PERFORMANCE ESTIMATION OF A COOPERATIVE BRAIN-COMPUTER INTERFACE BASED ON THE DETECTION OF STEADY-STATE VISUAL EVOKED POTENTIALS

Hubert Cecotti

School of Computing and Intelligent Systems
Magee Campus, University of Ulster
Londonderry, UK

Bertrand Rivet

GIPSA-lab UMR5216
CNRS - Grenoble Universities
Saint Martin d'Heres, France

ABSTRACT

To be better suited for the expectations of healthy people, new brain-computer interface (BCI) paradigms should be proposed. To tackle this problem, we investigate the emerging field of cooperative BCIs, which involves several users in a single BCI system. Because combining trials over time improves performance, combining trials across subjects can significantly improve performance compared to a single user. However, cooperative BCIs can only be used in particular settings, and new paradigms must be proposed to efficiently use this approach. To show the advantages of a cooperative BCI based on steady-state visual evoked potentials (SSVEP), we evaluate and compare the performance of combining decisions across subjects, and over time. By considering a reliable brain response such as SSVEP responses, this study represents a first evolution for the combination, and the choice of the combination method for creating a cooperative SSVEP based BCI. The results suggest that six people would be enough to obtain a perfect accuracy within one second of EEG signal.

Index Terms— Brain-Computer Interface, Steady-State Visual Evoked Potentials, Cooperative BCI

1. INTRODUCTION

Pioneer works such as the P300 speller [1] have opened the research field of brain-computer interface (BCI). In the last decades, BCIs have always been used as a way to interact with different devices in relation to the subject's will. The goal is typically to send commands to a computer, *e.g.,* spelling words [2], and in assistive technologies [3], where it is possible to move a wheelchair or robot. Due to the low signal-to-noise ratio (SNR) in the electroencephalogram (EEG) signal, many strategies have been employed to increase the SNR by enhancing the signal through filtering or by combining several trials over time [4]. Whereas combining EEG signals over time has been used since the first P300 speller, where the stimuli appearing on the keys of the virtual keyboard are repeated several times, combining EEG signals across subjects has not been efficiently used in BCI until recently for decreasing reaction time [5] and for 2-D pointer control with two users [6]. In this type of BCI, several users are implicated for enabling a command. In this study, we propose to evaluate the performance and the number of subjects that are needed for a BCI combining several users, and based on the detection of steady state visual evoked potentials, which is a robust brain response [7, 8, 9].

1.1. Combination strategies

The literature of multi-classifier systems offers a large choice of methods for combining the signals and/or the decisions from the different subjects [10, 11]. Increasing the number of subjects can have a different impact on the performance. By adding more subjects, it is possible to add subjects who have had bad performances: while their single performance does not allow them to use a BCI for a single user, their brain features can still contribute to the global performance. The choice of the combination strategy may depend on how possible it is to find subjects that can complement each other, particularly for adaptive combination methods. In this study, we limit the applications to strategies for binary classification.

Let O_i denote the output value of a binary classifier for subject i at a particular trial. O_i represents the confidence value of assigning the input to one of the two classes. Six classical combination scores over N_{sub} subjects are considered: the sum of the classifier outputs ($O_{sum} = \sum_{i=1}^{N_{sub}} O_i$), the weighted sum ($O_{w-sum} = w_0 + \sum_{i=1}^{N_{sub}} w_i \cdot O_i$), the product ($O_{prod} = \prod_{i=1}^{N_{sub}} O_i$), the maximum score ($O_{max} = \max_{1 \le i \le N_{sub}} O_i$), the minimum score ($O_{min} = \min_{1 \le i \le N_{sub}} O_i$), the majority voting (more than half of the outputs agree for the decision), and the consensus vote (the decision should be the same for all the outputs).

The maximum score (O_{max}) and the minimum score (O_{min}) are related to the notions of OR and AND in logic, respectively. The maximum score bases its decision with the most relevant subject. On the contrary, the minimum score bases its decision with the less relevant subject (with the lowest individual score): all the subjects have to be relevant to provide a high combined score. In addition, the sum (O_{sum})

and weighted sum (O_{w-sum}) scores can be seen as smooth or weighted OR, whereas the product score (O_{prod}) can be considered as a smooth or weighted AND.

1.2. Performance evaluation

To show the relevance and the type of performance that can be expected by combining the decision from several individuals, we report the Area Under the ROC curve (AUC) [12] and the ITR [13] in bits per minute (bpm) defined by $ITR = \frac{60}{T} \cdot \vartheta$ where

$$\vartheta = log_2(N_{out}) + Plog_2(P) + (1 - P)log_2(\frac{1-P}{N_{out}-1}) \quad (1)$$

and P being the probability of the good detection, *i.e.* the accuracy, N_{out} being the number of possible different outputs, and T being the time in seconds of recorded EEG signal that is required to take the decision among the N_{out} outputs.

2. EXPERIMENTAL PROTOCOL

SSVEP responses are not time locked to an event; an SSVEP response is evoked when the subject pays attention to a flickering visual stimulus. The efficiency of the classification of an SSVEP response depends on the time that is required to detect a response. A challenge is to reduce the time segment for the SSVEP response detection because it allows to decrease the required time for enabling a command. To qualify the effect of combining brain responses from several subjects, we evaluated the effect of combining several subjects for improving SSVEP response detection with a time segment of 1s.

The experiment consisted of a visual task where 10 healthy subjects (age=27.2 $\pm$ 2.4 years old, two females) were paying attention to a series of different flickering lights (6.66Hz, 7.50Hz, 8.57Hz, 10.00Hz and 12.00Hz) on a computer screen (diagonal size=15.4 inches, vertical refresh rate=60Hz, luminance= $180.0 cd/m^2$, with an estimated contrast of $280 : 1$). Subjects were sitting in a comfortable chair at about 60cm from the computer screen, in a non shielded room. Each stimulus had a luminance of about 0.46cd. The experiments were carried out sequentially. The order of the flickering boxes on which the subject had to pay attention was identical across subjects. In this study, we focused on the detection of a single flickering light that was at 7.50Hz. The EEG was recorded with a g.USBamp EEG amplifier from g.tec with a sampling rate of 128Hz. The electrodes were placed on AF_Z for ground, C_Z for the reference and PO_3, PO_4, P_Z, O_9, O_{10}, and O_Z for the input electrodes.

3. SIGNAL PROCESSING AND CLASSIFICATION

An analog bandpass filter between 2 and 30Hz, and a notch filter around 50Hz (main frequency in Europe) were applied directly inside the amplifier during the EEG acquisition. For the classification, we consider the responses corresponding to the following frequencies: 7.50Hz for the target class, and 6.66Hz, 8.57Hz, 10.00Hz and 12.00Hz, for the non-target class. The frequency corresponding to the target class also corresponds to the frequency of the flickering light. The frequencies for the non-target classes are used as frequencies that should not be detected when the subject is looking at the flickering light corresponding to the target. The following method has been used for the detection of the SSVEP responses [14, 15]. We consider a visual stimulation flickering at f Hz. The signal $y_i(t)$ as the voltage between the electrode i and a reference electrode at a time t is considered as:

$$y_i(t) = \sum_{k=1}^{N_h} a_{i,k} \, sin(2\pi kft + \Phi_{i,k}) + b_{i,t} \quad (2)$$

where N_h is the number of considered harmonics. The signal is divided into two parts: the SSVEP response and the remaining EEG activity, which is considered as noise. The first part corresponds to the evoked SSVEP response signal, which is composed of a number of sinusoids with frequencies in relation to the stimulus frequency and a number of N_h harmonic frequencies. Each sinusoid is defined by its amplitude and phase: $a_{i,k}$ and $\Phi_{i,k}$. $b_{i,t}$ corresponds to the background EEG activity. The detection of an SSVEP response on an EEG signal requires a time segment of N_t samples of EEG signal, with a sampling frequency of F_s Hz:

$$y_i = X_f a_i + B_i \quad (3)$$

where $y_i = [y_i(1), \ldots, y_i(N_t)]^T$ contains the EEG signal for the ith electrode in one time segment. The SSVEP model of the frequency f, X_f, is contained in a matrix $N_t \times 2N_h$ defined by

$$X_f(t, 2k-1) = sin(2\pi kft) \quad (4)$$
$$X_f(t, 2k) = cos(2\pi kft) \quad (5)$$

with $1 \leq k \leq N_h$. The vector a_i of size $2N_h$ contains the amplitudes. For N_y electrodes, the signal is defined as:

$$Y = X_f A_f + B \quad (6)$$

where $Y = [y_1, \ldots, y_{N_y}]$ contained the sampled EEG signals from all the electrodes. A_f contains all the amplitudes for all the expected sinusoids for every electrode signal related the the expected frequency to detect. Spatial filters are used to enhance the SSVEP response in the signal. A spatial filter is represented by a linear combination of the signals measured by different electrodes. We denote by s, a linear combination of y_i, the EEG after a spatial filter:

$$s = \sum_{i=1}^{N_y} w_i y_i = Yw \quad (7)$$

where w_i is the weight for the ith electrode. Several components can be created by using several sets of weights w.

We note N_s as the number of channels. We first estimate the background activity by removing the potential SSVEP components from the signal. It is achieved by projecting the signal onto the orthogonal complement of the SSVEP model matrix (X).

$$\check{Y}_f = Y - X_f(X_f^T X_f)^{-1} X_f^T Y \qquad (8)$$

Spatial filters $\hat{W}_f$ that maximize the Signal-to-Noise Ratio are obtained though determining the generalized Rayleigh quotient that maximizes the following expression:

$$\hat{W}_f \quad = \mathrm{argmax}_W \frac{Tr(W^T Y^T Y W)}{Tr(W^T \check{Y}_f{}^T \check{Y}_f W)} \qquad (9)$$

We denote by $\hat{Y}_f = X_f^T Y \hat{W}_f$ the signal after spatial filtering. The power of the expected frequencies and their harmonics are calculated for the N_s components. For each frequency, the evaluation of the SSVEP response is defined by:

$$R(f) = \frac{1}{N_s \cdot N_h} \sum_{i=1}^{N_s} \sum_{k=1}^{N_h} \left(\hat{Y}_f(i, 2k-1)^2 + \hat{Y}_f(i, 2k)^2 \right) \quad (10)$$

In the experiment, $N_s = 6$ and $N_h = 3$, the detection of an SSVEP response is performed by selecting the frequency with the maximum associated value $R(f)$. The database contained 713 trials for each subject (238 for target: when the subjects paid attention to the flickering light at 7.50Hz, 475 for non-target: when the subjects paid attention to other flickering lights).

4. RESULTS

The performances for combining classification results across subjects or across trials are estimated by the AUC, the classification accuracy and the ITR. They are depicted in Figure 1, which represents the mean AUC, accuracy and ITR across all the possible combinations of n subjects (Fig. 1(a), 1(c), 1(e)), and across all the possible combinations of n trials from the same subject (Fig. 1(b), 1(d), 1(f)). For single-event detection, the mean AUC across subjects is 0.920 ± 0.079. The results show that combining trials across subjects can improve the AUC. With the sum combined score (O_{sum}), the AUC increases from about 0.92 with a single subject to a perfect detection with eight subjects. Yet, the AUC achieved with the product and minimum scores do not monotonically increase with the number of subjects: there is a maximum AUC with about four subjects. These combined scores require that all the individual scores are relatively high to achieve a good detection while increasing the number of subjects reduces this probability.

Performance curves present the same behavior depending on the number of combined scores between the evolution of the AUC as a function of the number of subjects (Fig. 1(a)),

and as a function of the number of trials (Fig. 1(b)). For the combination of trials from the same subject and the combination of trials from different subjects, the accuracy starts at about $85.41 \pm 7.6\%$ with a single-trial, and increases sublinearly until around $99.7 \pm 0.27\%$ with ten trials and the methods O_{sum} and O_{w-sum}. Figure 1(e) shows the theoretical ITR as a function of the number of subjects that are involved in the decision with $N_{out} = 5$. The best performance is obtained with the sum method with an ITR of 137.93 ± 1.36 bpm. The ITR was estimated based on the choice between five commands and a time of $1s$ for the selection of a command ($T = 1s$). The same way that the AUC increases with the number of involved subjects, the ITR also increases for most of the combination methods. The increase of the AUC is directly reflected on the ITR (Fig. 1(e)), while it is not the case when combining trials over time (Fig. 1(f)). The evolution of the performance indicates a continuous increase of performance with the addition of new subjects. With only two subjects, it is possible to obtain a performance that is reliable enough for single-event detection.

5. DISCUSSION AND CONCLUSION

Cooperative BCIs involve the creation of new paradigms to be used online. SSVEP-BCIs are known as being fast and reliable compared to other BCIs [16, 17, 18]. To further improve the performance, we have shown that it was possible to significantly increase the ITR with simple combination methods by considering the detection of SSVEP responses occurring on a group of six subjects. With cooperative BCIs, it would be more efficient to have several users with relatively cheap systems with few sensors, than to have a single user with a very expensive system, *e.g.,* with 32 channels. A cooperative BCI system should ideally require six persons to provide optimal results in the case of SSVEP based BCI. Therefore, each subject should wear an EEG cap and have a personal amplifier. By considering the EEG signals from different users, one issue could be the synchronization the different EEG signals coming from each user. However, with a system based on the detection of SSVEP response based only on the frequency of the stimuli, a jitter would not be an obstacle if the phase is not used [19].

We have shown that combining trials across subjects can be advantageously used in particular BCI applications. Cooperative BCIs represent a fundamental new approach compared to current BCI because the user does not interact independently, but he interacts with other people who share the same goal. Moreover, the performance does not depend on a single user, but on the group of users. This promising approach could be successfully implemented in cooperative video games or in other applications involving several people.
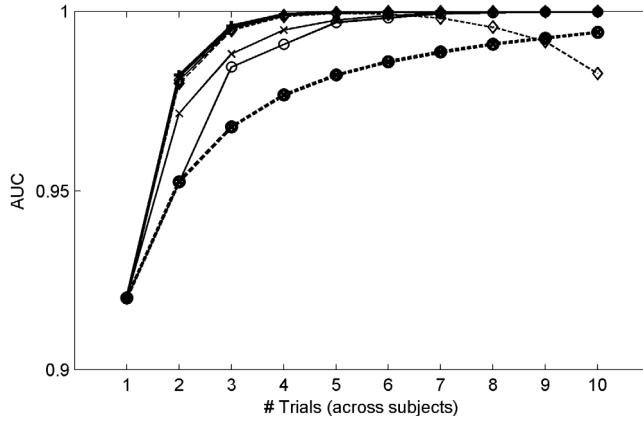
(a) AUC (single event)
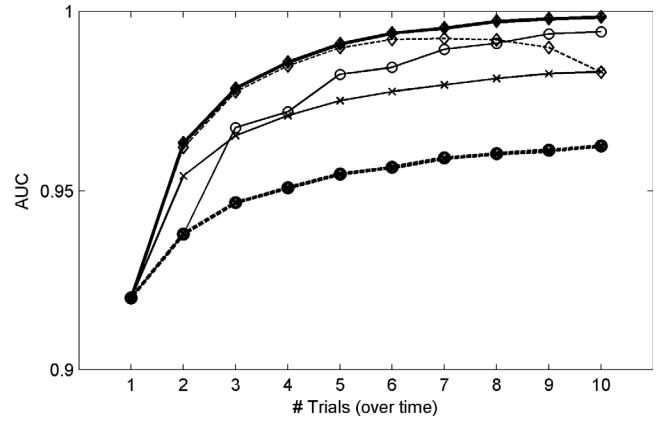
(b) AUC (single subject)
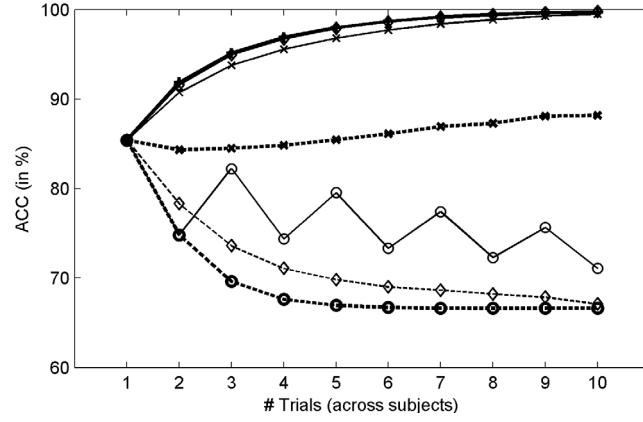
(c) ACC (single event)
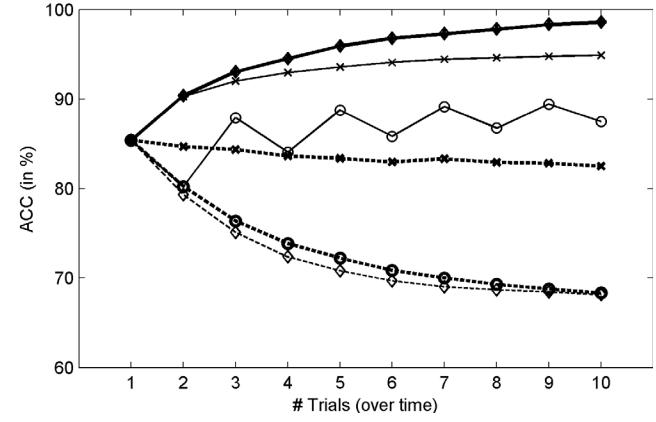
(d) ACC (single subject)
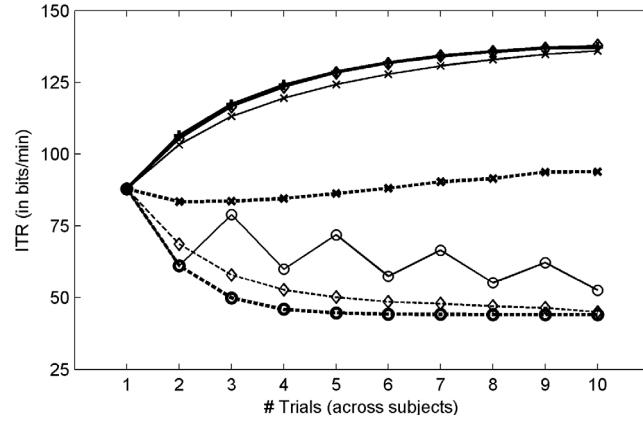
(e) ITR (single event)
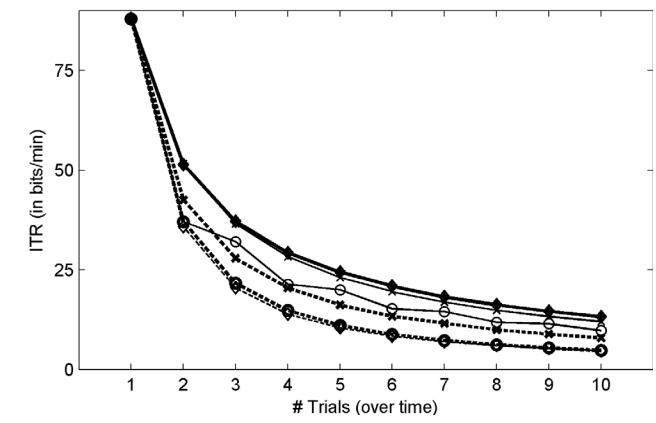
(f) ITR (single subject)

Fig. 1. SSVEP response detection results based on the AUC, the accuracy (ACC), and ITR. Figures 1(a) and 1(b): AUC for single-event detection as a function of the number of subjects and AUC for a single subject as a function of the number of trials, respectively. Figures 1(c) and 1(d): Accuracy for single-event detection as a function of the number of subjects and accuracy for a single subject as a function of the number of trials, respectively. Figures 1(e) and 1(f): theoretical ITR as a function of the number of subjects and the number of trials, respectively.

6. REFERENCES

[1] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, pp. 510–523, 1988.

[2] H. Cecotti, "Spelling with non-invasive brain-computer interfaces - current and future trends," *J. Physiology-Paris*, vol. 105, no. 1-3, pp. 106–114, 2011.

[3] J. d. R. Millán, R. Rupp, G. R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K.-R. Müller, and D. Mattia, "Combining braincomputer interfaces and assistive technologies: state-of-the-art and challenges," *Frontiers in Neuroscience*, vol. 4, no. 161, pp. 1–15, 2010.

[4] G. Townsend, B. K. LaPallo, C. B. Boulay, D. J. Krusienski, G. E. Frye, C. K. Hauser, N. E. Schwartz, T. M. Vaughan, J. R. Wolpaw, and E. W. Sellers, "A novel P300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns," *Clinical Neurophysiology*, 2010.

[5] Peng Yuan, Yijun Wang, Wei Wu, Honglai Xu, Xiaorong Gao, and Shangkai Gao, "Study on an online collaborative BCI to accelerate response to visual targets," *Proc. 34th Int. IEEE EMBS Conf.*, pp. 1–4, 2012.

[6] R. Poli, C. Cinel, A. Matran-Fernandez, F. Sepulveda, and A. Stoica, "Some steps towards realtime control of a space-craft simulator via a brain-computer interface," *Technical report CES-525, University of Essex.*, pp. 1–23, 2012.

[7] E. Lyskov, V. Ponomarev, M. Sandstrom, K.H. Mild, and S. Medvedev, "Steady-state visual evoked potential to computer monitor flicker," *Int. J. Psychophysiol*, vol. 28, pp. 285–290, 1998.

[8] F. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: focus on essential paradigms and future perspectives," *Prog Neurobiol.*, vol. 90, no. 4, pp. 418–438, 2010.

[9] Y. T. Wang, Y. Wang, and T. P. Jung, "Measuring steady-state visual evoked potentials from non-hair-bearing areas," *Proc. 34th Int. IEEE EMBS Conf.*, pp. 1806–1809, 2012.

[10] Ludmila I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, Wiley, 2004.

[11] Giorgio Fumera and Fabio Roli, "Performance analysis and comparison of linear combiners for classifier fusion," *Proc. of the Joint IAPR International Workshop on Structural, Syntactic, and Statistical Pattern Recognition*, pp. 424–432, 2002.

[12] Tom Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, pp. 861–874, 2006.

[13] C. E. Shannon and W. Weaver, "The mathematical theory of communication," *Urbana, IL: University of Illinois Press*, 1964.

[14] H Cecotti and B. Rivet, "Effect of the visual signal structure on steady-state visual evoked potentials detection," *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 657–660, 2011.

[15] O. Friman, I. Volosyak, and A. Gräser, "Multiple channel detection of steady-state visual evoked potentials for brain-computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, 2007.

[16] X. R. Gao, D. F. Xu, M. Cheng, and S. K. Gao, "A BCI based environmental controller for the motion-disabled," *IEEE Trans. Rehab. Eng.*, vol. 11, no. 2, pp. 137–140, 2003.

[17] G. Bin, X. Gao, Z Yan, B. Hong, and S. K. Gao, "An online multi-channel SSVEP-based braincomputer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, 2009.

[18] H. Cecotti, "A self-paced and calibration-less SSVEP based brain-computer interface speller," *IEEE Trans. on Neural Systems and Rehab. Eng.*, vol. 18, pp. 127–133, 2010.

[19] G. Garcia, "Detection of high-frequency steady state visual evoked potentials using phase rectified reconstruction," *Proc. of the 16th European Signal Processing Conference*, 2008.