Gaze Tracking and Scan Path Analysis in Where's Waldo Puzzles

Justin Billingsley
Clemson University
jabilli@g.clemson.edu

Michael Hauser
Clemson University
mhauser@clemson.edu

André Robinson
Clemson University
andre3@clemson.edu

ABSTRACT

Analyzing ways that humans scan through environments in order to find objects of interest to them allows for observation of scan methods in order to determine if similarities across samples occur. This experiment focused on subjects' scan paths to solving Where's Waldo puzzles dependent on restricted versus unrestricted viewing instructions. Scan paths between the two groups, unrestricted and restricted, were then compared using the Levenstein distance in order to determine overall similarity within the two groups and then compare across the two groups.

KEYWORDS

Scan path analysis, Levenstein distance, visual search

ACM Reference Format:
Justin Billingsley, Michael Hauser, and André Robinson. 2018. Gaze Tracking and Scan Path Analysis in Where's Waldo Puzzles. *ACM Trans. Graph.* 37, 4, Article 111 (August 2018), 5 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The time-classic puzzle solving game of Where's Waldo has been a constant in the adolescence of most children in Western Civilization since 1987. In this experiment, our main focus was using this popular and easily recognizable game to determine if there existed an average scan path method occurring across the pool of subjects utilizing the Levenstein distance in order to better understand how users visually scan pictures for objects they wish to find. The Levenstein distance seen in figure 1, is an algorithm which determines

$$\operatorname{lev}_{a,b}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0, \\ \min \begin{cases} \operatorname{lev}_{a,b}(i-1,j) + 1 \\ \operatorname{lev}_{a,b}(i,j-1) + 1 \\ \operatorname{lev}_{a,b}(i-1,j-1) + 1_{(a_i \neq b_j)} \end{cases} & \text{otherwise.} \end{cases}$$

Figure 1: Levenstein mathematical formula

the least amount of changes needed between string "a" and string "b" in order to have both strings be identical. A larger Levenstein distance calculated indicates greater variation between the two strings. In this study, we use this algorithm to compare the scan paths the subjects used to attempt solving the Where's Waldo puzzle. Upon initial discussion, the expected result of the experiment was that the users given no guidance regarding scan method to use would result in faster solution times than the subject pool given explicit instructions(scan top-down, left to right). We determined this possible outcome with the idea that the explicit instructions mimicked a linear search algorithm. In most cases, performing a linear search based algorithm will yield less than favorable results in comparison to a search algorithm which includes degrees of randomness. Due to the complexity of the Where's Waldo puzzle, the exploration pattern becomes more stereotyped yielding a less random scan path method[Di Nocera et al. 2006] which supports our underlying hypothesis. We believe that, after calculating the average Levenstein distance within groups for each puzzle, upon comparing these two distances they should be approximately the same. This would indicate that both directed and experimental groups search using some consistent scan path method.

Authors' addresses: Justin Billingsley, Clemson University, jabilli@g.clemson.edu; Michael Hauser, Clemson University, mhauser@clemson.edu; André Robinson, Clemson University, andre3@clemson.edu.

0730-0301/2018/8-ART111 $15.00
https://doi.org/10.1145/1122445.1122456

2 BACKGROUND

In recent years, viewer interest shown through regions-of-interest and visual point-of-regard measurements have been used to provide keen insight on how viewers examine images and scenes. Clustering algorithms have also been utilized in order to detect and analyze the key regions of interest when viewing scenes/images without any constraints towards the viewer[Santella and Decarlo 2004]

The Levenstein distance has commonly been used to determine similarities between two strings which can represent countless experimental areas of interest, for example scan paths. Fixation mapping and utilizing the Levenstein distance which uses pre-established areas of interest(grid squares) seemed to be the best avenue to acquire quantifiable data regarding how subjects view/solve Where's Waldo puzzles. This was then used to make conclusions regarding whether subjects, given no constraints, scan these images similarly to each other inherently.

2.1 Areas of Concern

We expected the calculated Levenstein distances between subjects in the same group to have small variances due to the presence of saccades deviating from the subject's desired scan path unknowingly.[Port et al. 2016]

We also remained aware that, although we may have dictated the desired scan path method to subjects in Group 1, this did not necessary mean that these individuals would comply to said guidelines. This would then result in the skewing of data and relative Levenstein distance calculations due to outliers present. We also had to maintain awareness that subjects had the possibility of believing they had found Waldo when instead they had identified a false image present in the puzzles placed to confuse users where the figure was incredibly similar yet not actually Waldo. In order to mitigate this concern, we instructed all users to stare directly at the figure they determined to be Waldo for 3 seconds prior to pressing the [spacebar] in order to allow them to be completely certain they were identifying Waldo as well as allow for the fixation to be properly identified when viewing the fixation maps following result analysis.

3 MATERIALS AND METHODS

3.1 Apparatus

The Gazepoint GP3 eye tracker is a standalone eye tracking device. The eye tracker is able to be positioned at the top or bottom of the computer screen; however, for this particular experiment all Gazepoint GP3 eye trackers were mounted at the bottom of the computer monitor and were unobtrusive to the user. The Gazepoint GP3 tracks where the user looked on the screen, with a degree of accuracy of half a degree, about 50 pixels. The Gazepoint tracked the gaze of both eyes, and the distance of the eyes from the screen at 60 Hz. The puzzles utilized were displayed on a Dell P2213 with a refresh rate of 59 Hz.

Figure 2: Gazepoint GP3 eye tracker

3.2 Stimuli

A selection of three full-color, 1600x900 pixel Where's Waldo images(red boxes indicating Waldo removed). The images were populated with distractions or similar characters to confuse or distract the subject from finding their target, Waldo.

Figure 3: Waldo Character

The three images were displayed in the following order(without red boxes indicating where Waldo was located).

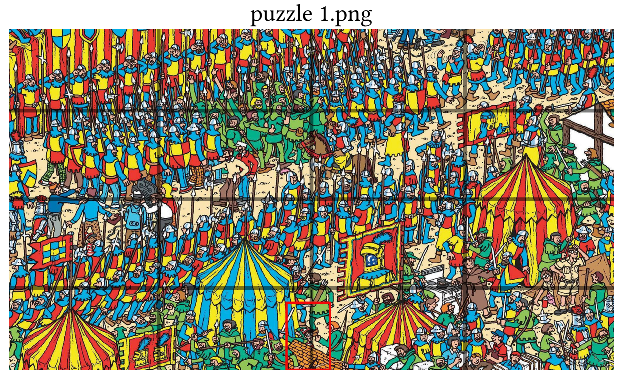

Figure 4: Waldo Image 1

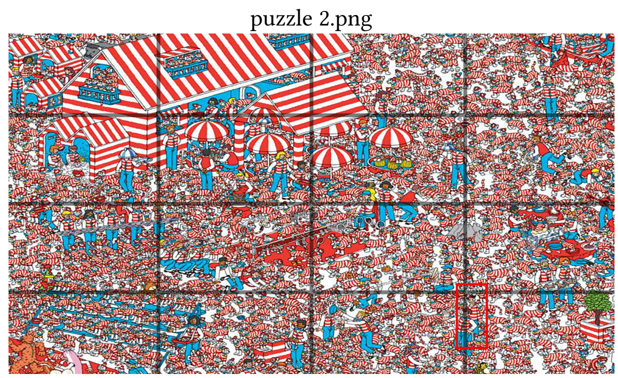

Figure 5: Waldo Image 2

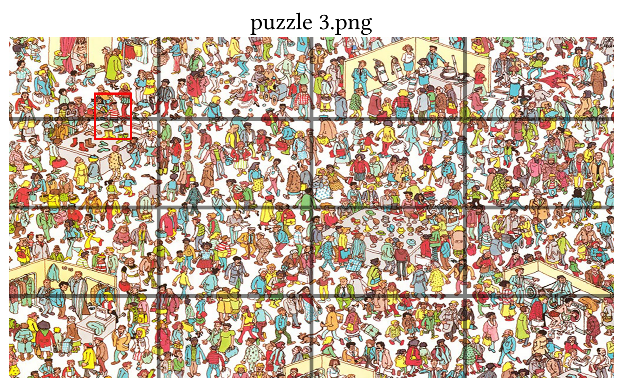

Figure 6: Waldo Image 3

3.3 Subjects

Eleven Clemson University undergraduate students, ranging from ages 19-26 With a mean age of 21, median age of 21 and Standard Deviation of 1.81, volunteered to participate in the experiment. All subjects had normal or corrected-to-normal vision with. One factor to remain aware of is that 3/12 subjects also indicated they have astigmatism. The sample was divided into two groups with Group 1 having knowledge of the desired scan path method to mimic while Group 2 was given zero guidance in regards to scan path method in order to receive the most natural search reaction to the stimuli presented.

3.4 Experimental Design

The experiment was performed as a 3x2 study with three Where's Waldo puzzles, one group of five participants(guided), and one group of six participants(unguided). The experiment focused on between-subjects data comparison to determine two desired data points. The first data point being the time that it takes the subjects to find Waldo in each of the images. The second being the Levenstein distance calculated between-subject groups to find the similarities of the scan paths used in the experiment. An idealized scan path was used in Group 1 to establish the directed group as the control since we expected the Levenstein distances inside Group 1 to be incredibly similar. The idealized scan path dictated to Group 1 is depicted in Figure 7 using puzzle 1 and represented with the alphabet A through P corresponding with a unique grid square.

1 AP.jpg

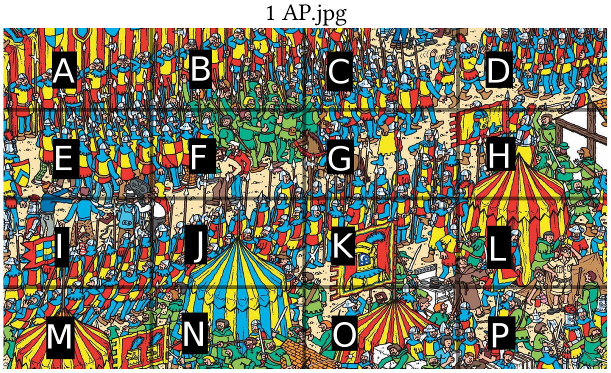

Figure 7: Original Puzzle w/ AOIs

3.5 Procedures

Subjects were unknowingly divided into one group of 5 and one group of 7. Group 1 being the directed group while Group 2 was the control group being given no instructions regarding a scan method to use when attempting to solve the Where's Waldo puzzles. Grid squares had been added to the puzzles in order to aid the directed group with scanning through each square in the order desired. Group 1 had been directed to scan each puzzle beginning at the top of the puzzle and following the grid path from left to right scanning to find Waldo until reaching the bottom right grid square, essentially "reading" the puzzle as if each row of grid squares was a sentence. Prior to the experiment, each subject sat in a chair in front of the monitor and, if part of Group 1, was told explicit scan path instructions prior to beginning the experiment.

Subjects were prompted stare at Waldo for 3 seconds then to press the [spacebar] upon discovering Waldo in order to indicate the task was completed. Each subject was given 2 minutes to find Waldo per puzzle for a total of 6 minutes maximum experimentation time per subject. Each subject was then shown each Where's Waldo puzzle in the order stated above and the resulting scan paths and elapsed time was recorded to be compared between subjects and subject groups.

4 RESULTS

The main intent of this experiment was to determine if Group 2, given no explicit scan path method, would follow a generalized scan path in accordance with the scan path created based on the directions given to Group 1. We also compared average completion times for each puzzle between the two groups in order to come to a conclusion as to whether Group 2 would indeed yield a faster completion time than the linear-search-esque scan path method given to Group 1. Utilizing the Levenstein distance, we were able to quantify the average statistical difference of the scan paths between subjects in the directed group and non-directed group.

As you can see in Figure 8 derived from our data, undirected puzzle scan paths for puzzles 1-3 depicted the highest level of variance with a greater Interquartile range(IQR) indicating that subjects in Group 2, on average, did not follow a scan path similar to the idealized scan path method dictated to Group 1.

4.1 Levenstein Distance comparisons

Avg Levenstein Distance by Trial		
Trial	Group 1	Group 2
Puzzle 1	22.533	24.1
Puzzle 2	26.4	25
Puzzle 3	23.2	13.5

The above calculations indicate the resulting mean Levenstein distances calculated from the experiment for each puzzle with Group 1 indicating Directed and Group 2 indicating Undirected respectively. In relation to the table above, it is clearly evident that Puzzle 3 influenced the Levenstein distances calculated for Group 2 strongly. This was determined to be most likely due to the fact that Waldo was present in the first grid square viewed by Group 1 and had the potential to be easily missed. It was determined that Group 1 subjects may have possessed some bias developed from the previous puzzles in the sequence where Waldo was present much farther along their designated scan path so they overlooked the first grid square('A') almost immediately.

The mean Levenstein distances were calculated by taking the recorded scan paths from each subject in a group for each puzzle and then comparing each scan path to each other using the Levenstein distance formula and then averaging up the determined Levenstein distances in order to determine the overall similarity estimate for how close the scan paths for every subject were to each other.

Avg Levenstein Summary Inc. Puzzle 3		
Group number	Avg L-Distance	Significance
1	20.87	0.026
2	24.04	
Avg Levenstein Summary Excluding Puzzle 3		
Group number	Avg L-Distance	Significance
1	24.55	0.949
2	24.47	

In relation to the table above indicating "Avg Levenstein Distance by Trial", in order to determine how much influence the Puzzle 3 results would have skewed overall data, we calculated the Average Levenstein distance observed in both groups then calculated the significance at the 95% level including and excluding Puzzle 3 in our calculations. When we included Puzzle 3 in our average calculations, the significance Pr(F>) = 0.026 indicating the likelihood of any difference between the two groups being due to random chance being extremely small or 2.6%. When Puzzle 3 was excluded,

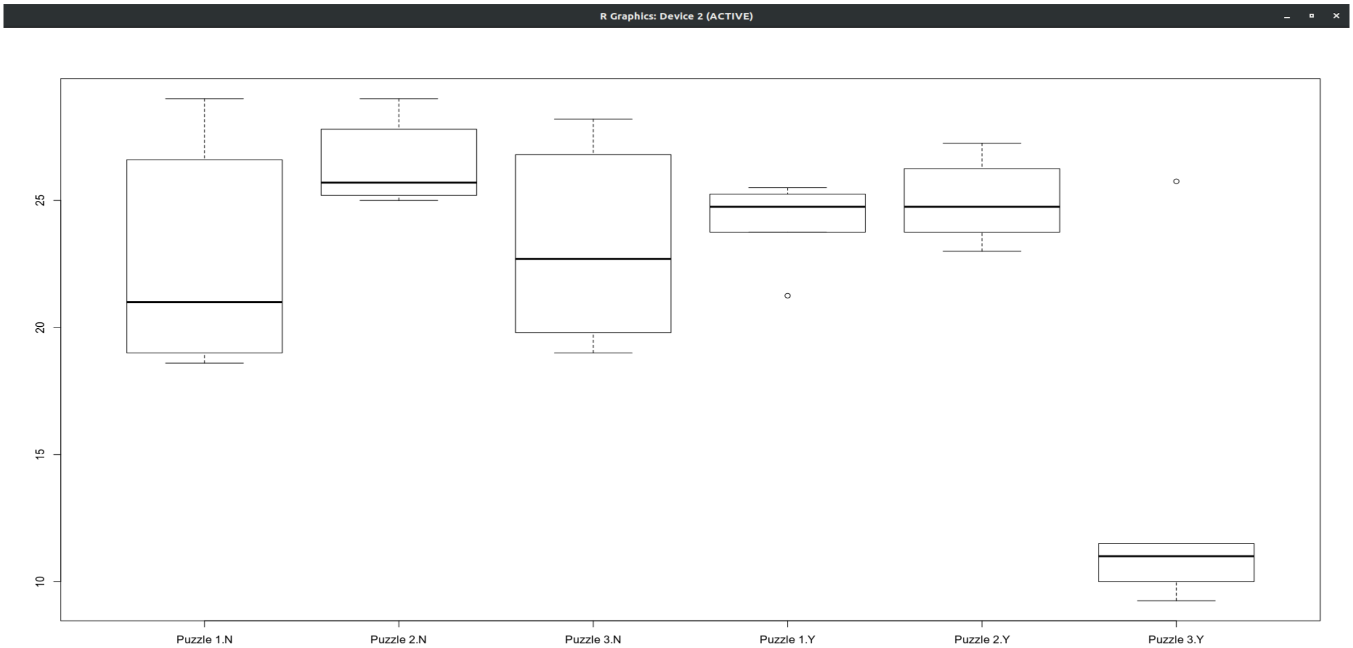

Figure 8: Box-whisker plots regarding Levenstein Distances

however, the Pr(F>) calculation = 0.949 or a 94.9% likelihood that any difference in the distances occurred by random chance.

4.2 Puzzle Time Completions

Avg Puzzle Completion Time by Trial		
Trial	Group 1	Group 2
Puzzle 1	29.35	13.04
Puzzle 2	63.03	49.64
Puzzle 3	25.29	48.15

The time results it took for each subject to complete each puzzle was recorded and averaged to achieve the average puzzle completion time each group had per puzzle. On Average, Group 2 completed each puzzle in significantly less time than Group 1. We expected the undirected group to accomplish the task of finding Waldo faster than their directed counterparts due to the fact that Group 1's scan path method mimicked a linear search which has O(n) runtime(assuming Group 1 members found Waldo along their search the first time) . While Group 2's search methods were completely up to whatever subject was performing the experiment at that time. We believed that randomized searches had a higher likelihood of finding Waldo either immediately or within shorter periods of time.

The only area which does not correlate with our hypothesis was Puzzle 3 as well with the puzzle completion time actually taking significantly (1.904x) longer for Group 2 to find Waldo. Also, it is important to remember that average representations may be skewed by outlier subjects finding Waldo either incredibly fast, never finding Waldo, or taking a much longer time to find Waldo than their group subject members.

4.3 Fixation Maps

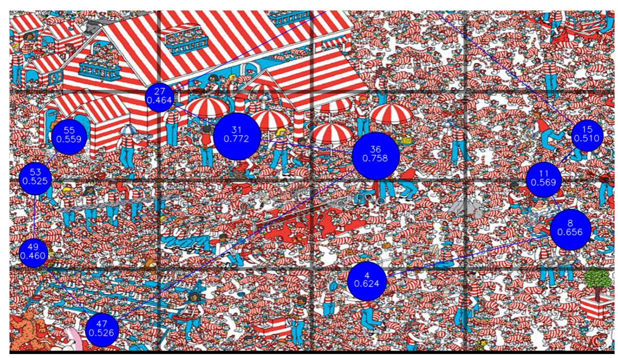

Figure 9: Fixation Map 1

Figure 9 depicts the fixation map compiled from Group 2 illustrating onto the puzzle the scan path this user chose. This Group 2 subject depicted a scan method appearing to climb up grid square columns and then down subsequent columns beginning on the right side of the puzzle, based on inferences made by the fixation map. The grid squares do seem to cause subjects given no explicit scan path directions to follow a path of some sort, possibly due to human nature/inherent societal biases relating to following within the lines of a perceived boundary.

Figure 10 depicts a fixation map compiled from Group 1 which depicts the ideal scan path method dictated to Group 1 prior to beginning the experiment. It is clear to see that the user follows the grid squares as if reading down the page scanning for Waldo as

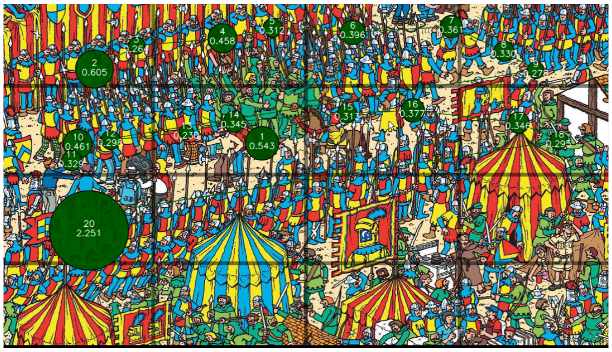

Figure 10: Fixation Map 2

they progressed further down the image. Fixation maps plotted the subjects's fixations as they scan through the established areas of interest which are indicated by the grid squares.

4.4 Areas to Improve

Due to the relatively large areas of interest, with all puzzles containing 12 grid squares(A-P respectively), it was challenging to collect significant statistical data regarding subjects' individual scan paths in comparison to each other. Upon further discussion, the inclusion of a larger number of areas of interests, in this situation by doubling the amount of column grid lines, would allow for more areas to be used in the Levenstein formula for comparison. In figure 10, the original image viewed by users is labeled A-P to indicate all areas of interest present and taken into account.

1 AP.jpg

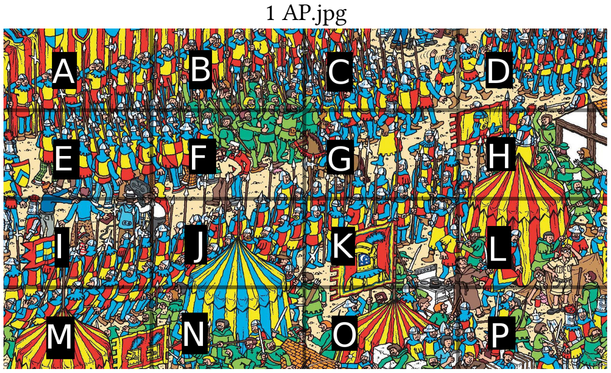

Figure 11: Original Puzzle w/ AOIs

Figure 12 represents a modified potential image which would have mitigated the likelihood of statistical insignificance by providing an A-AN grid square representation. The higher amount of areas of interest to be used in the calculation of the Levenstein distance would allow for a more refined answer to how close subjects in each group were coming to eachother in regards to their scan path methods.

An additional area of concern in compiling significant statistical data was the overall subject amount in the two groups. Our study

1 AAN.jpg

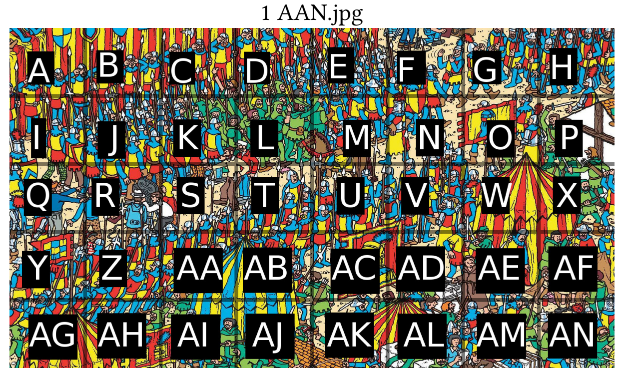

Figure 12: Revised puzzle w/ AOIs

compiled the results from 11 subjects participating in the experiment. 11 subjects is a small sample size and thus would not provide an accurate comprehensive representation of average scan paths for comparison due to the large influence of outliers over the study. One individual choosing to ignore the designated scan path instructions in Group 1 or any individual clicking the [spacebar] without actually finding Waldo in the picture could heavily skew data due to the limited amount of data points being taken in.

REFERENCES

Francesco Di Nocera, Michela Terenzi, and Marco Camilli. 2006. *Another look at scanpath: distance to nearest neighbour as a measure of mental workload.* 295–303.

Nicholas L. Port, Jane Trimberger, Steve Hitzeman, Bryan Redick, and Stephen Beckerman. 2016. Micro and regular saccades across the lifespan during a visual search of "Where's Waldo" puzzles. *Vision Research* 118 (2016), 144–157. https://doi.org/10.1016/j.visres.2015.05.013

Anthony Santella and Doug Decarlo. 2004. Robust clustering of eye movement recordings for quantification of visual interest. *Proceedings of the Eye tracking research applications symposium on Eye tracking research applications - ETRA2004* (2004). https://doi.org/10.1145/968363.968368