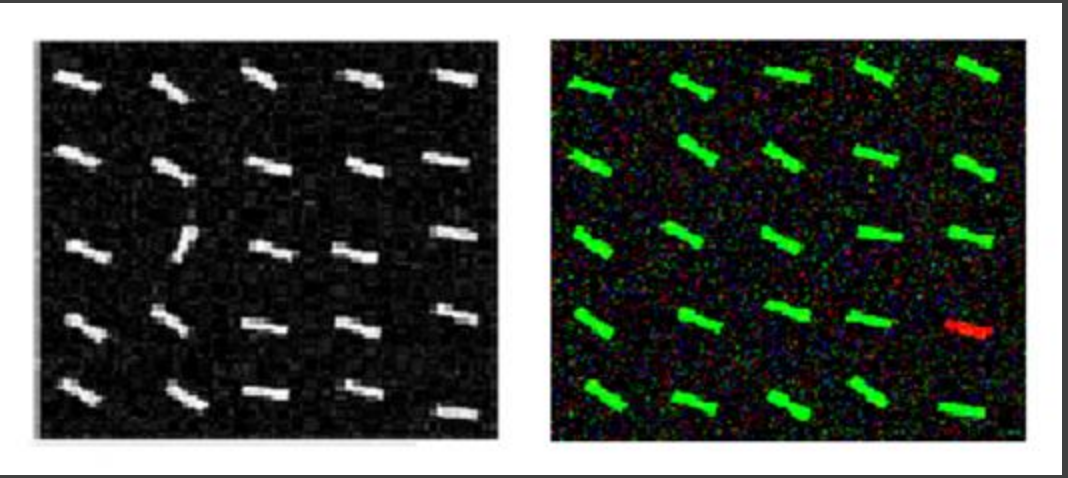
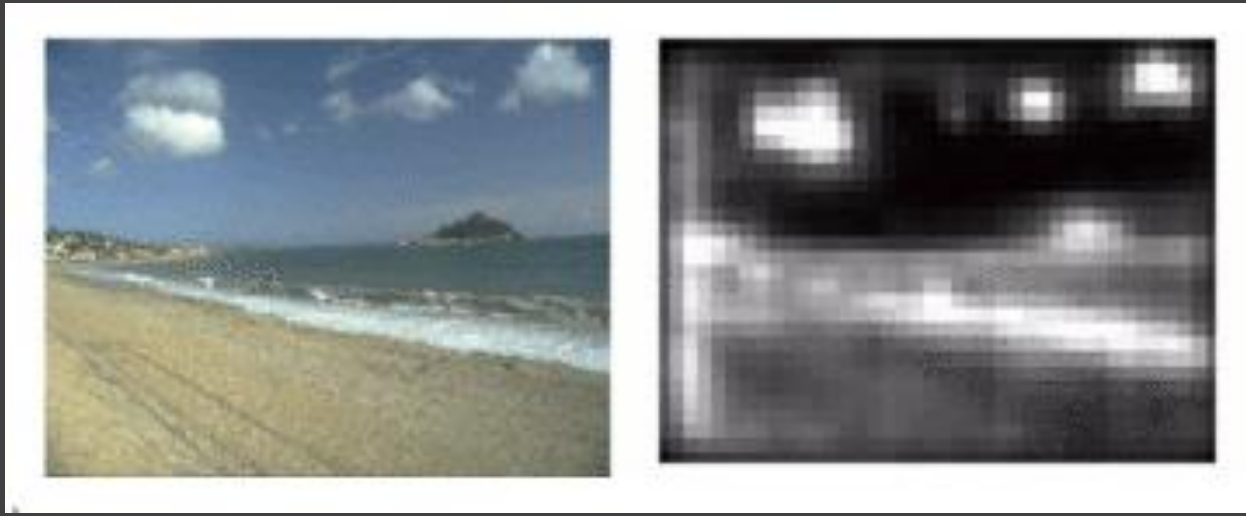
Top Down Attentional Guidance in Visual Search

Bottom Up Saliency

- Attention mechanism is guided by regions of contrast in an image i.e. features like color, intensity and orientation.
- Does not take into account the pre-acquired knowledge of objects.
- Task and goal-independent.

Salient features :
orientation & color





Saliency map
of an image

Top Down Influences

- Attention is guided by knowledge of the visual appearance of objects (how and where objects appear), or features, etc. of the target.
- Scene context guides our attention to regions having high probability of containing target objects.
- Task and goal-dependent.

Bottom-Up Vs Top-Down

- It is clear that top-down information plays a role in guiding our attention.
- The extent of Bottom-Up or Top-Down influence is not clear.
- There is some evidence that Bottom-Up saliency does not drive attention directly but through its correlation with objects(Nuthman & Henderson,2010)

Related Work

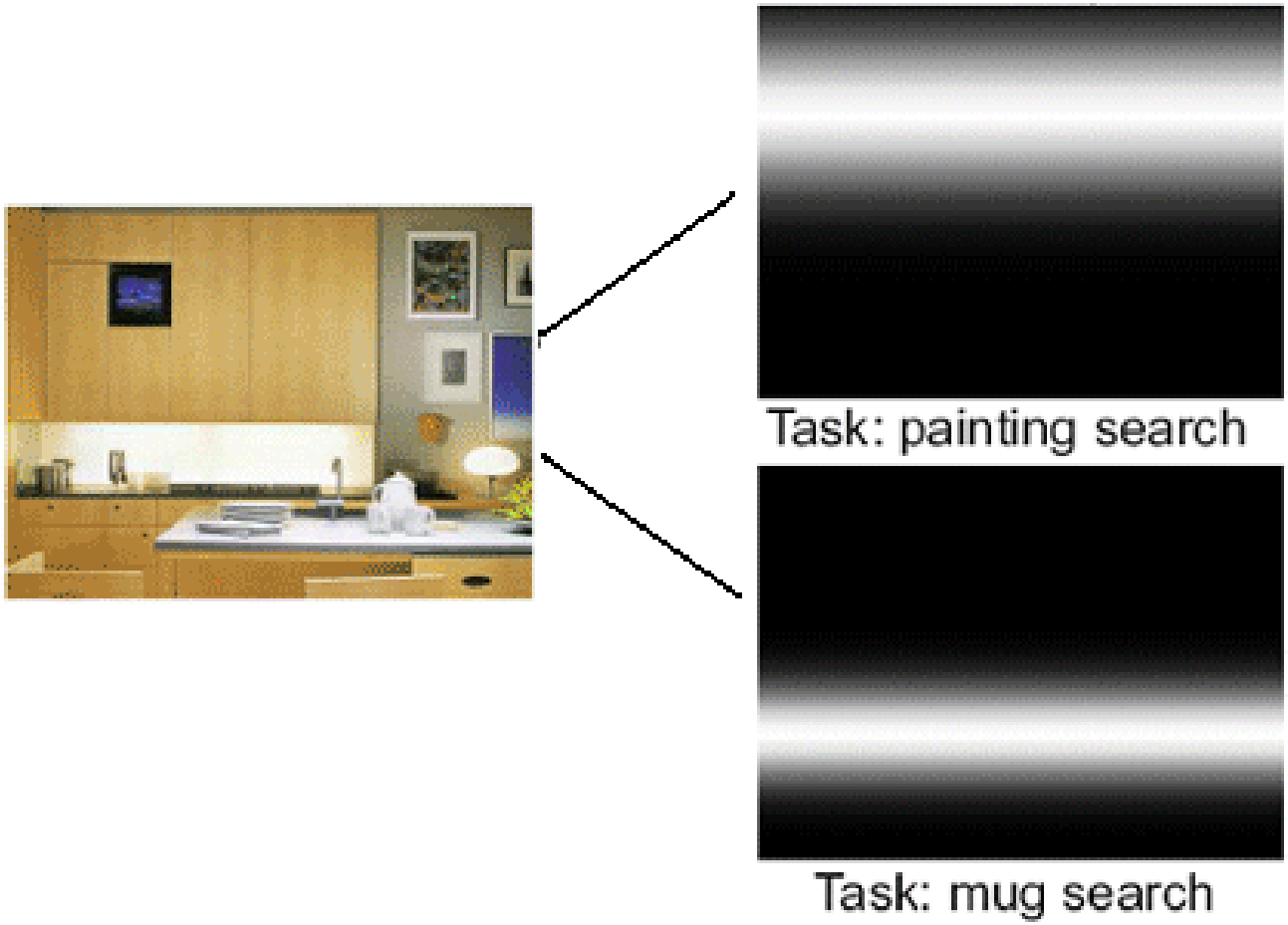
- Bottom up Computational model of Visual attention (Koch & Itti,2001).
- Contextual Guidance of Attention (A Torralba, A Oliva, MS Castelhano, JM Henderson,2006).
- Top-Down Saliency using Natural Statistics

  (C Kanan, MH Tong, L Zhang, GW Cottrell,2009).

Contextual Guidance

- The *gist* of the scene is acquired during the first few hundred milliseconds after the image onset.
- Visual system uses scene context to guide eye movements for exploring the target.
- Regions in the scene that have a higher probability of containing objects are paid more attention.

Contextual Guidance contd.



Top Down Object Based Information

- Specific template of target object is available in visual working memory that guides the search process.
- Visual system tries to match a representation of the target stored in memory against the scene.
- During the search process the regions that contain features related with the target template are fixated for a longer duration.
- Example – To search a mug in a scene, viewer would preferably attend to objects having features similar to that of a mug.

Problem Statement

A model of attention that combines -

- contextual based guidance, and
- top-down object based information

which predicts image regions that are likely to be fixated during visual search.

- Contexual Guidance(Oliva & Torralba,2006)

$$S = p(C=1, L \mid F, G)$$
$$= p(F \mid G)^{-1} p(F \mid C=1, L, G) p(L \mid C=1, G) p(C=1 \mid G)$$
$$\approx p(F)^{-1} p(L \mid C=1, G)$$

- Top-Down Object Based influence(Kanan,2009)

$$S_z = p(C=1 \mid F=f_z, L=l_z)$$
$$\approx p(C=1 \mid F=f_z) + const.$$

Experimental Evidences

- An experiment was done by Malcolm & Henderson to investigate how the visual system combines context based and template based top-down processes to facilitate search.
- Results showed that target template and contextual constraints combine additively to facilitating search.
- It also showed that visual system treats scene context and target template information independently.

Dataset

- Label Me Dataset

B. C. Russell, A. Torralba, K. P. Murphy, W. T. Freeman,

LabelMe: a database and web-based tool for image annotation. International Journal of Computer Vision, pages 157-173, Volume 77, Numbers 1-3, May, 2008.

Review of Objectives

- A comparative study of the saliency maps generated by our model and other competitive models.
- Interpretation of the results in terms of model's accuracy in predicting eye movements during visual search.

References

- *Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real world scene search. Journal of Vision, 10(2):4, 1–11, http://journalofvision.org/10/2/4/, doi:10.1167/10.2.4.*

- *Kanan C., Tong M., Zhang L., Cottrell G. (2009). SUN: Top-down saliency using natural statistics. Visual Cognition, 17, 979–1003.*

- *Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. Psychological Review, 113, 766–786.*

- *John M Henderson & Antje Nuthman(2010), Object-based attentional selection in scene viewing,Journal of Vision(2010), 10(8):20, 1-19*