An Error Aware SSVEP-based BCI

Fotis Kalaganis[1,2], Elisavet Chatzilari[1], Kostas Georgiadis[1,2],
Spiros Nikolopoulos[1], Nikos Laskaris[2,3], Yiannis Kompatsiaris[1]
[1] Information Technologies Institute, Centre for Research and Technology Hellas, Greece
[2] Artificial Intelligence and Information Analysis Lab, Dpt. of Informatics, Aristotle University of Thessaloniki, Greece
[3] NeuroInformatics GRoup, http://neuroinformatics.gr
Email: fkalaganis@iti.gr, ehatzi@iti.gr, kostas.georgiadis@iti.gr,nikolopo@iti.gr,laskaris@aiia.csd.auth.gr,ikom@iti.gr

***Abstract*—Error-Related Potentials (ErrPs) have been used lately in order to improve several existing Brain-Computer Interface (BCI) applications. In our study we investigate the contribution of ErrPs in a Steady State Visual Evoked Potential (SSVEP) based BCI. An extensive study is presented in order to discover the limitations of the proposed scheme. Using Common Spatial Patterns and Random Forests we manage to show encouraging results regarding the incorporation of ErrPs in a SSVEP system. Finally, we provide a novel methodology (Inverse Correct Response Time) that can measure the gain of a BCI system by incorporating ErrPs in terms of time efficiency.**

***Index Terms*—BCI; error related potentials; Information Transfer Rate; SSVEP;**

I. INTRODUCTION

Making mistakes is part of the human nature. The ability of recognizing and correcting the erroneous actions is crucial for human beings. There are plenty of neuroscientific studies regarding the ability of the human brain to recognize errors [1]. The distinct neuronal responses that are produced by the human brain during the perception of a mistake are referred to as Error Related Potentials (ErrPs). During the rapid growth of BCI technology over the last years, ErrPs have been used in this context in order to enhance such systems by facilitating error-aware and thus faster interaction.

BCIs are often developed to provide alternative communication solutions to people suffering from neuromuscular disorders such as amyotrophic lateral sclerosis and spinal cord injury [2]. Within this context steady state visually evoked potentials (SSVEP) proved to be an invaluable tool especially for patients with locked-in syndrome [3]. SSVEP are brain signals which occur in response to visual stimulation. When the retina is stimulated by a flashing light of a certain frequency, the brain produces electrical activity at the same (or multiples of this) frequency mainly recorded over the occipital regions [4]. By having multiple boxes flickering at different frequencies, each one representing a different interface option, the user can select the desired box using SSVEPs and in this way interact with the interface.

While SSVEP, being one of the widely used brain-based interaction paradigms, has managed to achieve significance reliability in detecting the user intentions, it has been mostly tested in controlled environments with ideal settings (e.g. well sound-proofed rooms, with optimal lighting conditions, etc.). However, this is not to be expected in a real world setting, where such optimal conditions are not practical for the everyday living of the users. Furthermore, it is common knowledge that the performance of SSVEP is highly subject-dependent (e.g. the accuracy of a SSVEP-based BCI can range from 30% to 100% for different subjects [5]). Overall, in a real world setting, detecting brain commands is still an error-prone procedure, forcing the users in unintentional interaction errors. Naturally, the question that rises is if we could leverage the ErrP signals in order to create a more responsive interface that would enhance the user experience.

II. RELATED WORK

It was back in the 1990, when Falkenstein et al. associated an electrophysiological event, called error-related negativity (referred to as Ne for convenience), with the ability of humans to monitor and detect erroneous actions in certain tasks [6]. The Ne is an electrical brain signal that has a frontocentral scalp distribution and peaks, approximately, 50 - 100 ms following incorrect responses [7]. The most prominent findings regarding the localization of this event indicate that the anterior cingulate cortex (ACC) is mostly activated during erroneous processes [8], [9]. Besides Ne, a second component, the error-related positivity (Pe), is associated with awareness of erroneous actions and consists of a large magnitude peak. This component is subsequent to the Ne and is most distinguishable in the centro-parietal brain area. Several variations of this special type of event-related potential (ERP) that occur when the subject perceives an error have been examined during different tasks [10].

Towards automatically detecting the ErrPs, in BCI applications we are mostly interested in single trial analysis since trial averaging, widely used in neuroscience, is only a way to confirm neural findings and extract statistical information. Until recently, due to the strong signal-to-noise-ration (SNR) of event related potentials (ERPs) most classification methods were based on simple domain knowledge rather than complex and advanced pattern recognition techniques [11]. The simplest algorithm used to detect ErrPs was by setting a threshold and comparing EEG values to this threshold. The threshold classification technique was first introduced for ErrP detection in [12]. In their scenario subjects were trying to move a cursor to the target word, which was "YES" and "NO", using mu and beta band modulations. The user would get feedback from the system on whether they were moving the cursor to the "YES" or the “NO” word, which would elicit ErrPs in case it was not the expected move. During this study, the overall performance of the BCI improved by using the aforementioned simple ErrP detection technique in the Cz electrode. Furthermore, this study confirmed that ocular artifacts don’t seem to affect the ErrPs since the Ne component was found in both eye blink-free trials and in cleaned ones.

The first effort in the single trial ErrP analysis that made use of more sophisticated machine learning techniques than just using a threshold was proposed in [13], where the authors focused their research in avoiding false positives (the classification of a correct response as an erroneous). The experiment was conducted following the ”d2 attention test” during which the participant is given a page full of letters and has to mark any letter ”d” with two marks around above it or below it in any order. The features, that the authors used, relied on subsampling of the original timeseries (from 100Hz to 20Hz) and were modeled by a Gaussian distribution. Therefore the classifier they used was a conjunction of Fisher’s Discriminant and Neyman-Pearson Lemma. Their reports were promising since they managed to improve the BCI performance fairly by designing classifiers that are capable of bounding false positives (FP), which would classify correct responses as errors. Also utilizing a Gaussian classifier, the authors in [14] present a BCI for controlling a robotic arm. Actually to make things simpler for the users they asked them to control the arm using keys towards a predefined direction. The trick here was that the interface deliberately misinterpreted the user’s action in order to elicit ErrPs. The results that were presented during their study was 79.9% average accurate detection of error trials and 82.4% on correct trials. We must state here that they also were based on the FCz and CPz electrodes using the subsampling technique for feature extraction.

When ErrPs were established in the neuroscientific community, more studies started to investigate how they could be exploited in order to improve the existing BCIs. One type of applications relates to the idea that ErrPs could automate the process of calibrating the BCI systems. In this direction, self calibrating BCI systems have made their appearance [15] showing that having a usable BCI control from the beginning of the experiment without any prior calibration is achievable. In the same vein, the authors of [16] used ErrPs to adapt a classifier online. The problem of missing labels, which is the main problem in such approaches was solved by detecting the presence of an ErrP during the training of a code-modulated visual evoked potential (c-VEP) BCI.

A second type of applications, closely related to this work, includes the utilization of ErrPs as an error-awareness mechanism that could facilitate the creating of more responsive interfaces. In this case the erroneous actions that need to be corrected are typically caused by the BCI system itself rather than the user. For example, the authors in [17] utilized ErrP signals in a mind-based control interface of robots, indicating which commands were erroneous and should be discarded. A visual feedback was provided to the subject indicating the output of the classifier before the actual execution of the given command (e.g. turn left). In case the feedback elicited an ErrP, this command was ignored and the robot continued to execute the previous command. The reported results were encouraging regarding the application of ErrPs towards improving the reliability and accuracy of the existing BCI system. In the same direction, the authors of [18] took advantage of neural correlations with error awareness so as to achieve higher ITR in a P300 based BCI. The proposed approach in this work falls under this category of applications, since it utilizes the ErrP signals as an error-awareness mechanism of a BCI system. However, compared to the brain controlled interface and the P300 speller, our BCI system materializes as a web site interface that relies on SSVEP signals to distinguish between the multiple options that a website offers. To the best of our knowledge, this is the first work that combines ErrPs with SSVEPs, providing an error-aware web site management BCI.

III. METHODOLOGY

This section mainly outlines the tools that were employed during this study and were used during single trial analysis of the ErrPs. The typical pipeline for ErrP detection consists of filtering, extracting temporal features (i.e. subsampling), concatenating the features of multiple channels in a single feature vector and finally classifying the vectors in correct and error trials using Support Vector Machines as the classification scheme. In this work, besides the standard procedure we also test a number of additional and different algorithms. More specifically, in order to remove bad recordings we test an outlier detection algorithm (Section III-A). Furthermore, we investigate whether Common Spatial Patterns (CSP), a well-known algorithm for EEG signal analysis, can help in extracting representations that better separate the two classes (Section III-B). Finally, in order to reliably separate erroneous from correct signals several learning techniques were examined besides the typical SVMs (e.g. Random Forests and AdaBoost) (Section III-C).

A. Outlier Detection

Detecting outliers in multivariate data (as in multichannel EEG) is not as straightforward as in the two dimensions. When the dimension exceeds two, one cannot always rely on visual inspection of the data since multivariate outliers may not be visible on lower dimensional settings. The Minimum Covariance Determinant (MCD) estimator that was introduced by Rousseeuw [19] is exploited in our study to detect and hence remove outliers. MCD estimator searches for the subset of h data points whose covariance matrix has the smallest determinant. Then Mahalanobis distance is used to compute the distance for each observation to the center of the data while taking into account their shape. Long distanced observations (trials in our case) are removed [20].

B. Common Spatial Patterns

The CSP algorithm provides spatial filters of the original two-class signal matrices that maximize their variance. Considering two multivariate signals, in our case multivariance refers to multiple channels, X_1, X_2 of size $N \times T_1$ and $N \times T_2$ respectively where N is the number of channels and T_1, T_2 is the number of time samples (in our study $T_1 = T_2$), the CSP algorithm provides the optimal weights, $\mathbf{w}$, so that the signal can be separated into additive subcomponents which have maximum variance among the two classes (error and correct trials in our case). This is achieved by the simultaneous diagonalization of the averaged covariance matrices over all trials of each class $\bar{C_1}$ and $\bar{C_2}$, which is equivalent to the generalized eigenvalue decomposition. Eigenvectors that correspond to the largest eigenvalues maximize variance for the first condition and eigenvectors that correspond to smallest eigenvalues maximize the variance for the other [21].

C. Classification

In this work, we examine three different classification schemes; SVMs, Random Forests and Adaboost. SVMs are among the most popular classification algorithms and aim to find the optimal hyper-plane that separates the positive class from the negative class by maximizing the margin between the two classes. Random Forests and Adaboost fall in the ensemble learning category, which, based on the assumption that multiple weak classifiers can perform better than a single but more robust classifier, trains multiple classifiers based on the same learning algorithm (decision trees in our case). Random Forests combine the output of random multiple weak classifiers using their average contribution. On the other hand, AdaBoost combines the outputs of each weak classifier into a weighted sum. The weights are optimized during the training phase.

IV. Experimental setup

A. Experimental protocol

The experimental protocol relies on the SSVEP-based selection of 5 boxes in the context of the MAMEM site. The participants were asked to select one of the 5 magenta boxes flickering at different frequencies. The boxes would flicker for 5 seconds, then they would stop and a preview of the box that was selected by the system was shown for 2 seconds to the participant by turning the magenta color of the selected box to green (Figure 1). In the case that the previewed box was not the same as the one the participant was asked to observe, we would expect to create an ErrP in the recorded EEG signal a few ms (200-800) after the preview. In order to maintain a similar ratio of correct and error trials for each participant, we opted for a random classifier to select the boxes (with a ratio of 70% correct and 30% erroneous). The participants were not aware of the random classifier so as to get a natural response to the unexpected erroneous trials.

B. Dataset

The signals were captured with the EBN cap (64 electrodes, 128Hz sampling rate). 5 subjects participated in the study, all male, right handed and between 26-37 years old. Each participant performed a total of 100 trials (20 trials per box), out of which 70 were correct and 30 erroneous. After the start of the selection preview, which lasted 2 seconds, the EEG signal was recorded for these 2 seconds in order to acquire the potentials and then the system would redirect to the selected option so that the participant could move to the next trial.

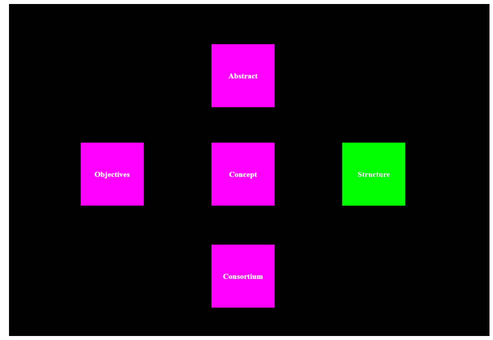


Fig. 1: Selection preview for the SSVEP-based interface

C. Implementation details

For the experiments we made the following choices. First zero-phase filtering (1-20Hz) was applied on the signals. For the ErrPs, the EEG signal between 0.2-0.8 seconds after the preview of the selection was used. For outlier detection, the function *robustcov* in Matlab was applied with default parameters. For the CSP, we used 6 electrodes (AF3, AF4, F7, F8, FCz, CPz) and kept the two most important components. Finally, 10-fold cross validation was used to extract the Accuracy, Precision and Recall of detecting ErrPs.

V. Experimental results

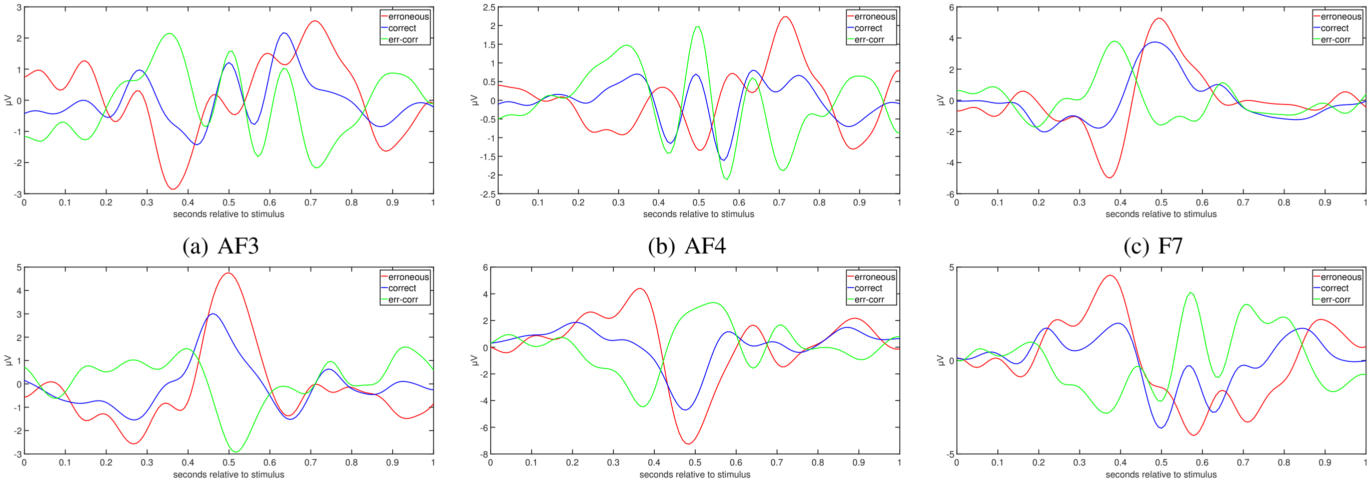
A. Visual Inspection

Figure 2 presents the average signal across all trials for one subject and for the 6 utilized electrodes. The red line corresponds to the average signal of the erroneous trials, the blue line to the correct trials and the green line is their difference. The contradiction in polarity is justified by the re-reference procedure we performed. With visual inspection we can easily find out two peaks at about 350 and 450 milliseconds. Comparing our results with the literature we observe that the two components are time delayed about 100 milliseconds which could be justified by the previous condition during which brain was performing a completely different task and habituated [22].

B. Maximizing ErrP detection rate

In order to come up with the optimal setup for the ErrP detection several settings have been investigated during our study. In this section we provide information of the most prominent. In order to assess the different variations a baseline configuration was decided. The filtered and sub-sampled signal was used as input to SVMs with 3rd degree polynomial kernel (this kernel was selected based on a set of preliminary experiments).

In order to assess the benefit of removing outliers, we compare the performance of the baseline configuration to the



(a) AF3 (b) AF4 (c) F7

(d) F8 (e) FCz (f) CPz

Fig. 2: Average responses for correct, error trials and their difference. Time (x-axis) is relative to preview onset.

TABLE I: Baseline configuration

	Accuracy	Precsion	Recall
S01	0.6800	0.3000	0.0667
S02	0.7700	0.8519	0.3000
S03	0.8000	1.0000	0.3333
S04	0.7200	0.6389	0.2000
S05	0.7600	0.7083	0.3333
Average	0.7460	0.6998	0.2467

one enhanced with the outlier removal technique presented in Section III-A. The results, in tables I and II respectively, show the impact of outlier removal in our study. It is worth noting that the most significant gain by this process comes from the precision metric, which is quite important in the case of ErrP detection. Thus, for the following experiments outlier detection is applied.

In order to assess the benefit of using the CSP method in the case of ErrP detection, we compare the previous configuration with the enhanced version of adding CSP-based filtering. In Table III we provide the results of the CSP-based configuration. Comparing Table III with Table II, we can see the significant improvement that CSP algorithm provides in both terms of accuracy and recall, without compromising the precision. So, from now on CSP filtering will be applied on the trials for the next experiments.

Finally, we investigate into various classification schemes; more specifically, SVMs with 3rd degree polynomial kernel, random forests and Adaboost (using decision trees as the baseline classifier). The results are presented in Tables III, IV, V for each algorithm respectively. We can see that SVMs and RFs perform significantly better than Adaboost, while there is no straightforward winner between them with respect to all three metrics. However, we can see that RFs provide a much higher recall rate, with minor compromises in terms of precision. For this, we select this classifier for our error-aware BCI system.

TABLE II: Outlier removal + SVMs

	Accuracy	Precsion	Recall
S01	0.7452	0.8333	0.1500
S02	0.7482	1.0000	0.2333
S03	0.7625	1.0000	0.1833
S04	0.7278	0.5714	0.1833
S05	0.7625	0.6905	0.3667
Average	0.7492	0.8190	0.2233

TABLE III: Outlier removal + CSP + SVMs

	Accuracy	Precision	Recall
S01	0.7643	0.7500	0.3000
S02	0.8000	0.9028	0.4667
S03	0.8750	1.0000	0.5500
S04	0.7278	0.7500	0.0833
S05	0.7375	0.6762	0.4667
Average	0.7809	0.8158	0.3733

TABLE IV: Outlier removal + CSP + Random Forests

	Accuracy	Precision	Recall
S01	0.8071	0.7500	0.5000
S02	0.8143	0.8333	0.6500
S03	0.8875	0.8000	0.7167
S04	0.7972	0.8148	0.4167
S05	0.7875	0.6458	0.5333
Average	0.8187	0.7688	0.5633

TABLE V: Outlier removal + CSP + Adaboost

	Accuracy	Precision	Recall
S01	0.7571	0.7167	0.5500
S02	0.7339	0.7937	0.4333
S03	0.7982	0.6667	0.5000
S04	0.7389	0.5370	0.3500
S05	0.7250	0.5238	0.3667
Average	0.7506	0.6476	0.4400

C. Measuring the efficiency of the system

Our objective in this section is to showcase the benefit of having an error detection system in a web-site BCI, with respect to the ITR (i.e. the time that is required to perform an action on the web-site). For this reason, we compute the inverse of the average time needed for an individual to complete a correct SSVEP trial by taking into account the SSVEP system accuracy as well as the precision and recall values of the ErrP detection system. This quantity, that will be referred to as *Inverse Correct Response Time*, is monotonically related to the ITR of the system (i.e. information transfer per unit of time). Denoting the number of trials to be completed as s, accuracy of the SSVEP system as Acc, the data length of the SSVEP trial as t (i.e. for how many seconds will the boxes flicker), the recall of correct responses as $Re(c)$, the recall of erroneous responses as $Re(e)$ and the time needed for the user to transition from the erroneous state to the selection panel as d, we calculate the time needed to complete s correct trials in a simple SSVEP system (i.e. without ErrPs):

$$Time = s \cdot t + (d + t) \cdot s \sum_{i=1}^{\infty} (1 - Acc)^i \tag{1}$$

Eq. 1 sums the time for s trials plus the extra time needed to redo the erroneous ones till none erroneous is left. Although the way to calculate the time needed by a simple SVEEP system is straight forward the calculation of time required in a SSVEP-ErrP combination system comes from the addition of four subcomponents. Considering the first stage of a SSVEP system where the system classifies the user intentions with accuracy Acc. Then the ErrP system detects the errors with a true positive rate (TP), a false positive rate (FP), a false negative rate (FN) and a true negative rate (TN). In the first case (TP), the user selected the intended box (correct trial) and the ErrP system did not detect an error. These trials do not need to be repeated. In the case of FP, where the SSVEP system made an error but it was not detected, the user needs d time to go back and t time to redo the trial. In the case of FN, where the SSVEP system did not make an error but an error was detected by the ErrP system, the user needs t time to redo the trial. Finally, in the case of TN, where the SSVEP system made an error and it was successfully detected, the user needs t time to redo the trial. In all cases, there is an additional time e that is essential for the ErrPs to be elicited, which is added to the time t of each trial.

$$Time_{TP} = Acc \cdot s \cdot Re(c) \cdot (t + e) \tag{2}$$

$$Time_{FP} = (1 - Acc) \cdot s \cdot (1 - Re(e)) \cdot (t + e + d) \tag{3}$$

$$Time_{FN} = Acc \cdot s \cdot (1 - Re(c)) \cdot (t + e) \tag{4}$$

$$Time_{TN} = (1 - Acc) \cdot s \cdot Re(e) \cdot (t + e) \tag{5}$$

Equations (2), (3), (4), (5) describe the amount of time needed during the first pass of trials (after the operation of the SSVEP classifier) modulated by the ErrP system. In order to calculate the total time required by the SSVEP-ErrP system to successfully run s trials, we recursively compute the previous equations substituting s with the number of remaining trials. This recursive computation leads to Eq. 6.

$$\begin{gathered} Time = s \cdot (t + e) + [Acc \cdot s \cdot (1 - Re(c)) \cdot (t + e) + \\ (1 - Acc) \cdot s \cdot (1 - Re(e)) \cdot d] \cdot \sum_{i=1}^{\infty} (1 - Acc \cdot Re(c))^{i-1} \end{gathered} \tag{6}$$

Finally, ICRT is defined to be the number of trials times the inverse of the already calculated time ($ICRT = s/Time$). Intuitively, the use of ErrPs is justified when both conditions are met: a) the accuracy of SSVEPs is not perfect, and b) the robustness of ErrPs detection is significant.

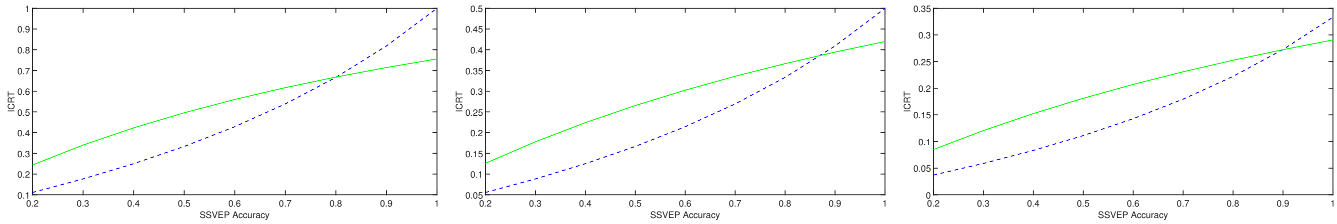
Since the SSVEP system highly depends on the duration of the provided signal we present figures that show the profit of our approach in three different scenarios that differentiate in time needed for the SSVEP detection to operate (e was set to 0.25sec in our case, which is the expected time for ErrPs to be elicited). In order to visualize the results, we plot the ICRT of the systems with respect to the accuracy of the SSVEP detection. The results can be seen in Fig. 3(a-c) for the three data lengths (t) of SSVEP. As shown by the figures, the ErrP system provides significant benefit in terms of ICRT in the case where the SSVEP-system accuracy is below a certain threshold, which depends on the SSVEP trial length (t). This point of change corresponds to the intersection of the simple SSVEP system (blue dashed line) and the SSVEP-ErrP system (green solid line).

VI. Discussion

In this paper we investigated the contribution of ErrPs in a SSVEP system under realistic conditions. The presented proof-of-concept shows that the limitation in such combination, between ErrPs and SSVEPs, is bidirectional. Systems utility is not only dependent from the SSVEP system accuracy but also from the capabilities of the ErrPs classifier. In order to enhance the performance of a SSVEP-based BCI we had to employ more sophisticated algorithms than the typically employed methodology including more delicate preprocessing procedure and capable classification methods. Finally, we present a configuration to enhance experience in BCI systems by trying to exploit neural responses associated with humans' critical ability to recognize errors. To this end, we show the limitations of a SSVEP-ErrP system based on a novel methodology to measure the efficiency of the system.

Acknowledgment

This work is part of project MAMEM that has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 644780.



(a) t=1 sec (b) t=2 sec (c) t=3 sec

Fig. 3: The ICRT of all participants (y-axis) averaged for the SSVEP-ErrP system (green solid line) compared to the simple SSVEP system (blue dashed line) plotted against the accuracy of the SSVEP system (x-axis).

REFERENCES

[1] A. P. Anokhin, S. Golosheykin, and A. C. Heath, "Heritability of frontal brain function related to action monitoring," *Psychophysiology*, vol. 45, no. 4, pp. 524–534, 2008.

[2] J. J. Shih, D. J. Krusienski, and J. R. Wolpaw, "Brain-computer interfaces in medicine," in *Mayo Clinic Proceedings*, vol. 87, no. 3. Elsevier, 2012, pp. 268–279.

[3] D. Lesenfants, D. Habbal, Z. Lugo, M. Lebeau, P. Horki, E. Amico, C. Pokorny, F. Gomez, A. Soddu, G. Müller-Putz *et al.*, "An independent ssvep-based brain–computer interface in locked-in syndrome," *Journal of neural engineering*, vol. 11, no. 3, p. 035002, 2014.

[4] F. Beverina, G. Palmas, S. Silvoni, F. Piccione, and S. Giove, "User adaptive bcis: Ssvep and p300 based interfaces." *PsychNology Journal*, vol. 1, no. 4, pp. 331–354, 2003.

[5] M. Nakanishi, Y. Wang, Y. Wang, and T. Jung, "A comparison study of canonical correlation analysis based methods for detecting steady-state visual evoked potentials." *PLoS ONE*, vol. 10(10), 2015.

[6] M. Falkenstein, J. Hohnsbein, J. Hoormann, and L. Blanke, "Effects of crossmodal divided attention on late erp components. ii. error processing in choice reaction tasks," *Electroencephalography and clinical neurophysiology*, vol. 78, no. 6, pp. 447–455, 1991.

[7] M. J. Crowley, "Error-related negativity," in *Encyclopedia of Autism Spectrum Disorders*. Springer, 2013, pp. 1159–1160.

[8] C. B. Holroyd, J. Dien, and M. G. Coles, "Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans," *Neuroscience letters*, vol. 242, no. 2, pp. 65–68, 1998.

[9] W. J. Gehring and A. R. Willoughby, "The medial frontal cortex and the rapid processing of monetary gains and losses," *Science*, vol. 295, no. 5563, pp. 2279–2282, 2002.

[10] R. Chavarriaga, A. Sobolewski, and J. d. R. Millán, "Errare machinale est: the use of error-related potentials in brain-machine interfaces," 2014.

[11] A. Selim, M. A. Wahed, and Y. Kadah, "Machine learning methodologies in brain-computer interface systems," in *Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International*. IEEE, 2008, pp. 1–5.

[12] G. Schalk, J. R. Wolpaw, D. J. McFarland, and G. Pfurtscheller, "Eeg-based communication: presence of an error potential," *Clinical neurophysiology*, vol. 111, no. 12, pp. 2138–2144, 2000.

[13] B. Blankertz, C. Schäfer, G. Dornhege, and G. Curio, "Single trial detection of eeg error potentials: A tool for increasing bci transmission rates," in *International Conference on Artificial Neural Networks*. Springer, 2002, pp. 1137–1143.

[14] A. Buttfield, P. W. Ferrez, and J. R. Millan, "Towards a robust bci: error potentials and online learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 164–168, 2006.

[15] I. Iturrate, J. Grizou, J. Omedes, P.-Y. Oudeyer, M. Lopes, and L. Montesano, "Exploiting task constraints for self-calibrated brain-machine interface control using error-related potentials," *PloS one*, vol. 10, no. 7, p. e0131491, 2015.

[16] M. Spüler, W. Rosenstiel, and M. Bogdan, "Online adaptation of a c-vep brain-computer interface (bci) based on error-related potentials and unsupervised learning," *PloS one*, vol. 7, no. 12, p. e51077, 2012.

[17] P. W. Ferrez and J. d. R. Millán, "Error-related eeg potentials generated during simulated brain–computer interaction," *IEEE transactions on biomedical engineering*, vol. 55, no. 3, pp. 923–929, 2008.

[18] M. Spüler, M. Bensch, S. Kleih, W. Rosenstiel, M. Bogdan, and A. Kübler, "Online use of error-related potentials in healthy users and people with severe motor impairment increases performance of a p300-bci," *Clinical Neurophysiology*, vol. 123, no. 7, pp. 1328–1337, 2012.

[19] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999.

[20] T.-D. Nguyen and R. E. Welsch, "Outlier detection and robust covariance estimation using mathematical programming," *Advances in data analysis and classification*, vol. 4, no. 4, pp. 301–334, 2010.

[21] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selelction in motor imagery based brain-computer interface," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*. IEEE, 2006, pp. 5392–5395.

[22] S. Schmid, D. A. Wilson, and C. H. Rankin, "Habituation mechanisms and their importance for cognitive function," *Frontiers in integrative neuroscience*, vol. 8, 2014.