Representing Indoor Location of Objects on Wearable Computers with Head-Mounted Displays

Markus Funk[1], Robin Boldt[1], Bastian Pfleging[1], Max Pfeiffer[2], Niels Henze[1], Albrecht Schmidt[1]
[1]University of Stuttgart, Pfaffenwaldring 5a, 70569 Stuttgart, Germany
[2]Leibniz University Hannover, Appelstrasse 9A, 30167 Hannover, Germany
{markus.funk, bastian.pfleging, niels.henze, albrecht.schmidt}@vis.uni-stuttgart.de,
boldtrn@studi.informatik.uni-stuttgart.de, max@hci.uni-hannover.de

ABSTRACT

With head-mounted displays becoming more ubiquitous, the vision of extending human object search capabilities using a wearable system becomes feasible. Wearable cameras can recognize known objects and store their indoor location. But how can the location of objects be represented on a wearable device like Google Glass and how can the user be navigated towards the object? We implemented a prototype on a wearable computer with a head-mounted display and compared a last seen image representation against a map representation of the location. While we only found a significant interaction effect, all objective and subjective measures generally favor the last seen image. Results suggest that a map representation is more helpful for gross navigation and an image representation is more supportive for fine navigation.

Author Keywords

Location Representation; Indoor Location; Visualization; Real-World Search; Wearable Computing; Augmented Reality

ACM Classification Keywords

H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces

INTRODUCTION

Probably everybody experienced losing an object without being able to recover it immediately. To prevent this, people tend to arrange objects according to the tasks the objects are used for [9] to have them at hand when needed. For instance, keys are often stored at a dedicated place at home, which is checked before leaving the apartment, to easily pick up the keys. But what happens if an object cannot be found immediately? A system could provide help by navigating the user towards the lost object.
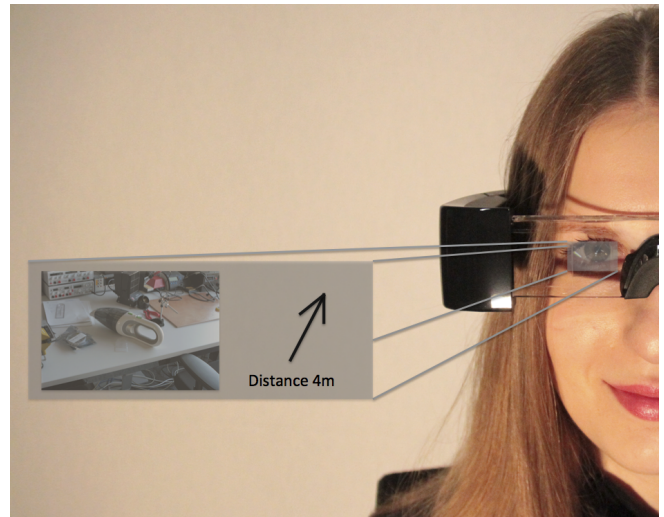
Figure 1. A user wearing our object finder prototype. The prototype displays the directions towards a lost object and a last seen image of it. The prototype comprises a wearable, sensor-equipped computer and head-mounted see-through glasses.

In 2002 Kindberg et al. [8] proposed separating physical entities into people, places, and things. To find the location of people and places there are already many different systems available. For finding people, mobile applications like Google+ on Android or Find Friends on iOS enable viewing the position of a friend on a map. In order to discover places nearby based on a user's GPS position, applications like Yelp or Foursquare assist users in finding a location. But when it comes to locating the position of ordinary physical objects out-of-the-box solutions are not yet available.

Although, systems for finding objects in the physical word have been proposed, representing their indoor location is a yet under-explored area. Traditional visual representations, e.g. marking an object's position on a map have proven to be a good choice for finding outdoor location [13], but might not be the best choice for indoor positions.

The contribution of this paper is twofold: (1) We introduce a "last seen image" representation of objects' locations for a wearable device (see Figure 1). (2) We conducted a user study to find the best representation of objects' indoor locations on a wearable computer with a head-mounted display.
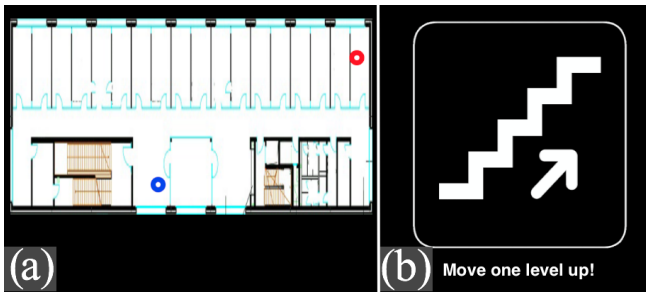
Figure 2. (a) The 2D map representation on the glasses. The blue dot is the user's position, the red dot is the target. (b) The icon that is displayed if the user needs to change the level.
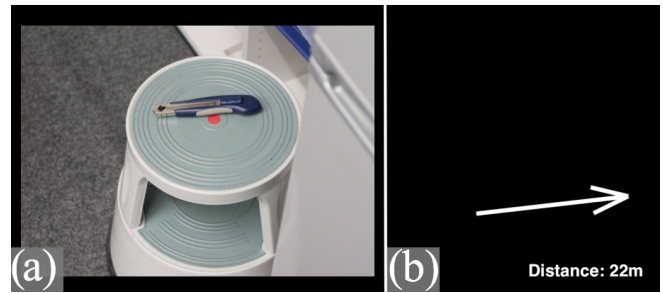

Figure 3. The last-seen image representation on the glasses. (a) Last-seen image of the sought object. (b) Arrow pointing towards the target when the user is more than 5 meters away form the target.

RELATED WORK

Previous research has proposed different approaches for systems that help a user with finding lost objects in the physical world. The challenges for building such systems are (1) identifying an object and retrieving its indoor or outdoor location, and (2) representing the results in a way the user can easily understand. Therefore, we divided the related work in two according categories.

Identifying Objects

There are several ways to automatically detect the presence of an object in an instrumented area. Previous work equipped important objects with barcodes [10] or AR-markers [4] to make them searchable. A camera is able to visually detect the markers and can thereby identify an object. Another technique is to equip searchable objects with radio frequency identification (RFID) tags. In Magictouch [12], a user is equipped with a wearable RFID reader, which sends an object's location to a system, whenever a tagged object is touched. However, those techniques are intrusive and not scalable as each object has to be equipped with a marker or tag. The Antonius system [5] equips the user with a wearable camera, which is constantly recognizing previously registered objects based on their visual appearance. Therefore the Antonius system can keep track of objects' positions without having to equip each object with a marker.

Representing Location

The representation of a location is important to being able to switch quickly from the digital representation to the real world. Representing location can be done *physically*, *symbolically*, *absolutely*, or *relatively* [7]. The representation of one or multiple *physical* or *absolute* locations on a zoomable map is the most common way to visualize location [13]. Usually, location-aware applications also represent the user's current position with a symbol. The Halo interface [3] represents locations that are outside the screen with a red semicircle. The radius of the semicircle indicates the user's distance to the target location. A large radius indicates a large distance and vice versa. A location can also be represented *relatively* to well-known or eye-catching reference points as for instance landmarks [1]. Also, a *symbolic* representation using a city name or an address is possible. Such representations are mostly used to represent outdoor-location. However, the scope of an object finding system is mostly indoors.

SYSTEM

We implemented a prototype to compare different representations of object's locations using an Epson Moviero BT-100[1] as a wearable computer with see-through head-mounted display. To manipulate the indoor location by a wizard-of-oz, we connected a Google Nexus S which streams latitude, longitude, floor level and azimuth [2]) using a simple UDP protocol. The head-mounted display then renders a visual representation from the streamed data. As the focus of this paper is representing the location, we assume to have visually identified the object before as suggested by Funk et al. [5] and have the indoor location present.

LOCATION REPRESENTATIONS

We implemented two representations of object's locations to run on the previously introduced system: A traditional map representation and a novel image-based representation. A stairwell icon telling the user to move upstairs or downstairs is displayed on both representations if the user is on the wrong floor (see Figure 2 (b)).

2D Map

The map representation is showing a 2D map of the current floor (see Figure 2 (a)). The target is marked on the 2D map with a red dot. The user's position is denoted as a blue dot. The user can observe the two dots while walking and thus find the right way to the sought object. This representation is both used for navigating the user towards the sought object and finding the object inside a room.

Last Seen Image

As an alternative representation, we introduce the last seen image. It displays the last captured picture in which the sought object was recognized by the camera of the wearable system (see Figure 3 (a)). The image also shows surrounding objects or furniture and thereby provides contextual information about the location of the sought object. As the last seen image works best when the user is inside the room where the last seen image was taken, we display an arrow pointing towards the direction of the sought object and displaying the distance (see Figure 3 (b)) if the user is more than 5 meters away from the object.

---

[1]**www.epson.eu/ix/en/viewcon/corporatesite/products/mainunits/overview/11373** (last access 2014-02-21)

[2]horizontal angle measured clockwise from true north

EVALUATION

We conducted a study to compare the last seen image representation to the 2D map representation on building-level.

Method

The study was conducted using a repeated measures design with the two location representations as independent variable. The conditions were performed in a counterbalanced way using Balanced Latin Square. As objective measures, we used the time the participant needs to find an object (task completion time). The task completion time was measured from the moment the participant left the starting point until the sought object was touched. For collecting subjective feedback, we used the NASA TLX [6] and the SUS [2] questionnaires after each condition. Additional qualitative feedback was collected after both conditions.

Apparatus

The study was conducted using the wearable prototype mentioned before. During a pre-study, we found out that the WiFi-based indoor positioning was not very stable at our location and might influence the user's perception of the representation. Therefore, the representation is manipulated by the experimenter in a wizard-of-oz manner, who was following the participant during the search task (see Figure 4).

Procedure

After the participants were introduced to the search task, they were equipped with our wearable prototype. When the participant felt comfortable with using the representation, the participant was instructed to find 3 objects using each representation one after another. The objects and representations were counterbalanced in each condition. The objects were distributed across a building consisting of three floors. The starting point was in the middle of the second floor. For each object the starting point was the same. To make the participant familiar with the sought object, the experimenter showed a picture of the sought object to the participant at the beginning of each search. As soon as the participant starts walking, the experimenter takes the time from the starting point until the moment the sought object is touched. After an object was found, the participant had to go back to the starting point. As the WiFi-based indoor positioning is too erratic and therefore might influence the user's perception of the representation, the gyroscope and the indoor location are manipulated by the experimenter as a wizard-of-oz. The experimenter accompanied the participant at all time to accurately react to position changes and to take the task completion time (see Figure 4). After the participants found three object using one representation, they were asked to fill the SUS and TLX questionnaires.

Participants

For our study, we recruited 16 participants (9 male, 7 female) aged between 15 and 42 years ($M = 25.31$, $SD = 6.55$) through our mailing lists and from our campus. Most of the participants were students from various fields comprising computer science, engineering and business administration. Additionally an administration secretary and a pupil took part in our study. 12 out of 16 participants stated that they are searching for a lost object more than once per week.
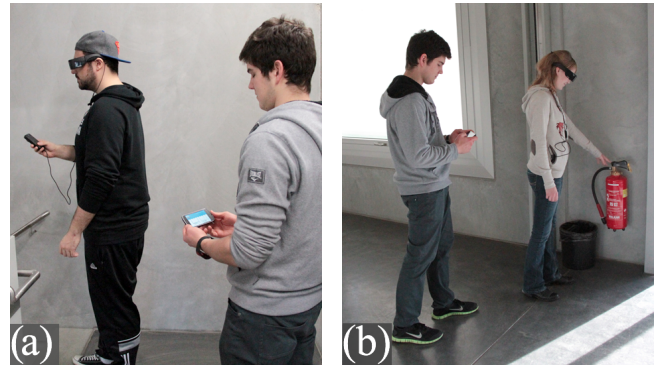
Figure 4. (a) The participant is using the wearable prototype and searching for an object. The experimenter imitates gyroscope and indoor location to create an optimal user experience for both representations. (b) Task completion time is collected until the participant touches the sought object.

Results

The study took about 45 minutes per participant. In the following we analyze the results of the Nasa TLX, the SUS and the task completion time.

The average subjective task load measured using the Nasa TLX is shown in Figure 5 (middle). Participants had an average task load of 30.25 (SD=16.87) using the 2D map and an average task load of 29.06 (SD=20.09) using the last seen image. A paired two-sided t-test did not reveal a significant effect of the two representations on the task load (t(15)=-0.25, p=.81).

The average SUS score for the two representations, used to assess their usability is shown in Figure 5 (left). The average SUS for the 2D map was 73.90 (SD=17.05) and the average SUS for the last seen image was 79.37 (SD=13.74). Again, we used a paired two-sided t-test to compare the two conditions and test did not reveal a significant effect of the two representations on the usability (t(15)=1.59, p=.13).

As the route length between the objects slightly differed, we expected that the time to find an object depends on the respective object. Thus, we consider the object as a random factor in the following. The efficiency measured through the task completion time for the two conditions is shown in Figure 5 (right). On average, participants needed 45.96 seconds (SD=24.46) to find an object with the 2D map and 41.37 seconds (SD=11.97) to find an object using the last seen image.

We used an Analysis of Co-Variance (ANCOVA) with object as a random factor to compare the two conditions. The analysis revealed no significant effect of the representation on the task completion time (F(1,5)=0.586, p=.48). Similarly, the object also did not had a significant effect on the task completion time (F(1,5)=4.208, p=.07). However, we found a statistically significant interaction effect of representation × object on the task completion time (F(1,5)=5.909, p<.001).

The qualitative feedback revealed that the participants liked the 2D map representation because the full overview of the current floor facilitates finding the shortest path to the target. In general, most participants were positive about the last seen image (e.g. that the last seen image was "*very easy to under-*

3
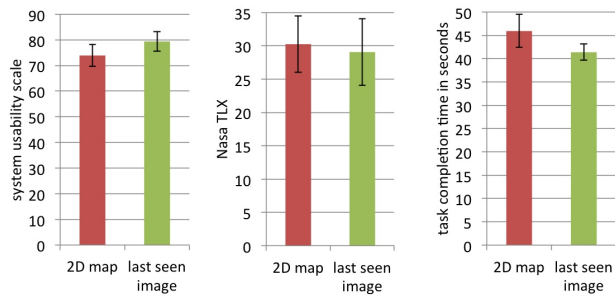
Figure 5. Average system usability scale (left) and average Nasa TLX (middle) and average task completion time for the two representations (right). Error bars show the standard error.

stand" and that *"The image at the end of the task, helps for quick localisation"*). Some participants found the arrow representation more cumbersome, because they struggled to find the correct room based on the arrow. However, all participants liked the last seen image of the sought object, as it provides a clear visual description of the object's position. Participants suggested a combination of both representations: a map representation to guide the way to the target and a last seen image when the participant is close to the target. In general, all participants liked the idea of a system which helps them finding lost objects.

Discussion

We conducted a study that compared two representations to represent the indoor location of objects. While we can not find a significant effect of the representation on the objective measures, in general, all object measures show a slight advantage for the image-based representation. We only find a significant interaction effect of representation $\times$ object on the task completion time. We assume that the significant interaction effect is mainly due to the fact that one object was more hidden and thus more difficult to find than the other objects. Qualitative feedback suggests that the last seen image representation leads to a faster result as it reveals the position of the hidden object immediately whereas the 2D-map representation only gives a vague idea of the target region.

CONCLUSION

In this paper, we compared an image-based representation of object's location to a map representation. We found that an image-based representation is easier when objects are hidden more difficultly. As for guiding the user into the direction of a target object on building-level, a map representation provides more contextual information for finding the way but also requires a floor plan of the building. An image-based representation, however, does not need any model of the environment and can compensate for non-precise indoor location information. We believe that by using an image-based representation of object indoor locations, a scalable and model-independent search engine for the physical world can be built.

ACKNOWLEDGMENTS

REFERENCES

1. Aretz, A. J. The design of electronic map displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society 33*, 1 (1991), 85–101.

2. Bangor, A., Kortum, P. T., and Miller, J. T. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction 24*, 6 (2008), 574–594.

3. Baudisch, P., and Rosenholtz, R. Halo: A technique for visualizing off-screen objects. In *Proc. CHI '03*, ACM Press (2003), 481–488.

4. Butz, A., Schneider, M., and Spassova, M. Searchlight – a lightweight search function for pervasive environments. In *Pervasive Computing*, vol. 3001 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2004, 351–356.

5. Funk, M., Schmidt, A., and Holmquist, L. E. Antonius: A mobile search engine for the physical world. In *Proc. UbiComp '13 Adjunct*, ACM Press (2013), 179–182.

6. Hart, S. G., and Staveland, L. E. Development of nasa-tlx: Results of empirical and theoretical research. *Human mental workload 1*, 3 (1988), 139–183.

7. Hightower, J., and Borriello, G. Location systems for ubiquitous computing. *Computer 34*, 8 (Aug. 2001), 57–66.

8. Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B., and Spasojevic, M. People, places, things: Web presence for the real world. *Mob. Netw. Appl. 7*, 5 (Oct. 2002), 365–376.

9. Kirsh, D. The intelligent use of space. *Artificial Intelligence 73*, 1-2 (Feb. 1995), 31–68.

10. Ljungstrand, P., Redström, J., and Holmquist, L. E. Websticker: Using physical tokens to access, manage and share bookmarks to the web. In *Proc. DARE '00*, ACM Press (2000), 23–31.

11. Nickels, J., Knierim, P., Könings, B., Schaub, F., Wiedersheim, B., Musiol, S., and Weber, M. Find my stuff: Supporting physical objects search with relative positioning. In *Proc. UbiComp '13*, ACM Press (2013), 325–334.

12. Pederson, T. Magic touch: A simple object location tracking system enabling the development of physical-virtual artefacts in office environments. *Personal and Ubiquitous Computing 5*, 1 (2001), 54–57.

13. Suomela, R., and Lehikoinen, J. Taxonomy for visualizing location-based information. *Virtual Reality 8*, 2 (2004), 71–82.

14. Yap, K.-K., Srinivasan, V., and Motani, M. Max: Human-centric search of the physical world. In *Proc. SenSys '05*, ACM Press (2005), 166–179.