classification algorithm implemented over the years. Therefore, Section iii. shows the obtained experimental results. Finally, conclusions are drawn.

II. MATERIALS AND METHODS

A. *System Description*

The architecture of the BCI instrumentation is shown in Fig. 1. An *AR Display* renders 2 flickering stimuli to elicit users SSVEPs. Then, three dry *Electrodes* are placed in *Oz*, *Fz*, and *A2* positions, according to the 10-20 International System [6], and capture the user EEG. The electrodes are connected to a portable *Acquisition Unit*, which sends the digitized EEG samples to a portable *Processing Unit*. The processing unit runs the SSVEP classification algorithm, and sends in real time the output command to the *BCI Application*, which actuates the received command and provides a feedback to the User in order to show the desired selection.

B. *Hardware*

The chosen XR device was the *Epson Moverio BT-200*. It is an Optical-See-Through device with a 23Â° diagonal field of view, and a nominal refresh rate of 60 Hz. The selected acquisition unit was the Olimex EEG-SMT, a 10-bit, 256 S/s, open source Analog-to-Digital converter. Finally, the adopted processing unit was the Raspberry Pi 4, a single-board PC connected via USB to the Olimex.

C. *Software*

The flickering icons on the Epson Moverio glasses were generated by means of an Android application realized with Android Studio. The XR environment consisted of two squares placed at opposite edges of the screen. The two squares reverse black and white according to the chosen flickering frequency (namely, 10 Hz and 12 Hz). Moreover, a software written in Python 3 on the Raspberry Pi 4 was used to (i) acquire the digitized signal via USB from the Olimex, (ii) process it, and (iii) send the output command to the specific target via TCP/IP [6, 12, 13]

D. *Rendering of the Flickering Stimuli*

With regards to the rendering of the flickering stimuli at 10 Hz and 12 Hz, Fig. 2 shows the implemented pixel alternations. Since 10 Hz is a submultiple of six of Moverio Refresh Rate (60 Hz), the length of the sequence that has to be repeated is equal to six frames (i.e., three frames white, and three frames black). Instead, 12 Hz is a submultiple of five of Moverio Refresh Rate, then the length of the sequence is equal to five frame (i.e., three frames white, and two frames black).

E. *SSVEPs Classification*

The algorithms implemented over the years were metrologically characterized by analyzing data related to an experimental campaign conducted on 20 untrained and healthy volunteers. For each volunteer, 24 signals were acquired. The chosen flickering frequencies were 10 Hz (rendered on the right side of the screen) and 12 Hz (rendered on the left). Each subject was asked to focus on one stimulus at time, for 10 s. Two metrics are used to evaluate the classification performance: (i) classification accuracy, and (ii) acquisition time. The classification accuracy is defined as the percentage of brain signal correctly classified, while the acquisition time represents the time duration of the signals considered.
In this work, three algorithms are considered.

- *Power Spectral Density Analysis:* The most intuitive approach used to detect and classify the elicited SSVEPs is based on a Power Spectral Density Analysis (PSDA) [14]. First, a Fast Fourier Transform (FFT) is applied to the user EEG. Then, a PSD is performed in the neighborhood of each frequency rendered on the display according to (1).

$$P(f_n) = \frac{1}{2k+1}\left[\sum_{i=k_n-k}^{k_n+k} A^2(i)\right] \quad (1)$$

Where: $P(f_n)$ is the PSD coefficient for the given frequency f_n, k_n is the corresponding bin in frequency domain, k is the number of neighbours to be considered, and A is the signal amplitude. Finally, the classification is performed based on the hypothesis that the observed stimulus is very likely to be that with the highest PSD [15]. However, this method requires a minimum time duration ΔT of the acquired EEG in order to correctly discriminate the harmonics, since an appropriate frequency resolution Δf is required [16], as explained in (2).

$$\Delta T = \frac{1}{\Delta f} \quad (2)$$

- *Canonical Correlation Analysis:* An alternative way to process SSVEPs is the Canonical Correlation Analysis (CCA) in time domain. It is a multivariate statistical method of correlating linear relationships between two sets of data [17]. CCA is performed between the EEG data and a set of sine waves having the same frequencies of the stimuli, and variable phase. A correlation coefficient ρ_{mn} is extracted for each stimulus frequency f_n. This relation is described by (3).

$$\rho_n = max_\phi \frac{cov(D, \Phi_n(\phi))}{\sigma_D\, \sigma_{\Phi_n(\phi)}} \quad (3)$$

Where D is the EEG data, Φ_n is the sine wave at the frequency f_n of each rendered stimulus, and ϕ is the phase ranging from 0 to 2π. Therefore, these coefficients are used for SSVEP classification. For the sake of example, in [17] the output of the classification was associated to the frequency with the highest correlation coefficient extracted. Alternatively, in [6, 13] the maximum value among the correlation coefficients ρ_n was compared with given threshold values: the signal was marked as classified only if the chosen correlation coefficient exceeded these thresholds. The classification performance achieved with the use of CCA are typically better than PSDA [15]. However, a band pass filtering for the EEG can be necessary during the pre-processing phase, due to the effect of spontaneous EEG activities not involved in SSVEP events.

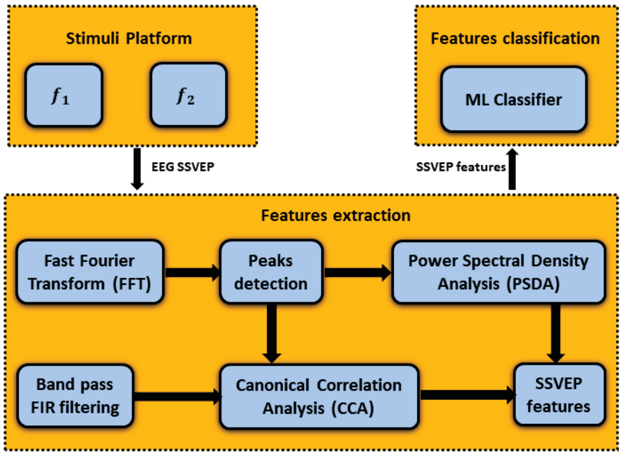

Fig. 3. Features extractions and ML-based classification

- *Features Reduction:* However, none of these two methods detects undesired frame per second (fps) variation of the XR headset during the generation of the flickering stimuli. Given the Moverio Refresh Rate (60 Hz), a variation of about 5 % inevitably leads to a shift of the rendered frequencies from 10.0 Hz up to 9.5 Hz or 10.5 Hz. This means that the elicited SSVEPs are shifted accordingly. Therefore, the classification performance of the proposed algorithms may decrease as the acquired signals and the reference ones are no longer consistent. For this reason, an adaptive strategy to find the actual SSVEP peaks has been employed. The main blocks are shown in Fig. 3. The *EEG Samples* are processed both in frequency and time domains, in order to obtain a reduced number of significant features.
  - In the frequency domain, a fast Fourier transform (FFT) is performed. Then, the actual SSVEPs *Peaks* are detected around the two rendered frequencies. In this way, the uncertainty introduced by the XR device during the generation of the flickering stimuli is mitigated

as the Power Spectral Densities (PSDs) P_1 and P_2 around the two detected peaks are more accurate.

– In the time domain, a *Band pass Filtering* between 5 and 25 Hz is applied by means of a Finite Impulsive Response (FIR) filter with linear phase response. Then, the Canonical Correlation Analysis (CCA) between the filtered signal and a set of sinewaves, having the frequencies of the two detected peaks and variable phase, is performed. As a consequence, also the two canonical correlation coefficients ρ_1 and ρ_2 obtained for each frequency are more accurate.

Ultimately, for a given brain signal of variable length, only four features are extracted and *Normalized*.
The *Classification* is carried out by means of three ML classifiers: in particular, Support Vector Machine (SVM), k-Nearest Neighbour (k-NN), and Artificial Neural Network (ANN) are employed. The FR algorithm was tested on this realized data set by means of Leave One Subject Out Cross Validation (LOSO CV). This validation strategy highlights the inter-individual reproducibility. It divides the data set in 20 folds, where each fold is constituted by a subject. Then, for each combination of the models hyperparameters, the process will run 20 times, each time with a different subject in the test set, taking the remaining ones in the training set.

III. RESULTS

In Table 1, the classification accuracy obtained by the proposed algorithm in function of the acquisition time T and the ML model is summarized. The uncertainty is evaluated at 2-σ. Clearly, increasing the duration of T leads to an increase of the classification accuracy as more information is given to the Features Reduction block. Overall, the best performance are obtained by ANN, but even a more simple classifier like k-NN reaches comparable accuracy levels.
In Table 2, a comparison between the results achieved by ANN is made with those obtained by the two classification algorithm previously developed (PSDA and CCA). As visible, the proposed ML-based algorithm provides a significant enhancement. The main contribution is given by the peak detection block, which allows to obtain more accurate features both in time and frequency domains, thus mitigating the uncertainty caused by unpredictable frame rate variation of the XR device. It should also be noted that both the CCA and PSDA are characterized by a worse inter-individual 2-σ uncertainty. Hence, the model proposed in this work offers a greater possibility to be generalized to every users.

Table 1. Accuracy Results in function of the Acquisition Time T for the three ML Models

T (s)	k-NN (%)	SVM (%)	ANN (%)
0.5	72.8 ± 4.1	74.8 ± 4.3	**75.0 ± 4.3**
1.0	80.7 ± 4.4	82.0 ± 4.4	**82.1 ± 4.4**
2.0	88.3 ± 2.6	**89.2 ± 2.3**	**89.2 ± 2.3**
3.0	93.3 ± 2.6	93.6 ± 2.3	**93.7 ± 2.5**
5.0	96.4 ± 2.1	96.4 ± 2.1	**96.7 ± 1.7**
10.0	99.0 ± 1.3	99.2 ± 1.3	**99.4 ± 1.2**

Table 2. PSDA, CCA, and FR Classification Accuracy in function of the Acquisition Time T

T (s)	PSDA [8] (%)	CCA [6] (%)	FR (%)
0.5	-	70.8 ± 4.5	**75.0 ± 4.3**
1.0	-	74.8 ± 8.1	**82.1 ± 4.4**
2.0	81.1 ± 7.6	84.9 ± 5.4	**89.2 ± 2.3**
3.0	87.7 ± 5.2	91.0 ± 4.2	**93.7 ± 2.5**
5.0	96.0 ± 2.6	95.4 ± 2.5	**96.7 ± 1.7**
10.0	98.9 ± 1.0	-	**99.4 ± 1.2**

IV. CONCLUSION

This work provides a review about the latest advancement on single-channel Brain-Computer Interfaces based on Steady-State Visually Evoked Potentials and Extended Reality. The proposed BCI instrumentation with the adoption of XR guarantees greater immersivity and engagement with respect to traditional setups. Over the years, three different algorithms were implemented to classify users SSVEPs. The current algorithm is based on a combined processing in time and frequency domains and on a ML classification. It reaches a classification accuracy up to 90 % with at least 2 s of acquisition time. These results outperformed the previous ones obtained with traditional processing strategies like Power Spectral Density Analysis and Canonical Correlation Analysis. Moreover, an additional advantage in using ML was the decrease in the inter-individual 2-σ uncertainty. Therefore, such approach can facilitate future developments of ready-to-use systems.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[2] A. Apicella, P. Arpaia, G. Mastrati, and N. Moccaldi, "Eeg-based detection of emotional valence towards a reproducible measurement of emotions," *Scientific Reports*, vol. 11, no. 1, 2021.

[3] A. Apicella, P. Arpaia, M. Frosolone, and N. Moc-

caldi, "High-wearable EEG-based distraction detection in motor rehabilitation," *Scientific Reports*, vol. 11, no. 1, 2021.

[4] L. Angrisani, P. Arpaia, E. De Benedetto, A. Esposito, N. Moccaldi, and M. Parvis, "Brain-computer interfaces for daily-life applications: a five-year experience report," in *2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1–6, 2021.

[5] A. Apicella, P. Arpaia, E. D. Benedetto, N. Donato, L. Duraccio, S. Giugliano, and R. Prevete, "Enhancement of ssveps classification in bci-based wearable instrumentation through machine learning techniques," *IEEE Sensors Journal*, pp. 1–1, 2022.

[6] P. Arpaia, L. Duraccio, N. Moccaldi, and S. Rossi, "Wearable brain–computer interface instrumentation for robot-based rehabilitation by augmented reality," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6362–6371, 2020.

[7] I. Martišius and R. Damaševičius, "A prototype SSVEP based real time BCI gaming system," *Computational intelligence and neuroscience*, vol. 2016, 2016.

[8] L. Angrisani, P. Arpaia, A. Esposito, and N. Moccaldi, "A wearable brain–computer interface instrument for augmented reality-based inspection in industry 4.0," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 4, pp. 1530–1539, 2019.

[9] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components," *Journal of neural engineering*, vol. 2, no. 4, p. 123, 2005.

[10] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain–computer interface (BCI) speller," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 6, pp. 1447–1456, 2014.

[11] P. Arpaia, N. Moccaldi, R. Prevete, I. Sannino, and A. Tedesco, "A wearable EEG instrument for real-time frontal asymmetry monitoring in worker stress analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 8335–8343, 2020.

[12] P. Arpaia, E. De Benedetto, L. De Paolis, G. D'Errico, N. Donato, and L. Duraccio, "Highly wearable SSVEP-based BCI: Performance comparison of augmented reality solutions for the flickering stimuli rendering," *Measurement: Sensors*, vol. 18, p. 100305, 2021.

[13] P. Arpaia, E. De Benedetto, and L. Duraccio, "Design, implementation, and metrological characterization of a wearable, integrated AR-BCI hands-free system for health 4.0 monitoring," *Measurement*, vol. 177, p. 109280, 2021.

[14] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao, "A practical vep-based brain-computer interface," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, pp. 234–240, 2006.

[15] G. Hakvoort, B. Reuderink, and M. Obbink, "Comparison of psda and cca detection methods in a SSVEP-based BCI-system," *Centre for Telematics & Information Technology University of Twente*, 2011.

[16] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface," *Proceedings of the national academy of sciences*, vol. 112, no. 44, pp. E6058–E6067, 2015.

[17] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for ssvep-based BCIs," *IEEE transactions on biomedical engineering*, vol. 53, no. 12, pp. 2610–2614, 2006.