Visual Salience and Memory Modulate Viewing Behavior of Natural Scenes in Monkeys, Seth D. König1,2,4 & Elizabeth A. Buffalo1,3,5

1. Introduction

When we view an image of a natural scene certain regions of the image attract our attention, and we move our eyes to fixate on those regions. Many models for predicting where we look have been proposed including a popular visual salience model by Itti, Kock, & Niebur (1998), which employs basic filtering properties of the early visual system to create salience maps. The majority of models use a deterministic method such as a winner-take-all algorithm to identify fixation locations, but deterministic methods do not account for the large variability in fixation locations across and within individuals. Here, we propose an additional layer to the visual salience models in which a biased, correlated random walk (BCRW) predicts fixations in rhesus monkeys during free viewing of natural scenes. In addition to bottom-up factors such as salience, top-down factors can influence where we look. Here, we also investigated how inhibition of return and memory for familiar images alter viewing behavior during a free viewing task.

2. Experimental Setup

Fixation and Saccade Detection: Cluster Fix

Visual Salience

Inhibition of Return (IOR) and Salience

Biased, Correlated Random Walk (BCRW) Model

The Biased, Correlated Random Walk (BCRW) model by Itti, Kock, & Niebur (1998), which employs basic filtering properties of the early visual system to create salience maps, indicates a different cluster. Clusters with the lowest velocity and acceleration are selected to be clusters consisting of global variability. Clusters with the lowest velocity and acceleration are selected to be clusters consisting of global variability. Clusters with the lowest velocity and acceleration are selected to be clusters consisting of global variability. Clusters with the lowest velocity and acceleration are selected to be clusters consisting of global variability.

3. Results

We defined pairs of pixels and generated locations for each fixation that were less than 1.5 pixels apart. The distance between prior and return fixations was significantly greater than the distance between prior and new fixations (p < 0.001, 2-sample Wilcoxon rank-sum test). The BCRW predicted distribution of fixation locations was significantly better at predicting fixation locations than the salience or image intensity models for familiar presentations (p < 0.05 & *** p < 0.001). Shuffling (cross-hatched pattern) of fixation locations significantly decreases the area under the ROC curve for all predictors. The mean area under the ROC curve was significantly greater for the BCRW model that the BCRW predicted fixation locations better than salience, and salience and image intensity (p < 0.05 & *** p < 0.001). We defined pairs of prior and return fixations as pairs of fixations that were less than 2 dva apart. The distance between prior and return fixations was significantly greater than the distance between prior and new fixations (p < 0.001, 2-sample Wilcoxon rank-sum test). The BCRW predicted distribution of fixation locations was significantly better at predicting fixation locations than the salience or image intensity models for familiar presentations (p < 0.05 & *** p < 0.001). Shuffling (cross-hatched pattern) of fixation locations significantly decreases the area under the ROC curve for all predictors. The mean area under the ROC curve was significantly greater for the BCRW model that the BCRW predicted fixation locations better than salience, and salience and image intensity (p < 0.05 & *** p < 0.001).

4. Conclusion

By incorporating realistic eye movements in a salience model in the form of a BCRW, we could predict viewing behavior better than salience alone.

- Inhibition of return regulates the salience at fixation locations and persists for approximately 4.5 seconds.
- Return fixations are ubiquitous and most often occur at salient locations.
- Memory for the scene decreases the relative importance of salience so that during the viewing of familiar scenes, monkeys make fixations in unvisited, less salient locations.
- The hippocampus may be an integral guide in helping behavior in familiar environments.

Future work will create a comprehensive model to predict viewing behavior during both novel and familiar presentations by incorporating experience into the BCRW model.

References


