A Comparison of a Transition-based and a Sequence-based Analysis of AOI Transition Sequences

Chia-Kai Yang
University of Zurich
Zurich, Switzerland
yang@ifi.uzh.ch

Tanja Blascheck
University of Stuttgart
Stuttgart, Germany
research@blascheck.eu

Chat Wacharamanotham
University of Zurich
Zurich, Switzerland
chat@ifi.uzh.ch

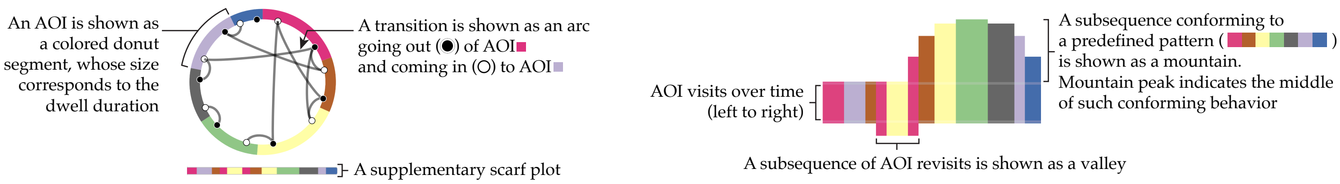

Figure 1: A sequence of Area of Interest (AOI) visits visualized with two methods. Left: the Radial Transition Graph (RTG) emphasizes transitions between AOIs. Right: Alpscarf visually emphasizes conformity to a subsequence.

ABSTRACT

Several visual analytics (VA) systems are used for analyzing eye-tracking data because they synergize human-in-the-loop exploration with speed and accuracy of the computer. In the VA systems, the choices of visualization techniques could afford discovering certain types of insights while hindering others. Understanding these affordances and hindrances is essential to design effective VA systems. In this paper, we focus on two approaches for visualizing AOI transitions: the transition-based approach (exemplified by the radial transition graph, RTG) and the sequence-based approach (exemplified by the Alpscarf). We captured the insights generated by two analysts who individually use each visualization technique on the same dataset. Based on the results, we identify four phases of analytic activities and discuss opportunities that the two visualization approaches can complement each other. We point out design implications for VA systems that combine these visualization approaches.

CCS CONCEPTS

• **Human-centered computing** → *Visualization design and evaluation methods*; *Visual analytics*; *Visualization techniques*.

KEYWORDS

Eye tracking, AOI, visualization, visual analysis, pattern detection

ACM Reference Format:

Chia-Kai Yang, Tanja Blascheck, and Chat Wacharamanotham. 2020. A Comparison of a Transition-based and a Sequence-based Analysis of AOI Transition Sequences. In *Symposium on Eye Tracking Research and Applications (ETRA '20 Short Papers), June 2–5, 2020, Stuttgart, Germany*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3379156.3391349

1 INTRODUCTION

Eye tracking studies can yield insights into patterns of viewing behavior. Prior eye tracking studies infer these patterns by analyzing transition sequences among areas of interest (AOIs) [Goldberg et al. 2002; Netzel et al. 2017; Yang 2012]. Sensemaking during these analyses requires contextual knowledge about the stimuli and tasks and, therefore, need a human analyst in the loop. To leverage innate human pattern recognition abilities, researchers often apply visualization techniques for analyzing AOI transition sequences [Blascheck et al. 2016a].

However, AOI transitions are challenging to visualize. Transition graphs and transition matrices can show transitions between pairs of AOIs, but using them to analyze sequences of transitions longer than two become difficult [Holmqvist et al. 2011; Olson et al. 1994]. Sequence-based visualizations such as scarf plots [Richardson and Dale 2005] can naturally show long sequences of transitions, but discovering occurrences of subsequences across participants is more tedious than transition-based methods. Some visual analytics (VA) systems, e.g., RTGCT [Blascheck et al. 2017], integrate both transition-based and sequence-based approaches, but such integration can be asymmetric: more functions are available for one visualization technique than the other. Therefore, knowledge on how these two visualization techniques afford types of insights and analysis strategies is essential in designing VA systems that effectively combine both techniques.

Our goal is to empower eye tracking researchers to discover insights about AOI transition sequences in their data. In this paper, we present a study that compares two analyses of the same dataset: one with a transition-based radial transition graph (RTG, Figure 1, left), and the other with a sequence-based Alpscarf (Figure 1, right). Based on the empirical findings about the insights and analysis strategies, we discuss important design implications for integrating transition- and sequence-based techniques for eye tracking analysis.

ETRA '20 Short Papers, June 2–5, 2020, Stuttgart, Germany

ACM ISBN 978-1-4503-7134-6/20/06…$15.00
https://doi.org/10.1145/3379156.3391349

2 TERMINOLOGY

Kurzhals et al. [2017] indicates that the role of an AOI transition analysis is to reveal the relations between AOIs or participants. These relationships are revealed though the presences and absences of transition patterns. Because a reference to a pattern can be interpreted in several ways (e.g., a regular expression pattern, a behavioral pattern that occurs across participants), we define the following terminology for this paper. We investigate *subsequences* within AOI sequences and higher-level patterns participants exhibit. First, we assume that the eye movement data of each individual participant is a sequence of AOI visits, in which each visit is an element of the sequence. Transitions between two AOIs connect these visits. Within such an AOI sequence, we can detect subsequences. A subsequence does not change the order of the individual elements but rather deletes zero or more elements. For example, if we take the AOI sequence *ABCABDA* we can extract the subsequence *ABCD*. Such subsequences can occur in many locations of an AOI sequence. Subsequence-searches can be easily automated.

We also define the term *pattern* as a subsequence of AOIs that a human analysts deem to be meaningful. For example, if we assume that our stimulus is a text with four lines defined as AOIs A, B, C, D. The pattern *ABCD* is a subsequence that an analyst would call a *linear reading pattern*. Once a human analyst define a pattern, it can be turned into a subsequence and expressed, e.g., in a regular expression, for the computer to search.

Pattern can be originated top-down or bottom-up. In a *top-down approach*, the analyst derives a pattern from the contextual knowledge about the stimulus or the task before searching its occurrences in the data. In a *bottom-up approach*, common subsequences are mined from the data, e.g., by finding common n-grams subsequences within the AOI sequences of all participants. In both approaches, analysts are required to make sense of the detected patterns, and to decide if the derived knowledge is adequately insightful for their research questions. Therefore, pattern-definition and subsequence-searching go hand-in-hand in an iterative process that combines a creative sensemaking by the human analyst with the efficiency of searches by the computer.

3 RELATED WORK

Automatic sequence mining algorithms can be applied to reveal AOI transition patterns [Ayres et al. 2002; Goldberg and Helfman 2010; West et al. 2006]. However, Eraslan et al. [2016] indicates that algorithms tend to detect only short patterns that are unhelpful for understanding viewing behavior. Several VA systems allows users to search by either regular expressions [e.g., West et al. 2006], a specification on a graphical user interface [e.g., sequence tool in Wu and Munzner 2015], or both [e.g., pattern editor in *VA*^{2}, Blascheck et al. 2016a]. Although *VA*^{2} allows powerful exact and fuzzy search, it visualizes one AOI per row, limiting the number of distinct AOIs that can be visually analyzed on the screen.

Scarf plots—a sequence-based visualization technique—provide a more compact presentation by encoding AOIs with colors [Richardson and Dale 2005]. Alpscarf extends the scarf plots by adding the vertical length to indicate conformity to a search pattern [Yang and Wacharamanotham 2018] (Figure 1, right).

Transition-based methods such as transition graphs and transition matrices have also been improved upon. For example, AOI hierarchies add trees onto the margins of the transition matrix as well as laying out of transition graphs as a tree [Blascheck et al. 2016b]. Blascheck et al. [2017]'s RTG technique visualizes transition graph with a radial layout, enabling multiple graphs to be overlaid to facilitate visual comparisons. Their system also allows multiple graphs to be arranged in a grid with their similarities and differences highlighted.

Our work extends the understanding of AOI transition analyses by providing empirical observations about how such an analysis unfolds in a sequence-based versus a transition-based technique. We choose Alpscarf as a representative for sequence-based visualization methods and RTG for transition-based.

4 ANALYSIS PROCEDURE

We aim to compare analysis strategies that transition-based versus sequence-based visualization techniques afford. Therefore, we collect data on analytic activities and insights from two analysts applying two different visualization techniques to inspect the same dataset. We *re-analyzed* the data from a study comparing how novices and non-novices read natural language text (NT) and source code (SC) (21 participants × 2 conditions ∈ {NT, SC} × 4 stimuli per condition = 168 AOI visit sequences) [Blascheck and Sharif 2019]. The AOIs correspond to each line of text or source code. The re-analysis goal is to identify occurrences of linear reading behavior[1] , which was the analysis focus in [Blascheck and Sharif 2019].

Because the original study was already visually analyzed with the RTG, a transition-based technique, in the first round analysis, one of the authors independently analyzed the data with Alpscarf[2] [Yang and Wacharamanotham 2018], a sequence-based technique. Then the Alpscarf-analyst discussed with the RTG-analyst (one of the authors of [Blascheck and Sharif 2019]) about their analytic strategies and insights. In particular, they were interested in the differences of the analysis results and why these differences occurred. Afterward, both analysts conducted further analysis using their individual visualization technique. In total, three further analysis iterations with RTG and four with Alpscarf were conducted. During this process, both analysts capture their individual and joint insights in field notes. In the end, both analysts went through them and extracted important themes.

5 RESULTS & DESIGN IMPLICATIONS

We present results of analysis process when analyzing the eye movement data using both transition-based and sequence-based techniques. We abstracted the analysis process into four phases. In each phase, each visualization technique showed its affordances and limitations. We discuss each of these and draw design implications for a VA system that aims to combine both techniques.

5.1 Phase 1: Operationalizing a Pattern

The first round of the analysis used a top-down approach because the focus on linear reading pattern was predetermined. In Alpscarf,

[1] Left to right, top to bottom for natural language text. In the source-code stimuli with multiple methods, the linear reading behavior was defined per-method.
[2] In duration-focus mode, normalized view.

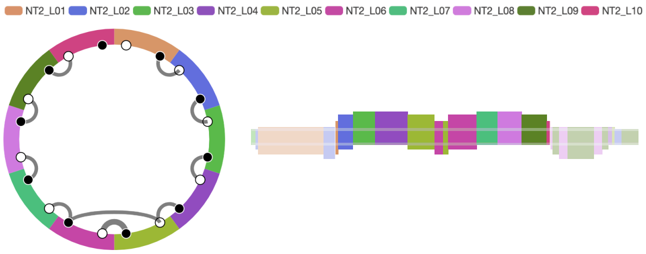

Figure 2: In the middle of linear reading sequence (shown in saturated colors in Alpscarf), the re-read (from L06 to L05) was mentally filtered out in RTG analysis.

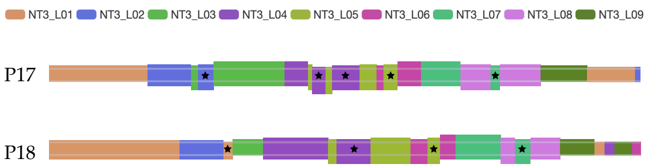

Figure 3: The revisits to previously-read neighbor AOIs (marked with ★) occurred in the middle of linear reading of P17 and P18 (incorrectly classified by RTG), resulting in no obvious mountains in Alpscarf.

this pattern was then operationalized as a linear sequence of AOIs as shown on top of Figure 2. For each of the matching subsequences, Alpscarf visualizes mountains centering at the middle of the match. In contrast, in RTG, patterns are operationalized mentally as occurrences of transition arcs. The linear reading pattern appears as small transition arcs that connect adjacent ring segments (AOIs) as shown in Figure 2. This mental-operationalization enables visual search on RTG to be more flexible. For example, the analysts can be lenient by allowing some transitions to be absent. However, the flexibility prevents the patterns from being automatically detected.

Design implication: *VA systems should accommodate both explicit and mental operationalization of search patterns.*

5.2 Phase 2: Adjusting Subsequence Definition

The pattern operationalization assumed that reading all AOIs (threshold = 100%) in the specified order is considered a linear reading behavior. This was the assumption the analyst applied for RTG analysis. However, with Alpscarf, this strict assumption would result in extremely low number of participants exhibiting linear reading behavior, e.g., 1 participant (P05) for NT2, compared with RTG analysis (19 participants for NT2).

If we further inspect other participants, we can observe two things: not fixating on the last AOI and re-reading of AOIs. For example, in NT2, the last AOI was a short line and, therefore, sometimes not focused by participants at all. In RTG, the linear reading pattern was accepted when the last AOI was not read. Such relaxation was implicitly applied by the RTG-analyst. The absence for the visit on one AOI is easily to detect in RTG, if an AOI has no outgoing nor incoming transition arc. Alpscarf, on the other hand, requires an explicit adjustment to accommodate this behavior. In Alpscarf, we lowered the threshold to 90% of AOIs that participants

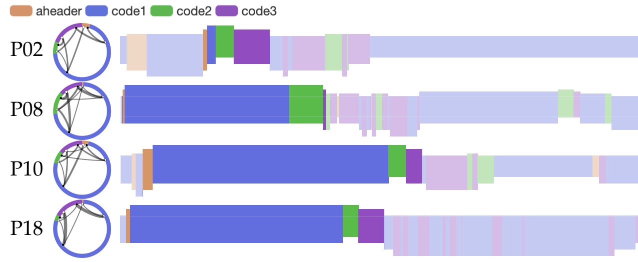

Figure 4: The subsequences of linear reading behavior (shown in saturated colors in Alpscarf) for P02, P08, P10, and P18 capture less than 50% of the total inspection time on stimulus SC5.

have to focus on to be considered as linear reading. This resulted in 3 participants (P01, P05, P23) in NT2 who read linearly.

In addition, we know that people may re-read a line if they did not understand it. If the re-read occurs only *once*, we still consider the subsequences as linear reading. In RTG analysis, such a re-read could be seen by a backward transition to the previous AOI and was mentally filtered out by the RTG-analyst (see Figure 2). In an extreme case for Alpscarf analysis, such a re-read may occur exactly in the middle of the subsequence (50% of the AOIs); therefore, for Alpscarf we define that for a subsequence to be counted as linear reading, at least 50% of the AOIs have to be read. This leads to 11 participants in Alpscarf showing linear reading behavior in NT2 (P01, P02, P03, P05, P06, P07, P10, P12, P17, P19, P23).

Nevertheless, we found this once-only re-reading tolerance was not strictly applied for some RTG analyses, resulting in several participants being incorrectly classified as linear reading even though they had multiple re-reads. The incorrectly classified participants were corrected by Alpscarf analyses results, with gaps between the mountains (see Figure 3).

Furthermore, some of the subsequences identified by Alpscarf only occupied a small portion of the total reading time, and these subsequences were not correctly detected during the RTG analysis. For example, in the analysis with SC5 (4 AOIs, function level), 9 participants exhibiting linear reading behavior in Alpscarf analysis, which are more than that in RTG analysis (5 participants). For the 4 participants two visualizations disagreed with each other (P02, P08, P10, P18), we found the linear reading subsequences identified by Alpscarf occupied a relatively small portion (≤50%) of the total reading time (see Figure 4). This shows the strength of Alpscarf in depicting short subsequences, whereas visually detecting short subsequences with RTG is mentally demanding and error-prone.

Design implication: *VA systems should accommodate to visually ignore small AOIs and to ignore re-reads to one previous AOI.*

5.3 Phase 3: Refining Subsequence Parameters

The Alpscarf analysis of NT2 results in 11 participants exhibiting linear reading behavior. This result is a subset of the RTG results (19 participants). However, for the 8 participants that RTG considered linear reading but Alpscarf did not (partly shown in Figure 5), they had short AOI visits to other AOIs disrupting the linear sequence. This could be considered as re-readings, or an artefact of

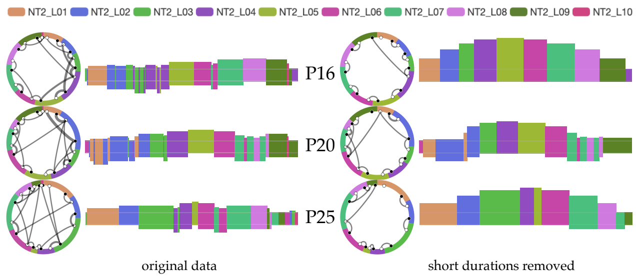

Figure 5: Comparison of RTG and Alpscarf visualizations for NT2. Left: without any removal of AOI visit, Alpscarf disagreed with RTG on 8 participants that RTG analysis reported as linear reading. The results of three participants are shown side by side in the figure. Right: in Alpscarf, linear reading patterns were revealed as mountains after removing short AOI durations (≤300 ms, ≤250 ms for P25)

the AOI definition, or just a calibration issue. It is reasonable that if the visits are short, they are not necessarily intentional reads and, therefore, in the RTG analysis these short AOI visits were mentally disregarded.

To filter out short AOI visits in Alpscarf, we defined a threshold of ≤300 ms (based on [Holmqvist et al. 2011, p. 377]). The application of this threshold results in a match between Alpscarf analysis and RTG analysis, except for P25. P25 read faster (9.2 s) than others (median = 25.4 s) and exhibited much shorter AOI visits. With a threshold of 300 ms, one AOI (L05) was discarded despite it belonged to the linear sequence. As a result, if we lowered the threshold to ≤250 ms, the linear reading subsequence for P25 would be preserved.

As shown in Figure 5, the transition arcs that connect adjacent ring segments (AOIs) are presented in RTG (both before and after the removal of short durations), while other transitions are mentally removed in RTG analysis when detecting linear reading subsequences. As a result, a transition-based inspection using RTG often leads to more participants depicting linear reading than Alpscarf. However, the analysis above showed that in many cases, a refined definition of subsequences with Alpscarf confirmed these patterns. Therefore, a VA system should allow an analyst to set a threshold to remove short AOI visits on demand, which are artefacts of the data collection, while accommodating different reading speeds of participants.

Design implication: *VA systems should enable analysts to refine the pattern operationalizations, e.g., with thresholds.*

5.4 Phase 4: Comparing Results of Multiple Parameter Settings

When we defined short AOI visits to be ignored, we found that some subsequences, which had been detected as linear reading before, were not shown as hits anymore or became shorter subsequences. In the second case, the steps taken in Phase 3 apply. This was the case for P25 (NT2) in Alpscarf analysis; whether this participant depicted linear reading depends on the threshold of short durations. In RTG analysis, P25 was considered as depicting a linear reading behavior because the analyst considered the two short visits to consecutive AOIs as (unintentional) re-reads. Because the sequence analysis of eye movement data is an iterative process, an analyst needs to compare the analysis results of different parameter settings, and then decides if the participant exhibits a certain behavior or not, by taking several factors (e.g., reading speed) into account. Additionally, it might also be appropriate to define a threshold range, for example ±50 ms, then Alpscarf highlights participants that have AOI visits within this range in a specific way for the human analyst to make a final decision. This is similar to how it was mentally done during the RTG analysis.

Design implication: *VA systems should allow analysts to compare the results of different thresholds and to define a threshold range.*

6 LIMITATIONS

Since this study compares the classifications for the same dataset resulting from Alpscarf analysis and RTG analysis, we deliberately focused on the detection of linear reading behavior which was reported in a previous analysis with RTG. Focusing on a different reading behavior, such as topic processing strategy or nonselective reading [Hyönä et al. 2002], might lead to different results and implications. Besides, the study had only two analysts analyzing the same data using two different visualization techniques. To take the interpersonal variability into account when comparing the derived insights qualitatively, future studies should involve more analysts and more datasets.

7 CONCLUSION

In this paper, we presented results found with the sequence-based analysis using Alpscarf. This was a re-analysis of a previous study conducted by Blascheck and Sharif [2019] with the RTG visualization technique. We showed that two analysts using two different visualization techniques and two different procedures—transition-based versus sequence-based visual analysis—lead to different results. We inspected these differences to understand them and to manually edit and curate the analysis. Based on these insights, we defined the design implications about the features that a VA system should fulfill for an exhaustive analysis combining transition-based and sequence-based analysis strategies.

REFERENCES

Jay Ayres, Jason Flannick, Johannes Gehrke, and Tomi Yiu. 2002. Sequential pattern mining using a bitmap representation. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 429–435.

Tanja Blascheck, Markus John, Kuno Kurzhals, Steffen Koch, and Thomas Ertl. 2016a. VA^2: A Visual Analytics Approach for Evaluating Visual Analytics Applications. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 61–70.

Tanja Blascheck, Kuno Kurzhals, Michael Raschke, Stefan Strohmaier, Daniel Weiskopf, and Thomas Ertl. 2016b. AOI hierarchies for visual exploration of fixation sequences. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*. 111–118.

Tanja Blascheck, Markus Schweizer, Fabian Beck, and Thomas Ertl. 2017. Visual Comparison of Eye Movement Patterns. *Computer Graphics Forum* 36, 3 (2017), 87–97.

Tanja Blascheck and Bonita Sharif. 2019. Visually Analyzing Eye Movements on Natural Language Texts and Source Code Snippets. In *Proceedings of the Symposium on Eye Tracking Research & Applications*. ACM. https://doi.org/10.1145/3314111.3319917

Sukru Eraslan, Yeliz Yesilada, and Simon Harper. 2016. Eye tracking scanpath analysis techniques on web pages: A survey, evaluation and comparison. *Journal of Eye Movement Research* 9, 1 (2016), 1–19.

Joseph Goldberg, Mark Stimson, Marion Lewenstein, Neil Scott, and Anna Wichansky. 2002. Eye tracking in web search tasks: design implications. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*. ACM, 51–58.

Joseph H Goldberg and Jonathan I Helfman. 2010. Scanpath clustering and aggregation. In *Proceedings of the 2010 Symposium on Eye-tracking Research & Applications*. ACM, 227–234.

Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost Van de Weijer. 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures* (1 ed.). Oxford University Press.

Jukka Hyönä, Robert F Lorch Jr, and Johanna K Kaakinen. 2002. Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology* 94, 1 (2002), 44.

Kuno Kurzhals, Michael Burch, Tanja Blascheck, Gennady Andrienko, Natalia Andrienko, and Daniel Weiskopf. 2017. A Task-Based View on the Visual Analysis of Eye Tracking Data. In *Eye Tracking and Visualization*, Michael Burch, Lewis Chuang, Brian Fisher, Albrecht Schmidt, and Daniel Weiskopf (Eds.). Springer, 3–22.

Rudolf Netzel, Bettina Ohlhausen, Kuno Kurzhals, Robin Woods, Michael Burch, and Daniel Weiskopf. 2017. User performance and reading strategies for metro maps: An eye tracking study. *Spatial Cognition & Computation* 17, 1-2 (2017), 39–64.

Gary M Olson, James D Herbsleb, and Henry H Reuter. 1994. Characterizing the sequential structure of interactive behaviors through statistical and grammatical techniques. *Human–Computer Interaction* 9, 3-4 (1994), 427–472.

Daniel Richardson and Rick Dale. 2005. Looking To Understand: The Coupling Between Speakers' and Listeners' Eye Movements and Its Relationship to Discourse Comprehension. *Cognitive Science* 29, 6 (2005), 1045–1060.

Julia M West, Anne R Haake, Evelyn P Rozanski, and Keith S Karn. 2006. eyePatterns: software for identifying patterns and similarities across fixation sequences. In *Proceedings of the 2006 Symposium on Eye Tracking Research & Applications*. ACM, 149–154.

Michael MA Wu and Tamara Munzner. 2015. SEQIT: visualizing sequences of interest in eye tracking data. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2015), 449–458.

Chia-Kai Yang and Chat Wacharamanotham. 2018. Alpscarf: Augmenting Scarf Plots for Exploring Temporal Gaze Patterns. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM.

Sybil Yang. 2012. Eye movements on restaurant menus: A revisitation on gaze motion and consumer scanpaths. *International Journal of Hospitality Management* 31, 3 (2012), 1021–1029.