Visual Analysis of Perceptual and Cognitive Processes

Michael Raschke¹, Tanja Blascheck¹, Marianne Richter², Tanja Agapkin¹ and Thomas Ertl¹

1Institute for Visualization and Interactive Systems, University of Stuttgart, Universitätsstrasse 38, 70569 Stuttgart, Germany
2EXC Simulation Technology, University of Stuttgart, Pfaffenwaldring 5a, 70569 Stuttgart, Germany

Keywords: Visual Analysis, Perception Processes, Cognition Processes, Visualization.

Abstract: The success of visualization techniques depends on their support of perceptual and cognitive processes to perceive the graphically represented information. Apart from measuring accuracy rates of correctly given answers and completion times in user studies, eye tracking experiments provide an additional technique to analyze perceptual and cognitive processes of visual tasks. This paper presents an interdisciplinary approach for studying structures of scan paths by visual means. We propose to annotate graphical elements with semantic information. This annotation allows us to analyze the fixation sequences on these annotated graphical elements with respect to reading processes, visual search strategies, and visual reasoning.

1 INTRODUCTION

A key factor for the success of a visualization technique is the efficiency of how users perceive information using this visualization technique. Apart from measuring completion times and recording accuracy rates of correctly given answers during task performances, eye tracking experiments provide an additional technique to analyze how the attention of a visualization observer is changing on a presented stimulus. Therefore, gazes on the stimulus are recorded and afterwards aggregated to fixations and saccades for measuring which areas on the stimulus have been focused on. These scan paths show fixations during a moment-by-moment processing of a visual task (Koerner, 2011), (Rayner, 1998) and therefore are most suitable to identify or differentiate between comprehension processes (Grant and Spivey, 2002).

To better study cognitive and perceptual processes during a visual task, we propose to annotate task relevant areas (areas of interest) on the stimulus with information about the semantic meaning of graphical elements inside these areas. As an appropriate visualization technique to study both declarative and procedural cognitive processes we are using the parallel scan path visualization technique (Raschke et al., 2012). By mapping semantic information to the area of interest axes in the parallel scan path visualization our approach allows us to study mental processes such as visual search, visual reasoning as well as cross-checking behavior.

2 RELATED WORK

Eye tracking is used in a wide field of user experiments in psychology, cognitive science, marketing, visualization, human-computer, and human-machine interaction. This paper will focus on the analysis of eye tracking experiments in the visualization research domain. One example of an eye tracking study in visualization research is the comparison of different types of graph layouts such as radial, orthogonal and traditional conducted by Burch et al. (Burch et al., 2011). We will later use this eye tracking experiment for a demonstration of our approach. Another example is the eye tracking experiment by Huang et al. Their results show that graphs are read following a geodesic-path tendency. As a result, links which go towards the target node are more likely to be searched first (Huang et al., 2009). Kim et al. investigate the influence of peripheral vision during the perception of visualizations (Kim et al., 2012).

Different approaches have been developed to analyze eye tracking data. One approach is to analyze eye tracking results with statistical methods to find correlations between different dimensions in the recorded eye tracking data (Holmqvist et al., 2011). These correlations can later be identified with metrics such as cognitive workload. The statistical analysis can be supported by techniques from visual analytics as discussed by Andrienko et al. (Andrienko et al., 2012). If the definition of areas of interest is possible, string editing algorithms can be used to find
patterns in eye movements between these areas of interest (Duchowski et al., 2010).

Classical visualization approaches for eye tracking data are heat map (Wooding, 2002) and scan path visualizations (Spakov and Miniotas, 2007). Different approaches of eye tracking data visualizations have been developed to satisfy various kinds of experimental set-ups and research questions. One implementation of an eye tracking data visualization is eSeeTrack by Tsang et al. that combines a time line and a tree-structured visual representation to extend current eye-tracking visualizations by extracting patterns of sequential gaze orderings (Tsang et al., 2010). A method, which can be classified in-between visual and statistical analysis techniques, is presented by Ponsoda et al. The authors use transition matrices for analyzing eye movement recordings during free viewing (Ponsoda et al., 1995). Opportunities and challenges for developing new visualization techniques for eye tracking data analysis have been discussed by Ramloll et al. (Ramloll et al., 2004).

Although many approaches have been developed to analyze eye movements, the interpretation of one or more scan paths with respect to focused areas on the stimulus and relating cognitive and perceptual processes is complicated and there is no common way of interpretation. To fill this gap between statistical results or visualizations of eye-tracking data and their interpretation, Convery et al. present a descriptive model of visual scanning (Convery et al., 2011). The authors have developed a method to describe an idealized scanning of visual representations. Where the assumption of an idealized scan path is useful for designing visualizations or simulation models of cognition processes, typical scan paths from eye tracking experiments do not show this ideal behavior. Therefore, common patterns of eye movements have to be extracted in order to formulate ideal scan paths. Additionally, these findings can also be used to set up models for simulating perceptual and cognitive processes of visual understanding (Pinker, 1990; Lohse, 1993; Shah, 1997). In our work we propose to annotate areas on a visualization with semantic information to better analyze the scan paths on the stimulus with respect to procedural and declarative knowledge processes. Besides using only simple word lists of identifiers for graphical elements our approach goes further and allows us to annotate areas on visualizations with more complex information from ontological models about graphical knowledge as presented by Pinker et al. in his visual search modeling algorithms.

Finally, an adequate visualization technique is missing, that both graphically shows fixation sequences on annotated areas on a stimulus and allows to study cognitive and perceptual processes. In this work we will use the parallel scan path visualization technique that maps gaze durations and fixations inside areas of interest to vertical axes in one diagram (Raschke et al., 2012). Areas of interest are defined on the stimulus and are mapped to vertical coordinate axes in the parallel scan path plot. The leftmost axis indicates time, starting from the bottom of the diagram with the start time of the eye tracking measurement. The orientation of the parallel scan path visualization is arbitrary. In the following we use a vertical time axis from bottom (start of the scan path recording) to top (end of the scan path recording) as introduced in the original work. The horizontal axis displays all selected areas of interest as independent values. Saccades between areas of interest are indicated with dashed lines. Ascending lines indicate fixations outside given areas of interest. Using the parallel scan path visualization for our approach, areas of interest axes in the parallel scan path visualization will be identified with semantically annotated areas on the stimulus.

3 MOTIVATION AND CONCEPT

Our concept is based on three assumptions according to findings during eye tracking user studies and to the literature:

1. To perform a given task using a visualization, task relevant areas on the stimulus are focused on in serial order. We assume that the sequence of fixations on these areas represent perceptual and cognitive processes (Koerner, 2011), (Rayner, 1998), (Grant and Spivey, 2002).

2. For the analysis of perception and knowledge processes we are dividing up knowledge into declarative knowledge and procedural knowledge according to standard classification of psychology and cognitive sciences (Anderson et al., 2004a).

3. Based on this definition we classify an information seeking method on a stimulus using two kinds of visual tasks: 1) Visual search and reading processes without drawing a conclusion - the semantic meaning of graphical elements and declarative knowledge is of interest - (cf. Figure 2a+d) and 2) tasks where the observer of a visualization has to draw a conclusion - the syntactical relations of graphical elements are of additional interest. Declarative knowledge as well as procedural knowledge play an important role (cf. Figure 2b+c+e+f).
Based on these three assumptions this paper contributes an analysis concept with three steps (cf. Figure 1):

1. Task relevant areas on the stimulus have to be identified.
2. These areas have to be semantically annotated with respect to the meaning of graphical elements inside these areas. This annotation can either be done classically with identifiers from word lists or with links to elements from an ontology that describes graphical elements and their relations for a given visualization concept.
3. Finally, saccades between these areas as well as fixation durations and summarized fixation durations outside these areas are visually analyzed.

In the following we will describe these three steps in more detail.

### 3.1 Identification of Task Relevant Areas

In most cases, areas with a high density of fixations of one or more participants indicate graphical elements which are relevant to perform a given task. These areas can simply be identified with heat maps (cf. Figure 1, middle). After relevant graphical elements of the stimulus are identified, areas of interest are defined which contain these graphical elements.

### 3.2 Annotation of Task Relevant Areas

Areas of interest are annotated with information about their semantic meaning (cf. Figure 1, top). The semantic information could be the type of the graphical element, its name, or a reference to another resource that contains further information about this graphical element depending on the research question. An adequate naming of the graphical elements is crucial for the syntactic and semantic analysis of the scan paths (Perini, 2005), (Pinker, 1990). If an ontology or a model of visualization knowledge is available, additional information such as the relationship of focused graphical elements can be used during the analysis.

### 3.3 Visual Analysis

Figures 2a-c show three exemplary parallel scan path visualizations for a) visual search and reading processes, b) visual reasoning, and c) mixed processes representing different types of perceptual and cognitive processes.

Figure 2a shows an idealized scan path for reading the position of a given point in a Cartesian coordinate system (Figure 2d shows the stimulus). An idealized fixation sequence could be the following: focus on the data point (1), on the x-axis to read the corresponding x-value (2), on the data point again (3), followed by a reading of the vertical position on the y-axis (4). At this moment, all task relevant graphical elements have been focused on at least one time to answer the task. Figure 2b shows the idealized scan path for the task “Why is the road wet?” (cf. Figure 2e) that could be to first focus on the road (1), then on the puddle (2), on the cloud (3), on the sun (4), and on the fire hydrant (5). Finally, the attention could move to the puddle again (6). This fixation sequence shows single steps of logical reasoning: The solution is to verify, if there is rain coming from the cloud (false), or if the fire hydrant is sprinkling (true), or if both is true (false). Finally, Figure 2c shows an ideal scan path for the task “Is the lowest temperature of all given temperatures in...”
southern France lower than the given temperature in northern France?” (cf. Figure 2f). The idealized scan path shows a first fixation on the $-10^\circ$C label in the south east of France (1), then on the $-4^\circ$C label in the south (2). The next step is to move the attention to the northern part of France to read the $0^\circ$C label (3). By memorizing the lower temperature from the south and comparing it with the one in the north the answer to the question can be given. This fixation sequence indicates that declarative knowledge has to be memorized and processed. Times for reading the temperature values and memorizing them as well as restoring them again from memory can be directly seen in the visualization.

4 CASE STUDIES

In this section we will demonstrate how the presented approach can be applied to four case studies using data from real eye tracking experiments. Due to our work aiming at demonstrating the presented approach, we will concentrate on showing fixation sequences of some exemplary stimuli and participants. The first three case studies are based on a pilot study with ten participants. With this pilot study we tested the experimental design and set-up for a future eye tracking experiment with forty participants to study perception and cognition processes in different visualization scenarios. For every scenario, we have designed a separate experiment with a within-subject study design. At the beginning of every scenario, participants were given a short tutorial with instructions to the following task. Afterwards the actual experiment was started. Before every measurement in a scenario, we calibrated the eye tracker system Tobii T60 XL with the calibration algorithm implemented in Tobii Studio 2.7. To further confirm this calibration of all stimuli during a run of a task, we showed a black cross in the center of the screen between every stimulus. Subjects had to focus on this marker. To start a task in a scenario, participants had to press a key on the keyboard. Then, the stimulus was presented. Subjects had to press a key on the keyboard again, if they were sure to answer the given question. The stimulus disappeared and the black cross was presented again. Then, the participants had to give their answer. The fourth case study is based on a an experiment conducted by Burch et al. Due to space limitation we will refer to the original work (Burch et al., 2011) for a detailed description of the experiment.

4.1 Fixation Loops during a Visual Task

Twenty-four different bar charts were used in the first task block and have synthetically been generated with randomized values for each category. To study different visual search strategies we have used four types of labels of the x-axis: alphabetically, non-alphabetically, with increasing numbers, and with randomized numbers. Participants had to perform two
classes of tasks: 1) read the value of a bar for a given category and 2) compare the values of two bars of two given categories with each other and state the value of the higher one. In this sub section, we will show results of an analysis of one stimulus and recorded fixations of one participant who has performed a task of the first task class. A demonstration of an analysis for a result from the second task class will be given in the next sub section. Figure 3a shows a stimulus presenting a bar chart. In this example we have used a numerical labeling of the bar categories with continuous, increasing, natural numbers. The participant had to read the value of the bar with category “7”. Figure 3b shows a heat map over all ten participants for this task. The annotated areas on the stimulus are presented in c). We have defined four areas of interest: one for the label indicating bar category “7”, one at the top point of the bar and one at the corresponding value label on the y-axis. A fourth area of interest was used to check, if the participant had focused on the black cross at the beginning of the task. Figure 3d shows the parallel scan path visualization of the fixation sequence for one participant. The completion time for this task and this participant was approximately 4.5 seconds. First, the participant focused on the black cross, then searched for the bar label. Afterwards, his attention alternately moved from the bar top to the bar value and back, followed by a short check of the correct category. At this moment, he has focused on all task relevant graphical elements on the stimulus and thus could have answered the task. However, he did not state the answer, but performed another reading of the bar value to cross-check his answer. This second loop of fixations can clearly be seen in the fixation plot starting at 3 seconds.

4.2 Visual Search

Figure 4a shows one stimulus for the second class of tasks of the first task block with a heat map of all participants (b), the areas of interest with the annotations (c), and the parallel scan path plot of one participant from the experiment (d). The task was to compare the values of the bars of category “H” and category “B” and to give the value of the higher one. The visualization shows, that the participant at first searched for the labels of category “H” and “B” starting from the label “I”. This process is shown by an ascending plot line from left to right in the parallel scan path plot starting from the bottom with two fixations on label “H”. The completion time for this visual search process is approximately 4.5 seconds. Once, the location of the bar for the categories “H” and “B” were found, the participant compared the height of the top points of the bar. This is graphically represented by an alternating attention between “top point bar B” and “top point bar H”. Finally, the participant read the bar value for bar “H” by moving his attention to the area of interest on the y-axis with the bar values.

4.3 Visual Analysis of Reasoning Processes

In a second task block we studied reasoning processes using Venn diagrams. We presented ten Venn diagrams with different numbers of elements to study different levels of logical complexity. One stimulus is shown in Figure 5a. In this task block participants had to answer logical questions such as “How many people have a high school graduation (in German “Abitur”) or got their dream job (in German “Traumjob”) solely by absolving an apprenticeship (in German “Ausbildung”)?” see Figure 5. These questions represent logical reasoning processes based on logical AND, OR, NAND, NOR, and XOR relations. A heat map over all participants’ fixations is shown in Figure 5b, annotated areas of interest on the stimulus are presented in c), and the parallel scan path of one participant in d). In total, we have defined seven areas of interest. Three areas of interest represent the three labels for the plotted categories on the right side of the stimulus, and four areas of interest show task relevant information such as the number of people with high school graduation, dream jobs, and apprenticeship (numbers inside ellipses).

The parallel scan path visualization shows, that in the first two seconds of the experiment, the participant focused on the labels on the right side of the diagram. Once these three categories were identified and mentally memorized the participant alternatively moved his attention between the areas of interest “Number of people with high school graduation and dream job”, “Number of people with a dream job”, and “Number of people with a dream job and apprenticeships”. This second phase represents the logical reasoning process. In the last third of the experiment the participant checked the category labels again and finally stated the answer.

4.4 Visual Analysis of Cross-Checking in a Complex Visualization Scenario

Burch et al. evaluated traditional, orthogonal, and radial tree diagrams and compared accuracy, completion times, and exploration behavior to test six hypotheses for the task: “Find the least common ancestor from a set of marked leaf nodes.” (Burch
Figure 3: We have studied how participants read values of a bar of a given category in our pilot study. This example shows the task “Please, read the value of bar ‘7’.”. The stimulus for this task is shown in a), a heat map of all participants in b), the annotated areas of interest on the stimulus in c), and finally, the parallel scan path visualization of one participant in d). The analysis shows, that this participant performed one “fixation loop” in the second half of the experiment before stating the answer.

Figure 4: This example shows results of the task “Compare the value of category ‘H’ with the value of category ‘B’ and state the value of the higher one.”. In d) the ascending parallel scan path from left to right in the first half of the experiment shows the fixations of the visual search process for finding the two labels “H” and “B” followed by comparing the bar heights.

Figure 5: In a second task block we have studied reasoning processes using Venn diagrams. These figures show results of the analysis of the task “How many people have a high school graduation (in German “Abitur”) or got their dream job (in German “Traumjob”) solely by absolving an apprenticeship (in German “Ausbildung”)?”. For a detail description of the analysis we refer to the text in section “Case Studies”.

Figure 6: These four figures show results of the fourth case study where we demonstrate how the presented approach can be used to analyze processes of cross-checking during performing a given task, as discussed by Burch et al.
et al., 2011). For a demonstration of our approach we will concentrate on the second part of their fifth hypothesis “[...] cross-checking is used more frequently in the radial layouts.”. In the case of their experiment “cross-checking” means, that participants find a solution for the task and verify this solution several times. In a follow up work Burch et al. proved this hypothesis using visual analytics techniques (Burch et al., 2013). We can verify their results using our approach. Figure 6a shows a radial tree diagram stimulus from the experiment together with the heat map of all participants (Figure 6b). Relevant graphical elements are the solutions (on the left, below the blue point which indicates the root node of the tree) given by the participants and three marked leaf nodes 1, 2, and 3 (marked by three red arrows in the original experiment), see Figure 6c. The parallel scan path visualization is presented in Figure 6d for the time interval between 0s and 24s. It shows three selected participants (blue, red, and green) from the experiment. The fixation sequence of the green subject shows, that he at first focused on leaf node 3, then on the solution, and then on leaf node 3 again. Next, his attention moved to the other two leaf nodes and the solution. He did not perform a cross-checking and answered the question after approximately 11s. At this moment he has focused on all relevant semantic graphical elements for this task. This eye movement pattern can be interpreted as an ideal scan path. The blue subject performed cross-checking after having focused on all relevant graphical elements during approximately the first 9s of the measurement. The red participant first shifted his attention between leaf node 3 and 1, then between leaf node 2 and 1.

5 LIMITATIONS AND IMPROVEMENTS FOR FUTURE WORK

A crucial question for analyzing eye tracking results is whether the recorded eye movements reflect mental processes? There has been a long discussion going on about this question and it is not yet clearly answered. Some works argue that mental processes and eye movements are not correlating (e.g. (Anderson et al., 2004b)), others argue that eye movements reflect mental activities which they call the “Eye-Mind Hypothesis” (e.g. (Just and Carpenter, 1980) and (Rayner, 1998)). Results of our eye tracking experiments show that fixations and sequences of fixations on a stimulus correlate with cognitive processes. We think, that this question cannot be answered with a definite “yes” or “no”. It depends on the complexity of the visualization, the visual task and the required mental processes. But we believe, that the presented approach can be helpful for discussing this question.

The second step of our analysis approach is to annotate graphical elements with information about their semantic meaning. In this step an adequate naming of the graphical elements is extremely important for a goal oriented analysis and correct interpretation of the fixation sequences. The question how an adequate naming can be found is an open question and has to be investigated in future work. In our work, we have proposed to either use identifier from word lists according to standard approaches in eye tracking research or to link areas of interest to elements of an ontology. This ontology can then be used during the eye tracking analysis and for the set-up of a cognitive model about declarative knowledge of graphical elements. Another issue about the annotation concept is, that the parallel scan path visualization technique only allows non-overlapping areas of interest. Thus, we cannot annotate graphical elements which are overlapping. This overlapping of several semantic annotation areas additionally leads to further questions about knowledge modeling as in our case it is not clear which semantic meaning of an area of interest on the stimulus is relevant.

During the analysis of the pilot study results we found limitations of the parallel scan path visualization technique. The fourth case study shows, that the higher the number of visualized participants is, the more difficult it becomes to find common eye movement patterns due to visual clutter of the parallel scan path visualization technique. To reduce this visual clutter, we propose to implement filtering, aggregation, and contextual focusing algorithms into the visualization tool. Additional pattern recognition algorithms such as Levenshtein distance can be used to pre-process the eye tracking data. Thus, only fixation sequences will be shown which are relevant for a certain research question. If two participants strongly vary in their completion times, their completion times could be scaled in order to plot them on the same axis length. This would make it possible to better compare different fixation time stamps with each other and thus, compare different visual search strategies. Another limitation is caused by the limited number of vertical axes which can be presented in the visualization. Additionally, an optimal ordering of these axes can positively influence the analysis. An automatic ordering could be done using information from transition matrices.
6 CONCLUSION

This paper contributes an approach to visually analyze eye tracking data with respect to perceptual and cognitive processes based on a semantic annotation of graphical elements inside the visualization. To be able to use the concept in practice we have implemented a prototype and have presented four case studies covering reading processes, visual search strategies, visual reasoning and cross-checking. Finally, we have discussed limitations of this approach and presented possible improvements for future work. Besides tackling limitations in future work, we are planning to use the analysis concept during the development of cognitive models for visual search and visual reasoning.

REFERENCES


