# Interesting Objects are Visually Salient

Lior Elazary and Laurent Itti
Department of Computer Science and Neuroscience
Graduate Program

## INTRODUCTION AND MOTIVATION

- How do we decide which objects in a visual scene are more interesting then others?
- High-level cognitive processes?
- Low-level Stimulus?


We found that the bottom-up saliency computations showed a 43% probability (chance 21%) of finding an interesting object within the first fixation, and over 76% probability (chance 43%) within the third fixation.

## STIMULUS/EXPERIMENTS

LabelMe dataset created by Russell et al. (2005).

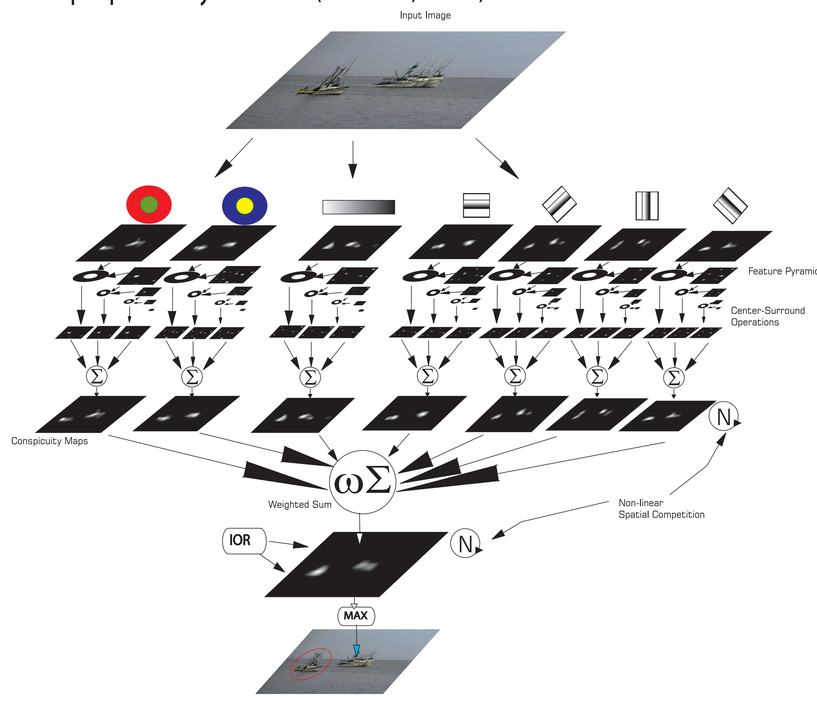
74,454 annotated objects in 24,863 scenes.

Dataset provides a good indication of what people would find "generically interesting"

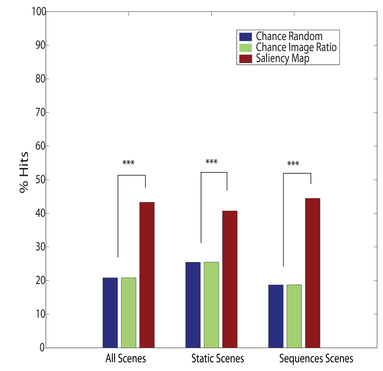
Determined Hit Rate, maximum/ratio between labeled objects and saliency value, and how subsequent fixations indicated a labeled object.

## MODEL

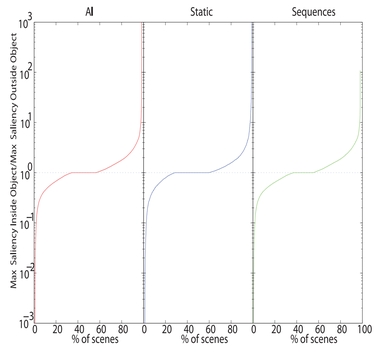
Saliency map was computed according to the algorithm proposed by Itti et al. (Itti et al., 1998).



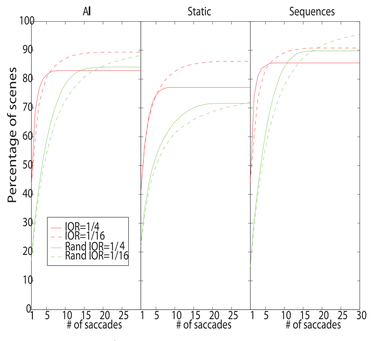

## RESULTS



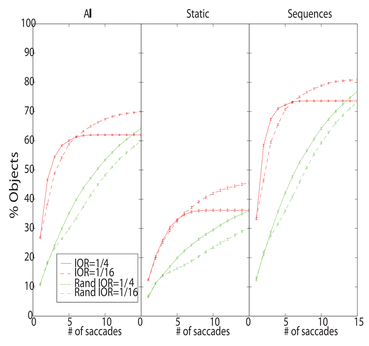

Percentage of hits in the LabelMe dataset. *** = z score > 30, p << 0.001.




Ratio of the maximum saliency value within labeled regions (human-selected objects) to the maximum saliency value within unlabeled regions (background).




Number of saccades taken to reach the first labeled object vs. the number of scenes. Within the third fixation there is a 76% probability of finding a labeled object.

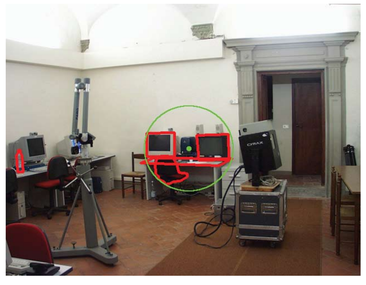



Percentage of objects found in the scenes vs. the number of saccades taken to find these objects.

## RESULTS

Example of a static scene and its saliency map with a labeled object found in the first saliency-guided saccade (to the maximum saliency value over the entire image). Image size: 2592x1944 pixels. IOR radius (green circle): 162 pixels (1/16 of the image width).

Example of a sequence scene where the first saliency-guided saccade resulted in a miss, but the second was able to hit the labeled object. Image size: 720x480 pixels. IOR radius (green circle): 45 pixels (1/16 of the image width).



Example of a static scene where the first saliency-guided saccade was a miss, and it additionally inhibited four labeled objects due to a large IOR. Image size: 330x272 pixels. IOR radius (green circle): 160 (1/4 of the image width).

## DISCUSSION AND CONCLUSIONS

- We therefore conclude that the saliency map is a strong indicator of what people chose to label in complex natural scenes.
- The saliency map showed a 43% probability of finding a labeled object within the first fixation (chance 21%), and over 76% probability within the third fixation (chance 43%).
- This means that even though choosing objects to label might seem like a "free" decision, humans are largely bound by bottom-up processing that influence their higher decision.
- The saliency map computations can be used for object detection and object recognition algorithms, among many other applications.

Itti, L., Koch, C., & Niebur, E. 1998. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11), 1254–1259.

Russell, B. C., Torralba, A., Murphy, K. P., & Freeman, W. T. 2005. LabelMe: a database and web-based tool for image annotation. MIT AI Lab Memo AIM-2005-025, September.