Visual Salience in Interface Design

Melissa Gallagher
Rensselaer Polytechnic Institute
Troy, NY 12180
gallam@rpi.edu

Bonnie John
Carnegie Mellon University
Pittsburgh, PA 15213
bej@cs.cmu.edu

ABSTRACT

For my summer project as part of the Computing Research Association's Distributed Mentor Program, I worked to see if visual salience significantly affected exploratory behavior in a user interface. I analyzed data collected from the InfoSearch task to see if in that task visual salience had any effect. If visual salience had been significant we would have extending the ACT-R model SNIF-ACT to include salience information in its behavior, to make a more complete model of information search

INTRODUCTION

Visual Salience

Visual salience is the subjective perceptual quality of an object that allows it to stand out from its neighbors and grab our attention [1]. Koch and Ullman [3] proposed a neurally plausible computational architecture for predicting bottom-up visual attention that involves creation of a saliency map. A saliency map is a topological map that includes the saliency value for every point in a scene. This form of bottom-up visual attention is computed in a pre-attentive manner and operates very rapidly. The top-down or goal directed visual attention is dependent on the task at hand. Both of these forms of visual attention can run in parallel [2].

Latent Semantic Analysis (LSA) and Semantic Fields

LSA is a statistical method for evaluating text that allows researchers to determine similarity between words and sets of words.[4] LSA is trained on a corpus which functions as a knowledge base. LSA-SF takes into account the layout and the semantic values of the elements of a webpage and creates a heat map based on a decay function [5]. This heat map is used to predict eye movement toward relevant information for a visual search task.

SNIF-ACT and CogTool

CogTool is used for cognitive performance modeling for interactive devices. It accurately predicts skilled user behavior in an interface using KLM and GOMS. There is a need to be able to predict the time for exploratory users in an interface. There is current work extending SNIF-ACT to create a more complete model of exploratory user behavior. SNIF-ACT operates on the information foraging theory but only looks at the information scent of links with little focus on the layout of the page.

METHOD

Participants

There were eight participants in the InfoSearch Task. There were 6 females and the average age was 18.75.

APPARATUS

Behavioral recording equipment

For this task an eye-tracker from LC Technologies system was used with a sampling rate of 16 Hz. The task was conducted in a modified version of the Firefox™ internet browser. The Visualization-Interaction Architecture (VIA), was used to record human-browser interaction. VIA has a custom data analysis tool to produce information like semantic relatedness maps and visual saliency maps.

Procedure

Participants were given eight reference questions to answer using only Wikipedia™. They were not allowed to use the search feature and could only navigate through Wikipedia by clicking on links. The question and answer form was contained in a one tab of the browser and the Wikipedia was displayed in a separate tab. The participants started their search on the portal of portals page every time (http://en.wikipedia.org/wiki/wikipedia:main_page/portal:contents/Portals).

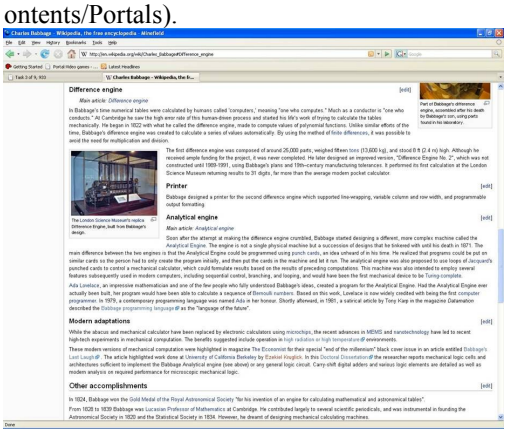

Figure 1- An example where the answer to the question was found

Figure 2- An example semantic map for the page where the answer was found

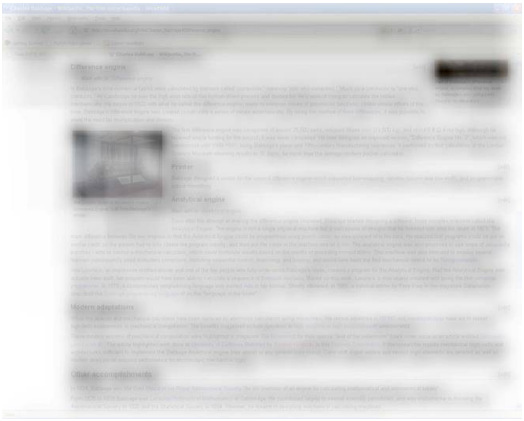

Figure 3- An example salience map for the page where the answer was found

RESULTS

The first and last fixations on the first and last page

The first set of analyses was on the semantic and salience values for the first ten fixations on the start page and the last ten fixations on the page the answer was found on were examined. These values are shown for two users in Figures 1 through 4. For the salience values on the start page the values remain low but do not vary greatly from each other. The semantic values on the start page increase from the first fixation to the second but do not significantly vary after that. For the last page the salience values increase consistently over the last ten fixations, and the semantic values decrease of the last ten fixations. This early analysis suggested that the participants were focused on completing the task at the beginning of the experiment, but then once they found the answer, the high salience points on the last page, they would then look at things that were visually interesting. This may or may not have been the case because to input the data the participant would have to click back to the other tab and the location to switch tabs is a more salient one. Due to this finding we used another method of data analysis to find test the hypothesis.

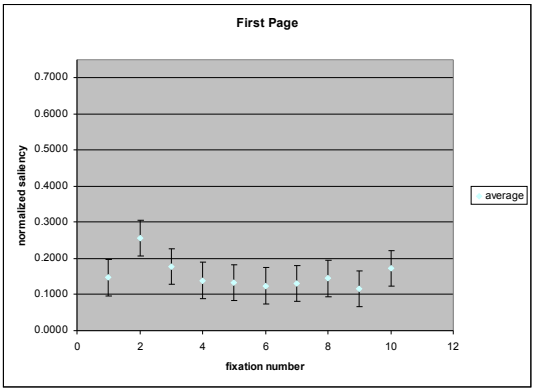

Figure 1- The average normalized Salience values for the first ten fixations on the first page

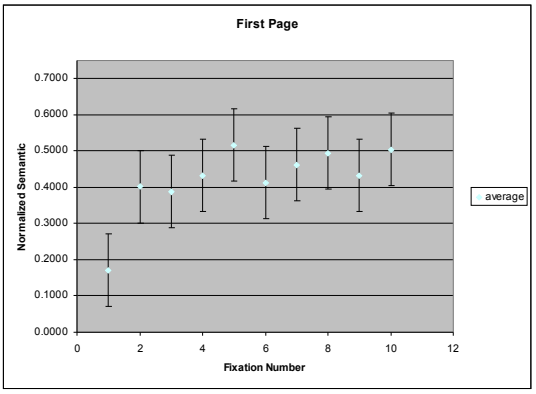

Figure 2- The average normalized Semantic values for the first ten fixations on the first page

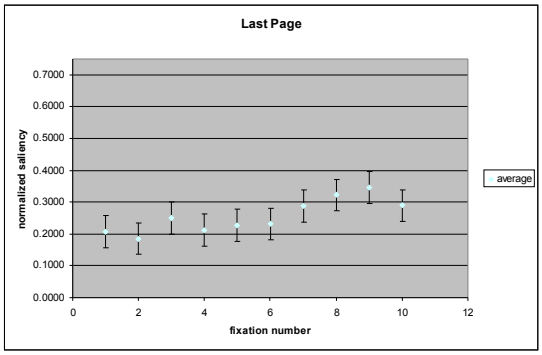

Figure 3- The average normalized Salience values for the last ten fixations on the page the participant found the answer

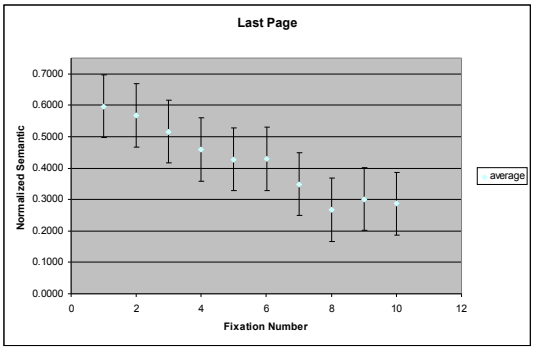

Figure 4- The average normalized Semantic values for the last ten

fixations on the page the participant found the answer.

First Viewport Analysis

To try and eliminate some of the effect that the browser interface had and to focus on the page layout we decided to look at the first five and the last five fixations on the first viewport. A viewport is the section of the page that is visible before the user scrolls the page or navigates away to another page. We looked at the first viewport for the first time a participant views a page so any effects on saliency for repeated viewing would be controlled for. We also selected only the pages that were on the path to the answer. Pages the participant deviated to and then went back to their path were omitted. We then differentiated the last fixations on the page by whether or not people were scrolling to a new part of the page or they were clocking a link to a new page. We did this because if a participant is clicking on a link it would be expected to have high semantic values where if they are scrolling away it should have lower. We only looked at first viewports that had more than ten fixations on them and had a length shorter than 16 seconds. We found that there were no trends for salience and semantic values for the first and last fixations on a viewport (Figure 5 and Figure 6). There was a small difference when we differentiated why the participant left the viewport according to the trend that was expected, though this difference was not statistically significant.

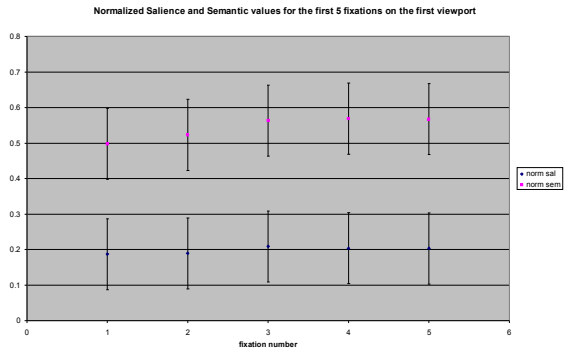

Figure 5- The average normalized Salience and Semantic values for the first five fixations on the first viewport

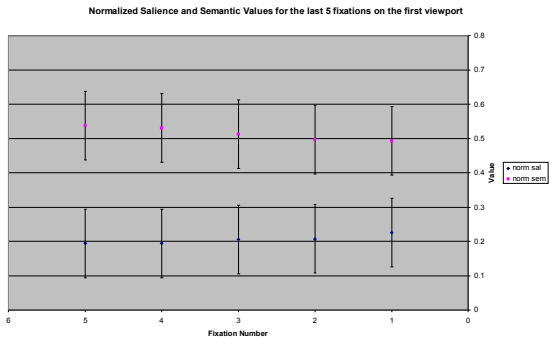

Figure 6- The average normalized Salience and Semantic values for the last five fixations on the first viewport

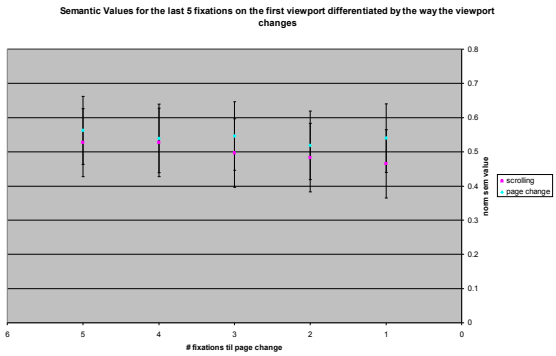

Figure 7- The average normalized semantic values of the last five fixations on the first viewport differentiated by method of exiting

INTEGRATION

Integrating Salience code with CogTool extension

To including salience values as a CogTool extension we utilized the iLab's visual salience code. The code is written in C++ and is designed to run on a UNIX based operating system. For the project we had to engineer the code so it would run on Windows. We then created a Java class that would communicate the salience values of objects to the SNIF-ACT model. Since we found the salience values to have little effect on exploratory user behavior this extension was not included.

DISCUSSION

While for this task we found no effect on salience or semantic value we feel that this is an area for future study in more interfaces, especially ones where salience is important. There were a number of reasons that we found no effect on salience. The first reason is that in Wikipedia most of the pages are very similar primarily consisting of text and not many objects. The second reason is that on the pages that do have pictures, the more salient points, the answers to the questions are not on the pictures so the users would tend to ignore them. The other problem is that since we normalized the values for the measures small differentiations in the data look much bigger. Once salience's effect on a user interface has been determined a more complete model of exploratory user behavior can be created.

ACKNOWLEDGMENTS

We thank the CRA-W DMP, the CogTool team, and the CogWorks lab for their help on this project.

REFERENCES

1. Itti, Laurent (2007) Visual salience. Scholarpedia, 2(9):3327.
2. Itti, Laurent & Koch, Christof. Computational modeling of Visual Attention. *Nature Reviews Neuroscience 2*, (2001), 194-203.
3. Koch, C. & Ullman, S. Shifts in selective visual

attention: towards the underlying neural circuitry. Hum. Neurobiol. 4, 219-227 (1985)

4. Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. Psychological Review, 104, 211-240

5. Stone, B., & Dennis, S. (2007). Using LSA Semantic Fields to Predict Eye Movement on Web Pages. In D. S. McNamara & J. G. Trafton (Eds.), 29th Cognitive Science Society Proceedings (665-670). Austin, TX: Cognitive Science Society.