EEG correlates of active visual search during simulated driving: An exploratory study

Hadrien Renold†, Ricardo Chavarriaga†, Lucian Gheorghe†,‡, José del R. Millán†
† Chair on Non-Invasive Brain-Machine Interface (CNBI), Center for Neuroprosthetics, School of Engineering,
Ecole Polytechnique Fédérale de Lausanne, Switzerland
{hadrien.renold,ricardo.chavarriaga,jose.millan}@epfl.ch

‡ Mobility Services Laboratory, Nissan Research Center
Nissan Motor Co., LTD., Atsugi, Japan
lucian@mail.nissan.co.jp

***Abstract*—Brain responses to visual stimuli can provide information about visual recognition processes. Several studies have shown stimulus-dependent modulation of the evoked neural responses after gaze shifts (i.e. eye fixation related potentials, EFRP) depending on the relevance of the fixated object. However these studies are typically performed on still images under constrained conditions. Here we extend this approach to study overt visual attention during a simulated driving task. Simultaneous analysis of eye-tracking and electroencephalography data revealed similar patterns than those previously reported. However, natural visual exploration yielded shorter fixations, which imposes constraints in the analysis of the elicited brain responses. Nevertheless, we found significant differences between EFRPs corresponding to target objects or non-object stimuli. These results suggest the possibility of decoding such information during driving, allowing better understanding of how drivers process the environmental information.**

I. Introduction

Brain-Computer Interfaces (BCI) –traditionally aimed at providing a communication channel for people with motor disabilities– can also provide insight into the cognitive processes taking place during daily tasks. This information can in turn be used to improve performance in such tasks. For instance, smart cars could use it to provide tailored support for each driver [1]–[3]. Existing works in these lines have focused on detecting anticipated and emergency braking [2], [4], steering actions [3] as well as workload and levels of attention [5]. Here we focus on identifying neural patterns related to visual processing and recognition during driving. Decoding of these patterns can provide information to the car about objects in the environment (e.g. traffic signs, detour panels or pedestrians) that have been perceived as relevant by the driver, as well as those that may have been neglected.

Previous work has reported such correlates in visual search tasks for static and simple images with and without eye movements [6]–[8], [9]. They suggest that evoked EEG responses can be decoded to identify when a subject perceives a target image amid a sequence of other stimuli. In these experiments, using the so called rapid serial visual protocol, the subject fixates on a screen while he/she observes images presenting at the same location [6], [7]. More recently, this type of analysis has been extended to cases where subjects are allowed to freely move their eyes to scan larger, more complex images [8], [9].

This study further explores these findings in a more ecological scenario where subjects are driving in a simulated environment. We report brain patterns elicited by perception and recognition of relevant and irrelevant visual stimuli along the road. Unlike traditional experimental protocols in well-controlled conditions, this work requires simultaneous analysis of eye movement patterns and electroencephalography signals. This allows us to identify stimulus-dependent modulation of the evoked EEG responses after gaze shifts (i.e. eye fixation related potentials, EFRP)

II. Methods

A. Experimental protocol

Six subjects (1 woman; aged between 24-30) took part in the experiment. All of them had normal or corrected to normal vision (2 subjects used glasses and two used contact lenses). The study was approved by the Research ethical committee of the EPFL Brain and Mind institute and all participants gave their informed consent. Subjects were asked to drive in a simulator following the directions given by arrows on the ground just before the intersections (Figure 1). While driving they were required to actively explore the visual scene to look for a specific visual stimuli.

The only instructions given to subjects were to count the number of boards on a road segment that contain a visual target that matches a previously shown cue. They had to report this number at the end of the trial after having passed the last board (i.e., the response does not reflect the time when the stimuli is perceived). The reporting was done by way of a pedal shift on the steering wheel. If there was no cue at the beginning of the trial, subjects were instructed to look freely around. Subjects performed 5 runs on the same map. For each run there was 10 trials of each search task (either capital E as a target or inverted capital E) and 8 trials with no task, all randomly ordered.

1) Simulated driving environment: The simulated driving environment is a regular grid city, with no other vehicles or signs. The vehicle speed is limited to 60 km/h. Each city block

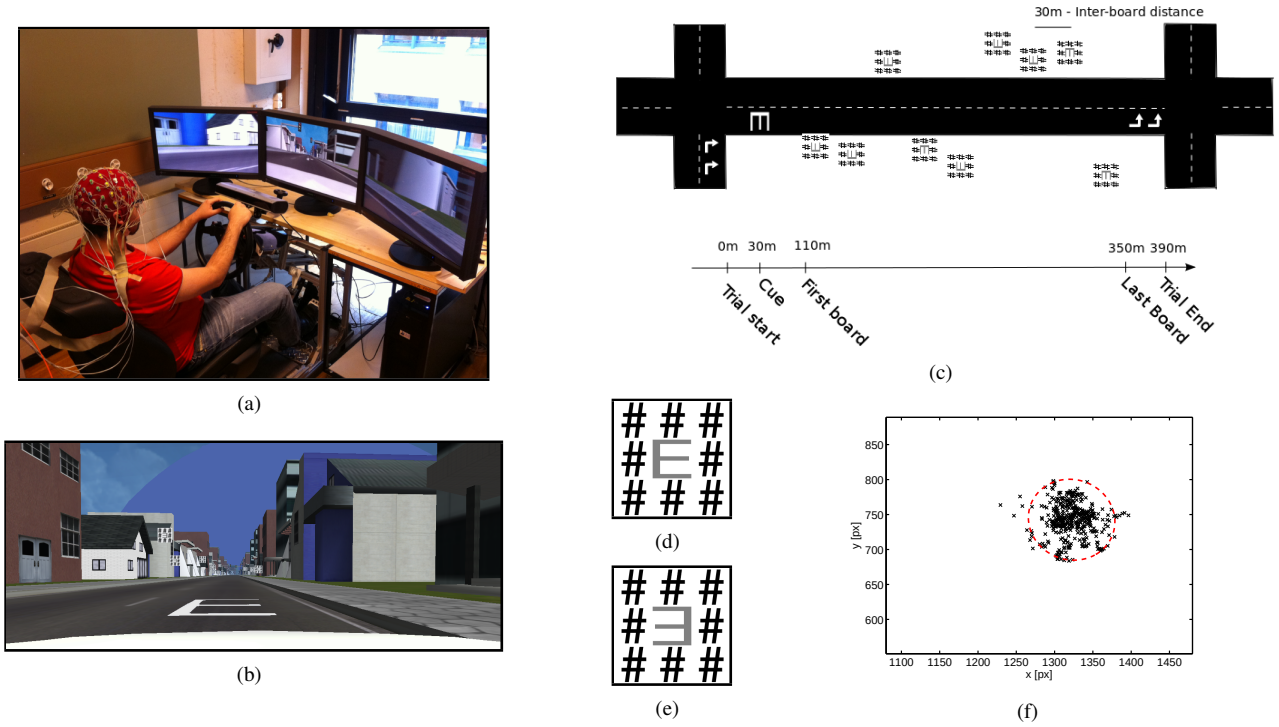

Fig. 1. Experimental protocol. (a) Experimental setup. (b) Example of the driver's view at the beginning of each trial. (c) Schematic representation of one protocol trial. (d-e) Stimulus boards. (f) Example of the calibration of gaze dispersion (Subject S2). Each point corresponds to the gaze location during the calibration period. The dotted red circle corresponds to the 95% confidence interval of the fitted two dimensional Gaussian. The scale is zoomed in around the center with the same ratio as the screen.

is considered a new trial. For each trial there are 9 boards on the side of the street of size 3 by 3 m. Each board contains either a capital E or inverted capital E surrounded by # signs, as shown on Figures 1(d) and 1(e). The symbols were chosen because they are geometrically and saliently equivalent. The surrounding # signs are there to create a crowding effect and force the foveating in the center of the board [8]. The boards are evenly spaced by 30 m, and placed randomly on either side of the street. Their horizontal and vertical positions were randomly attributed in the ranges 0-8 m from the street and 0-10 m from the ground, respectively, and adjusted to avoid overlaps on the screen. At the beginning of each block (i.e. trial) there is a cue on the ground indicating the target for the current trial. The number of targets per trial varied between 0-4 (average number across trials was 2.5), whereas the rest of the boards were distractors. This yields about 250 target signs and 650 non-target signs per subject. The placement of the boards ensures that the subject has to actively scan the image to search for the target board.

B. Experimental set-up

The experimental set-up consists of a car simulator composed of an adjustable car seat, steering, gas and brake pedals (no clutch) and three optical 3D screens. Subjects where seated at about 1.5 m from the screen. Car related data (e.g. steering angle, speed, position) from the driving simulator was recorded at 256 Hz, as well as the position on the screen of the visual stimuli.

Eye movement data was recorded at 120 Hz via a USB RED eye tracker (SMI Vision). The tracker has a horizontal viewing angle of $\pm$ 30° and was placed at 60 cm from the subject in such a way that it did not obstruct the screen's line of sight. It uses the corneal reflection as well as head tracking to calculate the view point on the screen. During the experiment the eye movement tracking was limited to the center screen.

EEG data was recorded at 2048 Hz using a 64 channel BioSemi ActiveTwo System, in an extended 10/20 layout configuration. EOG was also recorded using electrodes positioned above the nasion and below the outer canthi of the eye as well as below and above the right eye.

Recordings were synchronized for off-line analysis using a 4 Hz square signal sent from the driving simulator to the eye tracker via UDP and to the EEG recording device via parallel port.

C. Eye Movement data preprocessing

We calibrate the eye tracker before each experiment, including a 9 point calibration of the eye tracker and a 4 point validation. The process was performed until the standard deviation of the distance between the gaze fixation point recorded and the position of the cross on screen was smaller than 1°.

During the experiment eye fixations were detected using the SMI Event Detector which relies on a dispersion algorithm.

It detects as a primary event fixations and blinks and the rest of the data is consider as saccades. The parameters of the algorithm are: the minimum fixation duration Δt_{min} and the maximum dispersion D_{max}. To detect fixations it uses a sliding window of length Δt_{min} and calculates dispersion over that window. If the dispersion is below D_{max} then the window length is increased until the dispersion crosses the threshold. The center of fixation is given by the centroid of the samples within this time window. Based on previous works we set the minimum fixation duration Δt_{min} to $100ms$.

The dispersion parameter was set independently per subject. For this purpose, they were asked to fixate at a cross at the center of the screen for at least 4 s. We fit a normal distribution on the gaze positions provided by the eye-tracker and set D_{max} as the dispersion of the iso-contour of 95% confidence interval of the fitted distribution. Figure 1(f) shows an example of this process. The normal distribution yielded a good fit for all subjects. Since one of the subjects (S1) didn't perform this phase, the dispersion threshold was set as the average value of other participants (i.e. corresponding to a circle of radius $\sim$ 60 px).

Finally, we estimate the maximal distance at which subjects could perceive the stimuli on the board. This value was used to ensure that fixations considered for analysis fell within the visual range of the subject. For that purpose, subjects first got familiar with the protocol by driving in the simulated environment. Then for 5 trials they were asked to press the button as soon as they could identify the furthest board, indicated by the experimenter, and report immediately the symbol in it. The viewing distance was set as the average distance over 5 trials.

D. EEG data preprocessing

EEG data was downsampled to 256 Hz, then spatially filtered using a 5 cm Laplacian. A zero phase fourth order Butterworth bandpass filter in the range [2,12] Hz was applied to remove slow oscillations and high frequency noise. EOG artifact correction methods were not applied since they may corrupt the EFRPs. Moreover, we analyse signals during the fixation period, where eye movements are expected to be absent, or very small (at most limited by the dispersion threshold, c.f. Section II-C). Signals were visually inspected in the temporal and frequency domain to detect noisy electrodes. Rejected electrodes were replaced by the average value of its neighbours. If more than 4 channels were noisy, or if those affected the areas of interest, the run was discarded from the study. In total, two subjects had one run discarded for this reason. After extracting fixation related epochs, we rejected epochs if any channel had an amplitude greater than 150 μV.

E. Epoch selection

We analyse eye-fixation related potentials (EFRP) to characterize target-dependent modulations. To this end, gaze shift information was used to identify fixations and extract corresponding EEG epochs (c.f. II-C). We then compute the probability that each epoch corresponds to a gaze fixation over a stimulus board. For that purpose we use the same normal distribution estimated at the calibration phase centred at the fixation location (points outside the 95% confidence interval were assigned a zero value). We integrate this function over the region overlapping with a given board to assess the probability that the driver foveated on it. Since the distance to the board changes due to the car movements, we choose the mean distance over the fixation to compute this probability. We discard fixations which overlap more than one board, as well as those where the fixated board was further than the viewing distance estimated at the calibration phase. Finally, we assign each fixation to the board which has a non-zero probability, and sort them into three classes: targets, distractors and non-objects (i.e., not a board). For the non-objects class we chose fixations that were not on boards occurring either during the trials (after the cue and before the intersection) or during the free eye movements trials.

TABLE I

LEFT. NUMBER OF EPOCHS USED PER SUBJECT FOR THE ANALYSIS. RIGHT. STATISTICAL SIGNIFICANCE OF THE DURATION OF FIXATIONS BETWEEN TARGETS (T), DISTRACTORS (D) AND NON-OBJECTS (N/O) FOR ALL SUBJECTS INDIVIDUALLY AND POOLED (LAST ROW).

	Number of epochs		Statistical tests		
	Before	After	T vs D	T vs N/O	D vs N/O
S1	192	31	0.65	0.07	**0.04**
S2	172	49	0.76	0.13	**0.04**
S3	255	90	0.3	**0.05**	$\sim$ **0**
S4	224	63	0.9	0.63	0.29
S5	204	126	**0.01**	**0.003**	0.4
S6	174	79	0.3	0.13	$\sim$ **0**
All	1221	438	0.1	0.08	0.9

We extracted EEG epochs corresponding to the signal elicited during the fixation. To ensure that the activity was not contaminated by the subsequent gaze shift we retain for analysis fixations longer than 300ms. Finally, we also exclude fixations preceded by saccades with amplitudes smaller than 70 px, so as to discard small corrective gaze shifts. We'll further discuss this issue in Section III.

Previous work has shown that the EFRP response is dependent on the saccade amplitude and duration. To avoid confounding factors, we restrict further analysis to distractors and non-objects epochs having similar characteristics with the target epochs. In general, there is a much larger number of fixations to non-object and distractor stimuli than for targets. We use the Mahalanobis distance over amplitude and duration to select an equal number of epochs for all classes. Therefore, we kept for the analysis those fixations that better match (pairwise) the characteristics of target epochs (c.f Table I).

F. Statistical analysis

To analyse the discriminability of EFRPs corresponding to targets, distractors and non-objects we run a statistical analysis both at a group and a subject level. In the first case, we run a paired T-test over all (channel, time) points. We then filter out the T-test in order to reduce the type I errors when

doing multiple comparisons, such that the p-value follows the following criteria [10]:

1) $p < 0.05$ for at least 6 consecutive time points (23.5 ms)
2) At least one neighbouring electrode also satisfied the previous condition.

For the subject-level statistical analysis we run a non-parametric Wilcoxon rank-sum test on all (channel, time) points. We then filter using the same criterion and define the h-value, where h = 1 if it is statistical significant and 0 otherwise. The sum of h-values across subjects is used to determine what are the common discriminant areas between subjects. For visualization, we report the average values over a 25 ms window.

III. Behavioral Analysis

We first perform a behavioural analysis characterizing fixations by their fixated distance and duration as wells a the saccadic amplitudes. These results are shown in Figure 2. We first look at the distance subjects were fixating at. All subjects exhibit a rather uniform distribution , suggesting no particular preference. Moreover, subjects also fixate at distances larger than the recognition threshold ($\sim$210) suggesting that they recognized the board but not the symbol within it and would return to it when it was closer.

Some subjects exhibit slightly longer fixation duration for targets and distractors than for non-objects, c.f., table I. For 5 subjects out of 6 the fixations for either the targets or distractors are significantly different from the non-objects. However, results are to be taken carefully as the datasets are unbalanced. As mentioned before, for the EEG analysis we select all fixations with a duration longer than 300 ms, which is a compromise between eliciting later components of the EFRP and the number of epochs.

The distribution of saccadic amplitude reveals two peaks, at 50 px and 150 px. The first one is hypothesized to be due to microsaccades, repositioning, or to be saccades to an object close by. We therefore remove fixations with a saccade amplitude smaller than 70 px (see section II-E) in order to be sure that the object has not already been recognized at the previous fixation.

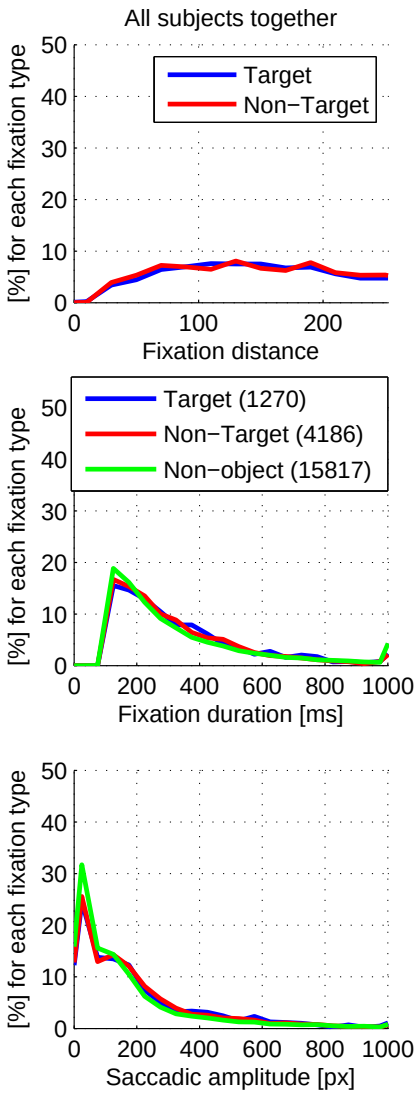

Fig. 2. Behavioral analysis. Top, Distribution of the distances (in meters) at which the objects were when the subject was fixating. The distribution was computed with 20 m bins. Middle, Duration distribution of those fixations below the viewing distance. The distribution was computed with 50 ms bins. The number in parenthesis are the number of epochs. Bottom, Amplitude distribution of the saccade preceding fixations. The distribution was computed with 50 px bins.

IV. Discriminability

A. Group analysis

We assess here the discriminant information of each channel both at the group and subject level. Figure 3 shows the topographic plot of the grand average ERP of all subjects. Results are similar to [9] although activity appears more localized, presumably due to the use of the Laplacian spatial filter. We see a positive activity (P100) at around 100 ms after fixation on O1 and O2. It then spreads laterally and is followed by a negativity over centro-occipital areas. Finally appears a positive activity over POz which lasts until 250 ms. This pattern is observable for the three types of fixations with small differences in amplitude. All subjects except S2 show the expected P100 response, and 4 out of the 6 subjects show the later activity over POz.

The corrected t-test of the pooled data showed discriminant activity (targets versus non-object) both over occipital areas: in the periods [110-150] ms (channels O1,Oz,O2) and [210-250] ms (channels O1,Oz). Furthermore, the first period also appears as slightly discriminant for distractors vs non-objects; while the second for targets versus distractors. We ignore any discriminability before 100 ms as this period is likely to be heavily affected by eye movement artifacts.

B. Subject-based analysis

We evaluate what is the common discriminant activity (channel,time) across subjects. For this purpose, we first compute the discriminability for each subject individually and

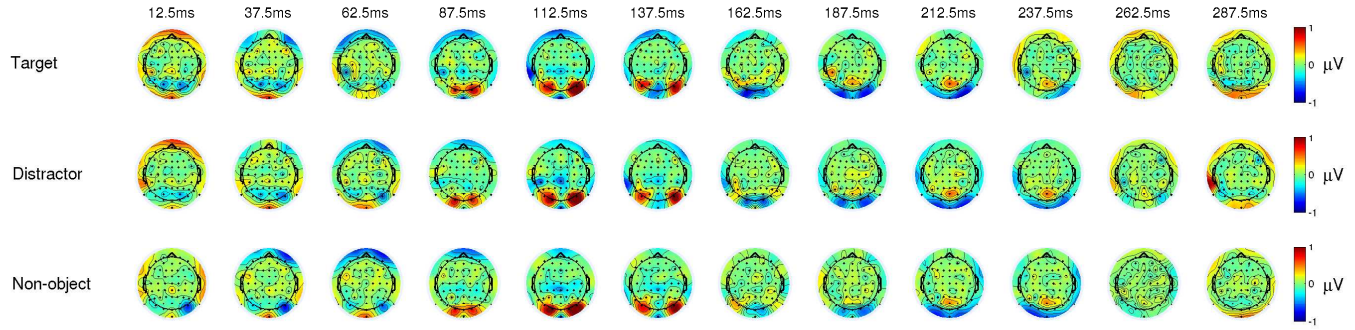

Fig. 3. Topographic plot of the grand averages of EFRPs binned over 25ms windows. The column header indicates the time sample corresponding to the center of each window.

then sum the h values. These values are presented in Figure 4. As expected, the common discriminant activity across subjects is mainly around the occipital sites. Moreover two subjects also present discriminant patterns on PO3, O1 and O2 for the three conditions, although more pronounced for target versus non-object.

Considering individual patterns, S3, S4 and and S6 present discriminability in the occipital area. Compared to the group analysis less regions are found to be discriminant, mainly due to the lower number of epochs for the statistical analysis (see table I). This is particularly the case for subject S1 and S2.

V. Conclusion

This study provides a first exploration of modulations of EEG activity due to overt deployment of visual attention in ecological environments. Despite the existence of a large amount of studies describing EFRPs, almost all of them are performed in well-controlled laboratory studies using simple stimuli [11]. More important, even those studies using more natural stimuli typically employ static images, avoid gaze shifts or instructed subjects to fixate during long periods (>300ms) [8], [12], [13]. This has allowed to identify discriminat late EFRP components [8], even at a single subject level [11]. Nonetheless, this leaves unanswered the question of how these signals are during realistic conditions and natural behaviour.

In this study, subjects were instructed to freely explore the scene for relevant stimuli while driving in a simulated environment. In consequence, signals will be affected by less controllable factors such as optic flow, and driver's movements. Allowing natural behaviour yielded short fixations periods. We found statistical differences in the fixation duration between targets and distractors, on one side, and non-objects on the other.

As in previous studies, we found consistent EFRPs patterns in 5 out of 6 subjects [8]. This pattern corresponded to a first positive activity over occipital electrodes at about 110-150 ms after fixation onset, and a second positive activity over midline parieto-occipital and occipital electrodes at about 210 to 250 ms. We reported results using a Laplacian spatial filter. Similar patterns were observed using Common Average Reference (CAR) or T7-T8 as reference. However these yielded more spread patterns of activity, and appeared more prone to EOG contamination at frontal areas.

Furthermore, despite the short fixation periods –which limited our analysis to a period of 300 ms after fixation onset–, we found these patterns to be discriminant between target and non-objects (c.f., Figure 4). This is in line with the hypothesis that target stimulus should elicit the strongest response with respect to other fixations. Analysis at a subject level is severely impaired by the number of available epochs (i.e. fixations meeting the criteria). Among the four subjects with enough epochs to perform statistical analysis (>60 epochs), two showed significant differences for target versus non-objects (c.f., Table I).

The possibility of achieving discrimination at a subject-level, i.e. for BCI applications, requires thus the possibility of gathering more epochs to properly model EEG responses to the different type of stimuli. Besides the obvious approach of recording more data, methodological approaches can be adopted to relax the inclusion criteria used in this study, without forcing subjects to perform longer fixations. One possibility is to apply artifact removal approaches to filter out EOG contamination in the later components of the EFRP. Moreover, combining the information about fixation duration and EEG responses could also increase the recognition accuracy. Future work will explore these alternatives.

Acknowledgment

This study was supported by Nissan Motor Co. Ltd., and carried out under the "Research on Brain Machine Interface for Drivers" project. The authors would like to thank Marija Uscumlic for useful discussions.

References

[1] C.-T. Lin, L.-W. Ko, and T.-K. Sheng, "Computational intelligent brain computer interaction and its applications on driving cognition," *IEEE Computational Intelligence Magazine*, vol. 4, no. 4, pp. 32–46, 2009.

[2] Z. Khaliliardali, R. Chavarriaga, L. A. Gheorghe, and J. d. R. Millán, "Detection of anticipatory brain potentials during car driving." *Conf Proc IEEE Eng Med Biol Soc*, vol. 2012, pp. 3829–3832, 2012. [Online]. Available: http://dx.doi.org/10.1109/EMBC.2012.6346802

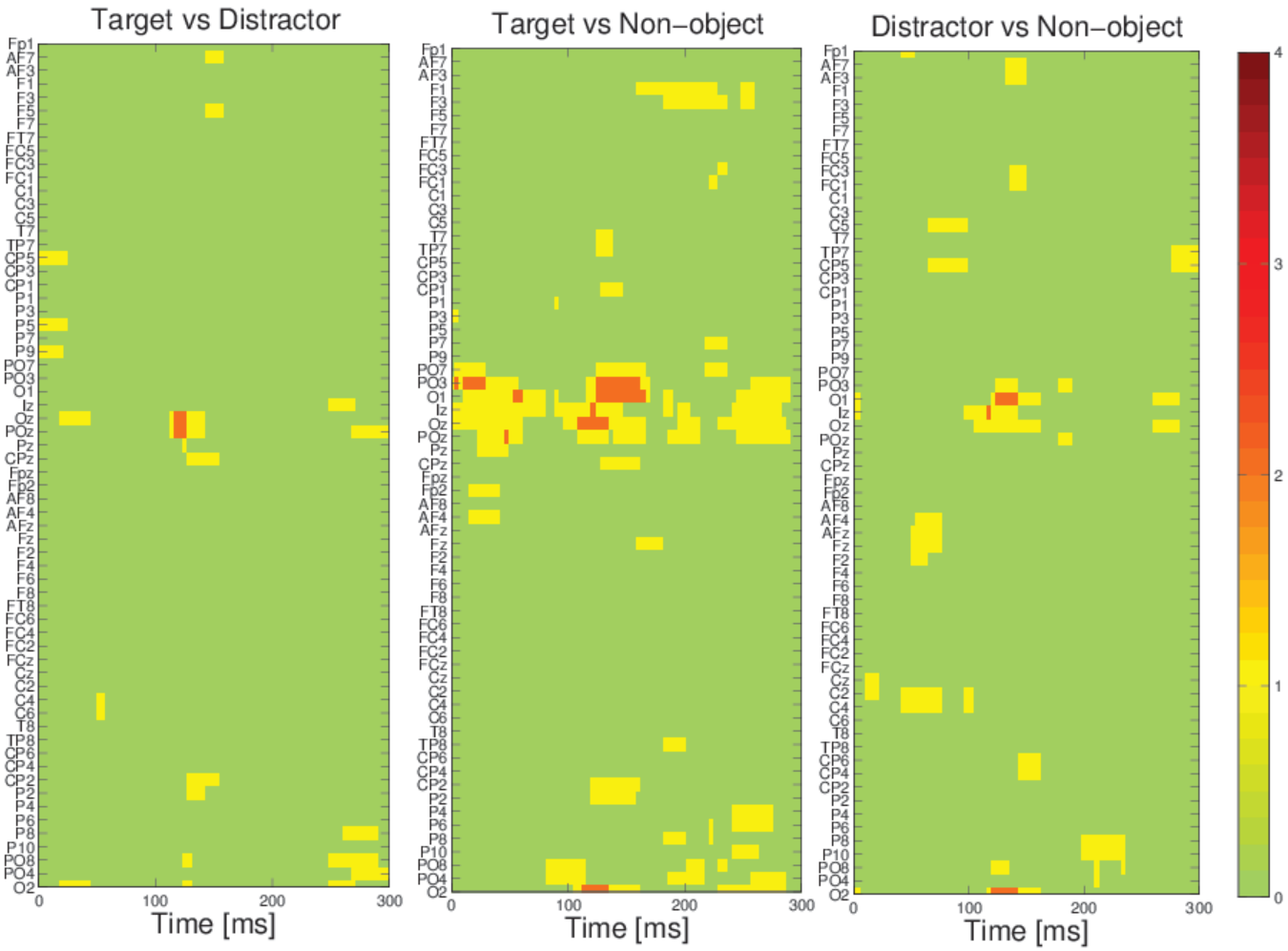

Fig. 4. Subject-level discriminability. Plots show accumulated h-value over subjects for the different conditions. Values are color coded corresponding to the number of subjects showing significant discriminability.

[3] L. A. Gheorghe, R. Chavarriaga, and J. d. R. Millán, "Steering Timing Prediction in a Driving Simulator Task," in *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013.

[4] S. Haufe, M. S. Treder, M. F. Gugler, M. Sagebaum, G. Curio, and B. Blankertz, "EEG potentials predict upcoming emergency brakings during simulated driving." *J Neural Eng*, vol. 8, no. 5, p. 056001, Jul 2011. [Online]. Available: http://dx.doi.org/10.1088/1741-2560/8/5/056001

[5] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness." *Neurosci Biobehav Rev*, Oct 2012. [Online]. Available: http://dx.doi.org/10.1016/j.neubiorev.2012.10.003

[6] P. Sajda, E. Pohlmeyer, J. Wang, L. Parra, C. Christoforou, J. Dmochowski, B. Hanna, C. Bahlmann, M. Singh, and S.-F. Chang, "In a blink of an eye and a switch of a transistor: Cortically coupled computer vision," *Proceedings of the IEEE*, vol. 98, no. 3, pp. 462–478, 2010.

[7] M. Uscumlic, R. Chavarriaga, and J. d. R. Millán, "An iterative framework for EEG-based image search: Robust retrieval with weak classifiers," *PLoS ONE*, vol. 8, p. e72018, 2013.

[8] J. E. Kamienkowski, M. J. Ison, R. Q. Quiroga, and M. Sigman, "Fixation-related potentials in visual search: a combined EEG and eye tracking study." *J Vis*, vol. 12, no. 7, p. 4, 2012. [Online]. Available: http://dx.doi.org/10.1167/12.7.4

[9] M. Uscumlić, "Exploiting recognition of natural visual stimuli for brain computer interfacing," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, Switzerland, May 2013.

[10] S. Dehaene, L. Naccache, L. Cohen, D. Le Bihan, J.-F. Mangin, J.-B. Poline, and D. Rivière, "Cerebral mechanisms of word masking and unconscious repetition priming," *Nature neuroscience*, vol. 4, no. 7, pp. 752–758, 2001.

[11] A.-M. Brouwer, B. Reuderink, J. Vincent, M. A. J. van Gerven, and J. B. F. van Erp, "Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials." *J Vis*, vol. 13, no. 3, p. 17, 2013. [Online]. Available: http://dx.doi.org/10.1167/13.3.17

[12] S.-T. Graupner, S. Pannasch, and B. M. Velichkovsky, "Saccadic context indicates information processing within visual fixations: evidence from event-related potentials and eye-movements analysis of the distractor effect." *Int J Psychophysiol*, vol. 80, no. 1, pp. 54–62, Apr 2011. [Online]. Available: http://dx.doi.org/10.1016/j.ijpsycho.2011.01.013

[13] P. Rämä and T. Baccino, "Eye fixation-related potentials (EFRPs) during object identification." *Vis Neurosci*, vol. 27, no. 5-6, pp. 187–192, Nov 2010. [Online]. Available: http://dx.doi.org/10.1017/S0952523810000283