**Proceedings of the 4th International
IEEE EMBS Conference on Neural Engineering
Antalya, Turkey, April 29 - May 2, 2009**

FrA1.4

Do We See Before We Look?

An Lou and Paul Sajda
Department of Biomedical Engineering
Columbia University
New York, USA
al2082@columbia.edu, ps629@columbia.edu

Abstract—**We investigated neural correlates of target detection in the electroencephalogram (EEG) during a free viewing search task and analyzed signals locked to saccadic events. Subjects performed a search task for multiple random scenes while we simultaneously recorded 64 channels of EEG and tracked subjects' eye position. For each subject we identified target saccades (TS) and distractor saccades (DS). We sampled the sets of TS and DS saccades such that they were equalized/matched for saccade direction and duration, ensuring that no information in the saccade properties themselves was discriminating for their type. We aligned EEG to the saccade onset and used logistic regression (LR), in the space of the 64 electrodes, to identify activity discriminating a TS from a DS on a single-trial basis. We found significant discriminating activity in the EEG both before and after the saccade. We also saw substantial reduction in discriminating activity when the saccade was executed. We conclude that we can identify neural signatures of detection both before and after the saccade, indicating that subjects anticipate the target before the last saccade, which serves to foveate and confirm the target identity.**

Keywords: EEG, eye-tracking, visual search, saccade, linear discrimination, target detection

I. INTRODUCTION

Do we need to "look" before we "see"? During visual search, evidence for a target or object of interest can potentially be accumulated from saccade to saccade, resulting in a build-up of evidence prior to the last saccade which foveates the target. Such evidence accumulation likely represents a combination of peripheral detection and prediction that a target will be located at the next fixation given that it has not been detected at previous fixations.

Surprisingly, there has been limited research exploring this and related questions within the context of EEG signatures analyzed relative to precise saccadic (or other eye-movement) time-locking. One exception is [1], where the authors introduced an EEG and eye-tracking system that tracked brain state during a single eye fixation and studied the early cognitive processes during reading. Compared to traditional methods studying eye-fixation-related potentials (EFRP), which record eye fixation information from electrooculograms (EOGs), their system has the advantage of coupling accurate time measures from event related potentials (ERPs) and the eye position relative to the stimulus. Similar to their study, in this paper we study EEG locked to eye-movements—i.e. saccades. We adopted stimuli similar to ones we used previously to study target detection in serial presentations of briefly flashed images [2]. Subjects performed a visual search task for multiple random scenes while we simultaneously recorded 64 channels of EEG and tracked subjects' eye position using an eye-tracker.

In the following sections we first introduce our EEG and eye-tracking system that enables the simultaneous recording of subjects' EEG and eye position. We then describe our experimental paradigm for studying visual search. We explain our methods for sampling the target and distractor saccade-locked EEG so that differences in saccadic features (e.g. saccade direction and duration) are eliminated, enabling us to focus on discriminating information in the EEG. We then use logistic regression to identify EEG components which maximally discriminate between target and distractor conditions and compute the significance of the discriminating activity via a resampling procedure. We show that subjects anticipate seeing the target before a saccade to the arget is made.

II. METHODS

A. *Subjects*

Seven healthy right-handed subjects (one female and six male, age from 25 to 36, mean age 30 yrs) participated in the study. All subjects had normal or corrected to normal vision and reported no history of neurological problems. Informed consent was obtained from all participants in accordance with the guidelines and approval of the Columbia University Institutional Review Board.

B. *Experimental Paradigm*

The task for the subject was to search for people in a large image consisting of a set of image chips. Each large image had eight clusters of image chips, with each cluster containing seven chips that were extracted from video sequences that were used in a previous study [2]. These image chips contained either one or more people, or were natural images without people. Within each cluster of image chips, the spatial correlation of the cluster could be characterized into one of three conditions: 1) all the image chips were similar (extracted from successive video frames) and the spatial correlation was high (high correlation and high clutter); 2) there was a single image chip in the cluster with the remaining chips being uniformly black (low correlation and low clutter); 3) all the image chips were extracted from unrelated video clips so they were spatially uncorrelated (low correlation and high clutter). Figure 1 shows sample images for the task. The size of each

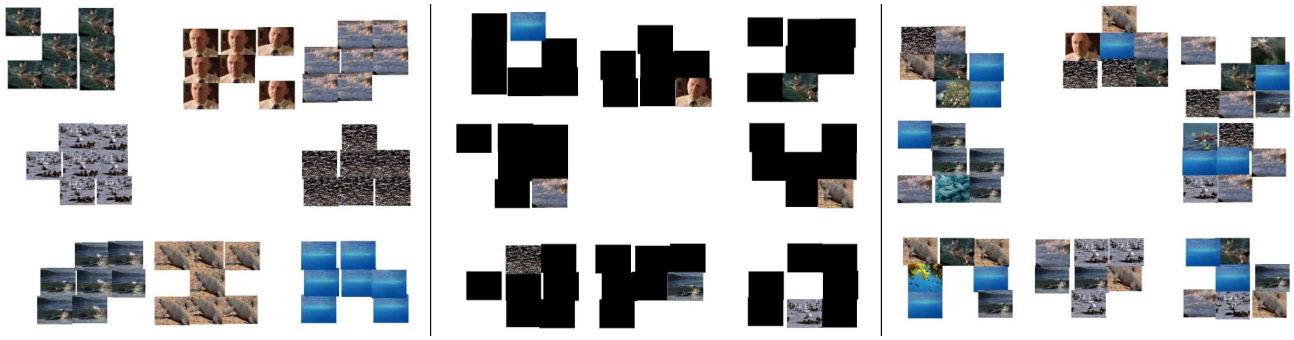

Figure 1. Example stimuli used in the experiment. Condition 1 –Image chips within a cluster were highly correlated (high spatial correlation, high clutter). Condition 2 – each cluster had a single image chip, with the remaining chips being uniformly black (low spatial correlation, low clutter). Condition 3 – image chips within each cluster were not spatially correlated (low spatial correlation, high cluster). In Condition 1, there was at most one cluster of target chips in each image, while in Condition 2 and 3 there was at most one target chip. Stimuli without target(s) are not shown.

image chip was 128 by 96 pixels which extended roughly 1.5 degrees of the visual field, and the size of each search image was 1500 by 1200 pixels. For each condition, 125 images contained targets and 50 images no targets. For images that contained target(s), the target(s) could only appear in one cluster. The appearance of targets in these images was randomly distributed across the eight clusters (positions).

Each trial began with a fixation image, which was a black circle on a white background. The fixation image lasted for a variable time, and then the image to be searched was presented to the subject. Once they had found a target or determined that there was no target chip in that image, they pressed a button (right and left index finger keys for a detected target and no target, respectively) on a joystick which was connected to the eye-tracking machine. Subjects were instructed to respond as fast as possible. A trial automatically timed-out if the subject did not make a button response within 10 seconds after the image onset.

There were 125+50=175 images for each stimulus condition, which led to 175×3=525 images (trials). These 525 trials were further divided into 5 sessions for the experiment, with each session including 105 trials. Trials were randomly presented to the subjects.

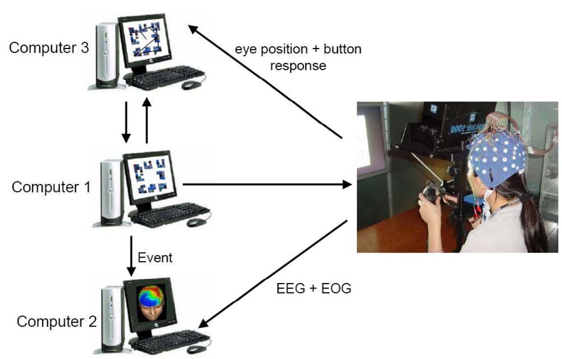

Figure 2. Experimental system overview. The system consisted of three computers, in which Computer 1 controlled the stimulus display and event signal generation; Computer 2 recorded and processed the EEG, and Computer 3 recorded and processed theeye position and button response data. Subjects performed the experiment in an electrostatically shielded room. A 30-inch cinema display (Apple, CA) was connected to Computer 1 and was placed inside the shielded room for presenting the visual stimulus to the subjects.

C. *Data Acquisition and Preprocessing*

An illustration of our stimulus display and data acquisition system is shown in Figure 2. An EyeLink 1000 (SR-Research, ON, Canada) with 1 kHz sampling rate was used for eye position recording and saccade detection. A 64-channel EEG system (Sensorium, VT) was used for our EEG recording. The sampling rate for the EEG was also 1kHz. After EEG acquisition, visual inspection was performed for each individual trial such that trials with substantial noise and artifact were removed from further EEG analysis, though these trials were included in the behavioral analysis (note that rejected trials accounted for less than 5% of the data). Following data acquisition, a software-based filter was used to remove DC drift and line noise. Eye-blink and eye-movement artifacts were recorded and later removed from EEG using a maximum difference method [3].

D. *Behavioral Data Extraction*

We extracted EEG data based on subjects' motor responses and eye position during the visual search. In our analysis only correct trials were included. Within each trial where subjects searched for a target, they made one or more saccades before the motor response. We term the saccade to a target as a "target saccade" (TS), and the saccade to a distractor a "distractor saccade" (DS). The pattern of eye movements during visual search can be very complex. For example, one might first make a fixation very close to a target, and then make one or two more saccades, before finally fixating on it. After fixating on a target, a subject might also direct their eyes to other locations and then return fixation to the target chip, responding with a button press.

We defined a TS to be the first saccade to a fixation location less than 100 pixels (roughly 1.5 degrees of visual field) away from the center of a target frame. This location could be outside the target frame, but it was within the visual angle subtended by the fovea and of the same order of the measurement error, so we assumed the subject was able to detect the target at this distance. If the fixation locations were never within this range in a correctly-responded trial, we relaxed our criteria to be 120 pixels away from the target chip center. If such a target saccade was still absent, we defined the TS to be the saccade to a fixation location that is the closest to a target. However, if this distance was greater than 150 pixels, we excluded this trial from further behavioral and EEG analysis because it is likely that the subject responded without detecting the target.

DS extraction was also needed for the EEG analysis. To make sure all the distractor trials were free of target detection, only saccades in correctly-responded distractor trials, where no target chips were contained in the image, were used for our EEG analysis.

To reduce eye-movement artifacts, saccades preceded by a fixation shorter than 100ms were not used for the analysis of EEG locked to saccade onsets. If the subject made a saccade outside the target cluster after a target saccade, that target saccade (fixation) was also excluded from the EEG analysis.

E. EEG Data Extraction

During a free-viewing paradigm the probabilistic characteristics of eye movements to targets, such as duration, direction and distance that a saccade traveled, may be different from that of distractors. To ensure that the discriminating activity was due to target detection in the EEG, and not due to the differences in these saccadic features between target and distractor conditions, we selected our data prior to the EEG analysis such that the characteristic distribution of saccades for TS and DS were matched.

The direction of a saccade is independent of its duration and travel distance. However, we found that the duration and distance were closely correlated. Thus we selected our trials such that the two-dimensional joint distribution of duration and direction of TSs matched that of DSs for our EEG analysis. More specifically, the saccade duration between 0 to 100 ms was equally divided into 4 bins and the saccade direction (from 0 to 360 degrees) was also divided into 4 quadrants for matching the our distributions. We selected our data for EEG analysis such that the number of samples of TSs equaled the number of samples of DSs for each duration bin and each direction quadrant.

F. Discriminating Components

We aligned the EEG to saccade onset and labeled these snippets (i.e. segments of EEG) as TS or DS, and used logistic regression (LR) in the space of the 64 electrodes, to identify discriminating components—i.e. we learned a linear projection in the space of the EEG electrodes that optimally discriminates a TS from a DS snippet. Specifically, we applied LR to a 50ms wide temporal window, sweeping the window across time and learning an optimal discriminator for each time window (see [3,4] for details of the methodology). This was done both before and after the saccade and thus the discriminators could be index by latency (positive and negative) relative to the saccade. We then used a leave-one-out (LOO) procedure to estimate the area under the receiver operating characteristic (ROC) curve (AUC, [5]), which we use as the measure of the strength of the discriminating activity. A significance level for AUC was determined via resampling, whereby the LOO procedure was repeated 30 times with a different randomization of the truth labels for each TS or DS snippet. Assuming a Gaussian distribution of the resulting AUC values, we estimated the $p = 0.05$ significance level. For each discriminating component we also estimated a forward model [4].

III. Results

A. Behavioral Results

Figure 3 shows group averages of the number of saccades for the correctly responded trials in the three conditions. We see that for all the correctly-responded trials containing a target, subjects found the target within a very few number of saccades. Even for Condition 3, where frames were highly uncorrelated and there was substantial clutter from distractor frames within each cluster, the number of saccades a subject made to locate the target was low. This suggests that subjects may know where the target is before the saccade to a target was made. In the following EEG analysis, we identify neural signatures that are indicative of this target detection process.

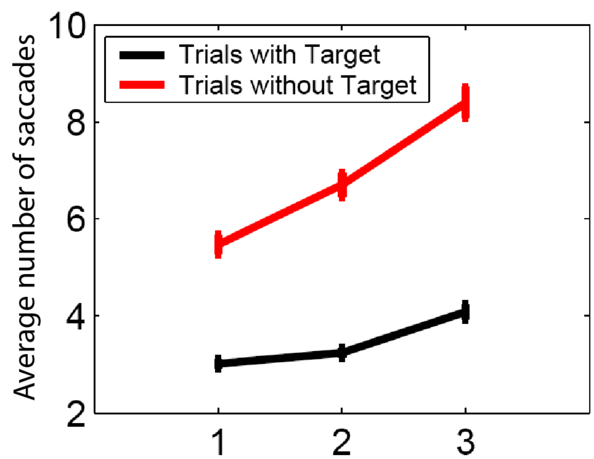

Figure 3. Behavioral results. Group averages showing the number of saccades for the correctly responded trials in the three conditions. The x axis indexes the three experimental conditions. Error bars indicate the confidence interval at the 95% significant level.

B. EEG Results

A summary of our EEG analysis results are shown in Figure 4. Shown are three latencies of the discriminating components, at -80ms, -20ms, and 60ms relative to saccade onset for the three conditions. The red dashed line shows the p=0.05 significant level of the discriminating activity. The discriminating activity is deemed significant if the AUC value is above this level. The scalp plots, generated via EEGLAB ([6]), show the group average sensor projection of discriminating activity at -80ms and 60ms relative to saccade onset.

From this figure we can see that the discriminating activity between TS and DS conditions is significant before the saccade was made, suggesting that subject anticipated seeing

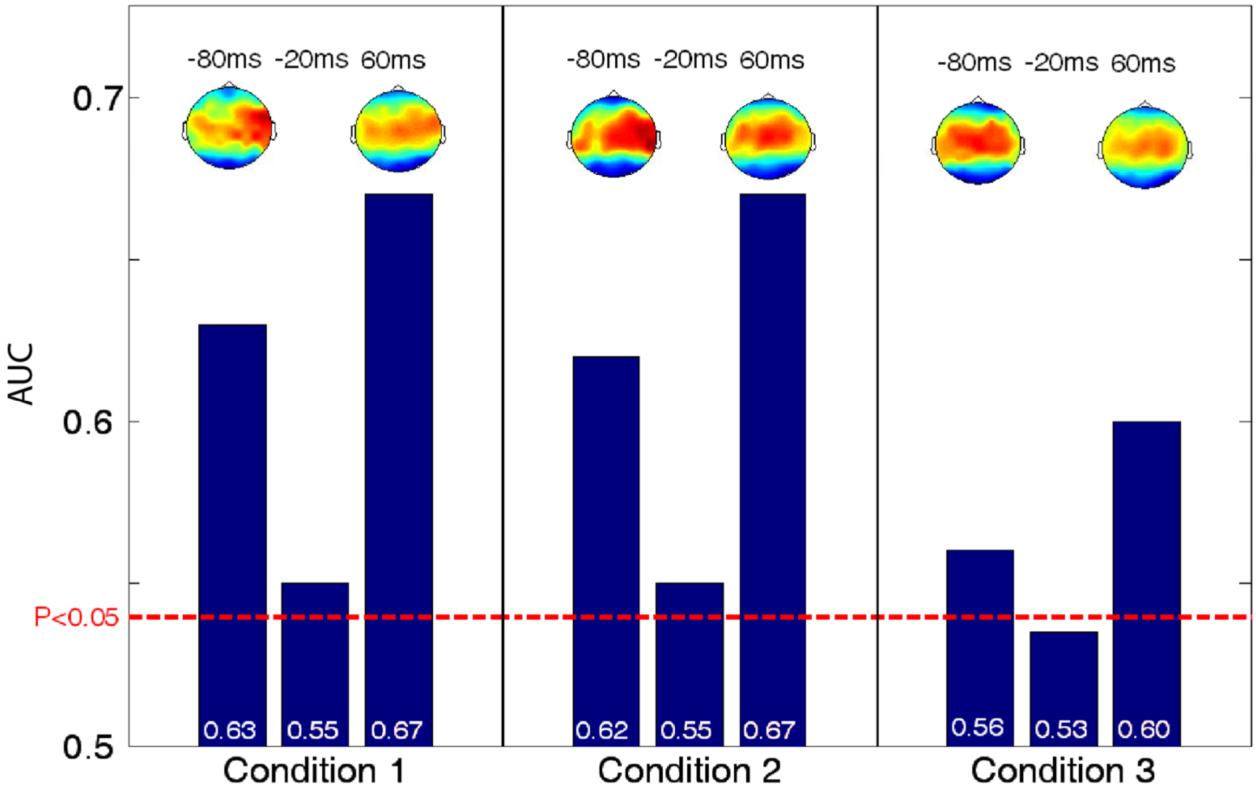

Figure 4. Group averages showing discriminating activity between saccade to targets and that of distractors for the three conditions. The blue bars show the AUC value, which is also marked in white at the bottom of each bar, at -80ms, -20ms, and 60ms relative to saccade onset. The red dash line shows the $p<0.05$ significant level of the discriminating activity. The scalp plots show the group average sensor projection of discriminating activity at -80ms and 60ms relative to saccade onset. Red represents positive correlation and blue negative correlation.

the target before they foveated it. This is also consistent with our behavioral findings that the targets could be found within a very few number of saccades.

We also find that the scalp topologies of the discriminating components are consistent with front-centro activity which would be associated with activity of the eye-fields. The frontal and supplementary eye-fields (FEF, SEF) have been implicated in saccadic planning and evidence accumulation.

Lastly, we see that the strength of the discriminating components for Condition 3 are substantially less than those for Conditions 1 & 2. This is expected since Condition 3 has the lowest spatial correlation and highest clutter and thus targets chips are the least salient and would thus likely result in lower evidence both pre and post target saccade, relative to the other two conditions.

IV. Conclusion

In this paper we used a simultaneous EEG/eye-tracking system to study target detection locked to saccades during visual search. We find discriminating components prior to the saccade which are predictive of whether the saccade will be toward a target or distractor. This suggests that during visual search we may "see" before we "look".

Acknowledgments

This work was supported by funding from the NIH (EB004730) and DARPA. We would like to thank Drs. Lucas Parra and Misha Pavel for fruitful discussions and suggestions.

References

[1] T. Baccino and Y. Manunta, "Eye-fixation-related potentials: insight into parafoveal processing," Journal of Psychophysiology, vol. 19, no. 3, pp. 204– 215, 2005.

[2] A. Luo and P. Sajda, "Behavioral and electrophysiological differences for target identification in natural sequence versus rapid serial visual presentation (RSVP)," Proc. Society for Neuroscience Annual Meeting (SFN'07), San Diego, CA, 2007.

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

[4] L. Parra, C. Alvino, A. Tang, B Pearlmutter, N. Yeung, A. Osman, and P. Sajda, "Linear spatial integration for single trial detection in encephelography," NeuroImage, vol. 17, pp. 223-230, 2002.

[5] J.A. Swets "ROC analysis applied to the evaluation of medical imaging techniques," Investigative Radiology, vol. 14, pp. 109-121, 1979.

[6] A. Delorme, S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," Journal of Neuroscience Methods, vol. 134, pp. 9-21, 2004