Proceedings of the 5th International
IEEE EMBS Conference on Neural Engineering
Cancun, Mexico, April 27 - May 1, 2011

A 3-D Immersive Environment for Characterizing EEG Signatures of Target Detection

David C. Jangraw and Paul Sajda, *Senior Member, IEEE*

Abstract— **Visual target detection is one of the most studied paradigms in human electrophysiology. Electroencephalographic (EEG) correlates of target detection include the well-characterized N1, P2, and P300. In almost all cases the experimental paradigms used for studying visual target detection are extremely well-controlled – very simple stimuli are presented so as to minimize eye movements, and scenarios involve minimal active participation by the subject. However, to characterize these EEG correlates for real-world scenarios, where the target or the subject may be moving and the two may interact, a more flexible paradigm is required. The environment must be immersive and interactive, and the system must enable synchronization between events in the world, the behavior of the subject, and simultaneously recorded EEG signals. We have developed a hardware/software system that enables us to precisely control the appearance of objects in a 3D virtual environment, which subjects can navigate while the system tracks their eyes and records their EEG activity. We are using this environment to investigate a set of questions which focus on the relationship between the visibility, salience, and affect of the target; the agency and eye movements of the subject; and the resulting EEG signatures of detection. In this paper, we describe the design of our system and present some preliminary results regarding the EEG signatures of target detection.**

I. INTRODUCTION

When a person loses his keys, she might rummage through her desk, opening drawers and pushing things aside. Each time she sees an object, she quickly decides whether it is the thing she is looking for and continues her search if necessary. This type of "target detection" is extremely important to the human experience. However, it has mostly been investigated using very simple experiments which bear little resemblance to our real-world experiences. For example, a subject in a laboratory experiment will commonly be asked to stay perfectly still and actively inhibit eye movements while simple, static stimuli are flashed onto a screen.

While there are good reasons behind the simplifications and controls of such experiments, the field of neuroscience could benefit from a study of more realistic situations. If we can characterize the effects of factors that most neuroscience experiments avoid - such as free viewing, subject and object motion, and partial occlusion of complex stimuli - we will move closer to being able to reliably decode brain activity as a person moves about her everyday life. Such progress could have important implications for wearable brain-computer interfaces (BCIs). The obstacles in the way of wearable BCIs are rapidly diminishing, as EEG systems become lightweight and wireless [1] and noise reduction techniques steadily improve [2].

In the following sections, we will introduce a system that allows us to record simultaneous EEG and eye tracking data from a subject as we present 3-dimensional dynamic stimuli to her in an immersive environment. We present results from a pilot study using this system to investigate the neural correlates of target detection in an immersive environment, and we outline some of the challenges and strategies that emerge from our analysis.

This work was supported by DARPA under contract N10PC20050.

D. C. Jangraw and P. Sajda are with the Department of Biomedical Engineering, Columbia University, New York, New York 10027 USA. dcj2110@columbia.edu, psajda@columbia.edu

II. SYSTEM OVERVIEW

A. Immersive Environment

We constructed several virtual 3-dimensional environments using Unity software (Unity Technologies, San Francisco, CA). Each consisted of a continuous hallway 5m tall and 5m wide. The subject's viewpoint was that of a virtual camera, placed 2m off the ground in the environment, that could be controlled either by a joystick or by pre-programmed navigation. Visual stimuli were placed in small "cubbies" on either side of the hallway, such that they gradually became visible as the camera passed them (Figure 1, top).

The visual stimuli were 3-dimensional models either constructed from basic shapes in Unity or adapted from models available online in databases such as Google 3D Warehouse (Google Inc., Mountain View, CA) or TurboSquid (TurboSquid Inc., New Orleans, LA). All models downloaded from these databases were freely available and free of copyright restrictions. Adjustments to the models (e.g., scaling and centering) were made in Google SketchUp (Google Inc., Mountain View, CA) or Maya (Autodesk Inc., San Rafael, CA).

B. Recording Paradigm

Subjects' EEG data were recorded using a Sensorium EPA-6 Electrophysiological Amplifier (Sensorium Inc., Charlotte, VT) with a sampling rate of 1kHz in an electrostatically shielded room (ETS-Lindgren, Glendale Heights, IL). The cap (Electro-Cap International Inc., Eaton, OH) has 79 Ag/AgCl electrodes, most of which are in a 10-10 montage (with several additional electrodes in occipital areas). However, other EEG systems that accept event markers via parallel port could be substituted without a change in the overall system. Subjects' eye position, blinks, saccades, and fixations were recorded using an EyeLink 1000 (SR-Research, ON, Canada) with a 1kHz sampling rate.

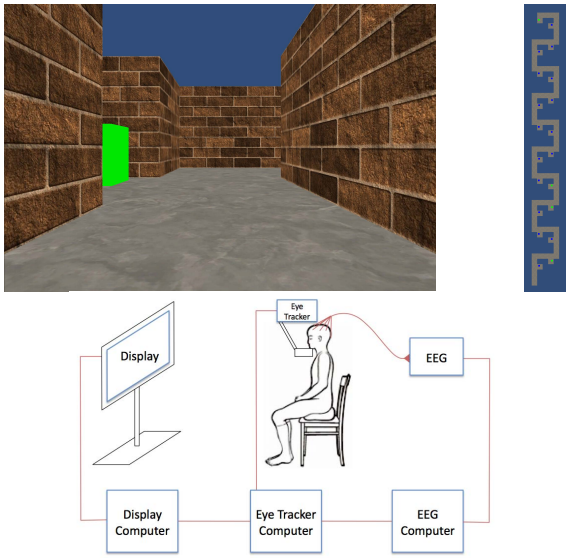

Fig. 1. A 3D virtual environment as viewed by the subject (top left), another as viewed from above (top right), and the system that presents them to the subject (bottom). The green and blue objects seen in the environment are some of the salient, solid-color objects used in our pilot study.

C. Experimental System Overview

Three computers were used to produce the stimuli and record the data (Figure 1, bottom). Computer 1 controlled the virtual environment and recorded the positions of the virtual camera in every frame. Computer 2 recorded eye position data and responses, and Computer 3 recorded the EEG data. Computer 1 controlled Computer 2 via a dedicated ethernet connection. Computer 2 was instructed to send a parallel port signal to Computer 3 every two seconds in order to re-synchronize the clocks of the EEG and eye position systems.

D. Timing information

After data had been collected, the system could perform a "replay" to accurately determine when each object first appeared on the screen. Each trial was played back using the position data recorded by Computer 1, and the objects were displayed in a color that was distinct from the background. Just before an object came into view, the replay was slowed to a fraction of its original speed, and the screen was checked pixel-by-pixel for the specific color of the object. The system logged the time when the first pixels of an object came into view. It also logged, for every frame, the rough outline of any visible object in screen coordinates. This can be compared with the eye position data to determine when the subject's eyes were fixated on each object. This information could be used to time-lock EEG analyses, post hoc, to the moment when an object first appeared, to the time when a saccade was first made to the object, or to times based on a more complex function of eye position and object position.

III. METHODS

To demonstrate the system's functionality, we constructed a simple experiment in which the virtual objects were large rectangular prisms of one of two solid colors. In each session, the subject was asked to count the objects of one color ("targets") while ignoring objects of the other color ("distractors"). Each subject observed an automatic navigation of 8 hallways, which contained 20 objects each. Each object was preselected as a target or distractor with a 25% chance of being a target. The target and distractor colors were rotated every two sessions to control for stimulus-specific effects.

EEG and eye position data were recorded from 3 healthy subjects (one female, ages 25 to 43 years) as described above. Informed consent was obtained from all participants in accordance with the guidelines and approval of the Columbia University Institutional Review Board. Channels were referenced to the left mastoid and forehead ground (except Subject 1, for whom the reference was placed on the nose due to a poor connection at the mastoid). In pre-processing the EEG data, we applied a band-pass filter (1-100Hz) and a notch filter (59-61Hz), and we down-sampled to 250Hz. The data was visually inspected, and five channels with excessive noise for at least one subject were removed from further analysis.

In our stimulus-locked analysis, we first time-locked (set $t = 0$) each trial to the first moment when any part of the object appeared on-screen. Baseline subtraction was performed using the last 200ms of data before stimulus onset. Trials on which the subject had blinked (between 350ms before and 1000ms after stimulus onset) were removed (70/480 trials). In our saccade-locked analysis, we time-locked to the end of the subject's first saccade to each object (defining this time point using the method described in [3]), and shifted our baseline and blink removal time windows to avoid electro-ocular contamination. Baseline subtraction was performed using data from 400 to 200ms before saccade offset. Trials containing blinks between 500ms before and 1000ms after saccade offset were removed (86/433 trials).

IV. RESULTS

A. ERP Analysis

Because our stimuli are so salient, we began with stimulus-locked analysis. In a trial-level average of all three subjects (Figure 2), the largest response to the distractor stimuli was a positivity starting approximately 350ms after the stimulus appeared. For target stimuli, we observed a distinct "target response" on top of this standard response which had three notable features. The first was a broad negativity with a peak 150-200ms after the stimulus appeared, a feature that was widely distributed over the scalp (Figure 3). This feature is consistent with the visual N1, and with evidence that selective negativities can be induced by attention to a certain stimulus color [4]. Second, the target trials' positive deflection at 350ms persisted well beyond that of the distractor trials, a feature consistent with the well-studied P300/P3b component seen in oddball tasks [5]. Finally, we observed a simultaneous uptick in voltage at frontal electrodes, but this component was transient while the P3b persisted (Figure 3). This likely corresponds with the P3f component described by [6].

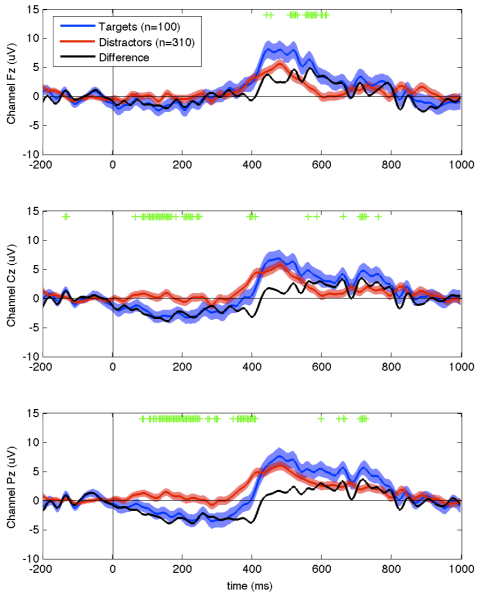

Fig. 2. Average voltage across all trials for electrodes Fz (top), Cz (middle), and Pz (bottom). Translucent patches indicate standard error (n=100 target trials, 310 distractor trials for the 3 subjects). Green crosses indicate that the target and distractor trials were significantly separated from one another at that time point (Wilcoxon rank-sum test, $\alpha = 0.05$). A 30Hz low-pass filter was applied for plotting.

We also noted a few discrepancies between our results and those of traditional studies of target detection. Most notably, the ERPs shown in Figure 2 have a very broad time course, and the P2 component is not visible in our results. Additionally, the P3b appears to be relatively small in amplitude. Some of these features can be attributed to the small number of subjects in our study - the amplitude, latency, and scalp distribution of P3b have been found to vary significantly with factors such as the age and gender of the subject [5], [7] - but we nevertheless sought to investigate possible confounds.

B. Time-Locking Investigation

The broad time-course of the ERPs could be the result of jitter between our $t = 0$ and the true onset of the visual evoked response. Analysis of the synchronization signals sent between the stimulus presentation computer and the EEG amplifier (not shown) led us to conclude that significant jitter was not produced by our system. We next investigated whether the visual evoked response might be time-locked to the first saccade to the object. Saccade-locked analysis yielded far less separation between target and distractor trials early in the trial (Figure 4). We concluded that stimulus-locked analysis was the better choice, while allowing that the true onset of the visual evoked response may be a more complex function of eye position and attention. Interestingly, saccade-locked analysis yielded more significant separation than stimulus-locked analysis from $t = 600ms$ to $t = 900ms$. In truth, there may be discriminating activity locked to both stimulus onset and saccade.

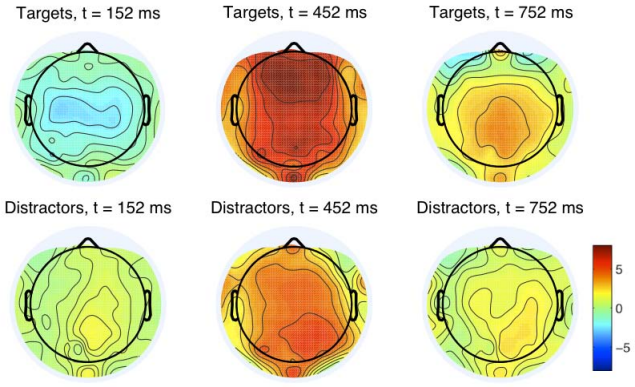

Fig. 3. Distribution of voltage (see colorbar for scale, in μV) on the scalp 152ms (left), 452ms (middle) and 752ms (right) after the onset of a target stimulus (top) and distractor stimulus (bottom) (average activity in 20ms time bin centered at given times, averaged across all trials for all subjects). The salient features of these plots correspond with an N1 (left), a P3b and P3f (middle), and a P3b (right).

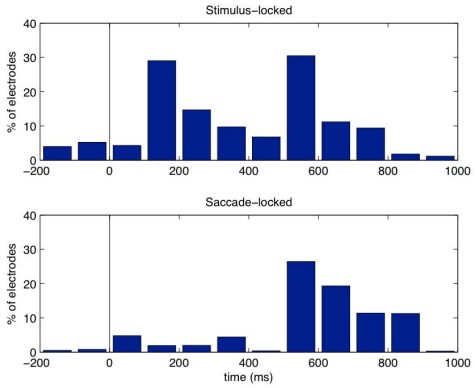

Fig. 4. Histograms of the times at which target and distractor trials were significantly separated (Wilcoxon rank-sum test, $\alpha = 0.05$), relative to stimulus onset (top) and the first saccade to the object (bottom). A value of 20 on the vertical axis, for example, means that during that time bin an average of 20% of electrodes had significant separation.

C. Electro-Ocular Contributions

A possible explanation for the frontal positivity that we observed is electro-ocular contamination in the EEG as a result of our free viewing paradigm. We noted that subjects were more likely to make saccades immediately after the object appeared if that object was a target (Figure 5). However, if this were driving our results, we would expect more frontal EEG separation between target and distractor trials early in the trial, when in fact the frontal activation does not emerge until some 300ms later. Also, the central vertical positioning of the object and hallway (Figure 1) meant that saccades were primarily horizontal. The EOG from such saccades would produce lateralized activations. Since our results were dominated by symmetric components, we concluded that electro-ocular contamination was not driving our results.

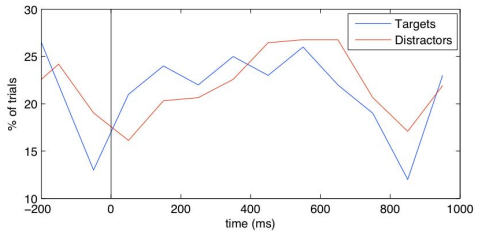

Fig. 5. Percentage of trials in which the subject made a saccadic eye movement in each time bin (locked to stimulus onset), averaged for all three subjects. Time bins are 100ms wide and centered at $t = i * 100 + 50ms$, where i is the set of integers. We see that the subjects are more likely to make a saccadic eye movement soon after an object appears if that object is a target.

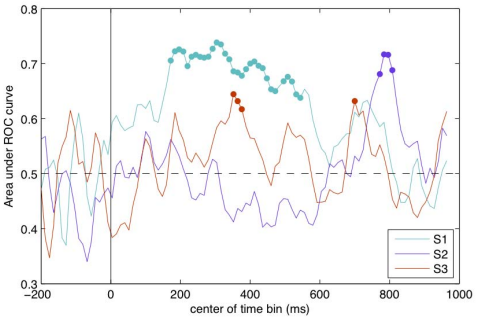

Fig. 6. Leave-one-out cross-validation for the results of logistic regression. For each 50ms window used as data, the time of the window center is plotted against the area under the ROC curve (AUC). Each line represents the AUC values for a different subject. Dots are plotted at time points at which the AUC was significantly above that of bootstrapping results (see Results section). The insignificant fluctuations before stimulus onset are likely results of coincidental correlations in the noise.

D. Logistic Regression

With a final BCI application in mind, we investigated the quality of the data for single-trial discrimination. We used a logistic regression algorithm to classify single-trial data from this system [2], [8]. Using a 50ms window sliding in 10ms increments, we learned a unique linear classifier for each subject at each window position. We then used leave-one-out cross-validation and calculated the area under the ROC curve (AUC) as the performance metric for each classifier. To assess significance, we used a bootstrapping technique to estimate the distribution of our classifier's AUCs given the same data with randomized truth labels (target or distractor). Specifically, we randomized the labels, ran our classifier, and calculated the leave-one-out AUC. This process was repeated 100 times for each subject at each window position, after which we computed the AUC distribution and the AUC leading to a significance level of $p = 0.01$. When assessed in this way, the classifier's results varied greatly from subject to subject (Figure 6). The inconsistent performance of this classifier supports the idea that more sophisticated time-locking is required to fully capture the target response in this dynamic, free-viewing paradigm.

V. DISCUSSION

The pilot study outlined in this paper observes a well-studied set of EEG components under a specific set of conditions (dynamic environment with predictable but gradual onset of stimuli) that are typical of normal human experience. The adaptation of other studies in visual neuroscience to this framework will shed light on neural components in ways that may make them more clearly related to behavioral deficits reported during a visit to a neurologist, potentially improving diagnoses. The capability of the system to allow control by the user will also facilitate the study of novel questions about the effects of navigation and agency on visual processing.

The permission of free viewing and the integration of eye tracking into our studies present challenges in time locking and trial consistency. We must be aware of confounds should we attempt to separate the effects of stimuli and related eye movements. But the elimination of eye movement constraints on the subject make the results distinctly valuable to applied neuroscience settings. This system could serve as a useful proving ground for BCI algorithms successful in hyper-controlled laboratory settings as they transition into wearable systems for healthy individuals.

In this paper, we have presented the incorporation of a computer-generated 3D immersive environment into a simultaneous EEG/eye-tracking system. We have demonstrated the ability of this system to elucidate questions of visual target detection, extending past studies into a largely unexplored domain. Future studies using this system will continue to advance towards a better understanding of the dynamic, immersive human experience.

VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of Ansh Johri of Columbia University's BME department to the library of 3D models and environments for the system.

REFERENCES

[1] Lin C-T, Ko L-W, Chang M-H, Duann J-R, Chen J-Y, Su T-P, Jung T-P. "Review of Wireless and Wearable Electroencephalogram Systems and Brain-Computer Interfaces – A Mini-Review." *Gerontology*, vol. 56 p.112-119, 2010.

[2] L. Parra, C. Spence, A. Gerson, and P. Sajda, "Recipes for the linear analysis of EEG," *Neuroimage*, vol. 28, pp. 326-341, 2005.

[3] A. Luo and P. Sajda, "Do We See Before We Look?" *Proceedings of the 4th International IEEE/EMBS Conference on Neural Engineering*, pp. 230-233, 2009.

[4] L. Anllo-Vento and S. A. Hillyard, "Selective attention to the color and direction of moving stimuli: Electrophysiological correlates of hierarchical feature selection." *Perception & Psychophysics*, vol. 58, p. 191-206, 1996.

[5] John Polich, Albert Kok, "Cognitive and biological determinants of P300: an integrative review." *Biological Psychology*, vol. 41(2), p. 103-146, 1995.

[6] S. Makeig, M. Westerfield, T-P Jung, J. Covington, J. Townsend, T. J. Sejnowski, and E. Courchesne. "Functionally Independent Components of the Late Positive Event-Related Potential during Visual Spatial Attention." *J. Neurosci.*, vol. 19, p. 2665-2680, 1999.

[7] R. Johnson, "On the neural generators of the P300 component of the event-related potential." *Psychophysiology*, vol. 30(1), p. 90-97, 1993.

[8] A. Gerson, L. Parra, and P. Sajda. "Cortical origins of response time variability during rapid discrimination of visual objects." *NeuroImage*, vol. 28, p. 342-353, 2005.