Gaze Sequences and Map Task Complexity

Fabian Göbel
Institute of Cartography and Geoinformation, ETH Zurich, Zurich, Switzerland
goebelf@ethz.ch

Peter Kiefer
Institute of Cartography and Geoinformation, ETH Zurich, Zurich, Switzerland
pekiefer@ethz.ch

Ioannis Giannopoulos
Department of Geodesy and Geoinformation, Vienna University of Technology, Vienna, Austria
igiannopoulos@tuwien.ac.at

Martin Raubal
Institute of Cartography and Geoinformation, ETH Zurich, Zurich, Switzerland
mraubal@ethz.ch

Abstract

As maps are visual representations of spatial context to communicate geographic information, analysis of gaze behavior is promising to improve map design. In this research we investigate the impact of map task complexity and different legend types on the visual attention of a user. With an eye tracking experiment we could show that the complexity of two map tasks can be measured and compared based on AOI sequences analysis. This knowledge can help to improve map design for static maps or in the context of interactive systems, create better map interfaces, that adapt to the user's current task.

2012 ACM Subject Classification Human-centered computing → Empirical studies in HCI

Keywords and phrases eye tracking, sequence analysis, map task complexity

Digital Object Identifier 10.4230/LIPIcs.GIScience.2018.30

Category Short Paper

Funding Swiss National Science Foundation Grant No.: 200021_162886.

1 Introduction

Maps are visual representations of spatial context that communicate geographic information and allow for spatial problem analysis [13]. The design of "better" maps is a key goal in cartography. However, the definition of "better" is vague and has been a topic of research for a long time. In his book, MacEachren provides an overview of how maps work at different levels and how design choices interact with the processing of information from a map [10].

Visual attention is a valuable source of information for cartographic design both when evaluating a map design or adapting the interface [5]. Tracking and analyzing visual attention on maps through eye tracking experiments has been proposed and used in cartography for quite some time (see [9], for an overview). Compared to other methods for evaluating a map design, such as a "think aloud protocol", eye tracking does not introduce additional cognitive load or affect the task. Research questions that have been addressed by eye tracking experiments range from testing the differences between expert and novice map users [12], evaluating cartographic design decisions [1], or analyzing task complexity and cognitive processes [11].

10th International Conference on Geographic Information Science (GIScience 2018).
Editors: Stephan Winter, Amy Griffin, and Monika Sester; Article No. 30; pp. 30:1–30:6
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

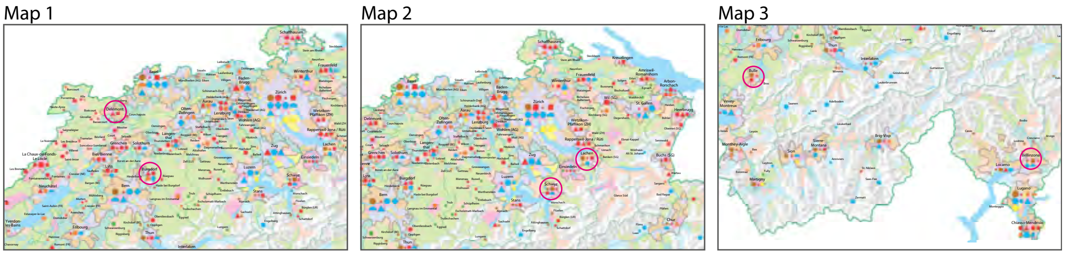

Figure 1 The three maps used in the experiment (legend excluded). The magenta circles on the maps indicate the cities that were subject of the tasks. These circles were not visible during the experiment. The map material is based on the economic map from the Swiss World Atlas[1].

Depending on the purpose of the analysis, different measures are commonly used for the analysis of gaze data collected during the interaction with maps. Some measures, such as average fixation duration [2], are not related to the map content and may provide general insights about the cognitive state of the user. Content-related measures, on the other hand, enable an analysis of which elements of the map or interface the user has paid attention to [8], thus allowing for a more detailed evaluation of the map or interface. For instance, Cöltekin et al. [3] used sequence analyses on Areas of Interest (AOI) to study individual and group differences for a geovisual analytics tasks on two different map interfaces.

In this short paper, we suggest to use compressed string analysis of eye tracking data to evaluate the impact of task complexity and different legend types on the visual attention of a user. The two gaze based legend types described in [6] and a traditional legend were tested on three different map extents. This result can help to improve map design or in the context of interactive systems, create better map interfaces, that adapt to the user's current task.

In this research we investigate whether the complexity of two map tasks can be measured and compared based on fixation sequences. In order to address this research question, we choose to analyze the mean fixation duration and perform a sequence analysis based on AOIs. The short paper is structured as follows: We first explain the experiment including an introduction to the task, the map and the legends used. Furthermore, we explain the procedure and the AOIs used. In the results section we report on average fixation duration and gaze sequences. Finally we discuss the results and provide an outlook on future work.

2 Experiment

We intended to test the search behavior and interaction with a map legend while performing a common comparison task. For this we chose three maps (Figure 1) with varying symbol density and the three legend types, one traditional and two that adapt to the users' gaze as described in previous work [6]. This results in a 3×3 within-subjects design, with three maps and three legend types. Each participant performed the task on each of the three maps extents once. Map extents, legend type and ordering were counterbalanced based on a Latin square. In the following, we explain the task in more detail.

2.1 Task, Map and Legend

The task of the user was to inspect two cities (A and B) on the map, and determine and name the industries that differ between the two. Visual inspection of the legend was required in order to interpret the meaning of the differing symbols. Before starting the actual task, the location of the two cities was presented to the participant in order to avoid search and only measure task-related gaze behavior. Panning and zooming was not possible.

Table 1 Number of symbols shown for the two cities that needed to be compared on the three maps. These numbers are taken as a measure for task complexity: the comparison task on Map 1 was less complex than that on Map 2, which in turn was less complex than that on Map 3.

	Number of symbols						symbol density	Visual angle between Cities
	in City A	in City B	in City A but not in City B	in City B but not in City A	total	different		
Map 1	2	4	0	2	6	2	high	5.7°
Map 2	5	5	2	2	10	4	high	9.1°
Map 3	4	6	2	4	10	6	low	34.3°

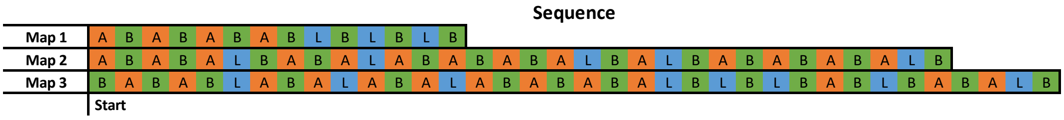

Figure 2 Example sequences of one participant's dwells on three different AOIs: the two cities whose symbol sets had be be compared (A, B) and the legend (L).

We expected the chosen approach to result in a very structured and predefined way of solving the task: first the participant looks at city A then at city B in search for symbols that differ. After finding at least one, the participant will search within the legend to determine its meaning. This structured approach allows us to break down the analysis to a sequence analysis on only three different AOIs (Figure 2).

As with this study we focus more on task difficulty and not on the design of the map itself, we employed an economic map from the Swiss World Atlas which had been designed by experienced cartographers to teach geography in schools[1]. We can identify four characteristics that among others, increase the search space and thus contribute to a higher task difficulty:

- Total number of symbols in a map extent
- Number of symbols per city
- Number of symbols that differ between two cities
- Distance between the cities

Based on this, we chose three map extents for our experiment (Figure 1). Map 1 and 2 feature a higher density of symbols compared to Map 3. The distance between the relevant cities is the shortest in Map 1, however, still exceeds the area that can be inspected with one fixation, followed by Map 2 and Map 3. Table 1 shows that the cities contained three to six symbols each, and that two to six differed between them. For instance, Map 1 has only two symbols that differ between the two cities. Furthermore, these symbols are all in city B. We assume that this makes the task the easiest on Map 1, followed by Map 2 and Map 3 as with them more map features differ.

The design of the legend was from the original map and a total of 26 symbols were shown (Figure 3). We tested a traditional legend and the two gaze based legend types described in our previous work [6] namely *fixed adaptive*, where the content of the legend is adapted to highlight the symbols that where visible within the last fixation on the map and *dynamic adaptive* which also adapts its placement to appear always at the bottom right position of the current field of view.

[1] `https://schweizerweltatlas.ch/en/`

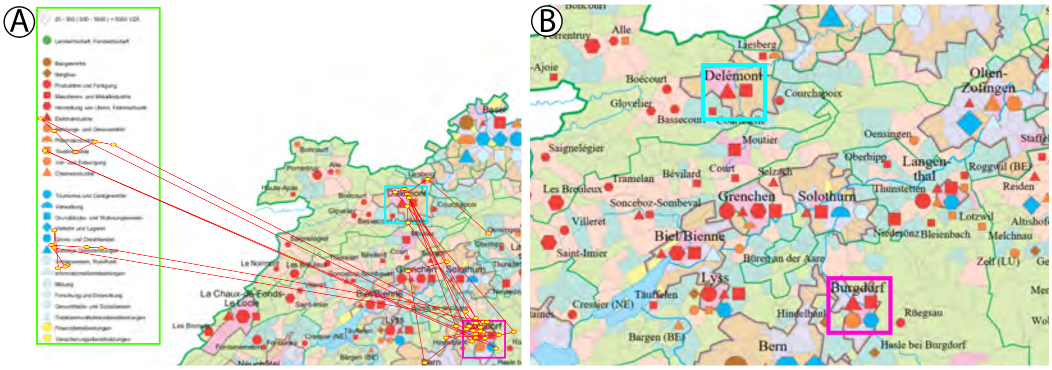

Figure 3 Ⓐ shows a part of Map 1 with the AOI for the legend in green, the AOIs for City A in cyan and for City B in magenta. The scan path is highlighted in red and fixations in yellow. Figure 2 shows the result as a sequence (see first row). Ⓑ shows the cities and symbols that need to be compared in detail.

In our previous experiment we could show that with the gaze-based legends, participants spent less task time on the legend compared to the traditional legend [6]. Here, however, we are interested in analyzing the impact of task difficulty on the gaze sequence and on usage of the legend.

2.2 Participants and Setup

18 participants (7 female) took part in our experiment with most of them having a professional background in Geomatics or Cartography. Their average age was 31.9 ($SD = 4.4$).

During the study, we collected gaze data using a Tobii TX 300 eye tracker. Additionally, we used a chin rest to keep the distance between participants and display (23″, 1920×1080 px) constant (60 cm). Before each run we performed a 9-point calibration.

2.3 Procedure

After filling out a demographic questionnaire, participants proceeded with a test run to familiarize themselves with the given legend type. Next, a preview map without the symbols was provided to show the locations of the two cities in question. When the participant indicated that she was ready, the actual task began, however, there was no time constrain to fulfill the task. These steps were repeated three times to test all different maps and legend types. This assured that each possible combination of map $\times$ legend was tested six times.

2.4 Area of Interest

As we are mainly interested in which sequence visual attention was spent on the map and the legend, for each task, we annotated the following three AOIs: Legend, City A and City B. In case of cities, the size of the AOIs comprised the city name and all related symbols (Figure 3). For the legend, the AOI was dynamically adapted to the size (and placement) of the legend which requires to track fixations in real time. Based on the gaze data coming at 300 Hz from the eye tracker, we used our online implementation of the I-DT algorithm first introduced in [4] to calculate fixations (80 px dispersion and 200 ms window size).

For deriving sequences from the gaze data, we denoted the fixations to AOIs in order of appearance. Consecutive fixations on the same AOI are handled as one visit, called dwell.

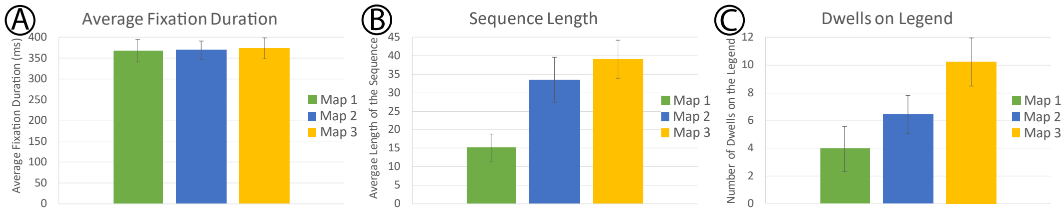

Figure 4 Results for mean fixation duration A, average length of sequences B, and average dwells on the legend C independent of legend type. Error bars indicate the 95% confidence interval.

3 Results

First, we calculated the average fixation duration (Figure 4 A). This is a measure commonly related to the task difficulty [7]. However, independent of the used legend type, a one-way ANOVA ($F(2{,}51) = .06$, $p = .940$) could not show a statistical significant difference between the three Maps.

From Figure 4 B we can see that in general, sequence length (i.e. number of dwells on an AOI) is shorter for Map 1 (15.1, $SD = 8.1$) followed by Map 2 (33.5, $SD = 6.0$) and Map 3 (39.0, $SD = 5.1$). A one-way ANOVA ($F(2{,}51) = 22.110$, $p < .001$) confirmed statistically significant differences between the Maps. All following results are Bonferroni adjusted ($\alpha = 0.017$). Post hoc analysis with a Tukey test resulted in a significant difference between Map 1 and Map 2 ($p < .001$), and Map 1 ($p < .001$) and Map 3 but not between Map 2 and Map 3 (textitp = .318). Furthermore, a one-way ANOVA showed no significant effect of legend types onto the length of the gaze sequence (Map 1: $F(2{,}15) = 1.261$, $p = .312$; Map 2: $F(2{,}15) = .400$, $p = .678$; Map 3: $F(2{,}15) = 3.530$, $p = .055$).

We also counted the number of dwells on the legend (Figure 4 C). As this data was not normally distributed we used a Kruskal-Wallis H which confirmed statistical differences ($\chi^2 = 29.832$, $p < .001$). Following the results of the Mann Whitney U post-hoc tests shows that participants dwelt significantly less often on the legend on Map 1 (mean = 3.94, $SD = 3.67$) compared to Map 2 (mean = 6.44, $SD = 3.05$, $U = 47.0$, $p < .001$) and Map 3 (mean = 10.22, $SD = 3.84$, $U = 16.0$, $p < .001$). Also the result between Map 2 and 3 is significant ($U = 56.0$, $p < .001$). If we compare these values with the number of different symbols in Table 1, we can see a correlation between number of symbols that differ between the two cities and participants' dwells on the legend. The ratio is between 0.51 and 0.62. Again, a Kruskal-Wallis H was applied to calculate the effect of legend types onto the number of dwells on the legend. However, legend type has no statistically significant effect with Map 1 ($\chi^2 = 4.258$, $p = .119$), but it has on Map 2 ($\chi^2 = 5.984$, $p = .050$) and on Map 3 ($\chi^2 = 7.645$, $p = .022$).

Furthermore, we analyzed the sequences before the legend was visited the first time. In average 5.8 switches between City A and City B have occurred before the gaze shifted to the legend. However, we could neither find statistical significance between the different maps ($\chi^2 = .917$, $p = .632$) nor did the legend ($\chi^2 = 7.743$, $p = .021$) seem to have an impact.

4 Discussion and Future Work

Although we could not find any significant differences in the fixation duration, evaluation of the sequence length indicates that more differences of symbols, which we take as an indicator for task difficulty, result in more focus switches between cities and legend. Furthermore, we could show that the number of dwells on the legend goes in line with the number of different symbols between two cities. The fact that the number of dwells on the legend was always higher than the number of different symbols is particularly interesting, as this suggests, participants mostly evaluated one symbol at a time, when visiting a legend and needed some more to reassure their answer.

One reason could be that the symbols consist of arbitrary shapes and colors (Figure 1). It could be that more decisive or iconic symbols are easier to remember and require less re-visits of the legend. Future work has to inspect AOI sequences in more detail. For instance, can correlation between sequences of different participants contribute to find common patterns for certain tasks? This knowledge can help to create better maps or in the context of interactive systems, create better map interfaces, that adapt to the user's current task.

References

1. Alžběta Brychtová and Arzu Çöltekin. An Empirical User Study for Measuring the Influence of Colour Distance and Font Size in Map Reading Using Eye Tracking. *The Cartographic Journal*, 53(3), 2016.
2. Henry W. Castner and Ronald J. Eastman. Eye-Movement Parameters and Perceived Map Complexity - I. *Cartography and Geographic Information Science*, 11(2), 1984.
3. Arzu Çöltekin, Sara I. Fabrikant, and Martin Lacayo. Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Science*, 24(10), 2010.
4. Ioannis Giannopoulos, Peter Kiefer, and Martin Raubal. GeoGazemarks: Providing Gaze History for the Orientation on Small Display Maps. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction (ICMI '12)*, New York, New York, USA, 2012. ACM Press.
5. Fabian Göbel, Ioannis Giannopoulos, and Martin Raubal. The importance of visual attention for adaptive interfaces. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI '16*, New York, New York, USA, 2016. ACM Press.
6. Fabian Göbel, Peter Kiefer, Ioannis Giannopoulos, Andrew T. Duchowski, and Martin Raubal. Improving Map Reading with Gaze-Adaptive Legends. In *ETRA '18: 2018 Symposium on Eye Tracking Research & Applications*. ACM, 2018, (to appear).
7. Kenneth Holmqvist and Richard Andersson. *Eye Tracking: A comprehensive guide to methods, paradigms and measures*. CreateSpace, Lund, 1 edition, 2017.
8. Peter Kiefer and Ioannis Giannopoulos. Gaze Map Matching: Mapping Eye Tracking Data to Geographic Vector Features. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, New York, New York, USA, 2012. ACM.
9. Peter Kiefer, Ioannis Giannopoulos, Martin Raubal, and Andrew Duchowski. Eye Tracking for Spatial Research: Cognition, Computation, Challenges. *Spatial Cognition & Computation*, 17, 2017.
10. Alan M. MacEachren. *How Maps Work Representation, Visualization, and Design*. Guilford Press, 2004.
11. Daniel R. Montello. Cognitive Map-Design Research in the Twentieth Century: Theoretical and Empirical Approaches. *Cartography and Geographic Information Science*, 29(3), 2002.
12. Kristien Ooms, Philippe De Maeyer, Veerle Fack, Eva Van Assche, and Frank Witlox. Interpreting maps through the eyes of expert and novice users. *International Journal of Geographical Information Science*, 26(10), 2012.
13. Terry A. Slocum, Robert B. McMaster, Fritz C. Kessler, and Hugh H. Howard. *Thematic Cartography and Geovisualization*. Pearson, Upper Saddle River, New Jersey, USA, 3 edition, 2009.