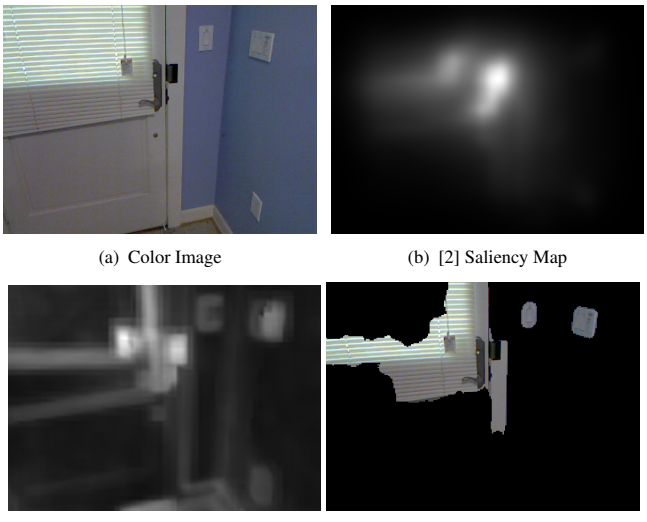
An In Depth View of Saliency

Arridhana Ciptadi
arridhana@gatech.edu
Tucker Hermans
thermans@cc.gatech.edu
James M. Rehg
rehg@gatech.edu

Center for Robotics and Intelligent Machines
School of Interactive Computing
Georgia Institute of Technology

Visual saliency is a computational process that identifies important locations and structure in the visual field. Most current methods for saliency rely on cues such as color and texture while ignoring depth information, which is known to be an important saliency cue in the human cognitive system. We propose a novel computational model of visual saliency which incorporates depth information. We compare our approach to several state of the art visual saliency methods and we introduce a method for saliency based segmentation of generic objects. We demonstrate that by explicitly constructing 3D layout and shape features from depth measurements, we can obtain better performance than methods which treat the depth map as just another image channel. Our method requires no learning and can operate on scenes for which the system has no previous knowledge. We conduct object segmentation experiments on a new dataset of registered RGB-D images captured on a mobile-manipulator robot.

Standard approaches to saliency use color, gradient, and intensity differences to distinguish unique regions from the rest of the visual field. We propose a novel method which incorporates depth measurements into the computation of visual saliency. Human subject studies have shown that depth is an important cue in determining salient regions in human visual processing [5, 6]. Depth measurements make it possible to separate objects which may be similar in appearance. In addition, shape information can be recovered from the depth channel and used to improve the discriminability of scene elements.



(a) Color Image (b) [2] Saliency Map

(c) Our Saliency Map (d) Resulting Segmentation

Figure 1: An example saliency map and resultant segmentation produced from the color and depth image pair. Note the two light switches in the upper right of the image, which our approach highlights as salient regions.

One motivation for saliency research is the development of generic object segmentation and detection capabilities [1, 3, 4, 7]. As an example, consider a personal robot that can move through a home environment and manipulate objects. During a clean-up task, the robot should be able to handle unfamiliar objects for which it has no prior experience. Saliency provides a basic mechanism for characterizing an unfamiliar scene and generating hypotheses about potential object locations. We use the task of generic object segmentation to quantify the effectiveness of our depth-based saliency method. We present a novel segmentation approach that incorporates saliency in an MRF model defined over superpixels. We demonstrate improved performance over several previous approaches.

Consider the images presented in Figure 1 captured on a robot operating in a home environment. The scene has multiple foreground objects of potential interest. Our saliency method scores all the relevant objects (door handle, light switches, window blind) more highly than the background elements of the scene. The resulting saliency-based segmentation successfully separates the foreground objects from the background producing a useful set of proto-objects for a robot to explore or manipulate.

We have collected a new dataset of color and depth images, which form the basis for our experimental evaluation. The dataset was collected using a mobile-manipulator robot in a real-world home environment. The dataset includes ground truth pixel-level segmentations of salient objects that we will release publicly. In summary, this work makes four contributions:

- We introduce a new method for estimating visual saliency, which combines color and depth measurements.
- We demonstrate that explicit 3D layout and shape features from depth measurements produce more informative saliency maps than approaches which simply treat depth as another channel of the image.
- We present an approach to saliency-based segmentation of generic objects based on a superpixel MRF and show promising results.
- We introduce a new dataset for depth-based saliency, including ground truth pixel-level segmentations of salient objects

[1] Stas Goferman, Lihi Zelnik-Manor, and Ayellet Tal. Context-aware saliency detection. In *IEEE Conference On Computer Vision and Pattern Recognition (CVPR)*, 2010.

[2] Jonathan Harel, Christof Koch, and Pietro Perona. Graph-based visual saliency. In *Neural Information Processing Systems*, 2006.

[3] Xiaodi Hou and Liqing Zhang. Saliency detection: A spectral residual approach. In *IEEE Conference On Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007.

[4] Xiaodi Hou, Jonathan Harel, and Christof Koch. Image signature: Highlighting sparse salient regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012.

[5] Congyan Lang, Tam V. Nguyen, Harish Katti, Karthik Yadati, Mohan Kankanhalli, and Shuicheng Yan. Depth Matters: Influence of Depth Cues on Visual Saliency. In *Proceedings of the European Conference on Computer Vision*, 2012.

[6] Jan Theeuwes, Paul Atchley, and Arthur F. Kramer. Attentional control within 3-d space. *Journal of Experimental Psychology: Human Perception and Performance*, 24(5):1476–1485, 1998.

[7] Dirk Walther and Christof Koch. Modeling attention to salient proto-objects. *Neural Networks*, 19(9):1395–1407, 2006.