Enhancing Handheld Navigation Systems with Augmented Reality

Alessandro Mulloni
Institute for Computer Graphics and Vision
Graz University of Technology
mulloni@icg.tugraz.at

Hartmut Seichter
Institute for Computer Graphics and Vision
Graz University of Technology
seichter@icg.tugraz.at

Dieter Schmalstieg
Institute for Computer Graphics and Vision
Graz University of Technology
schmalstieg@tugraz.at

MobileHCI 2011, Aug 30–Sept 2, 2011, Stockholm, Sweden.
ACM 978-1-4503-0541-9/11/08-09.

Abstract
We investigate the role of augmented reality (AR) as a new kind of handheld interface to enhance navigation. We integrate AR with other more common interfaces into a handheld navigation system, and we conduct an exploratory study to see *where* and *how* people exploit the AR interface. Based on previous work on augmented photographs, we hypothesize that AR is more useful as a support for wayfinding at static locations just before road intersections. In partial contradiction with our hypotheses, our results show that AR is used mostly while walking, usually shortly before *and* after road intersections. Our results help drawing considerations informing both the design of AR interfaces and the development of tracking technologies.

Keywords
Mobile augmented reality, navigation systems.

ACM Classification Keywords
H.5.1. Information interfaces and presentation: Artificial, augmented and virtual realities.

General Terms
Design, Experimentation.

Figure 1. (Top) Tilting the phone down, users access the map, glyphs and text instructions. (Bottom) Tilting the phone up, users access an augmented view with arrows that provide egocentric navigational hints.

Introduction

Studies that integrate map-based navigation systems with augmented photos show that photos enhance navigation. Chittaro *et al.* [1] demonstrate that augmented photos allow users to take correct decisions at road intersections significantly faster. While walking, most users rely on audio instructions and on glancing at the map. Hile *et al.* [3] highlight an "*80/20 rule*" followed by users, where 80% of the time is spent in the map mode and 20% ("*at critical points in the path*") using augmented photos. Walther-Franks *et al.* [6] confirm that users find it easier to orient with photos and that "*they [do] not consult the device all the time, but rather when it [is] necessary*". In general, these observations show that augmented views enhance navigation but users need to exploit multiple interfaces.

As also discussed by Hile *et al.* [3], photos do not always match the appearance of the environment due to its variability and they are rarely taken from the exact user's position. In contrast, *Augmented Reality* (AR) superimposes information by anchoring it directly to the live video from a camera. Feiner et al. [2] presented the first AR system for navigating a university campus. Since then, a number of head-worn AR navigation systems have been developed. More recently, handheld AR navigation systems appeared, such as MARA[1] or Wikitude Drive[2].

Since photos enhance navigation, we believe that AR can do at least as well – if not better – due to a closer match between the visualization on the display and the environment. However, AR requires highly accurate tracking of the handsets' position and orientation. Hence, correct augmentations are not always achievable in uncontrolled environments. Beyond, there is a lack of studies on how people actually use handheld AR for navigation. We lack knowledge of *where* people use AR, *e.g.* only at road intersections or anywhere, and *how* people use AR, *e.g.* standing still or walking. This knowledge is needed not only to inform the design of AR interfaces, but also to define the necessary improvements for underlying tracking technology.

In this work, we enhance a handheld navigation system with AR and we discuss results of an exploratory study, in which participants used our system to navigate a pre-defined outdoor path.

Implementation

We developed a multimodal navigation system (Figure 1). Similar to other navigation systems, we provide a *forward-up map* highlighting the user's position and the path to be followed. We also provide hints as *glyphs* and, to support eye-free usage, as *audio instructions*. Every new instruction is notified by the phone vibrating.

We additionally integrate an on-demand *AR interface*. We augment the environment with virtual arrows that indicate the direction the user should follow – when the arrow is outside the phone's camera view, we guide the user in turning the camera towards it. Similarly to our previous work [4], a tilting motion triggers a transition between map and AR: tilting the phone down shows the map, tilting it up transitions to AR. The system runs interactively on a smartphone. We track position using GPS and orientation using accelerometer and compass. In contrast to most commercial AR systems, if the user is standing still, we combine sensors and vision-based tracking [5] for a more stable orientation tracking.

---

[1] http://research.nokia.com/research/projects/mara

[2] http://www.wikitude.org/drivebeta

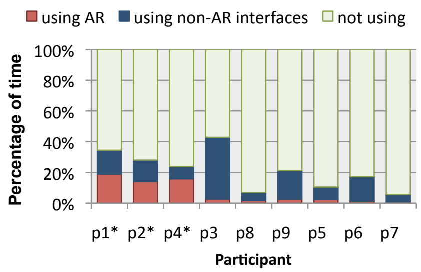

Figure 2. Percentage of time participants spent using the system (* indicates a participant with previous AR experience).

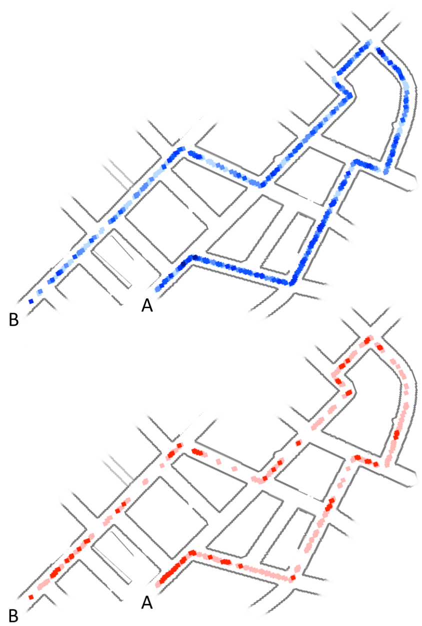

Figure 3. Locations in which non-AR (top) and AR (bottom) interfaces were used. Darker colors mean that more participants used the interface at the same location. Participants walked the path from A to B.

User study

We conducted an exploratory study of our system on a real-world navigation task. We looked closely at where and how participants exploit both non-AR and AR interfaces. In line with previous work on augmented photos, we hypothesized that AR, despite being available everywhere, is mainly used while standing still before road intersections (decision points on the path).

Nine people participated in the experiment (age 25–33, M=28.1, σ=2.6). Three had previous AR experience. No participant was familiar with AR navigation systems. Participants navigated a predefined path of 1.67 km (Figure 3). Throughout the study the GPS quality was usually excellent (*dilution of precision (DoP)* < 2) or good (2 < DoP < 5) and rarely moderate (5 < DoP).

After briefing participants on the study modalities, we walked them to the starting point of the path while they practiced with the system. We reminded participants to follow the path (no shortcuts) and use the device freely (not feeling forced to use it continuously). We followed and video recorded participants throughout the task. No participant exited the pre-defined path. After the task was completed, we collected subjective feedback through a semi-structured interview. We identified from the video recordings all *usage sessions* – sequences in which participants' eyes were continuously directed to the phone's screen. We then extracted all software logs that corresponded to the sessions. In the following, we present the results of our analysis of the software logs from all sessions.

Results

System usage averaged 21.2% (σ=12.6) of the overall task time. On average, 28.7% (σ=22.5) of the system usage was on the AR interface and 71.3% (σ=22.5) on non-AR interfaces. The average duration of a usage session, in which AR was used, was 4.8 seconds (σ=2.3), while sessions, in which AR was not used, lasted on average 1.8 seconds (σ=0.6). The latter sessions comprise usage of map, text instructions and glyphs. Figure 2 shows the percentage of usage time for each participant (p1-p9), distinguishing between the usage of AR and non-AR interfaces.

Previous AR experience. Participants with previous AR experience (p1, p2 and p4) exploited AR throughout the path (57.1% of system usage, σ=8.5), whereas all other participants used AR only a few times (14.5% of system usage, σ=7.5). Participants with previous AR experience justified the usefulness of the system for situations in which the turn to take was not clear (p1, p2), or when the signs with the street names were not visible (p4). Other participants commented that the map was sufficient (p6, p9) and more familiar (p5), it gave a better overview of the path (p3), or that the arrow visualization was not sufficiently stable (p8).

Where AR was used. Figure 3 shows where participants used the AR and non-AR interfaces. While the non-AR interfaces were used almost uniformly throughout the path, the AR interface shows less usage on straight path segments. In Figure 4, we look in higher detail at where participants with previous AR experience used AR and non-AR interfaces on average over all path segments – a path segment is one section of the path comprised between two consecutive road intersections. Both AR and non-AR usages increase when approaching the next intersection. Yet, while non-AR usage mildly increases in proximity of road intersections, AR usage shows a steeper curve with more usage just before an intersection (decision on the turn to take) and shortly after it (confirmation of being on the correct street).

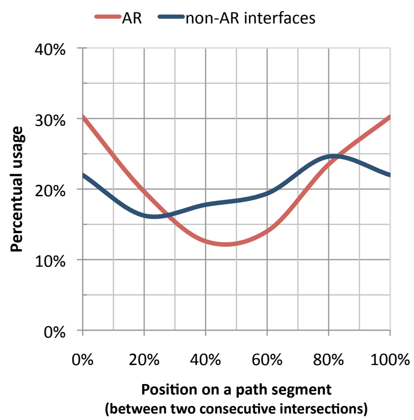

Figure 4. Average usage of AR and non-AR interfaces over path segments (between two consecutive road intersections).

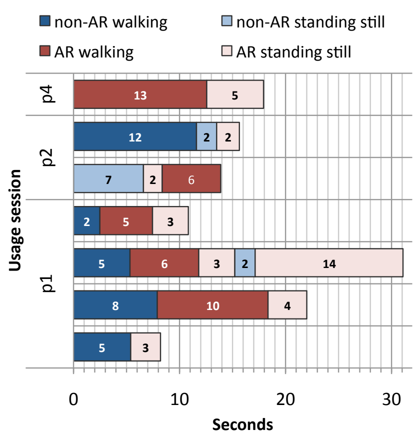

Figure 5. Usage sessions in which participants stopped walking and used the system while standing still (all timings are in seconds).

How AR was used. Both non-AR and AR interfaces were used almost only while walking. We observed a few cases of usage while standing still, presented in Figure 5. As can be seen, the system was usually used from a still position only after a failed attempt to use it while walking. Usage while standing still is probably only plausible at difficult intersections, in which the user does not succeed in making a decision on the fly while walking. In the observed cases, the interface used for making or confirming the final decision was always AR.

Affordance. During usage while walking, tracking relied on sensor data and caused visualization inaccuracies. Participants interpreted unintentional misplacements of the arrows as intentional instructions. For example, a participant interpreted a left-turn arrow with positional offset as an instruction to cross the street and turn left onto the opposite pavement. Participants interpreted errors in the orientation of the arrow as instructions to leave the pavement and walk on the street, or to move back from the street onto the pavement. Comments hint that the affordance of AR increased expectations on the accuracy of the visualized information.

Conclusion

In contrast to augmented photos, AR prompts usage while walking. AR was rarely used while standing still, usually after a failed attempt to make a decision while walking. Supporting a walking user with more accurate tracking is thus important, but as continuous, accurate tracking is a known hard problem, a more applicable solution is to inhibit usage while walking at the interface level. In general, our results show that users exploit AR mostly in proximity of road intersections: these are therefore the most important locations to support with accurate tracking. This can for example be achieved by feeding a tracker such as [5] with pre-recorded panoramas. Finally, tracking accuracy must be clearly communicated by the visualization, *e.g.* showing confidence intervals. Our results confirm that AR can be integrated with other common interfaces into a handheld navigation system. Yet, they also show that AR still needs more added value before it can, for inexperienced users, overtake other more common interfaces. We are positive that providing accurate tracking specifically at road intersection will enhance the value of AR and prompt more AR usage.

Acknowledgements

This work was supported by the Christian Doppler Laboratory for Handheld Augmented Reality.

References

1. Chittaro, L. and Burigat, S. Augmenting audio messages with visual directions in mobile guides. *Proceedings of MobileHCI 2005*, (2005), 107.
2. Feiner, S., MacIntyre, B., Höllerer, T., and Webster, A. A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Personal Technologies 1*, 4 (1997), 208-217.
3. Hile, H., Vedantham, R., Cuellar, G., et al. Landmark-based pedestrian navigation from collections of geotagged photos. *Proceedings of MUM 2008*, (2008), 145–152.
4. Mulloni, A., Dünser, A., and Schmalstieg, D. Zooming interfaces for augmented reality browsers. *Proceedings of MobileHCI 2010*, (2010), 161–170.
5. Schall, G., Mulloni, A., and Reitmayr, G. North-centred orientation tracking on mobile phones. *ISMAR 2010*, (2010), 267-268.
6. Walther-Franks, B. and Malaka, R. Evaluation of an Augmented Photograph-Based Pedestrian Navigation System. *Proceedings of SG 2008*, (2008), 94–105.