Playing it Real Again: A Repeated Evaluation of Magic Lens and Static Peephole Interfaces in Public Space

Jens Grubert and Dieter Schmalstieg
Institute for Computer Graphics and Vision
Graz University of Technology, Austria
[grubert | schmalstieg]@icg.tugraz.at

ABSTRACT
We repeated a study on the usage of a magic lens and a static peephole interface for playing a find-and-select game in a public space. While we reproduced the study setup and procedure the task was conducted in a public transportation stop with different characteristics. The results on usage duration and user preference were significantly different from those reported for previous conditions. We investigate possible causes, specifically the differences in the spatial characteristics and the social contexts in which the study took place.

Author Keywords
Augmented Reality; Static Peephole; Magic Lens; Field Trial; Proxemics; Social Context; In-the-Wild; Space; Place

ACM Classification Keywords
H5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous; H.5.2 [Interfaces and Presentation]: User Interfaces - Benchmarking

INTRODUCTION
The use of mobile Augmented Reality (AR) metaphors like magic lens (ML) for handheld devices has become increasingly popular for interacting with information situated in the physical world. Recently, researchers have begun to investigate the potentials and challenges of mobile AR outside of fully controlled laboratory conditions e.g., for understanding usage patterns [5], informing the design of novel interfaces [9], or conducting usability studies [2]. However, further evaluations are needed to better understand implications of the interplay between unique characteristics of mobile AR interfaces and performances in public space. For example, investigations of usage patterns and participant-audience interaction for mobile AR systems in public spaces share characteristics with evaluations of other interfaces such as mobile multimodal systems or large public displays.

MobileHCI '13, August 27 - 30 2013, Munich, Germany

ACM 978-1-4503-2273-7/13/08…$15.00.

Figure 1: The location of the study was a public transportation stop in Vienna, Austria.

Specifically, the effects of spatial configurations in public spaces and social contexts on the usage of interactive systems, which require rich spatial interactions (such as AR), need further exploration. Recently, a series of studies of an interactive installation [1] indicated the importance of social context over spatial factors in interacting with an installation in various locations. With this paper we add to the growing number of mobile system evaluations in-the-wild which take place under various public settings.

ORIGINAL AND REPEATED STUDY
In the original study we presented an evaluation on the usage of ML and static peephole (SP) interfaces for playing a find-and-select game in a public transport area and reported on the reactions of the public audience to participants' interactions [3].

One motivation of the original study was to investigate if and how individuals use a ML interface in public space if they can use an established SP interface instead. Both interface metaphors (ML and SP) have potential strengths and weaknesses regarding user acceptance which we discuss next. A potential benefit of using ML interfaces is the direct and expressive interaction with digital information connected to physical artifacts. This can lead to an increased engagement of users which is manifested in initial user feedback of first generation AR browsers [6] and the growing number of AR advertisement apps. However, this physical interaction can come at the costs of more frequent system errors (due to tracking failures), higher fatigue compared to screen based interaction, and a higher visibility of the interaction. From a spectator point of view this can be described as the discrepancy of visible rich spatial interactions (manipulations) of the user with a mobile device resulting in hidden effects [7]. This could lead to social acceptability issues [8] and unexpected (and unwanted) interactions between users and audience e.g., passers-by unknowingly walking into the interaction space of a user who is manipulating an augmented physical artifact. In contrast SP interfaces are an established way of interacting with purely digital content and are less subject to this manipulation–effect discrepancy as the manipulations (screen gestures), not only the effects, are less visible to spectators. However, they might engage people less in using an interactive system.

Further motivations of the original study included investigating audience reactions (and potential interactions with the users) and differences compared to a laboratory setting. To address these questions a find-and-select game was designed which allowed users to switch between the ML and SP interface at any time. Both quantitative and qualitative evaluation methods were used, including device logging (interaction durations, touch points, camera pose tracking), self-assessment questionnaires (for preference, ease-of-use, social presence, intrinsic motivation, future use), video-based interaction analyses (for analyzing interactions of passers-by and participants), field notes and semi-structured interviews. A between-subject design was used with location as the independent variable (levels: laboratory, public space) with relative usage duration (ML duration relative to the overall interaction duration) and preference as dependent variables of primary interest. Eight volunteers participated in each location. All participants of the public space condition were locals who had used this space before.

Findings of the study included that the ML interface was used significantly longer than the SP interface both in the public and in the laboratory condition. Furthermore, enjoyment and preference were significantly higher for ML compared to SP. Also preference rating for ML was significantly higher in the public group compared to the laboratory group. No other significant differences regarding usage duration or preferences between the two locations were reported. Furthermore, the interactions between passers-by and participants were very rare with 98% of people not noticing or briefly glimpsing while walking past the interacting participants.

We repeated the experiment at another public transportation stop in Vienna during two days in July 2012 (see Figure 1). The study design, procedure, and evaluation tools were reproduced (the camera location for recording participants and environment had to be adapted). Ten volunteers (5 females, 5 males, aged 19 to 37) participated in the study. They were acquired through social media channels and a social media company. Participants were locals from Vienna who used this transportation stop before to get to a nearby popular club. They received a small gift for participating.

FINDINGS

We report on our findings of the public condition in Vienna (PUV) and relate them to previous findings of the laboratory (LAB) and public (PUG) condition in Graz. For the between-subjects design (with location as factor with three levels), the collected data was not normal distributed (and could not be transformed to normal distributed data), thus we employed non-parametric null hypothesis testing. Two participants solely used the SP interface in PUV (one in LAB). Note, while we report based on data from all participants, also with those participants treated as outliers there were still main effects of location on the reported dependent variables.

Static Peephole was used most of the time

In the PUV condition the SP interface was used 56% of the time (over all participants) (PUG: 24%, LAB: 32%). A Kruskal-Wallis rank sum test indicated a main effect of location on usage duration (χ^2=26.72, p=1.5e-6). Pairwise comparisons using Wilcoxon rank sum test with Bonferroni correction indicated that ML was used significantly less in PUV (MD=0.36) compared to both PUG (MD=0.98, p=3e-6, r=.41) and LAB (MD=0.86, p=.001, r=.29) (see also Figure 2).

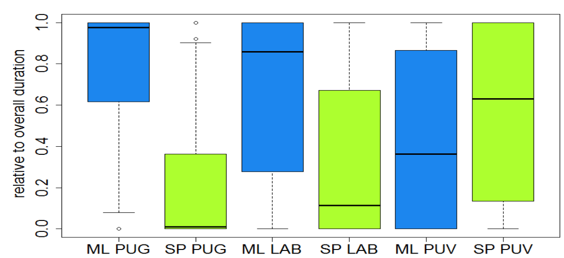

Figure 2: Relative usage durations for the magic lens (blue) and static peephole (green) interface for PUG, LAB and PUV.

Preference

A Kruskal-Wallis rank sum test indicated a main effect of location on enjoyment (χ^2=6.24, p=.04, “I enjoyed using the ML view in the environment”). Pairwise comparisons using Wilcoxon rank sum test with Bonferroni correction indicated that the rating was significantly lower in PUV (MD=3.5 on a 1 to 5 Likert scale) compared to PUG (MD=5, p=.027, r=.57).

Tracking Errors

Tracking errors result in a loss of augmentation and do occur even with state of the art tracking systems. In all conditions they occurred throughout the usage of the system. In total 147 (PUG 105, LAB 162) tracking errors occurred (14% of the overall gaming duration in ML mode, PUG 9%, LAB 7%). A Kruskal-Wallis rank sum test indicated no main effect of the location on the number of

Figure 3: Participant looking at two drunken men sitting on a nearby bench, who are chatting and watching the scene.

tracking errors per minute or per level but on the duration of tracking errors (χ^2=45.96, p=1e-10). Pairwise comparisons using Wilcoxon rank sum test with Bonferroni correction indicated that the durations of tracking errors were significantly different in PUV (MD=1.6 seconds) compared to PUG (MD=2.3, p=5e-4, r=.88) and LAB (MD=1.3, p=3e-11, r=.79).

Interactions with Passers-by

In the PUG condition only few interactions between passers-by and participants were observed. In PUV 241 (PUG 641) passers-by interactions were identified through video-based interaction analysis (using open and axial coding). From those, 50% (PUG 68%) were related to passers-by not noticing the participants or the recording setup. Twenty-two percent (PUG 30%) and 9% (PUG 2%) were staying and watching the participants actions for more than 5 seconds. However, in contrast to PUG, 15% of the interactions related to intrusions of the *social* space, as proposed by Hall [4] (see Figure 3), and 5% intrusion of the *personal space* (see Figure 4) of participants in PUV.

Reasons for Using ML and SP

Two participants used the SP interface exclusively in the main gaming phase in PUV (one in LAB). For one of those participants two men were sitting in the social space around him and were talking to each other during the whole duration of the game (see Figure 3). During the post-hoc interview he mentioned “this is not the prettiest and calmest environment” but then also mentioned that the ML mode is “a bit troublesome” due to the tracking errors. For the other participant a group of 6 men were standing in the public space (~10m away) and watching her while she was turning her back on them. However, in the post-hoc interview she argued that she found the ML interface “a bit more cumbersome … perhaps due to my height” [in relation to the poster]) not specifically mentioning the social context. Participants mentioned similar reasons for using the ML interface in PUV as in PUG and LAB conditions [3]. In addition, we employed a questionnaire similar to Rico and Brewster [8] to ask participants about locations in which they would use the interfaces.

Figure 4: Passer-by intruding the personal space of a participant, who is also watched by a woman sitting on a nearby bench.

No significant differences compared to the previous conditions were found. Also no coherent correlations between usage duration of the ML interface or preference could be identified.

DISCUSSION

The findings of the original study presented in [3] showed only minor differences in participants’ usage of the ML and SP interfaces for playing a find-and-select game between a public space and a laboratory setting. Both usage duration and users ratings indicated that participants preferred the use of the ML interface. However, this repeated study indicated significant different results both for usage durations and for preference rating compared to previous conditions, also when the participants who used SP exclusively were treated as outliers.

One challenge in conducting field studies is the potentially large number of confounding factors which can influence the evaluation outcomes. This impedes reliably identifying cause and effect relations for the outcomes of this study compared to previous conditions. However, while both public locations were transportation areas there were noticeable spatial and social differences between the two public locations which could have effects on the participants’ behavior.

The PUG condition was carried out a location primarily used as *transit area* for changing tram lines with the major waiting areas being more than 20m away (see Figure 5). It was located in a wide open space in the city center. People from all social contexts are using this place for changing trams. The general area is under video surveillance and the building at which the study took place was actively operated by the local tram company. It was a place with a high frequency of passers-by coming from several directions but only a few people were standing in the social space of the participants (rather walking behind the participants, see Figure 5).

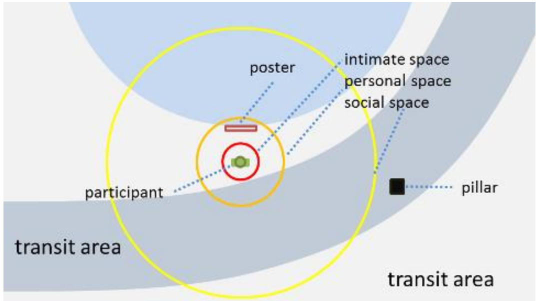

Figure 5: Schematic top down view of the space in the PUG condition with Hall's reaction bubbles [4] indicating the intimate (0.5m), personal (1.2m) and social space (3.6m) of the participants.

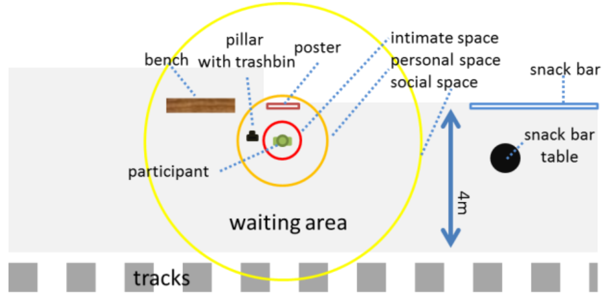

Figure 6: Schematic top down view of the space in the PUV condition. Participants were not constantly at a fixed position but eventually moved back and forth ~ 50 cm.

In contrast the location in the PUV condition was primarily used as *waiting* area for people coming from the exit of a near-by metro line (see Figure 6). It was located in a disadvantaged area (Vienna Leopoldstadt). Comments of participants about the "shabby" area and experimenter's observations indicate that there might have been a larger *social distance* between participants (mostly middleclass, students) and people with lower socioeconomic status present at the tram stop compared to PUG. In addition, while the number of people identified during the video analysis (in a similar timeframe) was 2.5 times less than in the PUG condition a larger amount of people were intruding the social space of participants for longer periods of time. Specifically, in the PUV condition 9 of 10 participants could see people sitting on a nearby bench (~2m away) in their periphery (and those people could watch them - see also Figures 3, 4, and 6).

Those differences between the locations could indicate that the social context in PUV could have inhibited the use of expressive, socially not common spatial gestures used in the ML interface, which is supported by the observations in Akpan et al. [1]. Still, there are other potential factors which could have influenced participant's behavior and ratings. They include fatigue, the perceived severity of tracking errors, the role of personality (e.g., intro- and extraversion), intrinsic motivation to use the interfaces, the novelty of the ML metaphor and demand characteristics.

CONCLUSION

We repeated a study on the usage of a magic lens and a static peephole interface for playing a find-and-select game at a public transportation stop. While the study design and procedure were reproduced the spatial characteristics and social context of the location of the study differed from the previous public condition. Significant differences both for the usage duration and the preference were found compared to previous runs of the experiment. Specifically, the magic lens interface was used significantly less and preferred less compared to a previous public condition. Qualitative data analysis indicated that the social context could have influenced the choice of interfaces.

Still, further repetitions of the study should be conducted to better understand individual factors which influence AR interaction in the public and interrelations between them. We also need more reliable combinations of evaluation methods targeting the study of expressive interactive systems in-the-wild.

REFERENCES

1. Akpan, I., Marshall, P., Bird, J. and Harrison, D. Exploring the effects of space and place on engagement with an interactive installation. In Proc. CHI '13, (ACM 2013), 2213-2222.
2. Dünser, A., Billinghurst, M., Wen, J., Lehtinen, V. and Nurminen, A. Exploring the use of handheld AR for outdoor navigation. In *C&G 36, 8 (2012)*, 1084-1095.
3. Grubert, J., Morisson, A., Munz., H and Reitmayr, G.: Playing it Real: Magic lens and static peephole interfaces for games in a public space. In *Proc. MobileHCI 2012*. ACM (2012), 231-240.
4. Hall, E. The Hidden Dimension. Anchor Books (1966) ISBN 0-385-08476-5.
5. Morrison, A., Mulloni, A., Lemmelä, S., Oulasvirta, A., Jacucci, G., Peltonen, P., Schmalstieg, D. and Regenbrecht, H. Mobile augmented reality: collaborative use of mobile augmented reality with paper maps. In *C&G. 35, 4 (2011)*, 789–799.
6. Olsson, T. and Salo, M. Narratives of Satisfying and Unsatisfying Experiences of Current Mobile Augmented Reality Applications. In *Proc. CHI 2012*. ACM (2012), 2779-2788.
7. Reeves, S., Benford S., O'Malley, C. and Fraser, M. Designing the spectator experience. In *Proc. CHI 2005*, (ACM 2005), 741-750.
8. Rico, J. and Brewster, S. Usable gestures for mobile interfaces: evaluating social acceptability. In *Proc. CHI 2010*, (ACM 2010), 887–896.
9. Sá, M. and Churchill, E. Mobile augmented reality: exploring design and prototyping techniques. In *Proc. MobileHCI 2012*. ACM (2012), 221-230.