34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Multiclass classification of single-trial evoked EEG responses

Hubert Cecotti [1], Anthony J. Ries [2], Miguel P. Eckstein [1], Barry Giesbrecht [1]

Abstract— **The detection of event-related potentials (ERPs) in the electroencephalogram (EEG) signal has several real-world applications, from cognitive state monitoring to brain-computer interfaces. Current systems based on the detection of ERPs only consider a single type of response to detect. Hence, the classification methods that are considered for ERP detection are binary classifiers (target vs. non target). Here we investigated multiclass classification of single-trial evoked responses during a rapid serial visual presentation task in which short video clips were presented to fifteen observers. Each trial contained potential targets that were human or non-human, stationary or moving. The goal of the classification analysis was to discriminate between three classes: moving human targets, moving non-human targets, and non-moving human targets. The analysis revealed that the mean volume under the ROC surface of 0.878. These results suggest that it is possible to efficiently discriminate between more than two types of evoked responses using single-trial detection.**

I. INTRODUCTION

The detection of evoked brain responses has evolved over the years for different types of application thanks to machine learning techniques that improve the detection of brain responses like event-related potentials (ERPs). The detection of ERPs is widely used for the creation of brain-computer interfaces, like the P300 speller [1]. In classical ERP based BCIs, a particular user interface is proposed to allow the selection of more than two commands. Indeed, ERP based BCIs do not assign the detection of an ERP to a command but they code multiple commands thanks to a sequence of several stimuli. For BCIs that offer the possibility to select more than two commands, it is important to distinguish BCIs that code several commands with the detection of a single brain response, *e.g.* the P300 speller, and BCIs that assign a command for each brain response, *e.g.* motor imagery based BCI. For motor imagery BCI, it is possible to consider a command for an imagery movement of different limbs, *e.g.* 4-class motor imagery BCI with left-hand, right-hand, foot and tongue actions [2].

Whereas it is possible to code several commands with the detection of a single type of brain responses for BCI, other types of paradigms, like rapid serial visual presentation (RSVP) tasks, require the detection of several types of brain responses to assign a visual stimulus to a particular class. Indeed, RSVP tasks have been successfully used for target detection [3], [4]. In this context, it is important to allow the possibility to detect several types of targets. To achieve such a task, the class of targets should be different enough to evoke ERP with different characteristics. After the presentation of a stimulus to a subject, it is typically possible to detect several peaks of different latencies and amplitudes in the EEG signal. A highly investigated component is the P300, which is a positive deflection of the ongoing EEG signal with a latency of more than 300ms to an event [5]. In addition to the P300, there exist other components like the N200, which is the most distinct response to motion-onset [6]. The different components of the ERP can depend on various parameters like the meaning of the stimulus (target vs. non target), the target probability, and the complexity of the choice [7], [5].

As the different components of an ERP may change in relation to the presentation of a particular type of target, they could be exploited as features for detection. In this study, we investigated single-trial detection of several types of targets using an RSVP task containing three types of visual stimuli: moving human, moving non-human, and stationary human. Classifier performance in this difficult task was precisely quantified using the volume under the ROC hyper-surface (VUS).

The remainder of the paper proceeds as follows. The second section details to the experimental protocol. The signal processing and evaluation methods are described in the third section. The ERP waveforms and the results for the binary and multiclass classifiers are presented in the fourth section. Finally, the interest of a the detection of several single-trial brain-evoked responses is discussed in the last section.

II. EXPERIMENTAL PROTOCOL

Fifteen participants (average age 39.5, 9 male) volunteered for the current study. Participants provided written informed consent, reported normal or corrected-to-normal vision and reported no history of neurological problems. Fourteen of the fifteen participants were right-handed.

Cognitive Technology Threat Warning System (CT2WS) video clips were used in an RSVP paradigm where observers made a manual button press when they detected a target (person or vehicle) presented among distractors. Video clips consisted of five consecutive images each 100ms in duration (500ms total duration). If a target appeared in the video clip it was present on each 100ms image. The distractor to target frequency ratio was 90/10. Distractors were background desert scenes without targets. Half the targets were people and half were vehicles. Within the two target classes, half were moving and half were stationary. RSVP sequences were presented in two minute blocks after which time observers were given a short break. Observers completed a total of 25

[1] Department of Psychological & Brain Sciences, and Institute for Collaborative Biotechnologies, University of California Santa Barbara, Santa Barbara, CA 93106-9660, USA

[2] U.S. Army Research Lab Bldg 459, Translational Neuroscience Branch, Aberdeen Proving Ground, MD 21005, USA

blocks. For the behavioral response, the subjects pressed a button on the keyboard with their dominant middle finger when they detected either a person or a vehicle.

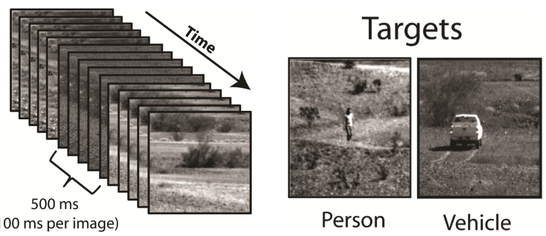

Fig. 1. RSVP task (left) and examples of targets (right).

Electrophysiological recordings were digitally sampled at 512Hz from 64 scalp electrodes arranged in a 10-10 montage using a BioSemi Active Two system. External leads were placed on the outer canthus and below the orbital fossa of both eyes to record EOG.

III. Methods

A. Signal processing

A set of features were extracted from the EEG signal to determine if an ERP corresponding to a particular class was detected or not. The goal was to find a set of features that enhanced the discrimination between two types of brain evoked responses. The EEG signal was first bandpassed filtered (Butterworth filter of order 4) with cutoff frequencies at 1 and 10.66Hz. The signal was downsampled by a factor of 16 to obtain a signal at 32Hz. For the following steps, we considered the observed signal over 612ms after the start of a visual stimulus (20 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. We consider the xDAWN method for spatial filtering [8]. It has been successfully applied in P300 based BCI and in RSVP tasks [9], [10]. An algebraic model of the enhanced signals XU is composed of three terms: the ERP responses on a target class ($D_1 A_1$), a response common to all stimuli, *i.e.* all targets and non-targets confound ($D_2 A_2$) and the residual noise (H), which are filtered spatially with U.

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

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 and superimposed evoked potentials, respectively. H is a real matrix of size $N_t \times N_s$. The spatial filters U maximize the signal to signal plus noise ratio (SSNR):

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

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

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

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 [8]. During the experiments, four spatial filters ($N_f = 4$) were used. The input vector for the classifier was obtained by the concatenation of the N_f time-course signals across spatial filters. The Bayesian linear discriminant analysis (BLDA) was used as a binary classifier [11], [12].

B. Evaluation methods

The Receiver Operating Characteristic (ROC) analysis technique has been applied to classification problems with two classes, *e.g.* for target detection with only one target [13]. Area Under the ROC Curve (AUC) has become a standard for evaluating the performance of binary classifiers. The ROC analysis includes several advantages, including the ability to determine classification accuracy based on the prior probabilities of classes. Given an application and a classifier, setting decision thresholds should be set accordingly to the problem, *e.g.* errors should be the same for every class, or a particular class should require no error. The choice of optimal thresholds can be critical for military or clinical applications where the reliability shall be optimized, *i.e.* the cost of some errors may have dramatic implications. Therefore, it is important to use a method for determining the thresholds that minimize the overall risk of an application. While this is typically done by taking into account the multiclass confusion matrix, because it introduces computational complexity when there are a high number of classes [14], [15], here we keep the AUC concept and extend it to the volume under the ROC hyper-surface (VUS). In this study we consider a 3-class problem by estimating the VUS (EVUS) by using three classifiers that adopt a one vs all strategy [16].

The operating point in the VUS is varied by changing the posterior output of the results by the vector $\Phi = [\phi_1, \phi_2, \phi_3]$, which corresponds to the classifier thresholds. $\phi_i > 0$, $i = 1, 2, 3$. The confusion matrices obtained from all combinations of Φ provide all the operating points in VUS. For the estimation of EVUS, we consider $N_\phi = 20$ thresholds in each class. Hence, we obtain a hyper-polyhedron with N_ϕ^3 vertices. The EVUS described by this hyper-polyhedron is then estimated by using the optimized QHull algorithm [17]. It is worth mentioning that for a problem with three classes, the EVUS for a random classifier is 0.166. In the next section, the EVUS for three classes is represented by a triangular surface plot. The class MTH, MTNH and NMTH denote the ERP responses associated to the presentation of

	MTH	MTNH	NMTH
MTH	109.78 ± 3.92	4.20 ± 2.95	14.09 ± 1.99
MTNH	15.71 ± 5.96	51.49 ± 20.14	15.49 ± 5.84
NMTH	15.53 ± 2.77	4.49 ± 2.24	101.40 ± 8.46

TABLE I

AVERAGED CONFUSION MATRIX

a moving target human, a moving target non-human, and a non-moving target human, respectively.

IV. RESULTS

In the present section, we first show the ERP waveforms for each stimulus, then the results of the binary classifiers and the multiclass classification. Figure 2 depicts the grandaveraged ERP waveforms for each stimulus class plotted with a baseline correction of -200-0 ms on the electrodes F_Z, C_Z, P_Z, O_Z, P_7 and P_8. These plots were created for each stimulus class and low-pass filtered at 30hz. Continuous, artifact free data were timelocked to stimulus onset and epoched from -200ms to 1000 ms. Only targets followed by a response within 200-1000ms or non-targets followed by no response were included in the analysis.

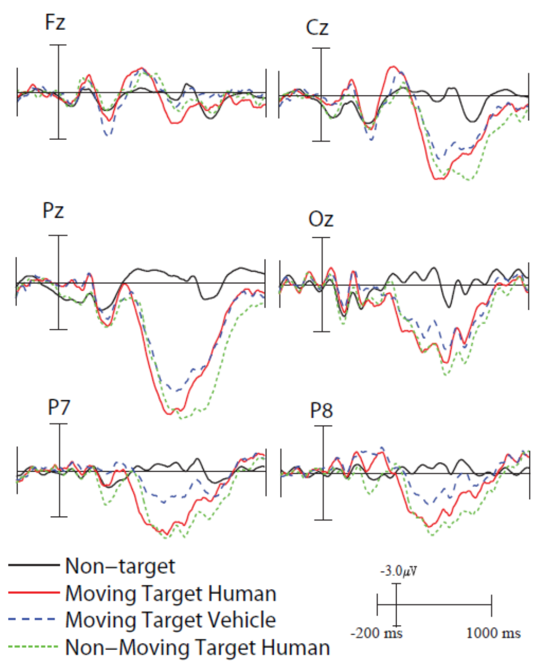

Fig. 2. Grandaveraged ERP waveforms for each stimulus.

The binary classification of each class with a one vs all approach was first evaluated. The AUC for these binary classifications is presented in Fig. 3. The mean AUC for the detection of a moving target human (MTH), a moving target non-human (MTNH), and a non-moving target human (NMTH) was 0.907, 0.855 and 0.914. The detection of MTH is easier than MTNH ($p < 0.05, t_{14} = 2.404$). Detection of NMTH was better than both MTNH ($p < 0.05, t_{14} = 2.589$) and MTH ($p < 0.05, t_{14} = 2.589$). These results suggest that it is easier to detect stationary human targets.

The confusion matrix Γ represents the main analysis for the performance of the multiclass classification. The diagonal elements of Γ represent the performance of each class, the other elements of Γ represent the different confusion errors between classes. The element $\Gamma(i, j)$ represents the number of evoked responses of class i that were detected as the class j. Table I presents the confusion matrix for the three classes.

Figure 4 presents the EVUS for each subject. The mean EVUS across subjects is 0.878 with a standard error of 0.0069.

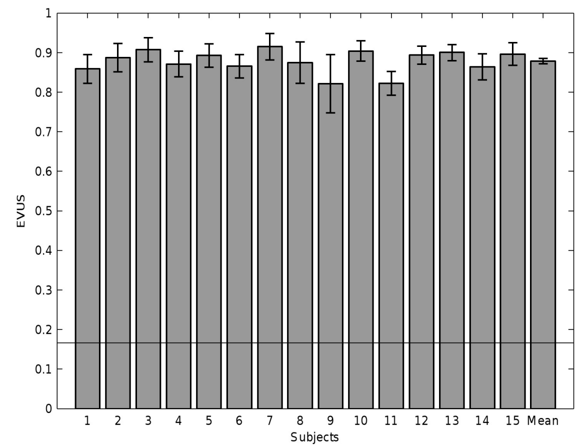

Fig. 4. EVUS for each subject. The error bars correspond to the standard error across sessions for each subject, and across subjects for the mean.

An example of a ROC surface corresponding to subject 1 is depicted in Figure 5. The curves represented on the sides of the cube represents the ROC curves associated to the pairwise analysis of the different classes, *i.e.* (MTH,MTNH), (MTNH,NMTH), and (MTH,NMTH).

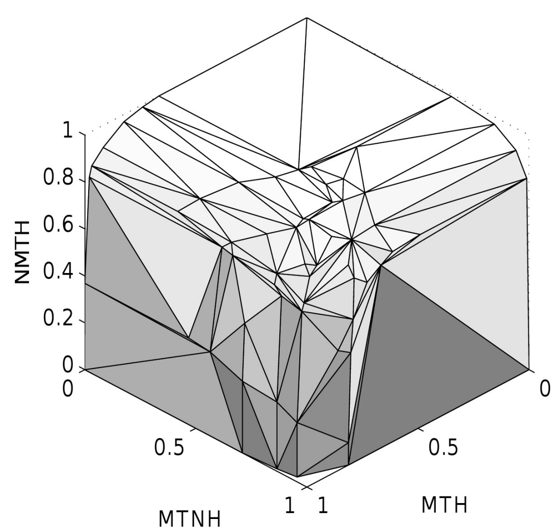

Fig. 5. Example of a ROC surface representing the performance of subject 1. (EVUS=0.9507)

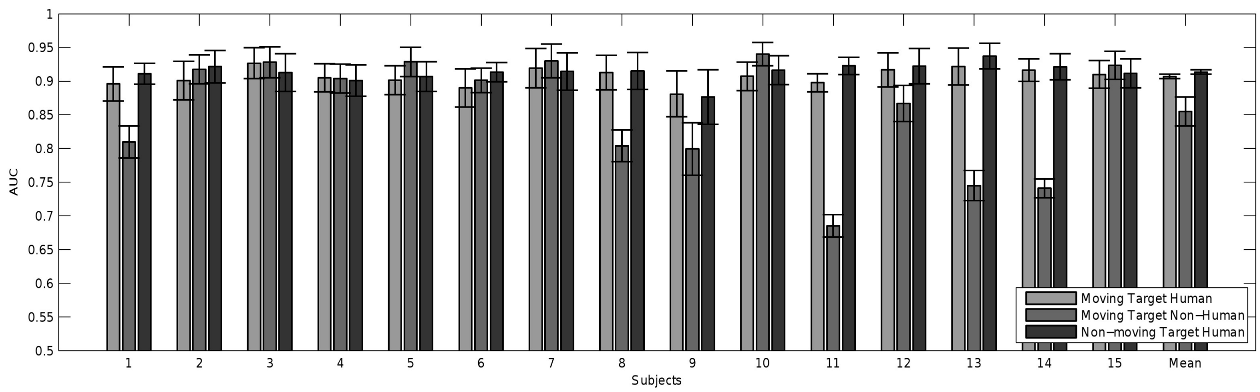

Fig. 3. AUC for each subject. The error bars correspond to the standard error across sessions for each subject, and across subjects for the mean.

V. Discussion and conclusion

In the ERP literature, it is well known that the target type can influence the different ERP components of the brain-evoked response. Those differences have yet to be fully exploited in the BCI field. For current ERP based BCI, the typical strategy relies on the detection of the presence of a brain response and not its type. This latter strategy does not fully explore the features of the different ERP components, *e.g.* the amplitude, latency and spatial distribution. Indeed, its main goal is to determine differences between the ERP waveforms evoked by the presentation of a stimulus corresponding to a target and one corresponding to a non-target. In this case, the problem corresponds mainly to the detection of the presence of a N200 and/or a P300 waveform in the signal. While this binary classification problem for single-trial remains difficult, we have shown that it is possible to successfully classify three different evoked brain responses using single-trial detection. This study has highlighted that it is possible to go beyond the level of presence/absence of an ERP component by considering differences between ERPs at the single-trial level.

These promising results could allow improving the throughput of target detection systems based on the detection of neural responses. In applications where potential targets can be presented in different ways like moving or stationary, it is important to categorize or rank the target type automatically. Indeed, target detection in natural images obtained from photos in real environments can contain many distractors that may look close to a target. Further works will deal with online multiclass classification.

Acknowledgment

This research was supported by the Institute for Collaborative Biotechnologies through contract W911NF-09-D-0001 from the U.S. Army Research Office. The authors would like to thank Gabriella Brick Larkin for her assistance with the data collection and Matthew Jaswa for programming the RSVP task.

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] Z. Y. Chin, K. K. Ang, C. Wang, C. Guan, and H. Zhang, "Multi-class filter bank common spatial pattern for four-class motor imagery BCI," *Proc of the IEEE Eng Med Biol Soc.*, pp. 571–4, 2009.

[3] E. A. Pohlmeyer, J. Wang, D. C. Jangraw, B. Lou, S. Chang, and P. Sajda, "Closing the loop in cortically-coupled computer vision: a brain-computer interface for searching image databases," *J. Neural Eng.*, vol. 8, p. 036025, 2011.

[4] P. Sajda, A. Gerson, and L. Parra, "High-throughput image search via single-trial event detection in a rapid serial visual presentation task," *Proc. of the 1st Int. IEEE EMBS Conf. Neural Eng.*, 2003.

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

[6] B. Hong, F. Guo, T. Liu, X. Gao, and S. Gao, "N200-speller using motion-onset visual response," *Clinicial Neurophysiology*, vol. 120, pp. 1658–1666, 2009.

[7] M. Falkenstein, J. Hohnsbein, and J. Hoormann, "Effects of choice complexity on different subcomponents of the late positive complex of the event-related potential," *Electroencephalography and clinical Neurophysiology*, vol. 92, pp. 148–160, 1994.

[8] 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.

[9] 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," *J. Neural Eng.*, vol. 8, p. 016001, 2011.

[10] H. Cecotti, R. Kasper, J. C. Elliott, M. P. Eckstein, and B. Giesbrecht, "Multimodal target detection using single trial evoked eeg responses in single and dual-tasks," *32nd International IEEE EMBC Conference*, 2011.

[11] 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.

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

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

[14] C. Ferri, J. Hernandez-Orallo, and M. A. Salido, "Volume under the ROC surface for multi-class problems," *Proc. of the 14th European Conference on Machine Learning*, pp. 108–120, 2003.

[15] T. C. W. Landgrebe and R. P. W. Duin, "Approximating the multiclass ROC by pairwise analysis," *Pattern recognition letters*, vol. 28, pp. 1747–1758, 2007.

[16] ——, "A simplified extension of the area under the ROC to the multiclass domain," *Proc of the 17th annual Symposium of the Pattern Recognition Association of South Africa*, 2006.

[17] C. B. Barber, D. Dobkin, and H. Huhdanpaa, "The quickhull algorithm for convex hulls," *ACM Transactions on Mathematical Software*, vol. 22, no. 4, pp. 469–483, 1996.