Where Is Waldo?

Visual Search Behavior in "Wimmelpictures"

Simone Liegl, Raphael Fritsch, Gabriela Tullius

Reutlingen University
Alteburgstraße 150, Reutlingen, Germany
{simone.liegl, raphael.fritsch}@student.reutlingen-university.de;
gabriela.tullius@reutlingen-university.de

Abstract

In order to explore an image, the human eye functions like a spotlight, scanning the content from one object to the next. This visual search behavior is implemented with the help of attention control. The following work surveys the visual search behavior in "Wimmelpictures", a special type of busy pictures. The research objective is to analyze different search strategies and to work out possible differences concerning age and gender. The university experiment is carried out by an eye tracker that records the fixations and saccades of the test persons. The results indicate three forms of search strategy: based on a pattern, based on feature selection, or a mixture of both. Our data shows the search for special features of the target is the most successful. Furthermore there are no differences concerning gender but some concerning age. All age groups need more time to locate the target with an increasing number of distractors in the image. The size of the target is also relevant as a larger target is found more quickly than the smaller one.

Keywords: Visual search behavior, "Wimmelpictures", Eye tracking

In: F. Pehar/C. Schlögl/C. Wolff (Eds.). Re:inventing Information Science in the Networked Society. Proceedings of the 14th International Symposium on Information Science (ISI 2015), Zadar, Croatia, 19th–21st May 2015. Glückstadt: Verlag Werner Hülsbusch, pp. 320–325.

1 Motivation

Cameras were built using the same principles as the human eye: the retina as a highly-developed complex electronic system and the lens performs focalizing. With these sort of cameras humans decide where to look and where to focus. They use the complex process of attention control to concentrate on relevant objects in one specific context while ignoring the rest. “Selective Attention is the ability to focus on only those items in the environment that are relevant to the present situation” (Cointin 1995: 3). This process can be observed in the so-called “Wimmelpictures”. These are busy pictures showing different scenes full of people. For the purpose of analyzing the visual search behavior with the help of eye tracking, we use an image created by Martin Handford. The aim is to find “Waldo” (see fig. 1). This form of visual search “refers to the process of visually scanning a scene and forming a conceptual “image” or notion of the scene as assembled in the brain” (Duchowski 2007: 222). The research objective is to analyse the search strategy employed by each subject and to determine differences in behavior based on gender and age.

Figure 1. Wimmelpicture with Waldo and eye fixation points

2 Theoretical grounding and derivation of hypotheses

The derivation of hypotheses is supported by theoretical knowledge of the eye, the visual system and attention control. In the following the focus is not on demonstrating the range of those areas but on using that knowledge to reason the various hypotheses.

Prior to the experiment realization, the research group had created five hypotheses. The first hypothesis predicts differences in search time due to the subjects' age. According to Duchowski (2007) the fixation times in pictures are influenced by factors like task, image size, image content and image type (2007: 222, 224). Cointin (1995) shows that independently of these factors, there are differences in visual search across lifespan. Children at the age of six and older adults require longer for search success compared to students. These observations can be explained by the development of the frontal lobe, which is responsible for attention control. This part of the brain evolves last and degenerates first (Cointin 1995: 2, 15). Our first hypothesis reassesses this statement: *(H1) Subjects between 20 and 30 years of age will solve the task in the shortest amount of time*. To test this hypothesis it is necessary to plot the variables *(V1) total time* and *(V2) age of subject*.

Background research indicates that gender is not a determining factor of search time. Due to this, we assume that there are no differences in gender. *(H2) Gender has no influence on the temporal result.* Variables (*V1)* and *(V3) Gender* are needed to verify H2.

The more complex an image, the more time is necessary to find Waldo. A subject requires more time to find Waldo in an image with a lot of distractors than in an image with less distractors (Thompson & Massaro 1989: 337–340). To verify this hypothesis the variables *V1* and *(V4) number of distractors* are needed for *(H3): The time to task solution increases with increasing numbers of distractors. V4* has two manifestations, depending on the number of distractors: few and many.

Treismann and Gelade laid the foundation of attention control with their feature integration theory of attention and other models like those from Itti et al. followed. Attention shifts more readily to objects that protrude because of their features, such as color or size (Treismann & Gelade, 1980: 97; Itti & Koch 2001: 196). The research group tests feature size with the following hypothesis: *(H4) A target with a larger size than the distractors is found*

more quickly than a target with the same size as the distractors. The variables *V1* and *(V5) size of target* (same size as distractors, bigger size as distractors) are used to test H4.

The last hypothesis concerns search strategy. The research group assumes the existence of different forms of search strategies: Either a search structure in pattern such as left to right; or a search structure based on a feature of the target; or finally a change of strategy during the search due the lack of success. *(H5)Three forms of search strategy exist: (1) pattern search, (2) feature selection search and a (3) mixture of both.* Therefore *(V6) route of fixations* is evaluated manually by watching each recording.

3 Outline of the study

The study was run at two open days at the Reutlingen University with 77 test subjects with a sample rate[1] over 60% of the ages between six and 60. The highest distribution with 55% is the age range between 16 and 25, with no subject between 36 and 41. 58% of the subjects were male, 42% female.

The experimental setup consisted of a computer with a Tobii X30 eye tracker. The experiment was composed of four levels: (1) small image and tall target – easy[2], (2) big image and tall target – medium, (3) small image and small target – difficult and (4) big image and small target – hard. For each level the target was placed in a different position. The subjects had 5 min 30 sec in total to solve all levels and 96% completed the task successfully. The average test time[3] was 2 min 51 sec, the average search time[4] is 1 min 53 sec. The best time was a search time of 28 seconds.

1 Rate of the acquired fixations

2 Names are in relation to hypothesis 3 and 4

3 Time of hole test including getting in the next level

4 Search time without getting in the next level

4 Results and discussion

Starting with a univariate analysis we could not confirm the first hypothesis as above as the distribution was spread widely. This result was confirmed by a statistical validation (ANOVA significance test with $p < 0.001$). The second hypothesis, *Gender has no influence on the temporal result,* could be confirmed using an unpaired t-Test with $p = 0.0437$.

Table 1. Paired t-test for the different levels

	medium	difficult	hard
easy	t(76) = 1,47 p = 0.1444	t(76) = 4,7 p < 0,0001	t(76) = -2,33 p = 0.02228
medium	--	t(76) = 4,06 p = 0.0001	t(76) = -2,66 p = 0.009365
difficult	--	--	t(76) = -3,14 p = 0.002354

To test hypotheses 3 and 4 (dependency on *distractors*), we analyzed each level with respect to the distractors. A similar pattern was determined for all search levels, comparing the sample means using a paired t-test with the results shown in table 1 confirmed our findings for hypotheses 3 and 4. The time in level three (difficult) is 31 seconds less than in level four (hard), the medium level is solved in ten seconds longer than level three (difficult). Furthermore the influence of the target size is emphasised by the fact that the average search time in the medium level (24 sec) is 21 sec faster than in level four (hard, small target).

For hypothesis 5 we derive from the eye tracker data that three strategies were used: a sequential pattern search, feature-based search and mixture of both. To summarize our findings we can state that subjects using a feature-based search (74%) were approx. 1 minute faster than a pattern type search (16%) and those changing their search strategy (10%). Furthermore it is evident that while performing the feature-based search the test persons were guided by the salient features. With this study we underline the research results in visual search behavior, such as Ehinger et al.

Our study indicates a number of conclusions. The most relevant is the influence of search strategy on time-based performance, which is important for the development of user interfaces for image retrieval. The search efficiency

could be enhanced by adapting users' interfaces to encourage feature-based searches for natural and artificial images instead of sliding-window techniques. Last but not least on a formative level the subjects rated the study design with Wimmelpictures very positive and motivating, hence we can encourage its use for similar works.

References

Cointin, E. H. (1995). Visual Attention Differences Across the Lifespan: A Study of Inhibition. Honors Project, Paper 96. Bloomington: Illinois Wesleyan. http://digitalcommons.iwu.edu/psych_honproj/96/ <4.1.2015>.

Duchowski, A. T. (2007). *Eye Tracking Methodology. Theory and Practice*. Second Edition. London: Springer.

Ehinger, K. A., Hildago-Sotelo, B., Torralba, A., Oliva, A. (2009). Modelling search for people in 900 scenes: A combined source model of eye guidance. *Visual Cognition* 17, 945–978.

Flowthow, S. (2009). Eye Tracking. Ein Überblick über Geschichte, Methoden und Anwendungen. Wiesbaden: Hochschule Rhein-Main. http://www.cs.hs-rm.de/~linn/fachsem0809/eyetracking/ <4.1.2015>.

Itti, L., Koch, C. (2001). Computational Modelling of Visual Attention. *Nature Reviews Neuroscience*, 2, 194–203. http://www.nature.com/nrn/journal/v2/n3/abs/nrn0301_194a.html <4.1.2015>.

Nielsen, J., Pernice, K. (2010). *Eyetracking Web Usability*. Berkeley: New Riders.

Thompson, L. A., Massaro, D.W. (1989). Before you see it, you see its parts: Evidence for feature encoding and integration in preschool children and adults. *Cognitive Psychology*, 21, 334–362.

Treisman, A. M., Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology*, 12, 97–113.