33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Multimodal target detection using single trial evoked EEG responses in single and dual-tasks

Hubert Cecotti, Ryan W. Kasper, James C. Elliott,
Miguel P. Eckstein, Barry Giesbrecht

Abstract— **The detection of event-related potentials in the electroencephalogram signal is a common way for creating a brain-computer interface (BCI). Successful detection of evoked responses can be enhanced by the user selectively attending to specific stimuli presented in the BCI task. Because BCI users need a system that performs well in a variety of contexts, even ones that may impair selective attention, it is critical to understand how single trial detection is affected by attention. We tested 16 participants using a rapid serial visual/auditory presentation paradigm under three conditions, one in which they detected the presence of a visual target, one in which they detected the presence of an auditory target, and one in which they detected both visual and auditory targets. The behavioral performance indicates that the visual task was more difficult than the auditory task. Consistent with the higher behavioral difficulty of the visual task, single trial performance showed no difference between single and dual-task for the visual target detection (mean=0.76). However, the area under the curve for the auditory target detection was significantly lower than the dual-task (mean=0.81 for single task, 0.75 for dual-task). The results support the conclusion that single-trial target detection is impaired when attention is divided between multiple tasks.**

I. INTRODUCTION

Brain-computer interfaces (BCIs) based on event-related potentials (ERP) typically require subjects to follow a specific task in order to produce a robust and detectable ERP in the electroencephalogram (EEG) signal, *e.g.* the P300 speller [1]. A critical requirement in these BCI tasks is that the user has to pay attention to key stimuli in order to elicit an ERP. An example that illustrates the potential effects of single vs. multiple tasks is the P300 speller. Studies have shown that the efficiency of the P300 speller depends on eye gaze [2], [3]. Besides, a common practice in the P300 speller for improving P300 responses is to count the visual stimuli appearing on the target. However, a person may desire to do several tasks at the same time, *e.g.* spell a word with a P300 speller and listen to the radio. This may impair the subject's ability to count visual stimuli. If a person does other tasks, the detection may be harder to achieve. Besides, BCI are often combined with other communication devices to provide an enhanced communication control. Based on the existing empirical evidence, the extent to which BCI performance is degraded by divided attention is unclear. Here we use a dual-task paradigm to investigate the extent to understand the impact of divided attention on BCI performance.

A dual-task paradigm is a procedure in experimental psychology that requires an individual to perform two tasks simultaneously, in order to compare performance with single-task conditions [4]. Two common findings are that *(i)* performance on both tasks is impaired compared to when each task is performed in isolation and *(ii)* if the tasks are not presented at exactly the same time, performance on the second task suffers the most severe impairment [5]. The main classes of explanations are capacity sharing, processing bottlenecks and cross talk. Capacity sharing assumes that people share a limited pool of attentional resources among the different tasks. Bottleneck theories, on the other hand, assume that the impairment arises because tasks are competing for a single limited capacity mechanism that processes information serially. Cross-talk explanations assume that the dual-task impairment arises because task-relevant information from one task interferes with the processing of the other task.

In dual-task experiments, the main issue relevant for the present work is that in addition to the drop in performance relative to a single task, there is also a reduction in the amplitude of the P300 ERP component. Because the P300 amplitude is sensitive to the amount of attentional resources engaged during dual-task condition, this leads to the hypothesis that single trial detection should be impaired under dual-task conditions [6], [7]. In this study, we tested this hypothesis using a rapid serial visual/auditory presentation (RSVAP) paradigm in which subjects had targets presented in the visual modality, targets presented in the auditory modality, or targets presented in both visual and auditory modalities.

II. EXPERIMENTAL PROTOCOL

Visual stimuli consisted of images of faces and cars embedded in random Gaussian noise fields. Participants sat 125 cm from the monitor in a darkened room. Images subtended a visual angle of 4.57 degrees. Images were presented at fixation at a rate of 2Hz, with each image on the screen for 500 ms and no inter-stimulus interval.

Auditory stimuli consisted of words from the military alphabet, *e.g.* "bravo", "charlie", recorded using a computerized text-to-speech function. Ten two-syllable auditory stimuli were used, with the average duration of playback equaling 499 ms and a standard deviation of 70 ms. Auditory stimuli were presented through two Dell speakers placed to the left and right of the monitor. Auditory simuli were

This research was supported by the Institute for Collaborative Biotechnologies through contract W911NF-09-D-0001 from the U.S. Army Research Office.

Department of Psychological & Brain Sciences, and Institute for Collaborative Biotechnologies, University of California Santa Barbara, Santa Barbara, CA 93106-9660, USA email: {lastname}@psych.ucsb.edu

presented every six images. The sequence of image and auditory presentation ran uninterrupted for 2 minute intervals with self-paced breaks in between.

Sixteen subjects (11 female) ranging in age from 18 to 44 (mean=21.5, sd=5.8) participated for either $20 an hour or school credit. All procedures were approved by the University of California Santa Barbara Human Subjects Committee. Each participant performed the RSVAP task under three attention conditions:

1) In the visual-only task, participants were instructed to monitor the images for the appearance of a face, and to press the spacebar key as soon as a face image was detected. No response to the auditory stimuli was required in the visual-only task.
2) In the auditory-only task, participants were instructed to monitor the auditory stream for a target word and press the return key when the target was recognized. The target word was randomly selected at the outset of the auditory-only task and was held constant throughout the task condition. During the auditory-only task, participants were instructed to leave their eyes open and to watch the screen, but to not respond to any of the images.
3) In the dual-task, participants were instructed to simultaneously monitor the visual and auditory streams for targets. The target word in the dual-task was selected randomly and separately from the target selection in the auditory-only task.

The order of the conditions was counterbalanced across participants. It is worth noting that the visual and auditory stimuli were identical in all attention conditions of the task, with the only difference being a focus on one modality over another. The target probability for both auditory and visual stimuli was set to 10%. A total of 4,800 images (480 faces) and 800 auditory words (80 target words) were presented in each condition. As one task is based on visual stimuli and the other on auditory stimuli, it precludes sensory-level conflicts for the targets. For the behavioral response, there exists no motor conflict between the two tasks. Thus, these parameters decrease any perceptual conflict that may arise.

A. Signal acquisition

The EEG signal was measured for each subject from 32 Ag/AgCl sintered electrodes mounted in an elastic cap and placed according to the International 10-20 system. Additional electrodes were above and below each eye, 1cm lateral to the external canthii, and placed at the right and left mastoids. The data were acquired at 512 Hz.

B. Signal processing

A set of features were extracted from the EEG signal to determine if an ERP has been effectively detected or not. The goal is to find a set of features that will enhance the discrimination between targets and non targets, *i.e.* ERP on the target and ERP on non-target. Because the target occurred on 10% of the trials, we expected that each target should evoke a robust P3 ERP component [8]. Therefore, we can estimate the presence of an ERP within 1s just after the presentation of the stimulus. The EEG signal was first bandpassed filtered (Butterworth filter of order 4) with cutoff frequencies at 1 and 10.66Hz. Then, the signal was downsampled to obtain a signal at a sampling rate equivalent to 32Hz. This new sampling rate corresponded to the sampling frequency used by the winning team of the competition in an international workshop on machine learning in signal processing (MLSP) [9]. For the following steps, we considered the observed signal over 812ms after the start of a visual stimulus (26 sampling points).

The next step consisted of enhancing the relevant signal by using spatial filters. Let us denote by $U \in \mathbb{R}^{N_s \times N_f}$, the spatial filters, where N_s is the total number of sensors and N_f is the number of spatial filters. The signal after spatial filtering is defined by $X_{filt} = XU$ where $X \in \mathbb{R}^{N_t \times N_s}$ is the recorded signal, N_t is the number of sampling points.

For spatial dimension reduction, we assumed that while the expected ERP is stable, the latency and amplitude of the ERP may vary over time for a given task, we assume a spatially stationary waveform of the ERP. This assumption permitted the consideration of a single set of spatial filters that can be applied to the complete signal. We consider here the xDAWN algorithm [10], [11]. This method has been already successfully applied in BCI for P300 detection in the P300 speller paradigm. An algebraic model of the enhanced signals XU is composed of three terms: the ERP responses on the target ($D_1 A_1$), a response common to all stimuli, *i.e.* targets and non-targets confound ($D_2 A_2$) and the residual noise (H), which are filtered spatially with U.

$$XU \quad = \quad (D_1 A_1 + D_2 A_2 + H)U. \tag{1}$$

where D_1 and D_2 are two real Toeplitz matrices of size $N_t \times N_1$ and $N_t \times N_2$ respectively. D_1 has its first column elements set to zero except for those that correspond to a target onset, which are represented with a value equal to one. For D_2, its first column elements are set to zero except for those that correspond to stimuli onset. N_1 and N_2 are the number of sampling points representing the target (the P300 response) and superimposed evoked potentials, respectively. H is a real matrix of size $N_t \times N_s$. We define spatial filters U that maximize the SSNR:

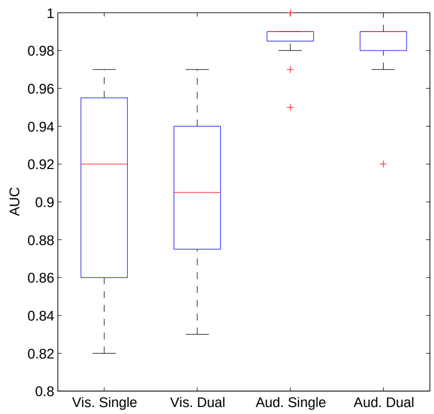
$$\text{SSNR}(U) = \text{argmax}_{\ U} \ \frac{Tr(U^T \hat{A}_1^T D_1^T D_1 \hat{A}_1 U)}{Tr(U^T X^T XU)} \tag{2}$$

where $\hat{A}_1$ corresponds to the least mean square estimation of A_1 :

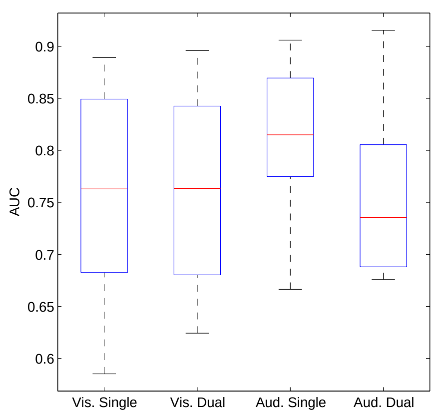
$$\hat{A} \quad = \quad \begin{bmatrix} \hat{A}_1 \\ \hat{A}_2 \end{bmatrix} = ([D_1; D_2]^T [D_1; D_2])^{-1} [D_1; D_2]^T X \tag{3}$$

where $[D_1; D_2]$ is a matrix of size $N_t \times (N_1 + N_2)$ obtained by concatenation of D_1 and D_2. Spatial filters are obtained through the Rayleigh quotient by maximizing the SSNR [11].

For the classifier input, we used four spatial filters ($N_f = 4$). For each input vector, the signals were normalized so that they had a zero mean and a standard deviation equal to one for each spatial component. Finally, the input vector was obtained by the concatenation of the N_f time-course signals across spatial filters. The Bayesian linear discriminant analysis (BLDA) classifier was used for the detection of the P300 wave [12], [13].



(a) Behavioral performance



(b) Single trial detection performance

Fig. 1. Behavioral and single trial detection performance in single and dual-task for auditory and visual target detection. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually.

III. Results

A. Behavioral performance

Behavioral performance was assessed by computing the estimated area under the curve (AUC) of the receiver operating characteristic (ROC) curve. This area was obtained as the normal cumulative distribution function of $d'/\sqrt{2}$ where d' is the sensitivity index $d' = Z(hit\ rate) - Z(false\ alarm\ rate)$ and $Z(p)$, $p \in [0, 1]$, is the inverse of the cumulative Gaussian distribution. This estimation allows us to obtain the same measurement for both behavioral and single trial classification. The mean hit-rate across subjects decreased for both the visual and auditory detection from the single task to the dual task, from 57.94% to 54.53% and from 85.70% to 81.17%, respectively. However, the only significant difference was for auditory target detection: paired t-test revealed that the hit-rate in the dual task was lower than in the single task ($t_{15} = 2.382$, $p < 0.02$). The low hit-rate of the visual target detection shows the difficulty of this task compared to the auditory target detection.

B. Single trial detection

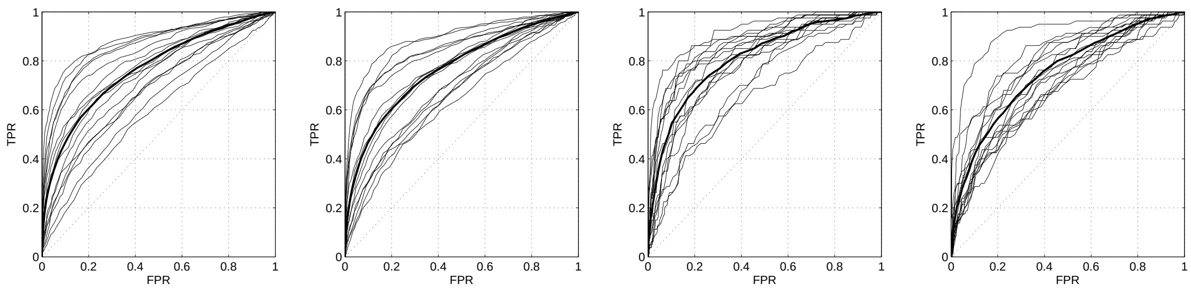
The different ROC curves for the detection of visual and auditory targets in single and dual-task conditions are presented in Figure 2. The mean AUC was 0.758, 0.759, 0.809, and 0.751, for single task visual target, dual-task visual target, single task auditory target and dual-task auditory target, respectively. For the four conditions, the standard deviation of the AUC is 0.090, 0.083, 0.071 and 0.067. The performance based on the AUC is resumed in Figure 1(b). These results show that it is possible to classify visual and auditory targets during single and dual-task. A pairwise t-test comparison indicates that there is no difference between the single and dual-task for visual target detection. For the auditory target detection, a pairwise one-tail t-test comparison shows that the single task provides a better target detection than with the dual-task ($t_{15} = 3.37$, $p = 0.0021$).

The average spatial distribution of each type of target detection and for each condition is shown in Figure 3. For the detection of visual targets, the activity is principally located in the occipito-parietal area. In the single task condition, the activity is higher in the right occipital area compared to the dual-task where the activity is more centered around P_Z. The activity related to the detection of auditory targets is present in the fronto-central area in the single task condition. In the dual-task, the activity is spread across the centro-parietal area. The difference between the single task and dual-task condition is more important in the auditory target detection, explaining the difference of accuracy between the conditions. The low difference for the visual tasks confirms the previous results showing no difference between single and dual-task.

IV. Discussion and Conclusion

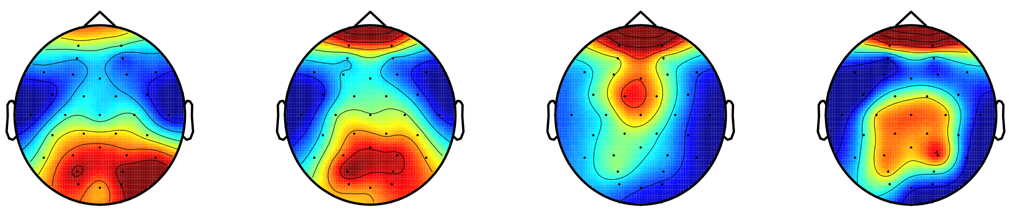
The presented method for target detection has been successfully evaluated on two different stimuli: visual and auditory. In every case, the AUC is superior to 0.75, proving that it is possible to detect a target relatively efficiently with a single trial of EEG activity. In the dual-task condition, the visual and auditory tasks had the same priority and we could have expected the same results as with the single task. However, visual target detection was a difficult task due to the high level of noise that was included in the images. The simultaneous auditory target detection had no impact on the visual target detection in the dual-task condition. For both the behavioral and single trial detection, there is no difference between single and dual-task. Yet, in the auditory task performance was degraded under dual-task conditions. Hence, single trial detection with EEG signal can be used as an efficient proxy for monitoring the performance of a subject between single and dual-task.

Transferring BCI paradigms outside of the laboratory while keeping a high reliability is a challenge. It is not only



(a) Single task - Visual target (b) Dual task - Visual target (c) Single task - Auditory target (d) Dual task - Auditory target

Fig. 2. ROC curves (the bold curve represents the mean across subjects).



(a) Single task - Visual target (b) Dual-task - Visual target (c) Single task - Auditory target (d) Dual-task - Auditory target

Fig. 3. Spatial distribution of the ERP based on the xDAWN algorithm [11]. (the red/blue color denotes a high/low activity.)

due to the material (EEG cap, amplifier) but also to the use of such system in real condition where the user has to achieve several tasks simultaneously. The present results underscore the importance of considering the real-world psychological context when evaluating the performance of any BCI. For instance, dual-task models and experiments may explain the differences between copy spelling and free spelling in some BCI spellers.

The present results not only have implications for the evaluations of BCIs designed to help disabled persons, but also for BCIs designed to enhance performance of healthy users. For example, BCIs have been developed for target detection is for detecting potential threats like in satellite images [14], [15], [16]. BCI could complement other communication devices, but only if these other tasks do not impair BCI performance. The present results suggest that these BCIs may be impacted if users are also engaged in other tasks. Future work is needed to define which BCI tasks are most robust to dual-task interference and changes in task difficulty. Further works will deal with the evaluation of the task difficulty and its impact on the EEG classification in dual-task conditions.

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] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, "Does the 'P300' speller depend on eye gaze?" *J Neural Eng*, vol. 7, no. 5, 2010.

[3] M. S. Treder and B. Blankertz, "(C)overt attention and visual speller design in an ERP-based brain-computer interface," *Behavioral and Brain Functions: BBF*, vol. 6, p. 28, 2010.

[4] H. Pashler, "Dual-task interference in simple tasks: data and theory," *Psychological Bulletin*, vol. 116, no. 2, pp. 220–244, 1994.

[5] H. Pashler and J. C. Johnston, "Attentional limitations in dual-task performance," *Attention*, pp. 155–189, 1998.

[6] J. B. Isreal, G. L. Chesney, C. D. Wickens, and E. Donchin, "P300 and tracking difficulty: evidence for multiple resources in dual-task performance," *Psychophysiology*, vol. 17, no. 3, 1980.

[7] J. Polich, "Updating P300: An integrative theory of P3a and P3b," *Clinical Neurophysiology*, vol. 118, pp. 2128–2148, 2007.

[8] R. J. Johnson, "A triarchic model of P300 amplitude," *Psychophysiology*, vol. 23, no. 4, pp. 367–84, 1986.

[9] J. M. Leiva and S. M. M. Martens, "MLSP competition, 2010: Description of the first place method," *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 112–113, 2010.

[10] H. Cecotti, B. Rivet, M. Congedo, C. Jutten, O. Bertrand, E. Maby, and J. Mattout, "A robust sensor selection method for p300 brain-computer interfaces," *Journal of Neural Engineering*, vol. 8, 2011.

[11] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, "xDAWN algorithm to enhance evoked potentials: application to brain-computer interface," *IEEE Trans Biomed Eng.*, vol. 56, no. 8, pp. 2035–43, 2009.

[12] U. Hoffmann, J. Vesin, K. Diserens, and T. Ebrahimi, "An efficient P300-based brain-computer interface for disabled subjects," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 115–125, 2008.

[13] D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, vol. 4, no. 3, pp. 415–447, 1992.

[14] N. Bigdely-Shamlo, A. Vankov, R. R. Ramirez, and S. Makeig, "Brain activity-based image classification from rapid serial visual presentation," *IEEE Trans. on Neural Systems and Rehab. Eng.*, vol. 16, no. 5, 2008.

[15] A. Gerson, L. Parra, and P. Sajda, "Cortically-coupled computer vision for rapid image search," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 174–179, 2006.

[16] L. C. Parra, C. Christoforou, A. D. Gerson, M. Dyrholm, A. Luo, M. Wagner, M. G. Philiastides, and P. Sajda, "Spatio-temporal linear decoding of brain state: Application to performance augmentation in high-throughput tasks," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 95–115, 2008.