Handheld Augmented Reality Indoor Navigation with Activity-Based Instructions

Alessandro Mulloni, Hartmut Seichter, Dieter Schmalstieg
Institute for Computer Graphics and Vision
Graz University of Technology, Austria
[mulloni | seichter | schmalstieg]@icg.tugraz.at

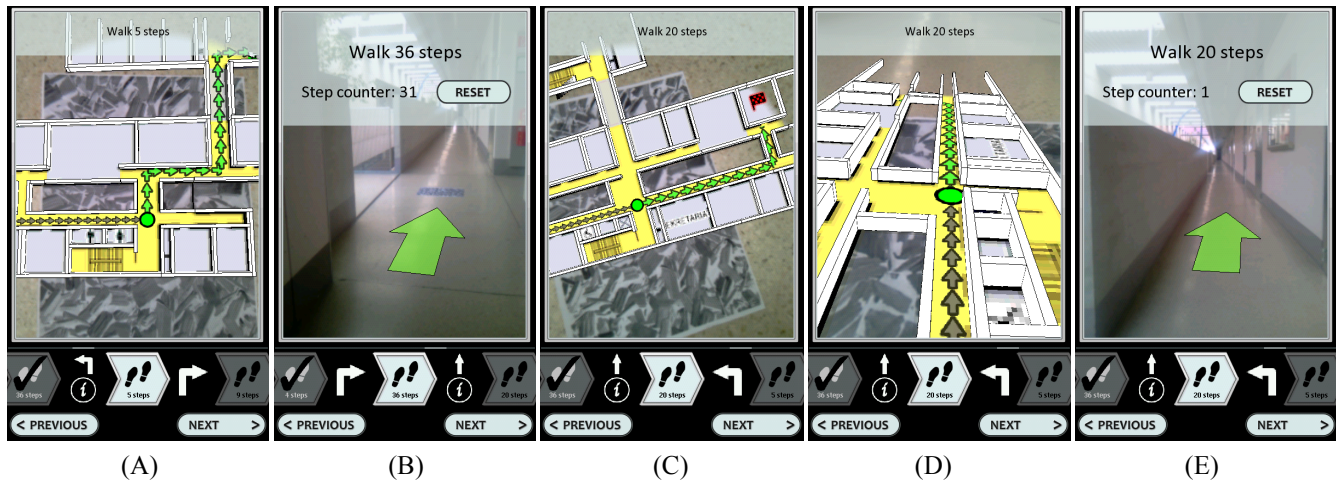(A)   (B)   (C)   (D)   (E)

Figure 1. Screenshots of our application. (A): As a user stops at an info point, detailed augmented-reality information is presented as a World-in-Miniature (WIM). (B): As she starts walking, the interface presents sketchy information on the current activity, and directional information as a perspective arrow. (C-D): As she stops at the next info point, the WIM is again shown. From afar (C), it appears as a 2D map. The target office room is now visible, marked with a red flag. From closer, tilted perspectives (D), it is possible to examine the path in 3D. After leaving the info point (E), the interface presents again sketchy information on the activity.

ABSTRACT

We present a novel design of an augmented reality interface to support indoor navigation. We combine activity-based instructions with sparse 3D localisation at selected *info points* in the building. Based on localisation accuracy and the users' activities, such as walking or standing still, the interface adapts the visualisation by changing the density and quality of information shown. We refine and validate our design through user involvement in pilot studies. We finally present the results of a comparative study conducted to validate the effectiveness of our design and to explore how the presence of info points affects users' performance on indoor navigation tasks. The results of this study validate our design and show an improvement in task performance when info points are present, which act as confirmation points and provide an overview of the task.

Author Keywords

Augmented reality, indoor navigation, activity-based navigation.

ACM Classification Keywords

H.5.1 [**Information interfaces and presentation**]: Multimedia Information Systems – *Artificial, augmented, and virtual realities.* H.5.2 [**Information interfaces and presentation**]: User Interfaces.

General Terms

Design, Experimentation.

INTRODUCTION

Most augmented reality (AR) systems for indoor navigation are based on the assumption of continuous localisation of the user and require either a significant effort to instrument the environment with the necessary infrastructure, or sensor-based estimates of user movement in the environment. A large amount of research in AR focuses on improving the localisation technology, for example by increasing accuracy or reducing drift. In contrast, few studies investigate the feasibility of indoor navigation with sparse localisation. For practical reasons, cost and time effective deployment of sparse infrastructure allows one to exploit such a solution in large-scale indoor scenarios. Yet, while AR indoor navigation systems that use sparse localisation provide detailed augmented cues at selected info points, they often outsource the burden of navigating between info points to the user [18].

In a recent work, Brush *et al.* [4] introduce activity-based navigation. Navigation instructions are formulated in terms of a series of user activities – *e.g.*, "walk 10 steps north" and then "walk 2 floors down". Like in turn-by-turn navigation, the instructions are user-centric, and the results of a user study show that it is a viable solution to help users navigate in a building, in the absence of continuous localisation. However, their solution is not robust against users' deviations from the pre-defined path, due to the complete lack of localisation.

In this paper, we present an interface that enhances AR indoor navigation based on sparse localisation with activity-based instructions (see Figure 1). We place a number of *info points* in an indoor environment to accurately locate the user. Additionally, at info points we provide contextual AR instructions accurately registered in 3D with the physical environment. We exploit activity-based instructions to guide users from one info point to the next, thus effectively eliminating the need for continuous localisation.

We define four main objectives for the interface design. Firstly, we want it to be robust to user's failures and path deviations. Secondly, we require low instrumentation of the environment. At sparse info points, the users can reset and re-orient themselves, while the required amount of instrumentation of the environment remains low. Thirdly, we want to adapt the interface depending on the localisation accuracy and on the user activity. When users are walking, we present sketchy activity-based information, as we do not know their accurate location and we do not want to overburden their attention. When users are approaching info points, we increase the complexity and precision of the information presented by showing details of the user's location and the path. Fourthly, the interface should interactively flow from one activity to the next with minimal user input: users progress to the next activity with a single tap on the screen. We refined and validated this design with pilot studies.

Finally, we conduct a study with 10 external participants to validate the effectiveness of our interface and to explore the impact of the presence of info points on navigation performance. We compare a condition in which info points are absent to a condition in which they are present, on tasks of indoor navigation. We cross-examine our findings using triangulation between quantitative and qualitative data. Our results validate the proposed design and show that the presence of info points improves task performance. While info points do not reduce the workload perceived by participants, we observe that participants value them as both confirmation points and as sources which provide an overview of the task.

RELATED WORK

Localising a user within the environment is important not only for navigation systems, but also in general for any location-aware application. Unfortunately, GPS cannot be used indoors because the satellite signal is usually unavailable when inside a building, unless special sensors such as HSGPS are used. A large body of work investigates alternative technologies for localising users when they are indoors. An extensive coverage of technologies for indoor localisation is beyond the scope of this paper. We therefore refer the reader to the detailed survey by Hightower and Borriello [10] for further information.

Different types of sensing infrastructure have been used as a basis for indoor localisation. For example, the Cyberguide project [1] uses an array of infrared beacons, the BAT system [2] uses ultrasonic waves, whereas Chittaro and Nadalutti [6] use radio frequency (RFID tags). Wagner and Schmalstieg [24] use computer vision to localise users by detecting fiduciary markers mounted on the walls. All these solutions require instrumentation of the environment and knowledge about the location of emitters/receivers or of artificial markers in the environment. An alternative approach is to exploit the wireless networks already present in the environment, as done in PlaceLab [17]. This approach does not require placing new infrastructure in the environment, but needs an up-to-date database with the locations of a large set of network access points which is not always possible for practical reasons.

A second approach often used for localisation is to incrementally measure users' movements, as shown for example by Kourogi and Kurata [13] with self-contained sensors and a wearable camera. The approach of calculating one's position as an increment over the previous position is called *dead reckoning* and it typically suffers from drift, losing accuracy over time. Good accuracy levels are achieved with wearable sensors over short distances. For example, Foxlin [7] achieves a drift of less than 1 meter for a 332-meter indoor track, with a number of shoe-mounted sensors. Dead reckoning is often used in hybrid approaches that try to combine the advantages of different technologies. For example, Merico and Bisiani [20] and Löchtefeld *et al.* [16] combine dead reckoning with manual recalibration performed by users to reduce drift. Golding and Lesh [8] use belt-worn sensors and combine dead reckoning with sensing infrastructure, using a pre-recorded database of signal intensities within the building. At present, however, dead reckoning is generally not viable for end-user applications for handheld platforms. As handheld devices are not tightly bound to a specific part of the human body (*e.g.*, the shoes or the belt), too many assumptions must be made on how the user will hold the device. For example, we cannot always tell that the compass' orientation will be the direction of movement.

User interfaces vary based on the type and precision of the localisation technology that is used. There is an inherent need to keep the user informed about the actual state of the system and to reflect this in the interface. For instance, with dense localisation available, the interface can present continuous feedback to the user. Chittaro and Nadalutti [6] automatically update the position and orientation of a

viewpoint in a virtual 3D model of the building, according to the user's physical position and orientation. All users were able to navigate a given path using this 3D model. Using this method requires instrumentation of the environment with RFID tags. Löchtefeld *et al.* [16] continuously update a map view highlighting the current position of the user, with minimal need for instrumenting the environment – only a number of "you-are-here" maps. Their approach is flexible, as it works with arbitrary maps in unprepared buildings, but it requires manual user intervention at initialization time and occasionally while in use. User intervention may be unpractical for end-user applications but it can be avoided by supplementing their solution with sparse localization, as proposed in our approach. Continuous localisation is also often used in AR to visualise egocentric navigational hints. For example, Reitmayr and Schmalstieg [22], Höllerer *et al.* [9], and Wagner and Schmalstieg [24] all augment the physical environment in real time with an arrow that guides the user. Höllerer *et al.* [9], similarly to our work, show an interface that adapts to localisation accuracy by transitioning between AR and world-in-miniature (WIM) [23] when the accuracy is respectively high or low, but they do not discuss the case in which there is no localisation at all.

The effort of instrumenting the environment can be reduced by installing a sparser infrastructure. For example, Mulloni *et al.* [21] localise users with a number of fiduciary markers at selected locations in a building. Whenever a marker is visible to the camera of the device, the view is automatically updated to show the location and orientation of the user on the map. When a marker is not available, users must manually use a map without any additional localisation support. Sparse localisation is often used in AR to augment props with exocentric navigation instructions. For example, Müller *et al.* [18] localise users with a number of "you-are-here" maps in the environment. Their system augments paper maps with the current location of the user and the path that the user is supposed to follow. When a map of the building is not visible, the system does not provide any information to the user. These two examples suggest, through their user evaluations, that sparse localization can support indoor navigation. Yet, they both lack support for users in between the sparse localisation points.

Most work on indoor navigation supported by AR focuses on improving continuous localisation technology and on exploiting such technology to present always-on egocentric AR navigation instructions, typically in the form of an arrow or a red line. Some work also shows that providing navigation instructions at sparse locations can support navigation, letting the user cope by herself with the navigation from one location to another. In this work, we look at leveraging the need for infrastructure in the environment for AR indoor navigation with sparse info points (see Figure 2) and exploring *activity-based navigation* to support users between those information points.
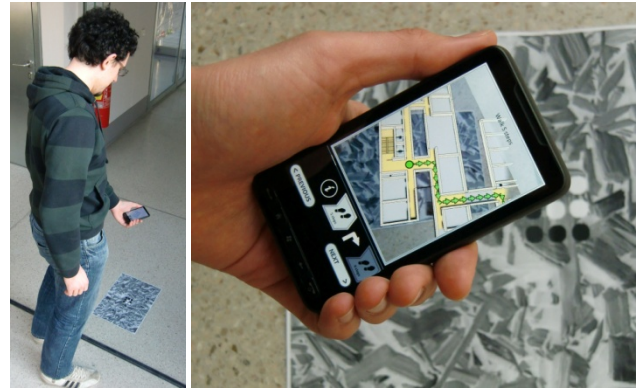
Figure 2. Info points are placed on the floor of the building. (Left) User accessing an info point. (Right) View on the phone.

Activity-based navigation is a novel approach by Brush *et al.* [4]. They provide users with a list of activities to be performed – *e.g.*, walking a number of steps in a certain direction, or going up a number of floors. The results of an evaluation show evidence that activity-based navigation is a viable solution. Yet, due to the complete lack of localisation their system is not robust against users who deviate from the predefined path, take shortcuts, or get lost. In this paper, we explore the design possibilities given by applying activity-based instructions to AR indoor navigation systems with sparse localisation. Inspired by the previous work on AR for indoor navigation, we provide both egocentric and exocentric navigation instructions in our interface design, transitioning between the two depending on the context.

INTERFACE DESIGN

We repeat the four main requirements to inform the design process of our interface:

1. Robustness to user failure such as path deviations.

2. Minimal instrumentation of the environment.

3. Adaptiveness to localisation accuracy and user activity.

4. Interactive flow of activities with minimal user input.

We recruited experts and external users for a pilot task to inform and refine the design of our prototype. The pilot task consisted of collecting a box from an office room and delivering it to another nearby office room and required a number of turns and one floor change, over a distance of about 100 steps.

This section presents the choices that led to the proposed design. We detail the space of design possibilities and we explain why we chose one over the other.

Robustness and minimal instrumentation

Interfaces for indoor navigation that are designed under the assumption of continuous localisation require a dense infrastructure to be deployed in the environment. In contrast, when dead reckoning is used to continuously localise the user, no infrastructure is typically needed, but the localisation suffers from drift and incrementally loses robustness and accuracy over time. Sparse localisation is

more applicable to indoor navigation, also in large-scale buildings, because the required infrastructure can be based on common navigation aids already present within buildings, *e.g.*, information boards or "you-are-here" maps.

We choose to combine sparse localisation at selected *info points* in the building with activity-based instructions for users to navigate from one info point to the next. The info points are AR tracking targets glued to the floor of the building (see Figure 2). The list of activity-based instructions exploits human dead reckoning capabilities over short paths to compensate for the lack of localisation.

Combining sparse localisation with activity-based instructions lowers the amount of infrastructure that must be placed in the environment, while still providing robustness against user failures or path deviations. Users can deviate from the pre-defined path, make mistakes and get lost, or simply try to rejoin a path at an intermediate info point. Additionally, this approach allows for re-routing the user based on her capabilities or the environmental constraints, for instance if an intermittent obstacle appears or if the user needs to use the elevator.

In contrast to an approach exclusively based on activities, the presence of a few localisation points gives both the user and interface a chance to reorient in case of disorientation. The info points can also act as spots where users can ensure they are on track. We observed that some pilot participants went straight to the nearest info point to reorient their interface and to obtain an updated list of activities when they were unsure about the next steps. We noticed that it is important to provide info points in proximity to decision points. Hence, the density of info points is dictated by the building layout and its occurrence of decisions points. Those decision points are usually near staircases or in general where multiple paths propagate.

Adaptive interface
The on-screen information is divided into two main areas, sharing the screen space with each other by using a sash-window metaphor to deal with overlap. The *activity view* (see Figure 1, lower part of the screenshots) presents the user with a list of activities that she must perform in order to reach the next info point, or the final destination. The *AR view* (see Figure 1, upper part of the screenshots) augments the building with AR navigation instructions.

Activities that are iconographically represented in the activity view are: *walk, change floor,* and *reach office*. The set of activities can be easily extended with further activities (*e.g.*, cross the bridge). Activities are shown as a sequence of arrow-shaped elements. Users can manually scroll through them using the touch-screen of the device. Tapping on any element sets it as the current activity. Users can also tap on the two buttons at the bottom of the screen to switch the current activity to the next or the previous one. Turns and info points between the activities are also visible.

The AR view supports the user with egocentric in-situ and exocentric contextual information regarding the activities to be performed. We adapt the visualisation depending on the localisation accuracy and on the type of activity the user is engaged in (either walking or standing still).

Whenever the system detects that the user is walking we automatically enlarge the textual information in the AR view (see Figure 1, B, E). Walking users must divide their attention between the physical movement and the use of the interface, and we avoid overloading them with information. We therefore provide easy-to-read, sketchy information that details exactly what the user must do in order to perform the current activity. The enlarged view provides details only on the current task that the user must perform. We also embed an indicative step counter in the view for the user's convenience. In this case, the AR view provides directional information in an egocentric frame of reference: we overlay the video with a perspective arrow always visible in the centre of the view, spatially oriented to lie flat on the ground and to point towards the current walking direction. While this arrow is overlaid on the video, it is not fully an AR interface as it does not strictly follow the definition of Azuma [3]: it is not registered in 3D with the environment, but only registered in orientation to a body-centred frame of reference. Its purpose is to give egocentric feedback on the walking direction, when the system does not have knowledge of the location of the user. This visualisation is analogous to floor signs commonly found in public areas to direct people's navigation.

Our system continuously scans the live video from the back camera of the device, for any visible floor-mounted posters. Whenever the system detects that the user stops over an info point we shrink the textual information to make more room for the AR visualisation (see Figure 1, A, C-D). Info points provide an exact location of the user in the building. Therefore, at info-points the system can recalculate automatically the activity list from the current location to the destination. By using this method we can cope with any type of path deviations by the user.

As the user is not walking while accessing the info point, we assume that more attentional resources are available and we provide more detailed and complex information. Floor plans are not easily readable for a layperson, partly due to their visual abstraction and partly because of spatial displacement from the local situation. One of the novel parts of our interface is the utilisation of floor-mounted AR tracking targets to provide a spatially registered overview at the info points as a WIM. Firstly, this uses an accepted affordance provided by existing navigation paradigms such as floor markings seen in department stores or sport venues. Secondly, we use the "through the lens" AR approach to spatially orient the view correctly for the user and to provide dynamic path information. The WIM appears as correctly aligned with the real building. The alignment between WIM, info point and building is pre-set manually.

We overlay the WIM with dynamic information on the location of the user (green circle) and the path that the user must follow (in green). If we do not detect a deviation from the previously computed path, we also show the direction from which the user approached the info point (smaller, in gray). Finally, if the target office room is near the info point, we also highlight it (red flag). We designed the WIM extruding walls and stairs from a 2D map, so that it appears as a 2D map from afar but if needed it can be explored in 3D from closer, tilted points of view (see Figure 1, C-D). The rationale is that a 2D map can quickly convey route information, whereas a WIM supports landmark recognition, as shown by Kray *et al.* [15]. When seen from above, the floor texture of the WIM provides further details, such as toilets, wall shadows and the location of doors.

Flow of activities
It is important that the interface interactively communicates the state of the user and shows previous and upcoming activities clearly. The interface should also require minimal user interaction for common tasks: switching to the next or the previous activity, and a roll back of a sequence of activities. We adopt a design based on a scrollable list of activities. Users can switch to the next or the previous activity with a single tap – either on a button or on the activity itself – and they can move forward or roll back by scrolling through the list.
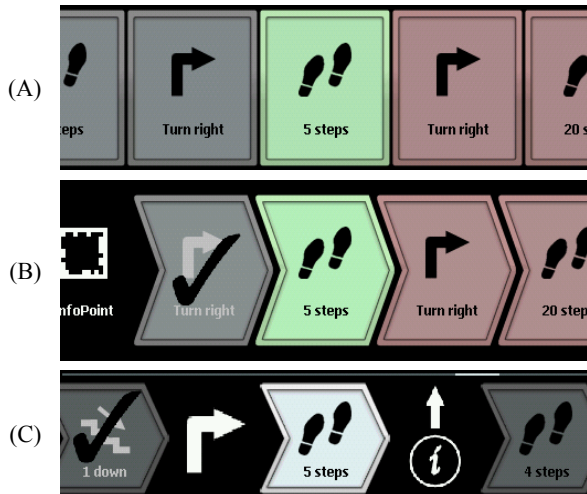
Figure 3. First (A), second (B) and last (C) prototype of the activity view. In (B-C) we convey clearly the flow of activities from left to right. We also add the position of the info points between activities. In (C) we do not show turns as activities, to eradicate the need for a double tap at each turning point.

Figure 3 shows all consecutive design iterations of the activity view. After the first pilot study, we adapted the design of the list of activities based on two issues we observed. Firstly, we observed that all activities were perceived as equally relevant, hence we tried to communicate the flow from left to right more clearly, to distinguish the already performed activities (left) from the current activity (middle, highlighted) and from the activities

to be performed next (right). We support the idea of flow by changing the design of our activity buttons, shaping them as arrows that point towards the activities to be performed next. We use checkmark icons to clearly label already performed activities. Secondly, we add the info points to the activity list. During the pilot study, we noticed that users had problems tracing back the activities they had performed after departing the last info point. Adding the info points to the activity list helps users to trace back what they did after the latest "safe position". After a further pilot study, we decided not to represent turns as activities, as this requires an unnecessary second tap on the next button at each turning point: first to advance to the turning activity and finally to advance to the following walking activity.

In our first design, we aimed at automatically detecting if the user had completed an activity, using information from a step counter and from the phone's sensors. Our objective was to let the list of activities flow automatically as the user completed consecutive activities. As a fallback, users could manually advance or return to an activity. After the first pilot, we removed this automatism from our design. We only provide visual feedback on the step count and the walking direction, as users found any automatic switch confusing when triggered erroneously by unpredicted causes (*e.g.*, longer or shorter stride lengths, magnetic influences on the phone's digital compass). We reason that the adopted one-tap solution to advance activities is a valid compromise between feedback given by the interface and minimal requirements of manual user interaction.

IMPLEMENTATION
We implemented our interface in an application running on an HTC HD2 smartphone at interactive frame rates (20–30 frames per second). The integrated sensors of the phone (an accelerometer and a digital compass) assist in estimating the device's orientation and counting the user's steps. The device's orientation is estimated using the gravity vector measured by the accelerometer and the north vector measured by the digital compass. We use a linear Kalman filter [12] to reduce jitter in the sensors' measurements. We furthermore use the local magnitude variations of the accelerometer measurements to count the user's steps [11].

At each info point we place a poster on the floor containing a pattern that can be detected and tracked using computer vision technology. We track the position and orientation of the device with respect to the posters using the natural-feature tracking approach of Wagner *et al.* [25]. This approach allows us to track the position and orientation of the device accurately also in the case of reflections on the poster, or when the poster is only partially visible. In the centre of the poster we encode a unique identifier for the info point as a 9-bit BCH code (4 redundancy bits and 5 data bits). The proposed design of the info point can be part of the overall building design, or it can coincide with a sufficiently distinguishable floor texture. The unique identifier can also be made completely transparent to the user by replacing it with wireless technology (*e.g.*, RFID tags).
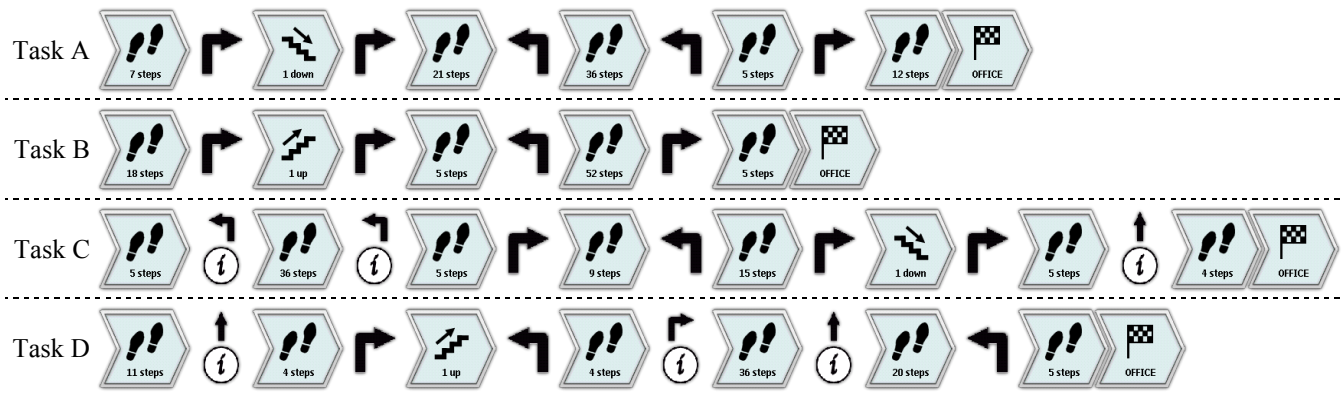
Figure 4. Tasks designed for the user study. Tasks A and B were designed for the NoIP condition, tasks C and D for the IP condition. All tasks have comparable length and difficulty. IP tasks appear longer as straight path segments are split in two by info points.

We supply the application with a graph of the corridors and office rooms in our department's buildings. A module of our application uses this graph to calculate the path between any pair of connected locations using the Dijkstra algorithm. This module can dynamically recalculate the path to any target destination whenever the user reaches an arbitrary info point.

USER STUDY

We conducted a user study to validate our interface design on tasks of indoor navigation. We used a comparative study design to explore how the presence of info points affects the performance of users in tasks of indoor navigation.

Presence of info points

As we wanted to explore the impact of the availability of info points on the performance of indoor navigation tasks, we used the availability of info points in the building as an independent variable of our user study.

The condition in which info points are absent (**NoIP**) is a baseline condition. In this condition, the system provides users with a sequence of activity-based instructions that must be followed in order to reach the target office room. There are no info points accessible throughout the path.

In the condition in which info points are present (**IP**) the system provides users with the same sequence of activity-based instructions as in NoIP. Yet, in this condition the sequence of activities includes information on the available info points through the path. During the study, we physically placed a number of info points in the building. Users could reach for a nearby info point to re-orient the interface and themselves, obtaining detailed location-aware information on the upcoming path segments.

Overall, we expected the presence of info points to keep users more on track and to lower their perceived workload. We therefore set the following hypotheses: (H1) shorter walked distances in IP compared to NoIP; (H2) lower number of navigation errors in IP compared to NoIP; (H3) lower workload perceived in IP than in NoIP.

Design

For the user study we recruited 10 participants (5 male and 5 female) aged between 24 and 35 (median 28) through the newsgroups of two universities. Our participants were predominantly early adopters. All participants were not familiar with the buildings where our study took place and they were never previously involved in our research work. We compensated our participants for their time with a voucher for a local media store.

We asked all study participants to find four office rooms inside our department buildings. We reason that this is a typical scenario for users who must navigate an unknown indoor environment, *e.g.*, an office building or a hospital. Our department is composed of four buildings that are interconnected to each other by several bridges. All buildings have a strong cubature with in-situ concrete walls and internal patios, and they contain a large number of repetitive features with a general lack of clear landmarks. There is virtually no signage for departments or offices. We consider it a hard case for indoor navigation. We designed 4 tasks, 2 for each condition (see Figure 4). All tasks have comparable difficulty, containing a similar number of turns and always one floor change up or down. We recorded a ground-truth step count for each task using a commercial step counter, and verified that all task have a similar path length (127 ± 3 steps).

We designed the study as within-subjects: all participants experienced both NoIP and IP conditions. Participants were asked to perform all four tasks; we used a Latin square to balance the order of tasks between them. One study session lasted on average 45 minutes.

We began a study session by collecting demographic data from the participant and showing her an informed consent form that introduced the procedure of the user study. We then let the participant conduct a tutorial task that forced her to try all functionalities of the application, supported by a verbal explanation by the examiner. In particular – as AR is not a commonplace interface metaphor – we enforced AR

training of the participant on two info points during the test task. We therefore assume that all our participants were at least familiar with the operation of the AR interface before starting the subsequent four tasks. Finally, we conducted the participant to one of the four locations designated as a starting point for the tasks, we gave her the device and we asked her to find the target office room. This procedure was repeated for all four tasks.

As the goal was to both validate our design and measure the impact of info points on task performance, we triangulated a number of methods to collect both qualitative and quantitative data. On the device, we ran a software logger continuously recording the application status and selected events. We asked the participant to wear the commercial step counter and we recorded the number of steps needed to complete each task for comparison against the ground truth. An evaluator followed the participant and noted all observations and all spontaneous feedback given by the participant while performing the task. The evaluator also noted on a map all occurrences of navigation errors. After each task, we asked the participant to fill in a one-page NASA TLX questionnaire. After all office rooms were found, we interviewed the participant. First, we asked for subjective feedback – whether having or not having the info points changed anything in their navigation experience, or if it was just the same. The question was aimed at collecting feedback on whether users found info points useful, and if so why. We structured the rest of the interview around all noted occurrences of navigation errors, to collect subjective comments on the issues that occurred in each situation.

RESULTS AND DISCUSSION
This section focuses on our analysis of the impact of the interfaces on navigational performance. We further discuss the problems that users experienced, which will inform the next design iteration.

Task performance
All users completed all tasks successfully. As a rough measure of task performance, we looked at the *step difference* from a recorded ground truth, at the *task completion time*, and at *navigation errors* made by the users based on severity of deviations from the path indicated by the interface (see Table 1). Step count and task completion time are of course a rough measure of the performance in a navigation task, as they also depend on walking speed and stride length. Furthermore, we coded all navigation errors as either *soft* or *hard* errors. Soft errors denote when the user departed the path indicated by the interface, noticed the mistake, and recovered from the position in which she had departed from the path. Hard errors denote when the user could not recover from the position where she departed from the path, but she had to roll back a number of activities and repeat those activities.

The NoIP condition has a median step difference of 29.75 steps more than the pre-recorded ground truth, about 23% of the whole path length. A t-test shows that the difference

between IP and NoIP is a statistically significant one $(t(9) = -3.14, p = .01)$ and supports our hypothesis (H1). This result points to the fact that users in the IP condition followed on average an almost optimal route, whereas in the NoIP condition they usually deviated by several steps from the optimal route. While deviating more from the optimal route, the results show that users in the NoIP condition were slightly faster. This is not surprising, as participants spent some extra time browsing the info points while in the IP condition. A t-test shows that the difference in task completion time is not statistically significant $(t(9) = -.25, p = .80)$.

	Step difference	Time (seconds)	Navigation errors		
			Soft	*Hard*	Total
NoIP	29.75	135	*19*	*3*	22
IP	-2.25	142	*9*	*1*	10

Table 1. Task performance per condition: median difference in step count from a pre-recorded ground truth, median task completion time, and total number of navigation errors.

We recorded more errors in the NoIP condition than in the IP condition. A Wilcoxon test shows that the difference is not statistically significant $(Z = -1.796, p = .07)$. Our second hypothesis (H2) is not supported statistically, but there is a clear trend pointing in favour of it. From the interviews, we could identify a number of error sources (see Figure 5). We list them below indicating the number of occurrences for each condition as (NoIP, IP).
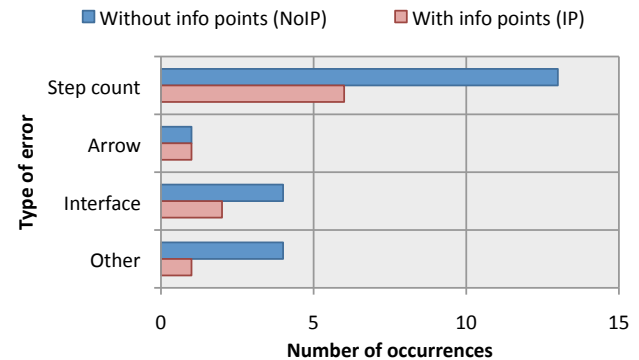

Figure 5. Number of navigation errors per condition and type.

Overshooting due to a wrong step count (13, 6). Inaccuracies in the step counter and errors in the mental count of the steps caused problems of users overshooting a turning point and erroneously going straight, in some cases performing the turn at the next intersection.

Confusion caused by the arrow (1, 1). Magnetic influences within the building caused the overlay arrow to point in an incorrect direction.

Issues with the design of the interface (4, 2). In four cases a participant (P1, P7, P9 twice) performed the wrong activity, either because she forgot to switch to the next activity, or because she clicked more than once on the next button. In

one case P2 confused one info point with another, and took the wrong turn. This is because info points are indistinguishable from each other in the activity bar, as shown in Figure 6 (A). In one case P9 was confused by arrows on both sides of the activity, as shown in Figure 6 (B). While the activity flow suggests a right turn in this case, we reason that the arrows, like the activities, should also be greyed out once the turn has been performed. Arrows in the interface have a high affordance signal and therefore easily override relevance of other parts in the interface.
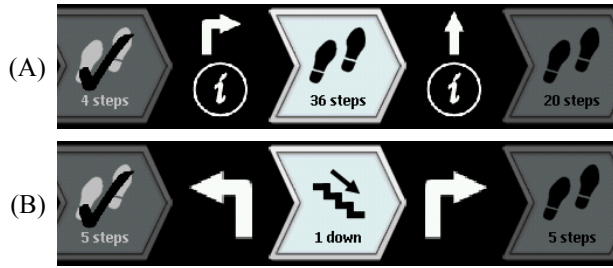
Figure 6. Interface ambiguities that caused navigation errors.

Other (4, 1). In one case P4 erroneously read "one floor down" in place of "one floor up". In another case P6 was distracted and missed one turn. Some other issues could not be explained or remembered by the participants.

In general, we observed that after a navigation error participants tried to match distinguishable building elements (i.e. landmarks) with the visualisation on their device in order to recover from the error. In particular, stairs acted as prominent landmarks because they also appeared as a clear checkpoint in the list of activities. For example, users walking past an intersection with the stairs, when the next activity was a floor change, were usually able to rapidly recover from the navigation error and go back to the stairs. For example, in the interview P3 said: "when I saw that the next activity was to go up the stairs, it was obvious to me that I had walked a few steps too many." Stairs were also used as a checkpoint for rolling back the list of activities in the case of hard errors. For example, P4 was disoriented in the NoIP condition, and rolled back activities to the last floor change, went back to the last set of stairs and then once again performed all activities from that point on.

Some participants also used more complex reasoning on the activities and the structure of the building to recover from an error. For example, P4 made a navigation error caused by the overlay arrow pointing in the wrong direction. In the later interview he commented: "In this case you reason more, like if you have to do 50 steps, you pick the longest corridor." Similarly, P7 got confused and rolled back to a walking activity with a large number of steps to perform, at the beginning of a long corridor that she remembered.

We observed that the turns given in the list of activities also helped in solving navigation ambiguities, for example in the

case of the overlay arrow pointing in the wrong direction. In a few cases we observed that users commented on the overlay arrow pointing in the wrong direction and decided to rely on the turn instruction in the activity list.

Overall, the presence of info points improved performance on the navigation tasks, keeping users more on track and reducing the number of navigation errors. In the case of navigation errors, users often tried to match the list of activities with prominent nearby landmarks, and to recover from these points. We now look at the impact of the tasks on the workload reported by the participants.

Workload
A Pearson correlation test shows that the TLX results (the weighted sum of all TLX indices) have a positive correlation to the number of navigation errors, which is statistically significant ($r = .563$, $N = 40$, $p < .01$). As shown in Figure 7, there is a slight tendency for a lower self-reported workload in the IP condition as compared to the NoIP condition. This is not a statistically significant difference, therefore our hypothesis (H3) is not statistically supported. Further, as the workload correlates positively with the number of errors, this difference in workload could be simply due to the fact that participants made more errors in the NoIP condition.
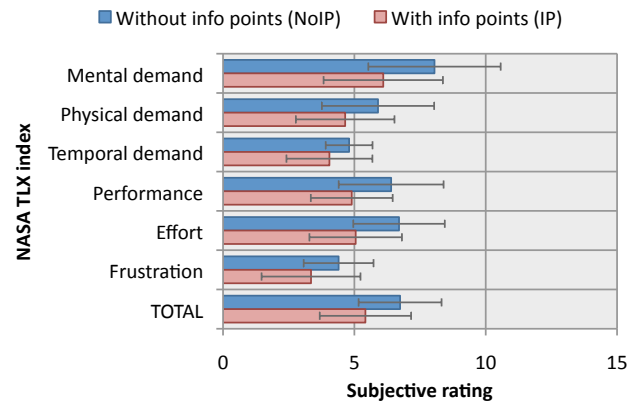
Figure 7. Average self-reported workload per condition (and 95% confidence intervals), for all indices of the NASA TLX questionnaire, on a scale from 1 (low) to 21 (high). TOTAL is the weighted sum as defined in the NASA TLX instructions.

All participants were asked if they felt that there was a difference in completing the task with or without the info points. 3 participants answered that it was also the same without them. Yet, participants also claimed that it was "easier" (P2, P9, P10), "more intuitive" (P5), "useful" (P3) and "reassuring" (P1, P3) to have the info points. Overall, the participants valued info points, but info points did not have an impact on their perceived workload.
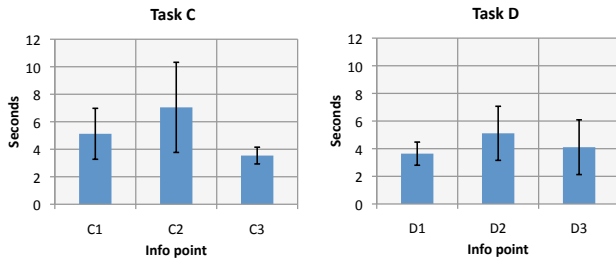
Info points
During the interviews we additionally questioned all participants on their perceived value of the info points. The answers from the participants reveal that the value of info points was twofold: both as *confirmation* and for *overview*.

Confirmation. Confirmation information is the second most needed type of information in navigation tasks, after wayfinding information [19]. From the interviews, it emerges that info points acted as checkpoints to obtain confirmation that the participant was on track. Info points were "good to check the position where you are located" (P6). At the info points "you have a point where you're sure you get information [about] where you are" (P9). As P10 remarked, "I was more sure that I'm on the right place, because with the [info points] I get feedback."

Overview. The info points also acted as spots where users could get an overview of the sequence of upcoming activities. For example, P1 and P8 stated that they were matching landmarks (*e.g.*, toilets, corridors) between the info point and the environment, and P1 added that the info point showed "more than one task, not only the next but also the one after the next". On the info point "you can look at the upcoming path: you get an overview, not only an arrow" (P3). "You see the direction where you have to go, not only the arrow. You get a good image of where you are" (P7). Overall, "you could do [the task] without the info points, but with the info points it was much more intuitive. You could see the way you had to go" (P5). This amount of overview was also contrasted to the arrow visualisation (P3, P7) that only showed information about the current activity.

Figure 8. Average duration of info point sessions and 95% confidence intervals for both tasks in the IP condition. Info points are numbered according to the order in which they appear in the task (see Figure 4).

The median number of info points used for each task is 3, thus almost every participant used each info point to complete the task. The average duration of an info point session (from the moment in which the info point was detected to the moment in which the participant stopped pointing the camera at it) is 3.72 seconds ($\sigma = 3.26$). As the standard deviation is high, we plot in Figure 8 the average duration of a session for each separate info point. Usage sessions of info points where a turn was needed (C1, C2, D2, see Figure 1, A) were longer than the info points where no turn was required (C3, D1, D3, see Figure 1, C-D). In particular C2 took the longest. We reason that this is because, in contrast to C1 and D2, C2 was not followed by a single long sequence of steps but by a set of smaller activities, and therefore the info point presented a larger amount of information. C2 is the info point shown in Figure 1, A. Differences are not statistically significant and they should therefore be considered only qualitatively.

In general, the sessions were only a few seconds long, supporting the observation that info points were used mainly as confirmation spots, in particular info points where users were not required to turn.

CONCLUSION

Overall, our interface was validated as an effective means to support indoor navigation with AR. Further, the addition of info points showed an improvement in the performance of navigation tasks, significantly reducing the step deviation from the optimal path and contributing to a reduction in the number of navigation errors. While the study participants did not perceive a reduction in workload when info points were present, they valued them for their twofold role of providing confirmation that they were on track and showing an informative overview of a number of upcoming activities.

Participants made a number of navigation errors in both conditions. We observed that participants often recovered from the errors by looking for matches between the visualisation of the activities and the structure and landmarks in the surrounding building. We reason that the overview provided by the info points was helpful in supporting this matching, as also stated by participants in some of the interviews.

In particular, participants contrasted the role of info points as overview providers with the arrow, which only showed information about the upcoming activity. This postulates an additional requirement for the next design iteration – supporting a match between activities and building structure. The objective is twofold: first, we want to support and enforce spatial reasoning also in between the info points; and second, we want to study how the support for spatial reasoning can make the interface more robust against navigation errors. Consequently, this hints towards a more informative visualisation also in between the info points. For example, we could look at increasing the information visualised when the localisation accuracy decreases, as proposed by Butz *et al.* [5].

Further, we will also check the type and quality of spatial knowledge that is developed using our interface, as this is a known problem for step-by-step navigation systems [14]. This points towards investigating the underlying cognitive processes, also in comparison to traditional navigation aids. However, doing so requires a mature interface design to avoid influence from external factors such as usability issues and user frustration.

In conclusion, our work has shown that a balanced mixture of AR and non-AR interfaces plays an important role in the progress of indoor navigation technology. Based on the results of our study we are very positive that pivoting the design around both the technology and the users'

capabilities is highly valuable for supporting people's needs in indoor navigation.

REFERENCES

1. Abowd, G.D., Atkeson, C.G., Hong, J., Long, S., Kooper, R., and Pinkerton, M. Cyberguide: a mobile context-aware tour guide. *Wireless Networks 3*, (1997), 421–433.

2. Addlesee, M., Curwen, R., Hodges, S., et al. Implementing a Sentient Computing System. *Computer 34*, (2001), 50–56.

3. Azuma, R. A Survey of Augmented Reality. *Presence 6*, 4 (1997), 355-385.

4. Brush, A.J.B., Hammil, K., Levi, S., et al. User experiences with activity-based navigation on mobile devices. *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI 2010)*, (2010), 73.

5. Butz, A., Baus, J., Krüger, A., and Lohse, M. A hybrid indoor navigation system. *Proceedings of the 6th international conference on Intelligent user interfaces*, ACM (2001), 25–32.

6. Chittaro, L. and Nadalutti, D. Presenting evacuation instructions on mobile devices by means of location-aware 3D virtual environments. *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, ACM (2008), 395–398.

7. Foxlin, E. Pedestrian Tracking with Shoe-Mounted Inertial Sensors. *IEEE Computer Graphics and Applications 25*, 2005, 38-46.

8. Golding, A.R. and Lesh, N. Indoor navigation using a diverse set of cheap, wearable sensors. *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*, (1999), 29-36.

9. Höllerer, T., Hallaway, D., Tinna, N., and Feiner, S. Steps Toward Accommodating Variable Position Tracking Accuracy in a Mobile Augmented Reality System. *In Proceedings of AIMS 2001: Second Int. Workshop on Artificial Intelligence in Mobile Systems*, (2001), 31--37.

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

11. Jimenez, A.R., Seco, F., Prieto, C., and Guevara, J. A comparison of Pedestrian Dead-Reckoning algorithms using a low-cost MEMS IMU. *2009 IEEE International Symposium on Intelligent Signal Processing*, (2009), 37-42.

12. Kalman, R.E. A new approach to linear filtering and prediction problems. *Journal Of Basic Engineering 82*, Series D (1960), 35-45.

13. Kourogi, M. and Kurata, T. Personal Positioning based on Walking Locomotion Analysis with Self-Contained Sensors and a Wearable Camera. *Mixed and Augmented Reality, IEEE / ACM International Symposium on*, IEEE Computer Society (2003), 103.

14. Krüger, A., Aslan, I., and Zimmer, H. The effects of mobile pedestrian navigation systems on the concurrent acquisition of route and survey knowledge. *Mobile Human-Computer Interaction–MobileHCI 2004*, (2004), 39–60.

15. Kray, C., Elting, C., Laakso, K., and Coors, V. Presenting route instructions on mobile devices. *Proceedings of the 8th international conference on Intelligent user interfaces - IUI 2003*, (2003), 117.

16. Löchtefeld, M., Gehring, S., Schöning, J., and Krüger, A. PINwI: pedestrian indoor navigation without infrastructure. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, ACM (2010), 731–734.

17. LaMarca, A., Chawathe, Y., Consolvo, S., et al. Place Lab: Device Positioning Using Radio Beacons in the Wild. In *Pervasive Computing*. 2005, 116-133.

18. Müller, H., Schöning, J., and Krüger, A. Mobile Map Interaction - Evaluation in an Indoor Scenario. *Workshop on Mobile and Embedded Interactive Systems, Informatik 2006 Gesellschaft für Informatik e.V.*, (2006).

19. May, A.J., Ross, T., Bayer, S.H., and Tarkiainen, M.J. Pedestrian navigation aids: information requirements and design implications. *Personal and Ubiquitous Computing 7*, 6 (2003), 331-338.

20. Merico, D. and Bisiani, R. Indoor Navigation with Minimal Infrastructure. *Positioning, Navigation and Communication, 2007. WPNC 2007. 4th Workshop on*, (2007), 141-144.

21. Mulloni, A., Wagner, D., Barakonyi, I., and Schmalstieg, D. Indoor Positioning and Navigation with Camera Phones. *IEEE Pervasive Computing 8*, 2 (2009), 22-31.

22. Reitmayr, G. and Schmalstieg, D. Location based applications for mobile augmented reality. *Proceedings of the 4th Australian user interface conference on User interfaces 2003*, Australian Computer Society, Inc. (2003), 65-73.

23. Stoakley, R., Conway, M.J., and Pausch, R. Virtual reality on a WIM: interactive worlds in miniature. *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM Press/Addison-Wesley Publishing Co. (1995), 265–272.

24. Wagner, D. and Schmalstieg, D. First steps towards handheld augmented reality. *Wearable Computers, 2003. Proceedings. Seventh IEEE International Symposium on*, (2003), 127-135.

25. Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., and Schmalstieg, D. Real-Time Detection and Tracking for Augmented Reality on Mobile Phones. *IEEE Transactions on Visualization and Computer Graphics 16*, (2010), 355–368.