360° Panoramic Overviews for Location-Based Services

Alessandro Mulloni[1], Hartmut Seichter[1], Andreas Dünser[2], Patrick Baudisch[3], Dieter Schmalstieg[1]

[1] Graz University of Technology   [2] HITLab NZ, Univ. of Canterbury   [3] Hasso Plattner Institute
Graz, Austria   Christchurch, New Zealand   Potsdam, Germany

[lastname]@icg.tugraz.at, andreas.duenser@hitlabnz.org, patrick.baudisch@hpi.uni-potsdam.de
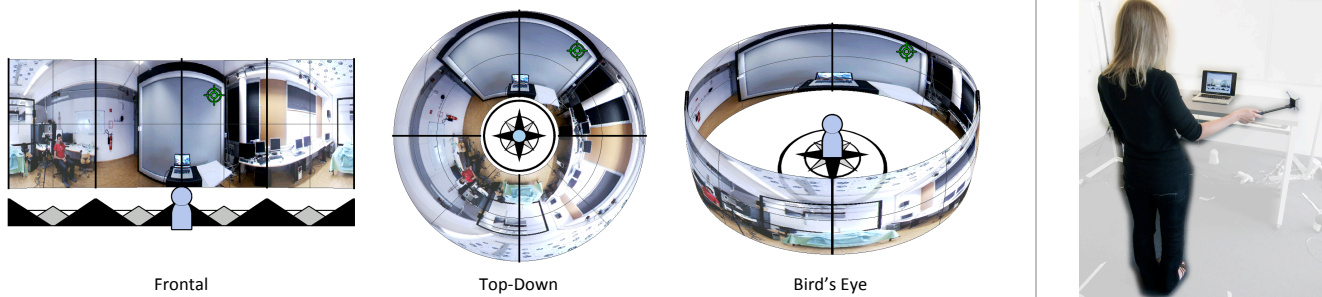
Frontal   Top-Down   Bird's Eye

Figure 1. Left: 360° panoramic overviews of a laboratory room. We consider three different shapes for visualizing the panorama. Right: the setup for our user studies consisted of a static display and a tracked wand (used for pointing in the environment).

ABSTRACT

We investigate 360° panoramas as overviews to support users in the task of locating objects in the surrounding environment. Panoramas are typically visualized as rectangular photographs, but this does not provide clear cues for physical directions in the environment. In this paper, we conduct a series of studies with three different shapes: *Frontal*, *Top-Down* and *Bird's Eye*; the last two shapes are chosen because they provide a clearer representation of the spatial mapping between panorama and environment. Our results show that good readability of the panorama is most important and that a clear representation of the spatial mapping plays a secondary role. This paper is the first to provide understanding on how users exploit 360° panoramic overviews to locate objects in the surrounding environment and how different design factors can affect user performance.

Author Keywords

Panorama; visualization; location-based service; magic lens

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces – Graphical User Interfaces (GUI)

INTRODUCTION

In this paper, we target users of location-based services (e.g., a navigation system) who are looking for objects within their visibility range. A typical scenario might require highlighting a building entrance. In such scenario, a panorama provides omnidirectional overview of the envi-

ronment. Further, since it is a first-person overview, users do not need to switch to other representations such as maps.

Panoramas can be rendered in any geometrical shape, but rectangles are typically used, similar to any other photograph. If we use the term *readability* qualitatively, as the ease of reading panorama elements, rectangles offer good readability of the panorama but do not clearly represent the fact that the environment surrounds the user: i.e., leftmost and rightmost points in the rectangle depict a part of the environment that is located behind the user. This paper is the first to investigate 360° panoramic overviews for users co-located with the panorama, rather than remote. In our scenario, users try to understand where points on the panorama are physically located in the environment. The spatial mapping between panorama and environment must be therefore clear.

In this paper we consider – in addition to *Frontal* – two shapes that provide a clearer representation of the spatial mapping: *Top-Down* and *Bird's Eye* (Figure 1, left). We conduct a sequence of user studies, asking participants to use panoramic overviews for locating points in the environment. Our results highlight that good readability of the panorama is the most important factor when a panorama is available, while a clear representation of the spatial mapping is advantageous when no panorama is available. We contribute to understanding how 360° panoramic overviews support users locating objects in their surroundings and how different design factors affect user performance.

RELATED WORK

Panoramas are widely used to explore remote locations, e.g. on 360 cities[1] or Google Street View[2]. Multi-perspective visualizations are also popular to gain overview on remote

---

[1] 360 Cities: http://www.360cities.net/

[2] Google Street View: http://www.google.com/streetview

paths. Zheng [3] presents route panoramas, a technique to join a sequence of images of a path into a single panoramic image. Similarly, Kopf et al. propose StreetSlide [4] where street-level images of a street's façades can be browsed as a single image strip. These works focus on the exploration of remote environments and do not aim at communicating the spatial mapping between visualization and environment.

First-person imagery can be advantageous in location-based services. Chittaro and Burigat [2] show that photos allow faster navigational decisions at road intersections, as compared to a map. Similarly, Google Maps for mobile[3] supports navigation with Street View images. Wither et al. [6] let users browse co-located panoramas through a *magic lens* – users physically turn the device in the environment to see corresponding portions of the panorama. Users successfully identify buildings in the environment from the panoramas, even if there is an offset between the users' location and the position in which the panorama was captured. In these works the panorama only provides a limited field of view, while our goal is to provide omnidirectional overview.

Baudisch et al. [1] target the necessity of overview for magic lenses in the context of online panorama capturing. They dynamically enlarge the field of view to fit the whole panorama, to help users capturing a gapless panorama. In a previous paper [5], we present a similar approach for augmented reality. We propose zooming between live video and an online-captured 360° panorama, showing that the overview is beneficial for complex tasks of spatial search in augmented reality. Similarly to this paper, both works provide omnidirectional overview for magic lenses. However, they visualize the panorama as a rectangle and give no clear indication of whether users understand the spatial mapping between panorama and environment from the rectangle.

DESIGN SPACE: CHOOSING ALTERNATIVE SHAPES

The choice of a shape for a panoramic overview is guided by two design factors: good *readability* of the panorama, and clear *representation of the spatial mapping* between panorama and environment. For our experiments, we considered three shapes (Figure 1). We also add further cues to each shape, visualizing an avatar, a wind rose and a grid.

Frontal. A rectangular shape provides good readability of the panorama, but no clear mapping to the environment.

Bird's Eye. A cylindrical shape maps panorama and environment in a direct way, but sacrifices readability. Due to the 3D view, the panorama is warped and occluded on the sides (we use 15% transparency to resolve occlusions).

Top-Down. A circular shape provides good readability of the panorama and direct mapping of the yaw. However, since the pitch is mapped to the distance from the center of the circle, the panorama is distorted for high pitch, and upside-down in the area of the panorama, which is behind the user.
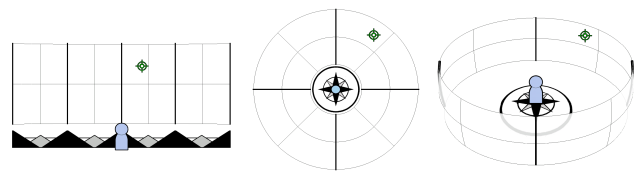

Figure 2. The three shapes in the condition with no panorama.

PILOT STUDY: USER STRATEGIES

We conducted a pilot study to gain first understanding on how users exploit panoramic overviews to locate points in the environment. We tested four conditions, defined by the variables Shape (*Bird's Eye* or *Frontal*) and Image (*without a panorama*, as in Figure 2, and *with a panorama* of the room, generated using a Ladybug camera[4], as in Figure 1).

We recruited 4 participants from our university. We showed them a sequence of panoramas, marked one point with a green crosshair, and asked them to point to the corresponding location in the environment. All participants used all conditions. We counterbalanced the order of conditions with a balanced Latin square and randomized the crosshair position.

To isolate and accurately measure comparative performance with the visualizations, we chose a controlled lab setup. We showed panoramas on a 13.3" screen, in a window of 1024×768 pixels. Participants pointed with a wand and then clicked a button on the wand (clicks were recorded by a laptop via a wireless receiver). We used a wand to acquire accurate motion and timing data via an infrared tracker[5].

Afterwards, we interviewed participants and asked them to describe the strategy they used, for each shape. From the interviews, we isolated two types of strategy*:* using grid and wind rose to determine *body-aligned directions* (left, right, front, back), and *visual matching* of objects in the panorama with corresponding objects in the environment.

We decided to conduct three experiments to isolate the effects of the two strategies: the first study with no panorama (no visual matching), the following two studies with a panorama (to support visual matching). During the three experiments, we collected questionnaire data on expertise with panoramas (browsing, capturing), maps and games (radars, 3D graphics). Expertise was balanced within the participants.

USER STUDY 1: ONLY THE SHAPE

In the first study, we evaluate how different shapes impact on the performance of the pointing task, assuming that no panorama is available (see Figure 2).

Independent variable. Shape: *Frontal*, *Top-Down*, *Bird's Eye*.

Hypothesis. We expected a difference in time: Top-Down < Frontal and Bird's Eye < Frontal, because Top-Down and Bird's Eye, in contrast to Frontal, provide a direct representation of the mapping between panorama and environment.
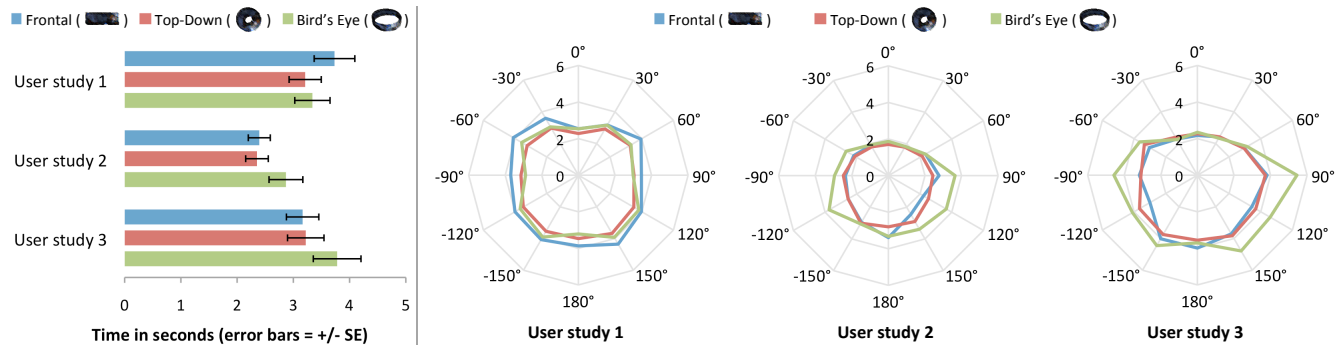

Figure 3. Left: mean time for each shape. Right: mean time for each shape and annotation angle (distance from center = seconds).

Experimental design. We selected 12 distinct intervals of size 30°, starting from [-15°, 15°]. The crosshair was assigned a random yaw angle within one of the intervals, and a constant pitch angle of 0°. Participants performed 5 repetitions of 3 (shape) × 12 (angle) trials, for a total of 180 measurements. We used balanced Latin squares to counterbalance the order of shape and angle intervals.

First, we introduced participants to the three experimental conditions. We clearly indicated to all participants the front, left, right and back lines in the visualization. Participants were allowed 6 practice trials for each shape. We instructed all participants to be as fast and accurate as possible but to give higher importance to accuracy. Finally, we conducted a short interview asking participants to describe the strategy they used to complete the task with each shape.

Apparatus. We used the same setup as in the pilot study.

Participants. We recruited twelve university students and staff (6m/6f), aged between 16 and 44 years (*mean* = 29.5, *sd* = 7.4). All participants used their dominant hand.

Results. For each of the 180 repetitions, we recorded task completion time (seconds) and unsigned error in yaw (degrees) between the target and the selection. For each shape × angle condition and participant, we calculated median time and error of the 5 repetitions.

Mean error values were 13.9° (*sd* = 4.0) for Frontal, 13.4° (*sd* = 2.6) for Top-Down and 14.1° (*sd* = 1.4) for Bird's Eye. A Friedman test did not show any effect of shape on error. In the following, we analyze time measurements under the assumption of comparable accuracy.

Mean task completion time (Figure 3, left) was 3.7s (*sd* = 1.3) for Frontal, 3.2s (*sd* = 1.0) for Top-Down, and 3.3s (*sd* = 1.1) for Bird's Eye. Since the data violates normality and sphericity (ANOVA requirements), we conducted a non-parametric Friedman test, which revealed a significant effect of shape on time ($\chi^2(2) = 16.67$, p < .001). Post-hoc Wilcoxon Signed Ranks tests with Bonferroni correction showed that all pair-wise differences are significant: Top-Down was significantly faster than Bird's Eye (p = .005) and Frontal (p = .003), and Bird's Eye was significantly faster than Frontal (p = .015). Bird's Eye was on average 3.5% slower than Top-Down; Frontal was on average 16.3% slower than Top-

Down and 12.3% slower than Bird's Eye. In Figure 3 (right), we can see that Frontal was generally slower than Top-Down and Bird's Eye, besides for target locations around 0°. A Friedman test shows no effect of repetition on time, highlighting no significant learning effect.

In the interviews, participants reported orientating to body-aligned directions (left, right, front, back) using grid and wind rose, and then refining the pointing direction. Participants reported imagining themselves "in the middle of the visualization" for Top-Down and Bird's Eye. For Frontal, more "thinking" with respect to body-alignment was required.

Discussion. Our results support our initial hypothesis, and highlight a further significant difference between Top-Down and Bird's Eye. In general, we see that, if no panorama is available, users rely on body-aligned reference lines to orient themselves and refine the orientation of annotations between these lines. Shapes which correctly represent the spatial mapping of the panorama to the environment result in significantly shorter task completion times.

USER STUDY 2: ADDING A PANORAMIC IMAGE
In the second study we evaluate how user strategies and performance change when a panorama is available.

Experimental design. We used the same setup and design as in study 1. However, in this study we used a panorama of the room in which the study took place (a quiet laboratory room, see Figure 1, left). The panorama was up-to-date with the room's appearance during the study. The panorama covers 360° in yaw, and [-45°, 45°] in pitch (in most real-world cases little information is present outside such range).

Hypothesis. We expected a difference in time: Top-Down < Bird's Eye and Frontal < Bird's Eye, because we expected good readability to be advantageous for completing the task and more advantageous than the strategy used in study 1.

Participants. 12 university students and staff (6m/6f), aged between 17 and 34 years (*mean* = 24.4, *sd* = 6.5). None of the participants had taken part in the previous study.

Results. In this experiment, we calculated error as the great-circle distance (to include error in pitch) in degrees between the target and the selection. Mean error was 16.5° (*sd* = 8.4) for Frontal, 15.7° (*sd* = 7.6) for Top-Down, and 17.1°

($sd = 8.0$) for Bird's Eye. A Friedman test showed no significant effect of shape on error.

Mean task completion time (Figure 3, left) was 2.4s ($sd = 0.7$) for Frontal, 2.4s ($sd = 0.7$) for Top-Down, and 2.9s ($sd = 1.0$) for Bird's Eye. A Friedman test revealed a significant effect of shape on time ($\chi^2(2) = 20.67$, p < .001). Post-hoc Wilcoxon Signed Ranks tests with Bonferroni correction showed that Bird's Eye was significantly slower than Top-Down (p = .002) and Frontal (p = .002). Bird's Eye was on average 21.9% slower than Top-Down, and 19.8% slower than Frontal. Friedman tests showed an effect of repetition on time, for all shapes (p < .001). However, Wilcoxon Signed Ranks tests only highlighted a significant learning effect between the first repetition and the others. This is not a problem for our analysis, since we perform it using the median values of the five repetitions.

In the interviews, participants reported looking for objects in the panorama and corresponding objects in the room (*e.g.*, a whiteboard, a fire extinguisher). For this strategy, most participants reported that Bird's Eye is inconvenient, as the sides are either not visible or warped and hard to see. One participant also reported issues with the backside of Bird's Eye, where the panorama appears mirrored. Figure 3 (right) illustrates this: we can see that Bird's Eye was slower than Top-Down and Frontal mostly around +/- 90°, in the warped or occluded regions.

Discussion. The results show that users adopt a strategy predominantly based on visual matching when a panorama is available, *independently* of the shape. As hypothesized, Top-Down and Frontal were significantly faster than Bird's Eye.

USER STUDY 3: VARYING PITCH ANGLE
In the third study, we aimed at corroborating the results from study 2 in the case of varying pitch angle.

Experimental design. Same as in study 2. In this study, the crosshair was assigned random pitch within [-45°, 45°].

Hypothesis. Same as in study 2: we expected a significant difference in task completion time: Top-Down < Bird's Eye and Frontal < Bird's Eye. This is because we expected visual matching to work also for non-zero pitch angles.

Participants. 12 university students and staff (6m/6f), aged between 23 and 50 years (*mean* = 30.1, $sd = 7.8$). None of the participants had taken part in the previous studies.

Results. Mean error was 15.0° ($sd = 5.2$) for Frontal, 16.6° ($sd = 4.7$) for Top-Down and 18.5° ($sd = 4.3$) for Bird's Eye. A Friedman test showed a significant main effect of shape on error ($\chi^2(2) = 12.17$, p = .002). Post-hoc tests showed a significant difference between Frontal and Bird's Eye.

Mean task completion time (Figure 3, left) was 3.8s ($sd = 1.5$) for Bird's Eye, 3.2s ($sd = 1.1$) for Top-Down, and 3.2s ($sd = 1.0$) for Frontal. A Friedman test revealed a significant effect of shape on time ($\chi^2(2) = 15.17$, p = .001). Post-hoc tests showed that Bird's Eye was significantly

slower than Top-Down (p = .002) and Frontal (p = .003). Bird's Eye was on average 16.1% slower than Top-Down, and 17.1% slower than Frontal. A Friedman test showed no effect of repetition on time.

In the interviews, reported strategies and issues with Bird's Eye were in line with study 2. Figure 3 (right) illustrates that Bird's Eye was again slower than Top-Down and Frontal mostly in the warped areas. Participants reported that Frontal was the easiest to find the pitch of points, whereas finding the yaw was considered harder. With Top-Down, participants had issues with the outermost third of the visualization (high pitch in the panorama), due to the strong distortion effect.

Discussion. The results of this study replicate the results of the previous study for the case of varying pitch angle.

DISCUSSION AND CONCLUSION
Our results show that, in the presence of a panorama, users mainly perform visual matching of objects in the panorama with correspondences in the environment. Consequently, good readability of the panorama has primary importance in the design of a panoramic overview. As a secondary strategy – and the main one when no panorama is available – users look for body-aligned directions within the visualization. A panoramic overview designed to represent the mapping between panorama and environment provides therefore a fallback, whenever a panorama is not available.

We are looking forward to apply this work to location-based services in outdoor scenarios. Here the environment is mutable and visual matching is not always feasible. Challenging scenarios are, for example, incomplete panoramas (as in our previous work [5]), outdated panoramas, or panoramas recorded from a slightly different location (as in Wither et al. [6]). In those scenarios, we will investigate how user strategies intertwine. Our results inform us on how design choices will affect support for visual matching and communicating body-aligned directions.

REFERENCES
1. Baudisch, P., Tan, D., Steedly, D., et al. Panoramic viewfinder: providing a real-time preview to help users avoid flaws in panoramic pictures. *Proceedings of OZCHI 2005*, (2005), 1–10.
2. Chittaro, L. and Burigat, S. Augmenting audio messages with visual directions in mobile guides. *Proceedings of MobileHCI 2005*, (2005), 107–114.
3. Jiang Yu Zheng. Digital route panoramas. *IEEE Multimedia 10*, (2003), 57–67.
4. Kopf, J., Chen, B., Szeliski, R., and Cohen, M. Street slide: browsing street level imagery. *ACM Transactions on Graphics (TOG) 29*, (2010), 96:1–96:8.
5. Mulloni, A., Dünser, A., and Schmalstieg, D. Zooming interfaces for augmented reality browsers. *Proceedings of Mobile-HCI 2010*, (2010), 161–170.
6. Wither, J., Tsai, Y.-T., and Azuma, R. Indirect augmented reality. *Computers & Graphics 35*, 4 (2011), 810–822.