

A standard cartographic generalization approach is used to cluster fixations that are in spatial proximity⁴.

Figure 1 (bottom) shows our implementation of the retrieval phase. In order to see all the fixation points on the map, that are aggregated by a given cluster, one can zoom into this region. When the zoom-in interaction crosses the zoom threshold, it leads back to the recording phase where previous fixations are shown unaggregated (as described above).

The multimodality of GeoGazemarks strongly couples the gaze based and zoom/pan interactions, since the zoom level (de)activates the recording and retrieval phase, and the panning influences the geo-location of the gaze fixations.

4. IMPLEMENTATION

We implemented a prototype of the GeoGazemarks interaction concept that was used in the experiment described in section 5.

4.1 Hardware

Our hardware consisted of the Ergoneers Dikablis mobile eye tracker⁵ with a gaze capture rate of 50Hz. The data was transmitted to a laptop via a coaxial cable. The laptop forwarded the data to a Samsung Galaxy Nexus smartphone where the prototype application was running. The connection between laptop and smartphone was established over WiFi in a closed network with a data rate of 25 Hz.

4.2 Software

4.2.1 Calibration and Recording (Laptop)

The Dikablis mobile eye tracking system comes with a series of software modules. We used the Dikablis Live Recorder software which allowed us to calibrate the users, as well as to start and stop the recording sessions. The software also provides real-time visual marker detection, returning marker dependent coordinates with respect to a marker attached to the smartphone screen. These marker dependent coordinates were transmitted over UDP to the smartphone.

4.2.2 GeoGazemarks Application (Smartphone)

The GeoGazemarks concept was realized as an Android application (Android 4.0). The graphical user interface is based on an OpenStreetMaps⁶ map service that includes standard panning and zooming functionality. The application implements implicit gaze interaction by computing fixations and visualizing them on the map.

From raw gaze data to fixations

The marker dependent coordinates received over UDP were first transformed into pixel coordinates with respect to the smartphone screen. As the human eyes are constantly moving, not all of these screen gazes imply perception: the gaze data need to be pre-processed by a fixation detection algorithm.

A fixation occurs when the eyes remain relatively still over some time. In the eye tracking literature it is generally assumed that information is only perceived during these fixations [24]. A fixation has a spatial ('relatively still') and a

temporal component ('over some time'). The mean fixation duration necessary in order to obtain information is task specific [24]. For instance, the mean fixation duration during scene perception is about 330 ms. After a preliminary test, we implemented our fixation algorithm with a radius of 30 pixels and a duration of 380 ms. In other words, if all gazes during a time span of 380 ms were within a radius of 30 pixels, they would form a fixation. If they were within this radius for 760 ms they would form a fixation with double duration, and so on.

Finally, the application transforms the fixation screen coordinates to respective geospatial map coordinates (latitude, longitude) using a method provided by the map service.

Cartographic visualization and clustering

The map service is enhanced by two overlays. One overlay is used for the visualization of the items that are only visible on the highest zoom level (i.e., fixation icons and points of interest), and the other overlay serves for the visualization of the fixation clusters which are visible on all other zoom levels. The fixation points computed in the previous step are used to set the fixation icons (See Figure 1, top).

The fixation clusters are computed by a grid clusterer. The grids are based on the screen density of the device, thus making the GeoGazemarks application also device independent. The clusters are formed according to the existing gaze fixations on the map. We used three sizes to visualize the clusters. The smallest cluster visualization was used for up to 3 fixations, the middle one for up to 8, and the largest one for more than 8 fixations. We selected a semi-transparent orange color their visualization, allowing the user to recognize the map characteristics underneath.

The proportions of the clusters as well as the proportions of the fixation icons were determined by a process known as perceptual scaling, with the following formula ([27], p. 308):

$$r_i = \left(\frac{v_i}{v_L} \right)^{0.57} \times r_L$$

r_i = radius of the circle to be drawn

r_L = radius of the largest circle on the map

v_i = data value for the circle to be drawn

v_L = data value associated with the largest circle

The fixation intervals for the clusters as well as the color selection for them were decided by a pilot user study and expert interviews (See 5.4).

5. EXPERIMENT

We used the implementation described in section 4 and conducted a user study in order to compare two conditions: a search task on a mobile phone screen using GeoGazemarks (Ggz), and the same task without (Std). The hypothesis we tested was that participants would be able to find a previously seen place on a smartphone map faster in the Ggz condition, i.e., when a gaze history visualization is present. The research hypothesis is stated accordingly (with respect to task completion times):

$$\begin{aligned} H_0 : \mu_{Ggz} &= \mu_{Std} \\ (H_1 : \mu_{Ggz} &< \mu_{Std}) \end{aligned}$$

⁴Refer to the implementation section for technical details.

⁵<http://www.ergoneers.com/>

⁶<http://www.openstreetmap.org/>

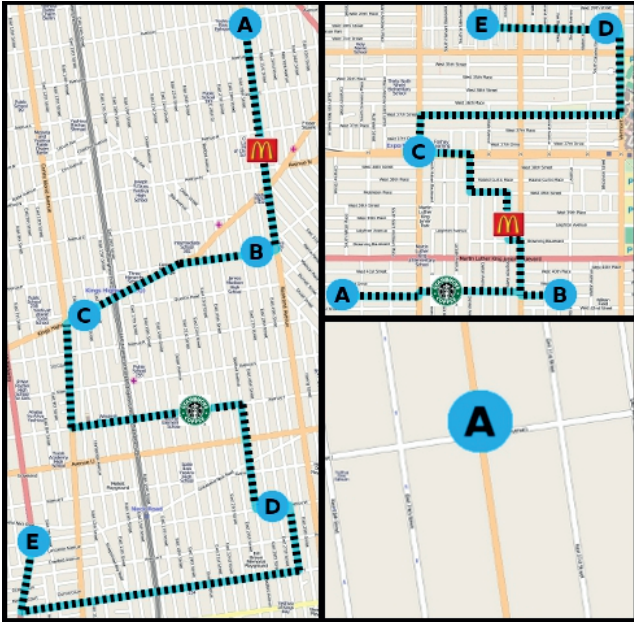


Figure 2: Point configurations for maps of New York (left) and Los Angeles (top right). The dotted line was not visible during the experiment. At most one point was visible at the same time on the screen (bottom right).

5.1 Setup

Participants were placed in front of a smartphone with a screen size of 4.6", fixed to an office desk using a car mobile phone holder (See Figure 3). They were told to choose the distance to the device naturally and keep this distance as constant as possible during the experiment.

Directly before the part of the experiment that requires eye tracking, the mobile eye tracker was mounted on the participant's head and calibrated to the distance of the device using the eye tracker's calibration software. In the parts of the experiment that did not require eye tracking, the participants did not wear the mobile eye tracker.

5.2 Design

A within subjects design was employed. Each participant had to perform the same search task in both conditions (with GeoGazemarks = Ggz; only standard map interaction = Std) with two different map areas. We chose one map section of New York, NY, USA (NY), and another map section of Los Angeles, CA, USA (LA). NY and LA were counter-balanced among the two test conditions. The reason we prepared two maps with distinct urban characteristics was to avoid confounding among variables. For the same reason we also switched the order (first/second) in which GeoGazemarks was tested. This yields the four cases listed in Table 1. Ten participants were tested for each case.

We placed seven point objects on each of the two maps: five of them were blue circles labeled from 'A' to 'E', two were famous logos of a restaurant chain and a coffeehouse chain (one McDonald'sTM, one StarbucksTM). The five blue circles determined a vector sequence the participants had to traverse (A → B → C → D → E). The vector sequences for LA and NY were chosen to be different, and to cover all



Figure 3: Experiment setup.

| | | Cases | | | |
|-------|-----------------|---------|---------|---------|---------|
| | | I | II | III | IV |
| Order | 1 st | NY, Std | LA, Std | NY, Ggz | LA, Ggz |
| | 2 nd | LA, Ggz | NY, Ggz | LA, Std | NY, Std |

Ggz = with GeoGazemarks, Std = without GeoGazemarks

NY = map of New York, LA = map of Los Angeles

Table 1: Four test cases of the experiment.

cardinal directions at least once. The two logo point objects were each placed in a way that one had to stumble upon them while traversing the vector sequences. The point layouts for LA and NY are displayed in Figure 2. The seven point objects were only visible on the highest zoom level – just as regular mobile map services would hide away unimportant points of interest when zooming out.

During each trial we logged all explicit user interactions with the phone (i.e., zooming, panning). In those trials with the Ggz condition we also collected implicit interactions (gaze).

5.3 Procedure

Each trial was composed of four steps. In the first step, the participants had to provide their demographic information as well as to state their experience level in using mobile maps. Steps 2 and 3 were the actual trials, with test condition and map depending on the test case (See Table 1).

Each trial started with a map of the respective city, centered at point 'A' on the highest zoom level. The participant was told to follow the navigation instructions given by a researcher. These instructions lead her along the vector sequence until she reached 'E'. Instructions included cardinal directions and linear features, such as '*Continue on Normandie Avenue heading west until you reach point B*'. These instructions were unambiguous and easy to follow. Up to this point, participants were not informed about the two logo points (McDonald'sTM and StarbucksTM) – they just stumbled upon them.

At point 'E' participants were asked to tap on the 'E' point, and then find their way back to McDonald'sTM (for NY), or StarbucksTM (for LA), as fast as they could. In the Std test condition they were allowed to use standard interaction, i.e., panning by swiping and two-finger zooming. In the Ggz test condition they could additionally see the GeoGazemarks history. Task completion time was recorded as the time between the 'E' tapping and a tap on the goal.

Participants could decide to give up on the task if they got totally lost on the map.

Just before the Ggz condition was tested, participants could try out the GeoGazemarks concept on a map of a European city (max. 3 minutes) in order to get used to it and to see at least once how the visualizations look.

In the last step, participants filled in another questionnaire and rated the GeoGazemarks concept in terms of usability. The overall duration of the experiment was between 20 to 30 minutes per participant.

5.4 Pilot Study

To optimize the test procedure and the visualization of our GeoGazemarks we conducted a small pilot study with 8 participants (5 male, 3 female) that had a professional background in Cartography and Geomatics. These users participated only in the pilot study.

The pilot followed the same procedure as described in section 5.3. In this version of the study, however, instructions included only cardinal directions, such as *‘Swipe to the South-West until you find B’*. The participants had severe problems in keeping these diagonal directions during swiping, probably because mobile screens are not quadratic. As we did not want participants to get lost in the instructions phase, we modified the instructions for the final study to refer to linear features, such as roads (See section 5.3).

The gaze clusters in the retrieval phase were initially not semi-transparent. The pilot study revealed that this was hiding too much of the map, so we determined a good transparency level through testing. Our expert participants also suggested surrounding the gaze clusters with a distinctive border to increase their visibility (refer to Figure 1, bottom).

In the pilot study only two participants could solve the search task in the Std condition. This gave us a first indication that our assumption that people would be much faster with GeoGazemarks might actually be true.

5.5 Participants

40 participants took part in the main study and were compensated with 15 CHF for their efforts. The mean age was 27.9 years, with 13 females and 27 males. They had different cultural and professional backgrounds. Half of the participants (20) had a professional background in Geomatics, GIScience, or Cartography. The participants rated their expertise on mobile maps on a scale ranging between 1 (no experience) and 7 (very experienced) with 3.73 (SD = 2.063). None of the participants had ever been to NY or LA, and none had any other previous knowledge about the two areas.

6. RESULTS

We analyzed the data collected during the user study in order to compare the differences in task effectiveness and efficiency for the standard interaction (Std) and GeoGazemarks (Ggz) conditions.

6.1 Effectiveness

Task effectiveness was analyzed by comparing the number of participants who completed the task successfully, and the number of participants who gave up: in the Std condition, 25 participants (62.5%) completed the task. With GeoGazemarks (Ggz condition) all 40 participants completed the task successfully.

6.2 Efficiency

Two efficiency measures were chosen for the analysis: task completion time and interaction sequence length. Both measures in the two conditions were compared using the non-parametric Wilcoxon signed-rank test.

6.2.1 Task completion time

As described in the procedure section, the task completion time was defined as the time difference between the tap on the ‘E’ point and the tap on the goal. For those 15 participants (37.5%) who gave up, the time difference between the ‘E’ tap and the give-up time was considered.

As predicted, participants could complete the search task significantly faster using GeoGazemarks than without (See Table 2).

| Task Completion Time | | |
|----------------------|--------------|--------------------|
| | Median | Standard Deviation |
| Ggz | 50.7965 sec. | 24.01 sec. |
| Std | 124.098 sec. | 85.36 sec. |

Ggz = with GeoGazemarks, Std = without GeoGazemarks
 $p < 0.001$, $Z = -5.027$

Table 2: Task Completion Time. The search task was completed significantly faster using GeoGazemarks.

6.2.2 Interaction sequence length

Interaction sequence length was defined as the number of interactions needed to complete the search task. For instance, if a participant completed the search task by performing the following interaction sequence *‘Swipe North - Swipe West - Zoom in’*, this was considered as an interaction sequence of length three. Again, the interaction sequence of those 15 participants who gave up was considered as finished at the time they gave up.

The findings for task completion time were reflected in the analysis of interaction sequences. The interaction sequences were significantly shorter using GeoGazemarks than without (See Table 3).

| Interaction Sequence Length | | |
|-----------------------------|--------|--------------------|
| | Median | Standard Deviation |
| Ggz | 25.5 | 17.17 |
| Std | 75 | 52.6 |

Ggz = with GeoGazemarks, Std = without GeoGazemarks
 $p < 0.001$, $Z = -4.799$

Table 3: Interaction Sequence Length. The interaction sequences were significantly shorter using GeoGazemarks.

6.2.3 Group differences

We were interested in whether task efficiency would differ significantly depending on user groups. We first looked at the group of expert users: we considered those participants as experts that have a professional background in Geomatics, GIScience, or Cartography (20 participants, see section 5.5), and all other participants as non-experts (20 participants). We suspected that people working with geoinformation on a daily basis might have superior cognitive mapping strategies. To separate professional background

from experience with mobile maps, we compared the participants' self-estimation on mobile map usage expertise (See section 5.5) for expert and non-expert users: non-expert users rated their expertise with 3.40 (SD = 2.010), expert users with 4.05 (SD = 2.114). A non-parametric Wilcoxon rank sum test revealed that the difference in their expertise is not significant ($p > 0.05$, $Z = -1.013$).

A non-parametric Wilcoxon rank sum test of the tracked data revealed that there is no significant difference in the task completion times between expert and non-expert users with GeoGazemarks ($p > 0.05$, $Z = -1.082$), as well as for task completion time in the Std condition ($p > 0.05$, $Z = -0.784$). The same holds also for interaction sequence lengths, once with GeoGazemarks ($p > 0.05$, $Z = -0.515$) and once without ($p > 0.05$, $Z = -0.595$).

We also did not find any significant differences between female and male participants, neither for the search condition with GeoGazemarks ($p > 0.05$, $Z = -1.314$) nor for the condition without ($p > 0.05$, $Z = -0.967$). The same is also reflected in interaction sequences (Ggz: $p > 0.05$, $Z = -1.099$, Std: $p > 0.05$, $Z = -0.838$).

6.3 Analysis of questionnaires

On a questionnaire participants were asked to rate GeoGazemarks. The rating scale ranged between 1 (strongly disagree) and 7 (strongly agree). The results showed that the visualization of clusters on the map was perceived as helpful: the relation between cluster size and the number of fixations aggregated by them was rated as helpful with a mean of 6.37 (SD = 1.005). The cluster visibility on the map was rated with a mean of 6.68 (SD = 0.730).

The mean of the overall rating for the GeoGazemarks concept (*'do you like the GeoGazemarks approach?'*) was 6.60 (SD = 0.591), and the task-dependent question (*'was the GeoGazemark approach helpful for performing the task?'*) resulted in a rating with a mean of 6.52 (SD = 0.847). Asked for their preference for GeoGazemarks over standard map interaction, participants showed a high preference (Mean = 6.53, SD = 0.877) for GeoGazemarks when dealing with situations such as the one in the experiment.

7. DISCUSSION AND CONCLUSION

In this paper we have introduced GeoGazemarks, an implicit gaze interaction concept for small display maps. It is based on the assumption that places on a map the user has spent visual attention to have become part of her cognitive map, and can thus be used as orientational cues on overview zoom-levels. This is relevant for any map task that requires the user to search for previously seen places. By highlighting places she has looked at before, GeoGazemarks prevent the map user from getting lost when zooming out, which is often the case for standard map interaction.

An analogy to the 'browsing history' of web browsers seems obvious: our system silently records the map usage history in the background, even though the user does not yet know at recording time whether the recording will be needed later. Intentionally setting a waypoint, on the other hand, can be compared to bookmarking a website. Still, GeoGazemarks are different because our visualization addresses spatial memory, not sequential memory (like the plain list in a browser history).

In our experiment we focused on a path integration task in which participants had to find back to a point of interest

they had stumbled upon before. The paths they had to follow were predefined for each of the two city maps to make experimental conditions comparable between participants. We expect that GeoGazemarks will also be beneficial for free map exploration tasks in which the user is free to choose her own path (such as Alice in the example in section 3.1), as this can also be reduced to path integration.

We chose two different cities and two different vector sequences to avoid confounding between variables. One might argue that this does not cover other geographic areas with completely different characteristics, such as the Sahara desert. However, as LBS are mostly used in cities, we believe that urban settings are the most relevant ones. Participants also told us that they did not focus much on city landmarks when GeoGazemarks were present. It is intuitive that the length and complexity of the vector sequence influence the performance of path integration. However, detailed research on this is out of the scope of this paper. We tried to choose sequences of realistic length and complexity for ordinary map search tasks. Two unfamiliar cities were selected in order to not give anyone the advantage of previous spatial overview knowledge. If the GeoGazemarks principle was applied to an area the user already has a cognitive map of, we still expect that the visual attention history would help to reduce the search space, although the efficiency advantage of GeoGazemarks might be a bit smaller.

Using monocular head-mounted eye trackers for small distances requires the participant to keep the distance to the device relatively constant. We instructed participants accordingly and observed an accuracy of approx. 5mm. These slight inaccuracies were smoothed out by the clustering. The eye camera (See Figure 3, right) has no serious impact on the field of view since after a few minutes, it is no longer perceived (recall that we had the training phase). With less intrusive hardware we would expect the GeoGazemarks concept to perform at least as good as in our current experiment.

8. OUTLOOK

An interesting aspect for future work where GeoGazemarks could be valuable, is to get a better understanding of people's information needs when interacting with maps by analyzing their browsing behavior. One question for future research is how mid- or long-term usage of GeoGazemarks will change the way participants use gaze. We might expect that they will use GeoGazemarks for explicit interaction, 'bookmarking' places by intentionally staring at them to create long fixations. An open question is also how to deal with multiple GeoGazemarks sessions in the same map region: can we just aggregate GeoGazemarks from different sessions? How helpful or confusing is gaze history recorded several days or weeks earlier?

Some participants suggested adding sequential information by connecting GeoGazemarks with lines. Future research will show whether such information could make GeoGazemarks even more efficient. However, it is yet unclear which visualization will prevent visual clutter due to frequently self-intersecting gaze tracks (on a high zoom-level) and self-intersecting map usage paths (on a low zoom-level).

9. REFERENCES

- [1] G. L. Allen. Spatial Abilities, Cognitive Maps and Wayfinding - Bases for Individual Differences in

- Spatial Cognition and Behavior. In R. Golledge, editor, *Wayfinding Behavior - Cognitive Mapping and Other Spatial Processes*, pages 46–80. Johns Hopkins University Press, 1999.
- [2] A. Aretz and C. Wickens. The Mental Rotation of Map Displays. *Human Performance*, 5:303–328, 1992.
 - [3] R. Biedert, G. Buscher, and A. Dengel. The eyebook - using eye tracking to enhance the reading experience. *Informatik-Spektrum*, 33(3):272–281, June 2010.
 - [4] R. Biedert, G. Buscher, S. Schwarz, J. Hees, and A. Dengel. Text 2.0. In *CHI extended abstracts on human factors in computing systems*, pages 4003–4008, New York, NY, USA, 2010. ACM.
 - [5] S. Burigat, L. Chittaro, and S. Gabrielli. Navigation techniques for small-screen devices: An evaluation on maps and web pages. *International Journal of Human-Computer Studies*, 66(2):78–97, 2008.
 - [6] J. R. Cauchard, M. Löchtefeld, P. Irani, J. Schöning, A. Krüger, M. Fraser, and S. Subramanian. Visual separation in mobile multi-display environments. In *24th annual ACM symposium on User interface software and technology (UIST)*, pages 451–460, New York, NY, USA, 2011. ACM.
 - [7] J. Dilleuth. Map Design Evaluation for Mobile Display. *Cartography and Geographic Information Science*, 32(4):285–301, 2005.
 - [8] A. T. Duchowski. *Eye Tracking Methodology: Theory and Practice*. Springer, London, 2nd edition, 2007.
 - [9] T. Gärling, A. Böök, and E. Lindberg. Cognitive mapping of large-scale environments: The interrelationship of action plans, acquisition, and orientation. *Environment And Behavior*, 16(1):3–34, 1984.
 - [10] G. Gartner, D. A. Bennett, and T. Morita. Towards ubiquitous cartography. *Cartography and Geographic Information Science*, 34(4):247–257, 2007.
 - [11] R. Golledge. Human wayfinding and cognitive maps. In R. Golledge, editor, *Wayfinding behavior: Cognitive mapping and other spatial processes*, pages 5–45. Johns Hopkins University Press, Baltimore, MA, 1999.
 - [12] S. C. Hirtle and P. B. Heidorn. The structure of cognitive maps: Representations and processes. In T. Gärling and R. Golledge, editors, *Behavior and Environment: Psychological and Geographical Approaches*, pages 170–192. Elsevier, 1993.
 - [13] S. C. Hirtle and J. Jonides. Evidence of Hierarchies in Cognitive Maps. *Memory & Cognition*, 13(3):208–217, 1985.
 - [14] R. J. K. Jacob. The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Trans. Inf. Syst.*, 9(2):152–169, 1991.
 - [15] D. Kern, P. Marshall, and A. Schmidt. Gazemarks: gaze-based visual placeholders to ease attention switching. In *Proc. of the International Conference on Human Factors in Computing Systems (CHI)*, pages 2093–2102. ACM, 2010.
 - [16] P. Kiefer, F. Straub, and M. Raubal. Towards location-aware mobile eye tracking. In *Symp. on Eye-Tracking Research and Applications (ETRA)*, pages 313–316, New York, 2012. ACM.
 - [17] R. M. Kitchin. Cognitive maps: What are they and why study them? *Journal of Environmental Psychology*, 14(1):1–19, 1994.
 - [18] B. Kuipers. The "Map in the Head" Metaphor. *Environment and Behavior*, 14(2):202–220, 1982.
 - [19] J. M. Loomis, R. L. Klatzky, R. G. Golledge, and J. W. Philbeck. Human navigation by path integration. In R. G. Golledge, editor, *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, pages 125–151. The John Hopkins University Press, Baltimore, Maryland, USA, 1999.
 - [20] P. Majaranta, U. K. Ahola, and O. Spakov. Fast gaze typing with an adjustable dwell time. In *Proc. of the International Conf. on Human Factors in Computing Systems (CHI)*, pages 357–360. ACM, 2009.
 - [21] M. Porta, A. Ravarelli, and G. Spagnoli. *ceCursor*, a contextual eye cursor for general pointing in windows environments. In *Symp. on Eye-Tracking Research and Applications (ETRA)*, pages 331–337. ACM, 2010.
 - [22] J. Portugali. The construction of cognitive maps. In J. Portugali, editor, *GeoJournal library*, volume 32, pages 321–344. Kluwer Academic Publishers, 1996.
 - [23] M. Raubal and I. Panov. A Formal Model for Mobile Map Adaptation. *Location Based Services and TeleCartography II*, pages 11–34, 2009.
 - [24] K. Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3):372–422, 1998.
 - [25] M. Rohs, J. Schöning, M. Raubal, G. Essl, and A. Krüger. Map navigation with mobile devices: virtual versus physical movement with and without visual context. In *Proc. of the International Conference on Multimodal Interfaces (ICMI)*, pages 146–153. ACM, 2007.
 - [26] A. Schmidt. Implicit human computer interaction through context. *Personal and Ubiquitous Computing*, 4(2/3):191–199, 2000.
 - [27] T. A. Slocum, R. B. McMaster, F. C. Kessler, and H. H. Howard. *Thematic Cartography and Geovisualization (3rd Edition)*. Prentice Hall Series in Geographic Information Science. Prentice Hall, 3 edition, 2008.
 - [28] S. Stellmach and R. Dachsel. Designing gaze-based user interfaces for steering in virtual environments. In *Symp. on Eye-Tracking Research and Applications (ETRA)*, pages 131–138, New York, NY, USA, 2012. ACM.
 - [29] B. Tversky. Cognitive maps, cognitive collages, and spatial mental models. *Spatial Information Theory: Theoretical Basis for GIS*, 716:14–24, 1993.
 - [30] T. Vildan and R. J. K. Jacob. Interacting with eye movements in virtual environments. In *Proc. of the SIGCHI conference on Human factors in computing systems (CHI)*, pages 265–272, New York, NY, USA, 2000. ACM.
 - [31] G. Weinberg, B. Harsham, C. Forlines, and Z. Medenica. Contextual push-to-talk: shortening voice dialogs to improve driving performance. In *Proc. of the 12th Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI)*, pages 113–122, New York, NY, USA, 2010. ACM.