Impact of Item Density on Magic Lens Interactions

Michael Rohs
Deutsche Telekom Laboratories
TU Berlin
Ernst-Reuter-Platz 7
10587 Berlin, Germany
michael.rohs@telekom.de

Johannes Schöning
German Research Center for Artificial Intelligence (DFKI)
Stuhlsatzenhausweg 3
66123 Saarbrücken, Germany
josc03@dfki.de

Robert Schleicher
Deutsche Telekom Laboratories
TU Berlin
Ernst-Reuter-Platz 7
10587 Berlin, Germany
robert.schleicher@telekom.de

Georg Essl
Deutsche Telekom Laboratories
TU Berlin
Ernst-Reuter-Platz 7
10587 Berlin, Germany
georg.essl@telekom.de

Anja Naumann
Deutsche Telekom Laboratories
TU Berlin
Ernst-Reuter-Platz 7
10587 Berlin, Germany
anja.naumann@telekom.de

Antonio Krüger
German Research Center for Artificial Intelligence (DFKI)
Stuhlsatzenhausweg 3
66123 Saarbrücken, Germany
krueger@dfki.de

ABSTRACT
We conducted a user study to investigate the effect of visual context in handheld augmented reality interfaces. A dynamic peephole interface (without visual context beyond the device display) was compared to a magic lens interface (with video see-through augmentation of external visual context). The task was to explore objects on a map and look for a specific attribute shown on the display. We tested different sizes of visual context as well as different numbers of items per area, i.e. different item densities. We found that visual context is most effective for sparse item distributions and the performance benefit decreases with increasing density. User performance in the magic lens case approaches the performance of the dynamic peephole case the more densely spaced the items are. In all conditions, subjective feedback indicates that participants generally prefer visual context over the lack thereof. The insights gained from this study are relevant for designers of mobile AR and dynamic peephole interfaces by suggesting when external visual context is most beneficial.

Categories and Subject Descriptors
H.5.2 Information Interfaces and Presentation: User Interfaces—*input devices and strategies, interaction styles*.

General Terms
Human Factors, Experimentation.

Keywords
Magic lens, dynamic peephole, small displays, mobile devices, camera phones, visual search.

1. INTRODUCTION
Mobile devices provide a convenient way to augment existing static information with dynamic and personalized content. For example, large paper maps are already available in public spaces but they only provide static information that is intended for broad use by a general audience. Mobile devices can provide specific content to the map dynamically and hence increase the value of such maps for navigation and exploration. Yet, mobile devices have limited screen space and hence do not provide good overview over large visual information areas. Combining the advantages of large scale paper maps and small dynamic displays has the potential to overcome both problems [6]. In this magic lens [2] approach the integrated camera of mobile phones is used to precisely track the device's position over the map and show additional information over the video stream in real time [7]. For example, a standard city center map could provide selected points of interest, such as nearby coffee places or museums, as well as parking lots and their associated parking costs (Figure 1).

MobileHCI'09, September 15 - 18, 2009, Bonn, Germany.
ACM 978-1-60558-281-8.

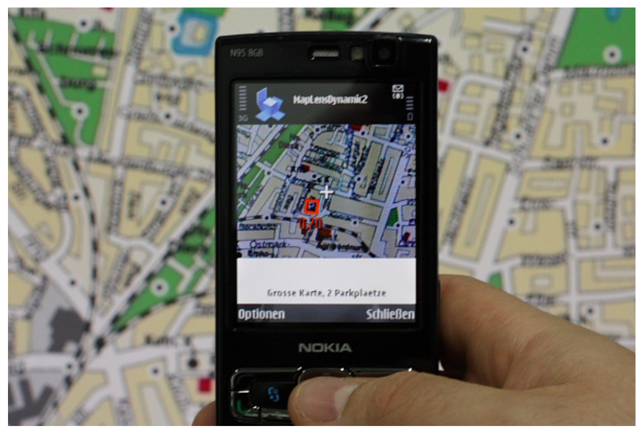

Figure 1. Camera view augmented with prices for parking lots. The blue parking signs are visible on the background map. The overlay graphics are generated by the phone.

A previous study on exploring maps with mobile devices [8] compared the performance of traditional joystick navigation (static peephole [3]), position-tracked navigation without visual context (dynamic peephole [3]), and position-tracked navigation with visual context (magic lens [2]). In the dynamic peephole case, panning is controlled by moving the device and the map information is only presented on the device display. In the magic lens case, map information is available on both the device display and the paper map. The magic lens provides video see-through augmentation of the external map. In the previous study, the two position-tracked interfaces outperformed the static peephole navigation method (joystick), but the magic lens interface (with visual context) surprisingly was not significantly faster than the dynamic peephole (without visual context).

The goal of this work is to explore the effect of visual context for different presentation sizes and item densities for this kind of handheld device interface. Our initial hypothesis was that the item density would affect to what extent users take advantage of the information that is provided in the background. Specifically we hypothesized that the effectiveness of visual context decreases as the density increases. We expected that for lower densities the magic lens condition (with visual context) would outperform the dynamic peephole condition (without visual context). The earlier experiment suggested that there is a density limit above which users will only use the device display and not switch their visual attention to the background. The results of this study can help to decide whether it useful to offer visual context in the background or just use a dynamic peephole.

2. RELATED WORK

Camera-equipped mobile devices can be used as a see-through tool [2] to augment background surfaces, such as paper maps, posters, or electronic displays. When the device is held above an object or surface, visual features in the scene are highlighted and additional information is overlaid in real-time on the device's display (Figure 1). The term magic lens [2] has been coined to describe this type of multi-layer interface in analogy to a reading or magnifying glass.

Whereas magic lens interfaces are based on the idea of real-time augmentation of the real world scene, dynamic peephole interfaces [3,9] denote a class of interface where the viewport of a mobile device is used as a window into a virtual space and no visual context is available outside the device display. As an example of a dynamic peephole interface, Yee [9] prototyped a spatially aware calendar application.

Magic lens interfaces with external context offer a particularly promising kind of interaction, since they allow for augmenting large-scale visual displays, such as paper maps, with private and up-to-date information on the personal device. Magic lens interfaces are becoming more and more ubiquitous. The *Wikitude* project (http://www.mobilizy.com/wikitude.php) is just one example of a magic lens that is now available for end-users. *Wikitude* is a mobile travel guide that uses location-based Wikipedia content. A user can apply the mobile device display like a magic lens and explore nearby geoferenced Wikipedia features overlaid on camera view of the mobile device.

In contrast to the more general work outlined above we investigate the effects of density and size of visual background for magic lens interaction. This study enables us to formulate guidelines for the usage of a magic lens interface and predict the performance gain that can be expected by providing visual context in a given situation.

3. USER STUDY

The user study investigated the effects of the presence of visual context, its size, and item density on completion time, error rate, and satisfaction in a basic search task. In the magic lens case, the items were visible on the background, but the attribute to look for was only available as a textual overlay generated by the phone. The aim was to simulate searching for dynamic information that is typically not present on a static background.

3.1 Participants and Apparatus

The study was conducted with 17 participants, 12 female, 5 male. They were university students aged 20-31 years (mean age 26.4 years). None of the participants was familiar with the city map. The experiment was performed on a Nokia N95 8GB camera phone. The client application showed the augmented view of the map and captured all user interactions and movements with timestamps. The background was displayed on a Barco LCN-42 LCD screen (42", 1920x1080 pixels, 934x527 mm). An abstract color pattern served as a tracking background for the condition without visual context. We designed it in such a way that the color and brightness range is similar to what one would find in a map (Figure 2, bottom). The pattern thus is a way to implement the dynamic peephole interface as a baseline condition. For the condition with visual context a colored city map was shown (Figure 2, top). For the large size the background filled the whole area (0.492 m^2) of the 42" display (Figure 2, right). For the small size half of the display area was used (0.246 m^2) (Figure 2, left). On these areas parking lot symbols were randomly distributed. The same real-time tracking method was used in all conditions [7]. It analyzes the video stream, computes the focus position on the background and the perspective mapping from the background coordinate system to the current camera frame. It provides graphical overlays with pixel-level accuracy, has an average frame rate of 8 Hz, and a delay of 170 ms. This provided good responsiveness for our purpose.

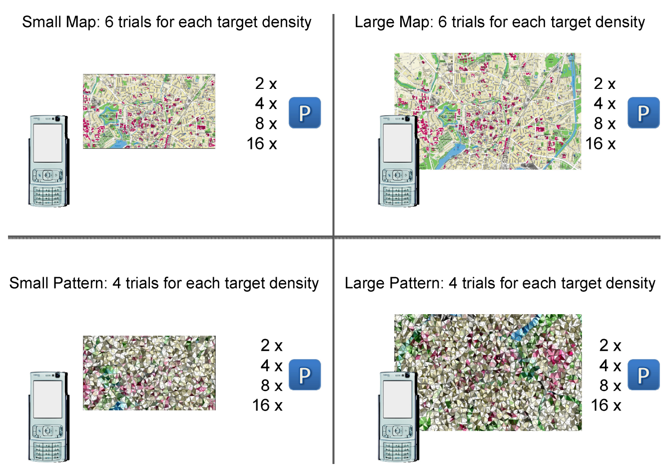

Figure 2. The experimental conditions: map or pattern as background, background size, number of P symbols.

3.2 Study Design

The study was set up as a 2x2x4 within-participants design with the following factors (Figure 2):

- Context information: (1) with visual context (city map) or (2) without visual context (abstract pattern)
- Size: (1) large (full area) or (2) small (half area)
- Number of items: 2, 4, 8, or 16

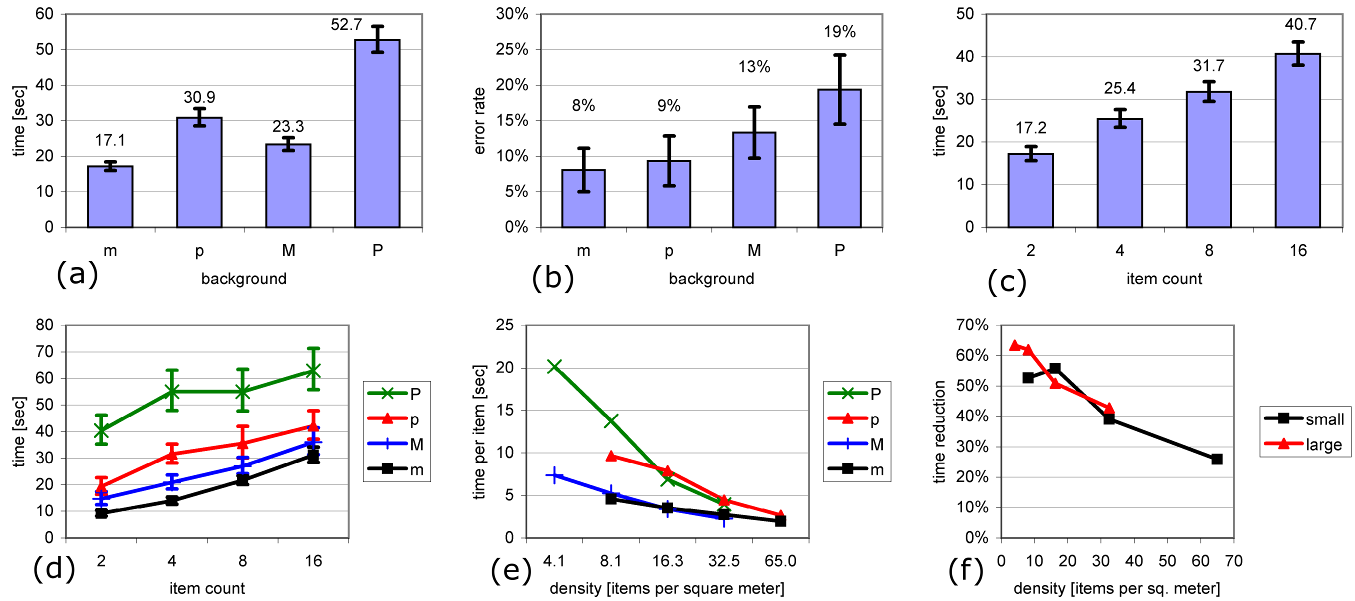

Figure 3. Results: (a,b) time and error rate by background, (c,d) time by item count, (e) time by item density, (f) time reduction by item density. m = small map, M = large map, p = small pattern, P = large pattern.

3.3 Tasks and Procedure

To cover a typical task for mobile map interaction we chose an object locator task, which is described as a fundamental task in the literature [5]. The general scenario for all conditions was that users had to find the cheapest among a given number of parking lots on the map (blue parking signs). For the conditions with visual context the parking signs were visible in the background, but the price for parking was only visible on the phone (Figure 1). The positions of the items as well as the parking prices (with varying minima) were randomly changed after each trial.

A single trial consisted of scanning the map in the defined condition and finally selecting the target. At any time the item closest to the screen's center was highlighted with a red frame and selected when the user pressed a button. After each selection the participants were informed about the success of their choice and the next trial could be started. After finishing six trials per item density in the magic lens condition or four trials per item density in the condition with no visual background, a screen informed the participants about the next condition. A previous study yielded that the search strategies for the pattern conditions are quite uniform and took longer than the map condition [8]. Hence, in order to keep the time for participants in a sensible range, we decided to use only four trials per pattern condition leading to an asymmetric design.

Initially, participants were given a short written description of the experiment and the instruction for the subsequent task. After that for each condition a 5-7 minute practice period followed for navigating with the mobile phone lens. After each block, participants were asked by the experimenter how they managed to use the phone lens and how they liked navigating with the phone lens. The order of blocks was counterbalanced.

3.4 Results

Participants had no difficulty using the provided interaction techniques to locate the items on the map and pattern. Trial times and error rates were the main performance measures taken. A histogram of trial time suggests that the data is log-normally distributed. Hence all means, confidence intervals, and ANOVAs are computed on the log-transformed data. For the sake of clarity the graphs show the retransformed means. Outliers of more than 3 standard deviations from the mean are excluded. 11 outlier trials were removed in this way.

The overall time per trial, measured from the start of a trial until a selection was made, was 26.7 s (95% confidence interval: 26.2-28.6 s). This measure includes correct and false selections. If the user did not select the cheapest parking lot in a trial, then this was counted as an error. The overall error rate was 12.4% (95% confidence interval: 10.5-14.2%). A repeated-measures mixed linear models ANOVA [4] (participants modeled as random effects, context, size, and count modeled as fixed effects) shows main effects for all factors: availability of visual context: $F_{1,216} = 141.2$, $p < 0.001$; background size: $F_{1,178} = 56.4$, $p < 0.001$; item count: $F_{3,367} = 59.9$, $p < 0.001$. Mixed linear models typically have larger denominator degrees of freedom than traditional ANOVA, but this does not lead to easier detection of significance due to wider confidence intervals. Figures 3a and 3b show the average trial times and error rates by background type. The times are pairwise significantly different (Sidak-adjustment for multiple pairwise comparisons). The error rates for the small map (8%) and pattern (9%) are comparable, those for the large map and the large pattern increase to 13% and 19%, respectively. Providing visual context for the small size thus reduces the search time by 44.4%. For the large size, the reduction is 44.2%.

Grouping the results by item count (Figure 3c) shows that the search time increases with the number of items. There is an interaction effect between the availability of visual context and the number of items ($F_{3,342} = 5.2$, $p = 0.002$). This suggests that the growth rate of search time with increasing item count depends on the availability of visual context. Figure 3d shows the search time per item count broken down by background type. As expected, the large pattern (“P”) takes longest, followed by the small pattern (“p”), the large map (“M”), and the small map (“m”).

The average density of items on the background (number of items divided by background size) was thus for the large background 4.1, 8.1, 16.3, and 32.5 items per m^2 and for the small background 8.1, 16.3, 32.5, and 65.0 items per m^2. Figure 3e shows that the search time per item decreases with increasing density. This is as expected, because the higher the density, the smaller the area the user has to scan in order to find an item. Irrespective of the background size, for the overlapping densities (8.1, 16.3, 32.5) the search times per item for the conditions with visual feedback (magic lens, "m", "M") are very close. This is also the case for conditions without visual feedback (dynamic peephole, "p", "P"). The overall times for the dynamic peephole are higher than for the magic lens, but their performances converge as density increases. ANOVA results support this interaction between density and availability of visual context.

For the highest density (65.0) there is only a small performance difference between magic lens and dynamic peephole. This result is in line with the abovementioned previous experiment [8], which did not find a significant difference in search times for a background of size A3 and a density of 137 items per m^2.

The interaction between density and visual context means that the time reduction that can be expected from using visual context (i.e. using a magic lens interface rather than a dynamic peephole interface) decreases as density increases. This is shown in Figure 3f. For densities below 20 items per m^2, the time reduction is above 50%.

A questionnaire subsequent to the experiment revealed that participants liked the conditions with external visual feedback better than without, regardless of background size.

4. CONCLUSION

This paper presented a study on the effects of visual context for magic lens and dynamic peephole interactions in a basic object locator task on paper maps. The main factors tested were the availability of visual context, the size of the context, and the number of items users had to investigate. In the case with visual context the items were visible on the background surface, but the attribute to look for was only available via the magic lens. Users in this case had the option to visually scan for the items on the background or to use the magic lens. For deciding whether the item was the right one, they had to inspect it on the mobile device.

We found that the usefulness of visual context for quickly finding items of interest does not primarily depend on context size, but on the density of the items. Visual context is most effective for sparse distributions. The denser they are distributed, the less clear the performance benefit that can be expected from providing visual context. One reason for this seems to be that for high densities it is more likely that the next item to inspect already appears on the display, hence making a switch of visual attention to the background unnecessary.

High densities also result in lower average distances between the items. Thus the next item may already appear in the visual periphery of the user, even though it is not yet located on the device display. In such cases there seems to be a tradeoff between shifting one's gaze from the device display to the background and moving the hand. For relatively close items moving the hand towards the item in the visual periphery may be the more efficient strategy, compared to switching visual attention from the device display to the background. Switching attention for visual search on the background incurs some cost, because of the need to refocus to the new layer of presentation at a different distance. Moreover, it has been found that people have difficulties to disengage attention from objects that are near their hands [1]. Further research is needed to clarify this tradeoff.

Both subjective (questionnaire) as well as behavioral data (search time and error rate) show that mobile navigation interfaces may benefit from a magic lens option to interact with public maps, especially if theses maps are supposed to cover a large area or items are distributed sparsely.

Given these results it is advisable to constrain the item density in magic lens interfaces. This can be achieved, for example, by performing suitable pre-filtering of information categories to limit the number of candidate items. When the density is too high, no significant performance benefits can be expected from external visual context, although visual context is preferred in this case as well.

5. REFERENCES

[1] Abrams, R. A., Davoli, C. C., Du, F., Knapp, W. H., and Paull D. 2008. Altered vision near the hands. Cognition, 107(3), 1035-1047.

[2] Bier, E. A., Stone, M. C., Pier, K., Buxton, W., and DeRose T. D. 1993. Toolglass and magic lenses: The see-through interface. In SIGGRAPH '93, 73-80. ACM.

[3] Mehra, S., Werkhoven, P., and Worring. M. 2006. Navigating on handheld displays: Dynamic versus static peephole navigation. ACM Trans. Comput.-Hum. Interact. 13(4), 448-457.

[4] Quenée, H. and van den Bergh, H. 2004. On multi-level modeling of data from repeated measures designs: A tutorial. Speech Communication, 43 (2004), 103-121.

[5] Reichenbacher. T. 2001. Adaptive concepts for a mobile cartography. Journal of Geographical Sciences, Acta Geographica Sinica, 11:43-53.

[6] Reilly, D., Rodgers, M., Argue, R. Nunes, M., and Inkpen, K. 2006. Marked-up maps: Combining paper maps and electronic information resources. Personal Ubiquitous Comput., 10(4):215-226.

[7] Rohs, M., Schöning, J., Krüger, A., and Hecht, B. 2007. Towards real-time markerless tracking of magic lenses on paper maps. Adjunct Proceedings of the 5th Intl. Conference on Pervasive Computing, 69-72.

[8] Rohs, M., Schöning, J., Raubal, M., Essl, G., and Krüger, A. 2007. Map navigation with mobile devices: Virtual versus physical movement with and without visual context. In ICMI '07, 146-153. ACM.

[9] Yee, K.-P. 2003. Peephole displays: Pen interaction on spatially aware handheld computers. In CHI '03, 1-8. ACM