Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia (MUM '16), December 12–15, 2016, Rovaniemi, Finland

Indoor Pedestrian Navigation Systems - Is More than One Landmark Needed for Efficient Self-Localization?

Christina Bauer
Chair of Information Science
Regensburg, Germany
christina2.bauer@ur.de

Manuel Müller
Chair of Information Science
Regensburg, Germany
manuel-tonio.mueller@ur.de

Bernd Ludwig
Chair of Information Science
Regensburg, Germany
bernd.ludwig@ur.de

ABSTRACT
Research examining pedestrian navigation systems that use landmarks to explain routes became popular in the past years. Nevertheless, it is still an open question how many landmarks should be depicted at once. In this paper a user study is presented that evaluates two different indoor navigation system designs that depict either one (N = 63) or four (N = 60) landmarks to guide the user. The time it took the participants to recognize where to go was captured as a dependent variable. Results show that the interface only depicting one landmark leads to faster self-localization. Therefore, it is argued that a pedestrian navigation system should mainly depict one highly salient landmark in a navigation instruction in order to keep navigation efficiency high.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation (e.g. HCI): Ergonomics; H.1.2. User/Machine Systems : Human information processing

Author Keywords
Pedestrian Navigation; Landmarks; Indoor; Wayfinding

INTRODUCTION
In any navigation scenario, referring to so-called landmarks is considered to be the most effective way to communicate a navigation instruction (see e.g. [5, 10, 16]). Landmarks support users during the wayfinding process, i.e. the repeating task of orienting oneself in the environment and navigating through space from one point to another [18]. Besides navigation in outdoor areas,indoor navigation became an important research field, since evidence exists that people tend to get lost easily in public buildings like airports or universities [6]. In this context, we identified an open research question concerning the design of indoor pedestrian navigation systems. Given that a comprehensive set of landmarks is available, different interface designs can be taken into account to guide a pedestrian. Existing approaches range from augmented reality depictions to plain text instructions. Nevertheless, most indoor pedestrian navigation systems only implicitly refer to landmarks or focus on the depiction of the users' surroundings. If landmarks are used, commonly several objects are referred to. We argue that it is furthermore important to examine how many landmarks should be displayed in a navigation interface. Therefore, we implemented two system prototypes that vary in the amount of depicted landmarks in a very reduced mobile map representation. Subsequently, we conducted a user study that focused on the question if it is useful to depict comparatively many landmarks. This research is based on the idea that depicting more landmarks supports users during the self-localization task as they can use the landmarks as "anchor points" and therefore could be potentially more successful in wayfinding. However, it is also important to note in this case that some studies showed that e.g. map interfaces can as well be distracting and a source of cognitive load [9, 12, 14]. As an interface showing several landmarks can be seen as a very reduced map, these UI elements could also be an obstacle for a fast and easy self-localization. Consequently, we compared a prototype showing only one landmark to a depiction of four salient objects in order to examine the relationship between this interface variation and task success. The remainder of this paper is structured as follows. First of all, we report on the related work concerning landmark depiction. In the following section our study on the display of landmarks is described and the results are presented and discussed. Finally, our findings are summed up and discussed.

MUM '16, December 12 - 15, 2016, Rovaniemi, Finland

ISBN 978-1-4503-4860-7/16/12…$15.00
DOI: http://dx.doi.org/10.1145/3012709.3012728

RELATED WORK
We analyzed the related work on indoor and outdoor pedestrian navigation with several questions in mind. First of all, we distinguished how the landmarks were selected, depicted and how many salient objects are referenced. Moreover, we noted how the different systems were evaluated.

In outdoor areas, most of the studies compare mobile map depictions to other designs such as augmented reality interfaces or virtual reality views. For instance, in [12] an augmented reality interface enhanced with directional arrows is compared against a mobile map depiction. Landmarks were only implicitly used by depicting street names. The results revealed that the map was used during the whole navigation task, whereas the augmented view mainly served for orientation at intersections. As the sample size in this experiment was quite small, data like the task completion time was only evaluated qualitatively. Another example for a study comparing a photo realistic depiction against two dimensional maps is presented in [14]. Once again, landmarks are only implicitly included by showing pictures of the current navigation scene or positioning markers on the map. The amount of depicted landmarks varies. This study mainly evaluates usability aspects of the map and the photo realistic system. The results showed that the latter provides a better user experience than mobile map interfaces. However, the authors also state that these elaborate depictions can be a source of cognitive load during the wayfinding task. In contrast to the work mentioned above, [1] evaluated pictures of landmarks to map depictions as well, but explicitly referred to one landmark at one particular navigation point (e.g. a building) in addition to depicting street names. In this study only qualitative data was gathered. In indoor environments, most of the studies examine the suitability of a mobile map depiction for the wayfinding task. Also in this context, landmarks are often only implicitly referred to. For instance, [4] realized a map prototype for a fair. Salient objects were depicted by showing the numbers of all stands. The time it took the participants to fulfill small wayfinding tasks such as finding an object in the environment was measured. The authors conclude that people had problems to identify the salient objects with their map interface. Similar findings were made by [15], who implemented a prototype for a shopping mall. In addition to [4], they depicted landmarks of a certain category (like shops or toilets) if the users had selected it. Moreover, the landmark selection was not user-centered but based on available data about shops and other facilities. Similar to outdoor environments, some research deals with the design of augmented user interfaces. [11] have implemented an application that depicts panorama pictures enhanced with directional arrows and achieved high usability values. The interface did not refer to landmarks.

All in all, the work presented above showed that the research focus in the design of pedestrian navigation systems is to vary either different mobile map depictions or implement photo realistic views. In this paper, we examine different mobile map representations to address our research question how many landmarks should be displayed. Whether our findings still apply for e.g. augmented reality interfaces is a topic of further research. Moreover, landmarks are either only implicitly displayed or the selection of the objects was done subjectively by the test supervisors. The amount of depicted objects varies from only one salient object to all possible landmarks in the users' surroundings. Therefore, we argue that the work presented in this paper is a first step to fill this research gap. In addition to usability ratings, most of the studies took the time participants needed to accomplish the whole task into account. In our work we examined the subtask of self-localization during wayfinding, which is described in detail in the next section.

STUDY ON THE DEPICTION OF LANDMARKS

The next subsections outline our study design and the results.

Interface and System Prototype

The route visualization of both system prototypes is based on an annotated graph that includes information about possible routes and available landmarks along the path. The sketch-like graph representation of both interfaces (see Figure 1) is based on the idea of [15], who argue for abstracting the depiction of

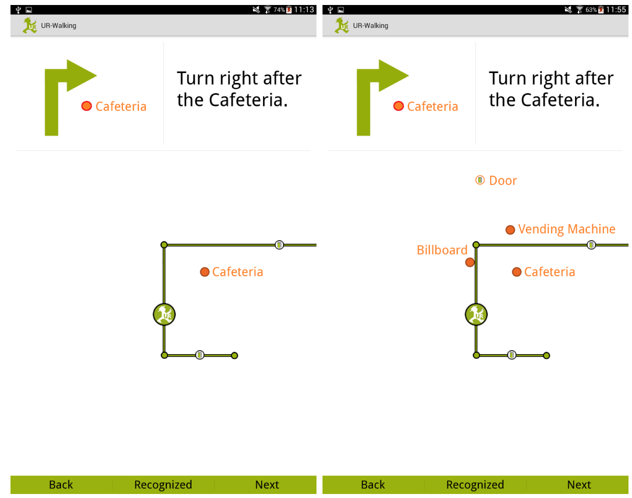

Figure 1. Screen shot of the interface prototype showing only one landmark (left) and displaying four landmarks (right). Participants received all instructions in German.

the surroundings of the users. Moreover, the interface is quite similar to the route depiction of [3], who showed that users effectively reach their destination with this depiction. Generally, the interface is divided into three parts. The main element in the middle of the screen shows the graph-like depiction of the particular navigation scene. We already conducted studies that showed that floor maps can be a source of distraction and therefore we did not include them in our prototype (see [13]). Either one or four landmarks are displayed in the graph-like fragment. The respective landmarks for each scene are displayed using an icon (for doors and stairs) or a circular marker labeled with the name of the particular object (see e.g. Figure 1 that shows the landmark "Cafeteria" and a picture of the scene in Figure 2). The estimated current position of the user is indicated by a green manikin located on the route segment. The zoom level for both interface designs was fixed to the same level. The text label of the landmarks was always 0.5 cm high to ensure readability. No localization technique was available for this experiment. The navigation instructions only switched when the "Next" button described below was clicked. Moreover, major difficulties occurred concerning the use of the initial compass to rotate the map according to the users current walking direction. Due to architectural specifics of the university buildings the sensor accuracy decreased to an unacceptable level. Therefore, the interface was oriented head-up according to the findings of [17]. A text instruction is shown at the top right of the UI. Each instruction was generated manually and was equally structured. It contains a reference to one landmark and a direction instruction, i.e. "right", "left", and "straight ahead". Next to the text instruction an arrow enriched with a landmark icon is displayed for every route section. The icon is positioned relatively to the arrow according to the current route segment. It is positioned at the arrow head if the user has to go through a door or up the stairs. This interface element is motivated by the observation of [9], who argue that this depiction sums up all information relevant to the current wayfinding task. The text instructions and the directional arrows did not differ among the user groups. Three buttons are

Figure 2. Route of the user study with assigned step number. The arrows indicate where the users saw the navigation instruction. The exemplary scene shows navigation step 2 with assigned landmarks.

positioned underneath the text (labeled "Back", "Recognized", and "Next"). They are important for experimental setup and therefore their detailed function is explained in the experimental set-up section. The experiment took place in the University of Regensburg. The test route included three level changes, four transitions of buildings and nine turns of direction (total length = 410 meters, see Figure 2). It contained different indoor areas such as small office corridors (steps 18-22), bigger hallways (steps 4-8,11,13-17) and halls (steps 1-3,9-10,12).

Participants and Devices

Students of an undergraduate course in software ergonomics were instructed to recruit test users for the study described in this paper. This resulted in a test sample of 123 participants. 66 male and 57 female persons participated in the study, most of them being students (82%). Their mean age was 22.8 years with a standard deviation of 3.1 (range:16-39). The tests were conducted with a Samsung Galaxy Tab 3 10.1. We chose a comparatively large display size (10 inches), so that the test supervisors could examine the actions of the participants at any time. 63 participants conducted the test with the interface showing only one landmark, whereas 60 test persons used the interface depicting 4 salient objects.

Landmark Selection

We identified 22 steps where a navigation instruction was needed (e.g. at intersections and floor changes.) Both interfaces included one "main" landmark which was derived from a pre-study with 87 participants (44 males, mean age = 23,1 years, 87% students). Participants had to name salient objects at each decision point along the test route and rate their salience according to the questions described in [8] on a 5-point Likert scale. First of all, we generated the subset of all landmarks that were rated 4.0 or higher. Afterwards, the landmark that was named by the majority of the participants was selected as the "main" landmark (e.g. "Cafeteria" in Figure 1). The main idea of our study is that depicting more landmarks could support users during self-localization. The prerequisite for this assumption is that all depicted landmarks are salient to users in the particular navigation situation. Therefore, the additional landmarks are also derived from the subset mentioned above. At several navigation steps only four suitable landmarks were available. Therefore, we decided to fix the amount of displayed landmarks to four for every navigation step. Consequently, the second interface depicts the four most salient landmarks including the "main" landmark. All in all, the average salience rating for the main landmark did not differ significantly from the salience of the additional landmarks (Wilcoxon-Mann-Whitney test, $p = 0.94$). Most of the main landmarks were doors or stairs. Five landmarks referred to rooms such as lecture halls that have an unambiguous name or a cafeteria. The additional landmarks for the interface depicting four objects also included signs, billboards and furniture.

Experimental Set-Up

The participants were randomly assigned to one of the prototypes. At the beginning, test persons were asked to fill in a form concerning their demographic data. After that, they were instructed how the prototypes work. For this purpose, they were guided to the starting point of the route and the test supervisor used the first screen to explain the functionality. Consequently, all data concerning the first screen was excluded from statistical analysis. In particular, it was emphasized that the participants had to indicate whether they understood where they have to go by clicking on the "Recognized" button. This time was captured as our main dependent variable. We logged the timestamps of every user interaction an could therefore calculate the recognition time, i.e. the time elapsed between seeing the instruction and clicking on the button. The recognition time for each step significantly correlates with the task completion time for the step (Spearman’s rank correlation, $r > 0.8$, $p < 0.001$). It has to be noted that data concerning the last step was not taken into account, as both interfaces only referred to the destination room number. Moreover, the test persons had to click a "Next" button to see the next navigation instruction. This button could not be clicked until the "Recognized" button was pressed. It was clarified that they had to click the "Next" button as soon as they had reached the destination of the current instruction. For instance, when the instruction *"Go ahead through the door"* appeared, the button had to be pressed immediately after the participant has passed the door. The "Back" button should only be pressed if the participants completely lost their way. The experimenter observed the participant from a distance, did not give any hints, and noted whether they understood the instructions and pressed the button at the right location. Normally, pressing the button at the wrong location caused that the participants lost their way. The participants performed the navigation task without any time constrains. Most of the participants were familiar with at least one of the buildings of the test route (mean 3.9 on a 7-point Likert scale with 1 = "not familiar at all" and 7 = "very familiar"). In order to avoid the possibility that they simply navigate to the destination without the help of the prototype, the participants were not told which room they are going to.

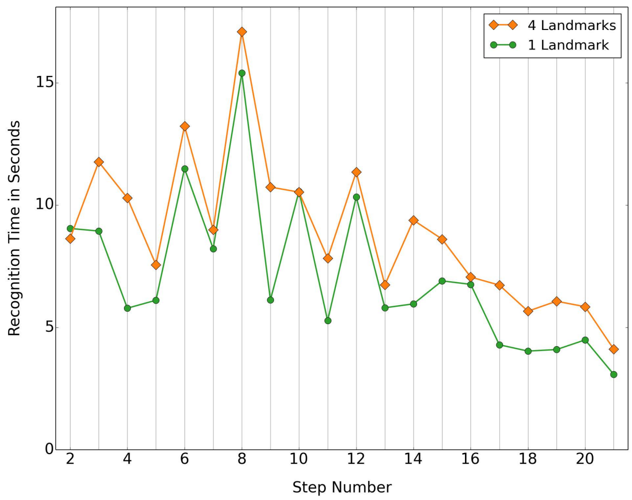

Figure 3. Means of the recognition times in seconds for each navigation step.

Results

The overall aim of this study is to examine whether indoor pedestrian navigation systems should display more than one landmark. As stated in the introduction, depicting more than one object could help to support the self-localization task necessary for wayfinding. Nevertheless, the related work also shows that displaying more information in a pedestrian navigation system could also enhance cognitive load and therefore decrease navigation efficiency. Consequently, we formulated an undirected hypothesis. First of all, we examined whether participants perform differently with the two interfaces in terms of overall task completion time. As already mentioned this is a popular procedure in similar studies (see e.g. [7]). On average it took the participants 6 minutes 22 seconds to accomplish the task (SD = 1.63). The data for both UI groups was not normally distributed, therefore a Wilcoxon-Mann-Whitney test was conducted. Concerning this variable no significant differences could be found ($p = 0.21$, mean 1 landmark = 6 minutes 12 seconds, mean 4 landmark = 6 minutes 34 seconds). Moreover, it was noted at which points participants lost their way, i.e. took a wrong turn or asked for help. In this case, the test supervisor guided the test person to the next decision point. 18 out of the 123 participants got lost at completely different steps (all except for 5,12,14,15,17,18). This happened nine times for both interface prototypes. Therefore, it cannot be stated if errors are caused by the interface or the particular navigation situation. The time it took the participants to recognize where to go to was captured as our main dependent variable. As already mentioned, this measurement is not influenced by walking speed and therefore is better suited to examine the suitability of different interface designs for navigation. The normality assumption had to be rejected for both groups using the Shapiro-Wilk test. A subsequent Wilcoxon-Mann-Whitney test revealed significant results ($Z = -5.966$, $p < 0.001$). Obviously, participants who navigated with the interface showing only one landmark could recognize faster where they had to go to (mean 1 LM = 7.0 seconds, mean 4 LM = 8.9 seconds). Figure 3 depicts the results for each navigation step.

CONCLUSION AND FUTURE WORK

In this paper, we examined whether depicting four landmarks in an indoor pedestrian navigation system better supports self-localization than only displaying one landmark. The selection of the landmarks for this examinations was based on a pre-study. Our results show that depicting only one landmark leads to faster self-localization. We used the time it took the participants to realize where to go to as our main dependent variable. For this purpose, participants had to click on a "recognized" button.

Nevertheless, it is still an open question whether depicting two or three landmarks compared to one has a different effect on self-localization. Based on our observations we propose the hypothesis that adding more landmarks is not helpful for navigation. The additional landmarks seem to be distracting and do not contribute to an efficient and effective self-localization. Moreover, it is not absolutely clear whether the interface element showing the landmarks was of use for the participants at all. For this purpose, we want to conduct similar studies and examine the visual attention of the participant using e.g. eye tracking. By this, we could also explore in detail whether the participants clicked the buttons needed for navigation at the intended location and if this has an influence on task completion. Furthermore, we could examine, for instance, how long test users looked at the text labels, the text instruction, the map, etc. The results also indicate that the steps differ in perceived difficulty (see Figure 3). It is a topic of further investigation which factor influence this user behavior. There may be a lot of potential paths to take an therefore, recognition time may increase. Moreover, different landmark salience levels may have an influence on self-localization efficiency. A navigation scene could only have very unsalient objects that are hard to recognize. On the other hand there may be a lot of very eye-catching landmarks that distract the user. This could not be examined in our study, since all objects chosen for the experiment were approximately equally salient. Other important factors are user characteristics. In our experiment, e.g., most of the participants were familiar with the building and therefore had knowledge about its basic structure. Therefore, the experiment should be repeated with absolutely unfamiliar persons. In this context, it may help to include more information in the interface. Moreover, the users' sense of direction might also have an influence (see e.g. [2]). For our study we chose to use a tablet as a test device in order to be able to observe the participants more closely. It is more likely that users would navigate with a smart phone. Consequently, future studies have to examine whether device size may have a influence on navigation efficiency and landmark perception. Moreover, it has to be analyzed whether our results apply for other interface designs like augmented or virtual reality views. Such interfaces generally depict more information than "traditional" map views. Therefore, referring to more than one landmarks could be helpful. Finally, we are currently working on confirming our results using a reliable indoor localization technique. This would make navigation without pressing the "Next"-button possible and therefore eliminate the influence of the possibility that participants receive the navigation instruction at different positions.

REFERENCES

1. Ashweeni Kumar Beeharee and Anthony Steed. 2006. A Natural Wayfinding Exploiting Photos in Pedestrian Navigation Systems. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '06)*. ACM, New York, NY, USA, 81–88.
2. Stefan Bienk, Markus Kattenbeck, Bernd Ludwig, Manuel Müller, and Christina Ohm. 2013. I Want to View It My Way: Interfaces to Mobile Maps Should Adapt to the User's Orientation Skills. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia (MUM '13)*. ACM, New York, NY, USA, Article 34, 9 pages. DOI: http://dx.doi.org/10.1145/2541831.2541841
3. Stephany Bigler, Annina Brügger, Fiona Utzinger, and Kai-Florian Richter. 2014. *Spatial Cognition IX: International Conference, Spatial Cognition 2014, Bremen, Germany, September 15-19, 2014. Proceedings*. Springer International Publishing, Cham, Chapter Up, Down, Turn Around: Assisted Wayfinding Involving Level Changes, 176–189.
4. Anders Bouwer, Frank Nack, and Abdallah El Ali. 2012. Lost in navigation: evaluating a mobile map app for a fair. In *Proceedings of the 14th ACM international conference on Multimodal interaction*. ACM, 173–180.
5. Claus Brenner and Birgit Elias. 2003. Extracting landmarks for car navigation systems using existing GIS databases and laser scanning. *International archives of photogrammetry remote sensing and spatial information sciences* 34, 3/W8 (2003), 131–138.
6. Beatrix Brunner-Friedrich and Verena Radoczky. 2006. *Visual Information and Information Systems: 8th International Conference, VISUAL 2005, Amsterdam, The Netherlands, July 5, 2005, Revised Selected Papers*. Springer Berlin Heidelberg, Berlin, Heidelberg, Chapter Active Landmarks in Indoor Environments, 203–215.
7. Luca Chittaro and Stefano Burigat. 2005. Augmenting Audio Messages with Visual Directions in Mobile Guides: An Evaluation of Three Approaches. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices &Amp; Services (MobileHCI '05)*. ACM, New York, NY, USA, 107–114. DOI: http://dx.doi.org/10.1145/1085777.1085795
8. Markus Kattenbeck. 2015. Empirically Measuring Salience of Objects for Use in Pedestrian Navigation. In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '15)*. ACM, New York, NY, USA, Article 3, 10 pages.
9. Christian Kray, Christian Elting, Katri Laakso, and Volker Coors. 2003. Presenting Route Instructions on Mobile Devices. In *Proceedings of the 8th International Conference on Intelligent User Interfaces (IUI '03)*. ACM, New York, NY, USA, 117–124.
10. Andrew J. May, Tracy Ross, Steven H. Bayer, and Mikko J. Tarkiainen. 2003. Pedestrian navigation aids: information requirements and design implications. *Personal and Ubiquitous Computing* 7, 6 (2003), 331–338.
11. Andreas Möller, Matthias Kranz, Stefan Diewald, Luis Roalter, Robert Huitl, Tobias Stockinger, Marion Koelle, and Patrick A. Lindemann. 2014. Experimental Evaluation of User Interfaces for Visual Indoor Navigation. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3607–3616.
12. A. Mulloni, H. Seichter, and D. Schmalstieg. 2011. User experiences with augmented reality aided navigation on phones. In *Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on*. 229–230.
13. Christina Ohm, Manuel Müller, and Bernd Ludwig. 2015. Displaying landmarks and the user's surroundings in indoor pedestrian navigation systems. *Journal of Ambient Intelligence and Smart Environments* 7, 5 (2015), 635–657.
14. Timo Partala and Miikka Salminen. 2012. User Experience of Photorealistic Urban Pedestrian Navigation. In *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '12)*. ACM, New York, NY, USA, 204–207.
15. Arto Puikkonen, Ari-Heikki Sarjanoja, Merja Haveri, Jussi Huhtala, and Jonna Häkkilä. 2009. Towards Designing Better Maps for Indoor Navigation: Experiences from a Case Study. In *Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia (MUM '09)*. ACM, New York, NY, USA, Article 16, 4 pages.
16. Martin Raubal and Stephan Winter. 2002. Enriching Wayfinding Instructions with Local Landmarks. In *Geographic Information Science: Second International Conference, GIScience 2002 Boulder, CO, USA, September 25–28, 2002 Proceedings*, Max J. Egenhofer and David M. Mark (Eds.). Springer, Berlin, Heidelberg, 243–259.
17. Nanja J. J. M. Smets, Guido M. te Brake, Mark A. Neerincx, and Jasper Lindenberg. 2008. Effects of Mobile Map Orientation and Tactile Feedback on Navigation Speed and Situation Awareness. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '08)*. ACM, New York, NY, USA, 73–80.
18. Stephan Winter, Martin Raubal, and Clemens Nothegger. 2005. Focalizing Measures of Salience for Wayfinding. In *Map-based Mobile Services: Theories, Methods and Implementations*, L. Meng, T. Reichenbacher, and A. Zipf (Eds.). Springer, Berlin, Heidelberg, 125–139.