

User Experiences with Activity-Based Navigation on Mobile Devices

A.J. Bernheim Brush, Amy K. Karlson, James Scott, Raman Sarin, Andy Jacobs, Barry Bond, Oscar Murillo, Galen Hunt, Mike Sinclair, Kerry Hammil, Steven Levi

Microsoft Research

One Microsoft Way, Redmond, WA 98052, USA

{ajbrush, karlson, jws, ramans, barrybo, oscarmu, galenh, sinclair, khammil}@microsoft.com

ABSTRACT

We introduce activity-based navigation, which uses human activities derived from sensor data to help people navigate, in particular to retrace a “trail” previously taken by that person or another person. Such trails may include step counts, walking up/down stairs or taking elevators, compass directions, and photos taken along a user’s path, in addition to absolute positioning (GPS and maps) when available. To explore the user experience of activity-based navigation, we built Greenfield, a mobile device interface for finding a car. We conducted a ten participant user study comparing users’ ability to find cars across three different presentations of activity-based information as well as verbal instructions. Our results show that activity-based navigation can be used for car finding and suggest its promise more generally for supporting navigation tasks. We present lessons for future activity-based navigation interfaces, and motivate further work in this space, particularly in the area of robust activity inference.

Categories and Subject Descriptors

H5.m. Information interfaces and presentation (e.g., HCI); Miscellaneous.

General Terms

Human Factors.

Keywords

Mobile applications, navigation, mobile user interfaces, sensor fusion, activity inference.

1. INTRODUCTION

Recent research has highlighted the usefulness of worn sensors to automatically detect and recognize a wide range of physical human activities using machine inference [2,7,14,22,28]. Inspired by the successes that have been achieved in discriminating among modes of human movement—such as between standing, walking, stair climbing, and riding an elevator—our work explores the role that various visual representations of these activities can play in helping users perform navigation tasks. We propose that by leveraging sensors that are easily integrated into a mobile phone, trails of the form “walk west 50 steps, go up 2 flights of stairs, walk north 20 steps” can be constructed automatically while a user traverses this path, and made available for later playback to help the phone’s owner (if they forget the path) or a different user

(so they can retrace the owner’s steps) find a previously visited location. We call this approach to re-finding places and objects “activity-based” navigation as it is based on human activities in reaching them (e.g. “take the stairs up one floor then walk north 50 steps”). Our goal is to better understand both the advantages of and challenges faced by activity-based navigation interfaces to motivate and guide the designs for this approach to navigation.

Activity-based navigation is an interesting alternative to map-based navigation used in GPS road map systems or indoor localization systems because it does not require that a map be provided or constructed, and it is not dependent on absolute positioning which may be unavailable in certain situations. Scenarios for which activity-based navigation may be appropriate include: finding a lost object by retracing one’s steps to the location in which it was last “seen” (e.g., in Bluetooth range) [9], following trails shared by other users to find an unfamiliar meeting room when a colleague is already there, finding a friend in a concert hall, dark cinema or sports arena, as well as in real-world pervasive games [4], for finding a hidden object or person.

In contrast to the considerable amount of research that has explored activity recognition, in this paper we focus on how people use and interact with activity-based navigation data once the activities have been inferred from raw sensor data. We specifically compare what types of sensed information help users navigate as well as explore how to present trails using an example application of finding a car in a garage. We designed and implemented Greenfield¹, an activity-based navigation user interface on a mobile device for following a trail to a car and conducted a user study of a car finding scenario with four conditions—three that presented different types of sensor information to the user using a GUI, and a fourth that used only verbal instructions. To examine how activity-based navigation both augments and is an alternative to GPS navigation, we conducted the study in two different parking garages, one with GPS signal and one without. Note that the trails in our study were manually constructed based on raw sensor data. We do not address the automatic creation of trails from sensed data, such as via inference techniques, in this paper.

During our study, our ten users successfully found eight cars each. While users preferred having absolute positioning information (GPS) and maps when they were available, users were able to find

¹ Greenfield is named for Greenfield Village, home to the Henry Ford museum.

cars without resorting to exhaustive search with as little information as a step count and relative floor number. We found many useful lessons for activity-based navigation interfaces on mobile devices, including the need to support users who intentionally stray from the provided path and pick up the trail somewhere along the way. Half of our participants preferred a combined visual presentation of step counts, floor changes, and compass information, and half preferred a non-visual verbal description that included only the garage floor number and quadrant in which the car was located. The relative success of the verbal condition highlights the value of providing high-level summary information in activity-navigation based interfaces. Of the three GUI-based conditions, providing compass directions was preferred over providing photos, which resulted in higher mental load and task times. Our findings also motivate further work in activity inference, including robustly determining changes between building floors and recognizing a user's direction of movement irrespective of device orientation.

2. RELATED WORK

Scientists studying personal navigation describe how people build cognitive maps by subconsciously remembering landmarks and how to move between them to reach a destination, e.g., [8,16]. In his book "Inner Navigation" [16], Erik Jonsson is particularly interested in when and why people experience breakdowns in spatial reckoning; examples include when we are distracted, on "auto-pilot" or when landmarks change. Inspired by recent research on activity recognition (particularly for elder care and health applications [7]), which suggests it is possible to recognize a variety of activities such as driving, walking, and using stairs and elevators [2,7,14,22,27,28], we wanted to explore whether using sensors to recognize activities could assist people when their personal cognitive maps are incorrect or incomplete, and particularly when navigation aids like maps and GPS may not be available.

Researchers have long been interested in using technology to assist with navigation. For example, Krüger *et al.* [18] built a personal navigation system that included desktop, car, and mobile PDA components and which made use of street, pedestrian and building maps. For landmark-based navigation, Beeharee and Steed [3] explored the use of images to augment route descriptions and found that photographs aided pedestrians in following routes, but that real systems face challenges in providing up-to-date images, such as matching the current time of day. Hile *et al.* [13] used computer vision techniques on an existing photo collection to provide automatically annotated images for pedestrian landmark-based navigation using GPS data. In contrast to these systems, we focus on the value of activity-based navigation when maps and/or GPS data may be partially available or not available at all.

For indoor environments and other places where GPS information is not typically available, many researchers have developed techniques to provide localization information based on radio beacons in the environment, such as RADAR [1]. Woodman and Harle use a foot-mounted inertial unit, particle filters, and a detailed building map to provide absolute positioning for a person inside the building without knowledge of wireless beacons or installed infrastructure [30]. Robertson *et al.* further describe the FootSLAM algorithm that could be used to collaboratively build

maps using foot mounted IMUs [29]. However, all of these systems require some sort of training or prior knowledge of the environment. In contrast, with activity-based navigation a trail can be created and retraced without prior knowledge or infrastructure.

Several research groups are working on inferring where devices are located on the body and their orientation [12, 20], while others are developing algorithms agnostic to position and orientation that detect changes in location [5,25]. NaviSeer [26] is a commercial belt-worn dead-reckoning device which claims to track users to an error less than 2% of the distance travelled by classifying and compensating for errors introduced by different movement types. Combined with activity recognition work, these advances suggest that it will be possible in the future to detect people's activities without constraining how they use and carry their mobile devices.

Most similar to our study are those that compared different presentation options for navigation using mobile devices. For example, Kray *et al.* [17] compared presenting route information using 2D and 3D maps and found participants preferred 2D maps. Froehlich *et al.* [10] compared different presentations of perspective, field of view, realism, and orientation in a restaurant finding service that provided virtual representations of participants' visible surroundings. While these studies provide valuable insights, they both augment or compare to street-level maps. Liu *et al.* [24] explored using photos, audio, and text to help people with cognitive impairments follow indoor directions to new locations, using a Wizard of Oz strategy that assumed the location of the participant was known. Participants suggested that photos might have been more useful if salient objects were clearly visible. In contrast to all of these studies, we focused on providing activity-based navigation information that de-emphasizes maps of the environment and can be used without relying on localization.

3. ACTIVITY-BASED NAVIGATION

We use the term *activity-based navigation* to refer to guiding a user to a destination using a *trail* of human movement activities such as walking a certain number of steps, going up or down a number of floors, etc. These activities may be obtained by using sensor data from a previous traversal of the trail by the same or a different user, who was carrying a suitable device such as a mobile phone with sensors. The device can conceivably automatically infer and record trail of activities based on sensor data gathered whenever the user is walking, thus obtaining trails without user assistance. The main advantage of this type of navigation over positioning-based navigation, such as using a GPS, is that it does not rely on pre-existing knowledge such as a map or 3D model of a building. Once a trail is created, it can be stored and later replayed to either the original user or, through sharing of the device or transmission of the trail data, a different user. The trail can also be reversed, so if a user walks from A to B (e.g., from their car to a shopping mall), they can be later guided from B to A.

There are several activities that can be determined today using physical sensors that can be incorporated into mobile devices. The most obvious of these is **step counting** using an accelerometer, which provides an approximate measure of distance, e.g., "walk 10 steps". A magnetometer can provide a **compass direction** for this distance, e.g., "walk 10 steps North". An inertial measurement unit (IMU) integrating accelerometer, gyroscope, and (sometimes) magnetometer information can be used to obtain

3D distances and directions, though IMUs suffer from higher power consumption and gyroscope-derived orientations suffer from drift.

While barometric pressure sensors drift over time due to weather, over timescales of a few minutes they can provide a reliable indication of small altitude changes such as a **change of building floor** [22]. Furthermore, combining barometric data with step counts/intervals derived from an accelerometer can disambiguate between modes of traversing floors, e.g., between an elevator/escalator vs. ramps/stairs, although disambiguating between these pairs (elevator vs. escalator) may be difficult.

While not in common use today, recent research has explored use scenarios for carrying an outward-facing camera, e.g., a SenseCam [15] or a cameraphone worn on a belt or lanyard. While much of this research has focused on stimulating episodic memory for people with memory impairments [15,21], the **photos** taken by this camera are excellent candidates for use in activity-based navigation, although it is important to note that researchers using photos as memory cues and for navigation have all described challenges in selecting appropriate photos (e.g., [3,13,21,24]). Finally, we must not overlook **absolute position** information collected from GPS location or, more generally, any positioning system that may be available. While activity-based navigation is likely to be most useful when GPS is not available, in many situations there may be the opportunity to present both types of information.

In this paper, we explore how the five types of activity information highlighted above—step counts, compass direction, floor changes, photos, and absolute position—can contribute to the user experience and performance in an example activity-based navigation application, namely that of car finding.

4. GREENFIELD: HELPING USERS FIND CARS

To explore activity-based navigation we designed Greenfield, an application that provides an activity-based trail to a parked car. This is useful both in situations where one has lost one's car and needs to find it, but also in finding a car parked by someone else, e.g. a private car parked by another family member, or a shared business vehicle. As Jonsson [16, pp. 53-55] observes, finding one's own lost car is a familiar example of a breakdown in spatial reckoning, a time when our cognitive map of where we parked our car fails, perhaps because a garage had a challenging layout, or another car we used as a landmark is missing.

In fact, several mobile phone applications already exist to help find cars (e.g., GParkS [11], and CarSpotter [6]). These use GPS to remember your car location and some allow people to manually record pictures or write down a parking number. The Garmin Nuvi 700 series of GPS products also automatically records the location if removed from the mount, allowing the user to navigate back to a parked car. While these applications highlight people's interest in assistance finding their cars, they make two assumptions that limit their usefulness. First, relying on GPS is not always practical as many places that cars are parked do not have GPS signal (e.g., underground garages), nor do GPS systems typically provide a way of identifying on which level of a garage the user has parked. Second, phone-based applications rely on the user to hit a button to record their GPS position or manually note



(a) (b)
Figure 1. The Greenfield user interface. (a) The overview screen for a trail of breadcrumbs. (b) The detail screen for an elevator breadcrumb.

their location, while the Garmin solution requires users to remember to take their GPS with them when leaving the car.

An activity sensing approach addresses both of these issues, since it does not rely on absolute positioning data, and the same sensors that determine activity types can be leveraged to identify when to start (e.g., accelerometer detects the end of a driving activity [e.g., 14,23], or Bluetooth connectivity with an in-car sensor is lost) and stop (e.g., Bluetooth connectivity with an in-office sensor is achieved, or a timeout is reached) the activity logging. Avoiding manual effort is a key point particularly in the case of finding one's own lost car; as Jonsson highlights, occurrences of spatial memory failure often involve user fatigue and distraction—which are likely the same times when a user might forget to perform an explicit action to remember their car location.

We developed Greenfield as a Silverlight application on the Menlo V1 platform (Figure 3d). Menlo is a prototype mobile device with a capacitive touch screen (4.1" diagonal, 800x480) running Microsoft Windows Embedded CE 6.0 R2 which incorporates a Bosch BMA150 3-axis accelerometer and Bosch BMP085 digital pressure sensor (barometer). The application provides users with a series of activities in a trail, which we call breadcrumbs. Each breadcrumb consists of: walking for a specified number of steps (optionally specifying a compass direction), going up/down stairs for a specified number of floors, or taking an elevator for a specified number of floors. The Greenfield overview screen (Figure 1a) shows a list of breadcrumbs, each representing an action to perform (e.g., walk up [x] flights of stairs) en route to a destination. Tapping a particular breadcrumb opens the detail screen (Figure 1b) which can show additional information including camera images and a map with GPS locations (if available), and allows users to navigate between breadcrumbs.

5. USER STUDY

To explore how users would make use of an activity-based navigation system that relies on the five activity types described in Section 3, we conducted a user study with Greenfield in which ten participants were asked to find eight cars starting from a workplace office and ending in a parking garage. We asked

participants to find “a colleague’s car”, i.e. following the scenario of finding a vehicle parked by someone else. We did not attempt to ask participants to find their own lost cars due to the difficulty of recreating such circumstances in a controlled manner. We reasoned that if participants can follow someone else’s activity trail, using their own to jog their memory will likely be easier.

While our activity-based navigation approach was originally inspired by situations in which GPS is unavailable, there are also occasions when GPS is available, but insufficient for completing a task (e.g., determining which floor a car is parked on). Thus we included two different garage conditions in our study (Figure 2a): (1) G_{GPS} , a free-standing garage with 4 above ground levels with GPS coverage and a neighboring multi-level office building; and (2) G_{NoGPS} , a 3-level underground garage directly below its associated office building with no GPS coverage. In each garage, participants experienced all four presentation conditions (Figure 2b): three different activity trail presentations using Greenfield, as well as a verbal description.

5.1 Activity Presentation Conditions

In designing the three Greenfield presentation conditions to study, we decided that users would require a minimum of step counts and floor changes (by stairs or elevator) to avoid needing to search the entire garage for the target car, which we saw no value in asking our participants to do. We call this baseline presentation **Steps/Floors** (Figure 1a, Figure 3a). We also always provided a map representation in the detail view. For trails in G_{GPS} the map displayed a 2D path superimposed on the map (Figure 3c), however in G_{NoGPS} the map only displayed an overview of the area with a large circle coarsely indicating the last known GPS position (Figure 3a).

The step count offered in Steps/Floors allows users to know whether the car is relatively “near” to or “far” from the elevator, but in the absence of GPS information, it does not give users enough directional information for them to follow a deterministic path to the car. We therefore created two other presentation conditions to compare the relative value that other types of sensor information might offer users; both augmented the baseline Steps/Floors (SF) presentation with directional information either in the form of compass directions, **SF+Compass**, (Figure 3b) or

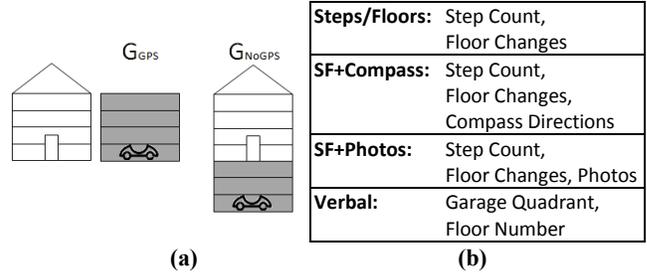


Figure 2. Study garage (a) and presentation (b) conditions.

photos, **SF+Photos**, (Figure 1b, Figure 3c). For the SF+Compass condition, we split walking breadcrumbs when the trail changed direction, and included absolute compass coordinates (N, E, S, W) with the step count. We favored this approach to offering relative instructions (turn left/right) since users who make a mistake are likely to compound their error with relative directions, and can more easily correct themselves with an absolute reference. While our device did not have a magnetometer, it had a side-mounted compass to provide real-time directional feedback (Figure 3d).

In addition to the three conditions showing sensed data (Steps/Floors, SF+Compass, SF+Photos) we also included a **verbal** condition as our control. In the verbal condition, we gave participants a verbal description of the garage level and quadrant (e.g., “3rd floor, NW quadrant”) where their friend’s car was located. This condition controlled for the occasions when people do not have any technology assistance. We had planned to use an existing car finding application as a further control condition, but after trying out several we found that their reliance on GPS data meant that the applications provided inaccurate or overly general information that would have required participants to search the entire garage for each car. The four presentation conditions that we explored in our study are listed in Figure 2b.

5.2 Creating Activity Trails

To create the trails to the parking spots, for each garage-building pairing we selected 4 starting offices in different quadrants of the building (NW, SW, NE, SE) and 4 target parking spots located in different quadrants of the garage. We distributed the start and end locations across different floors in both the building and garage,

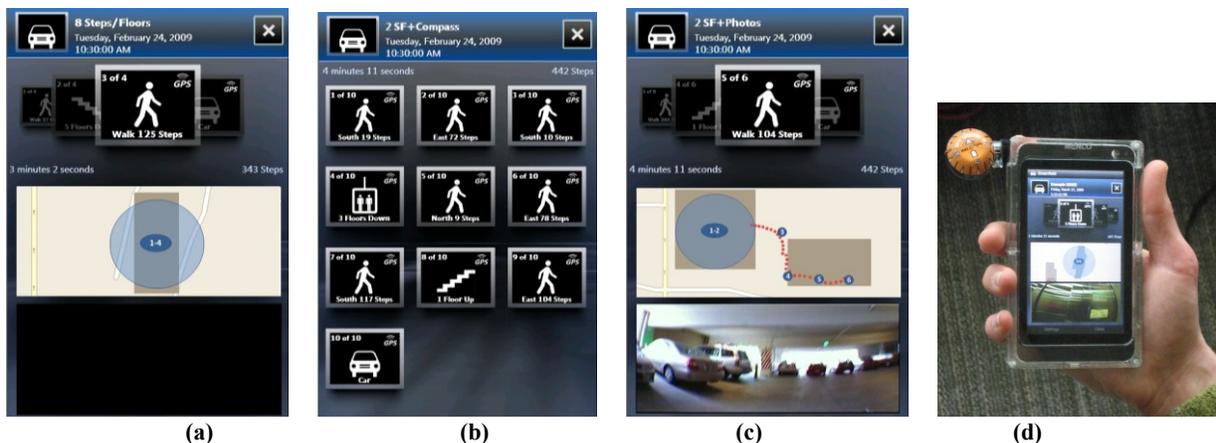


Figure 3. Example study conditions. (a) Detail screen of a Steps/Floors condition in G_{NoGPS} (note the lack of a GPS trace in the map). (b) Overview screen of a SF+Compass condition in G_{GPS} . (c) Detail screen of a SF+Photos condition in G_{GPS} . (d) Study hardware.

and rotated through different combinations of using stairs (S) and elevators (E) for each trial (e.g., SS, SE, ES, EE).

The Steps/Floors, SF+Compass, and SP+Photos presentations were hand-crafted for each trail based on data collected by two experimenters. To collect data one researcher walked from each parking spot to the associated office and back, carrying three Menlo devices in a purse, pocket and hand, and additionally wearing a pedometer, a RoyalTek RBT-2300 GPS logger as well as a SenseCam wearable camera [15] that captured environment images every 5 seconds to collect ground truth. A second researcher followed the first researcher and captured in real time the data collector's current activity {walking, stairs, elevator}, the start and end times for the activity, the number of steps or floors traversed, and associated major compass directions using a custom tablet PC application.

We had initially intended to use the internal accelerometer and barometer in Menlo to automatically obtain the Steps/Floors information to construct the study trails, adding the other information manually. However, when we analyzed the accelerometer and barometer data, we discovered that the location of the device in a pocket, bag or hand had a significant effect on the steps counted and floor determination which simple heuristics could not reconcile with reasonable accuracy. We therefore decided to use the ground truth data, and leave these issues for future advances in activity inference to address (see Section 2). For the SF+Compass condition, we chose the compass direction (e.g., N, S, E, W) corresponding to each change in the walk trajectory. Since all buildings and garages were aligned to compass points, we did not need to use directions such as SE or NW. For the SF+Photos condition (Figure 3c), we hand-selected a representative photo for each activity. For the trails to parking spots in G_{GPS} , we created a map for the detail screen from GPS data and annotated it with numbered waypoints corresponding to the associated breadcrumbs (Figure 3c). Note that while the breadcrumbs are presented to the user in the office-to-car direction (as if it were auto-reversed by the application), the pictures are in the car-to-office direction, mimicking a real use scenario where users are trying to find a car using a trail captured from car to office.

One effect of using ground truth instead of automatic inference was that we could not build real-time feedback into Greenfield, e.g., auto-advancing through the trail. Users instead browsed through trails using on-screen controls. Ways of providing automatic feedback suggested by the study responses will be discussed in Section 7.

5.3 Study Design

We recruited ten participants (5M, 5F), ranging in age from 26 to 53 (median 34) who worked in one of the two office buildings used in our study and who typically drove to work so they would be familiar with at least one of the parking garages. Five worked in the building associated with G_{GPS} (3M, 2F) and five worked in the building associated with G_{NoGPS} (2M, 3F). All participants worked at our institution, but were not involved in the research. Participants received a US\$15 café coupon for their time.

We used a within-subjects design in which participants found four parking spots in each of the two garages for a total of eight spots found; participants experienced each of the four presentation conditions in each garage (3 Greenfield and one verbal). Overall, participants used each presentation condition twice, once in each garage. To control for learning affects and familiarity with the garages, participants began by finding all four spots in the garage they parked most regularly, with half starting in G_{GPS} (parking spots 1-4) and half starting in G_{NoGPS} (spots 5-8). In each garage, the order in which the parking spots were visited was fixed across participants, but the administration of the 4 presentation conditions was counterbalanced across participants to ensure each parking spot was found using each of the 4 presentation conditions by at least two participants.

Study sessions began by gathering the participant's demographic data and personal car finding experiences and challenges, followed by a short tutorial on the Greenfield interface. For each of the 8 tasks, we led the participant to the office associated with the start of the trail, handed the participant a piece of paper with the color, make, model and license plate number of a car and told her "Your friend who works in this office needs you to get something from his car. He has given you this phone which recorded an activity trail on his way in from the car, and which you can now follow to find his car." In conditions other than verbal, we handed the participant a Menlo loaded with the appropriate trail and reminded the participant of the type of data that was available in that condition. In the verbal condition we provided only a verbal instruction (e.g., "4th floor, NE quadrant").

Participants then attempted to find the car, with a researcher following them and noting problems encountered and comments made by the participant. If needed, we prompted participants to help them get them back on track by providing the least amount of help possible (14 occurrences, analysis in Section 6.3). Once the car was found we recorded the total time and the number of steps taken based on a pedometer that participants wore. We then asked the mental effort, physical, effort, and frustration questions from the NASA Task Load Index (TLX) and how helpful participants found the information for finding the car. All interactions with Greenfield were logged by the software. At the end of the study, participants filled out a final questionnaire that asked them to compare their experience in the four different conditions, as well as an ETS Cube Comparison Test to assess their spatial orientation abilities. Total study time was about 2 hours.

The study administrators visited the 8 parking spots on the morning of each study day and recorded the details of the car parked there. In the few cases where the original car had moved, the administrator provided the updated car information to the participant once the researcher noticed the new car in the spot. We do not believe that these instances had a large impact on the results (e.g., slightly smaller step counts or time) because participants had followed the directions to get to the location and were quite close to the target spot when the change was discovered.

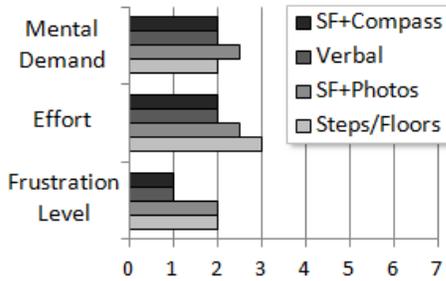


Figure 4. Median participant NASA TLX ratings after finding a car in each condition (1=“Very Low”, 7=“Very High”).

6. RESULTS

Participants’ current parking habits suggest an activity-based navigation system like Greenfield could be beneficial. In initial interviews, six participants reported losing their cars at least monthly with two reporting misplacing their car once per week. Even for participants who rarely lose their cars, when they do it can be very frustrating. For example, P10 said that when she does misplace her car: “it’s catastrophic.” To avoid misplacing their cars, most (8) try to park in the same spot or area, particularly when parking at work, and 3 participants mentioned similar strategies for parking at shopping malls.

6.1 Qualitative Comparison of Conditions

During the study each participant eventually found every car. However, the systematic variations in presentation condition and environment to which they were exposed led to notably different experiences.

Steps/Floors. Overall, participants did not find the step count information very helpful on its own. As Figure 4 shows, NASA TLX responses (1=“Very Low”, 7=“Very High”) collected after each task found that Steps/Floors scored worst on effort and tied with SF+Photos for most frustrating. Given that Steps/Floors offered no directional information, it is unsurprising that participants had more difficulty with Steps/Floors when no GPS data was available. Comparing participants’ experience between the two garages, a non-parametric Wilcoxon test showed that participants found Steps/Floors in G_{NoGPS} to be significantly more frustrating (2.5 vs. 1, $z=-2.54$, $p=0.011$), mentally demanding (3 vs. 1, $z=-2.26$, $p=.024$), less helpful (5 vs. 3, $z=-2.69$, $p=0.007$), and required more effort (3 vs. 2, $z=-2.41$, $p=.016$), than when in G_{GPS} .

Without the GPS-based map, participants used a variety of strategies for finding the cars, such as reasoning about the relative distance from the stairs or elevator based on step counts and walking in a grid pattern until finding the car. P2 commented for parking spot 7 that she was “basically thinking about walking a 53 step radius” and P10 said about the same spot, “well that sucks. I’ve got to figure out 53 steps.” P8 commented while finding parking spot 8: “no other information than 125 steps? I sense frustration.” Several participants wished explicitly for more information than was provided by the Steps/Floors condition. P7 commented “3 or 4 directions would have helped.” P1 said she needed “to know the direction out of the elevator, even NW would have been helpful.” Given that the Steps/Floors condition offered users the least information among our constructed

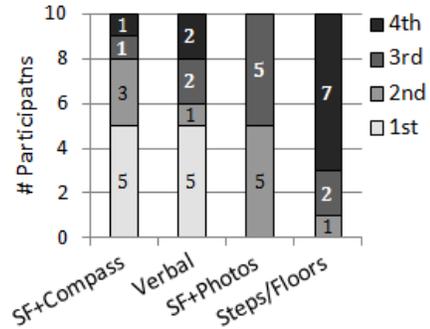


Figure 5. Total number of participants who ranked each study condition as 1st, 2nd, 3rd, and 4th preferred.

interfaces, it is unsurprising that it was ranked least useful (4th) by 7 participants at the end of the study as shown in Figure 5.

SF+Compass. Many participants clearly valued the addition of compass information, as shown by the five participants that ranked it as the most useful of the four presentations (Figure 5). SF+Compass also tied for lowest frustration level, effort and mental demand in the NASA TLX responses (Figure 4). However, comments from our participants indicated that the inclusion of directional information also led to some usability issues, primarily due to that fact that SF+Compass had the longest trails of any condition—each change in walking direction during a trail was conveyed with its own breadcrumb (Figure 3b). In SF+Compass, five participants deliberately did not follow the Greenfield trail, choosing to make their own way at the beginning and pick up the trail later. They described skipping breadcrumbs to locate the ones they needed and we observed other participants act similarly, particularly in G_{GPS} where they jumped to the garage breadcrumbs. P3 said “I’m skipping the first few directions because I know where the garage is,” P9 said “I actually don’t care whether the person took the elevator or stairs, I only care about steps 8 and 9,” while P1 asked “do I have to follow the breadcrumbs?” Although skipping was not unique to the SF+Compass condition, it was most noticeable. P10, who ranked SF+Compass as least useful, was very clear that she was overwhelmed by the number of breadcrumbs in the SF+Compass condition. Two participants voiced the opinion that instead of compass directions they would prefer more situated information. P2 asked for “Go up elevator and turn right instead of due east” and P5 said she would “rather have ‘walk to elevator’ than North or South, perhaps with an ID of the elevator, e.g., North elevator”.

As with Steps/Floors we saw evidence that the experience differed between the two garage environments, with participants finding that they had to work significantly harder using the Compass in G_{NoGPS} than when in G_{GPS} (2.5 vs. 1.5, $z=-1.99$, $p=.046$). This suggests that even in the presence of directional compass information in G_{GPS} , participants were using cues from the GPS paths to guide or confirm their progress.

SF+Photos. In contrast to the obvious directional value that compass information provided to participants, users were less enthusiastic about the value provided by photos for navigation, with SF+Photos consistently ranked in the middle of the pack (2nd or 3rd) among all representations (Figure 5). Participant interaction logs indicate all participants looked at the photos and we noted at least five trails during which photos were pivotal to the discovery of the car, especially in G_{NoGPS} where no other

directional cues were present. For example, P7 deduced the direction that his “friend” must have walked to the elevator from parking spot 6 and found the car easily, and P10 eventually determined that she was on the wrong level for parking spot 8 by examining the photo.

Despite these successes with photos, we also observed three occasions in which participants were either confused by the photos or spent considerable time trying to deduce locations from photos. Several people asked what time of day the photos were taken and realized they needed to focus on building landmarks since the cars might have changed. P8 had a particularly striking use of the photos. He walked backward in an effort to try to match what he was seeing in the photo to his surroundings, and described needing to “turn images to figure [them] out backward in his head.” Such examples of extra effort taken by users to incorporate the photos into their car-finding strategies were reflected in SF+Photos being rated as having the highest mental demand, tied with Steps/Floors for most frustrating, and rated as more effort than either SF+Compass or Verbal (Figure 4). Participants’ reliance on photos in G_{NoGPS} unsurprisingly lead to the perception of increased effort over the use of photos in G_{GPS} (3 vs. 2, $z=-2.21$, $p=.02$), where they could use both photos and GPS data.

The most frequent complaint about the SF+Photos condition was that it did not effectively convey directional information which was mentioned explicitly by 4 participants. For example, P2 complained about “lack of directions,” P5 said “this is where it doesn’t tell me what direction,” and P1 told us “compass dir[ections] would have been helpful.” On the final survey we asked participants whether it was confusing to see photos taken when walking from the car to the office. The median response was 2.5 on a 7 point scale (1=“Not Confusing” to 7=“Very Confusing”) suggesting that overall, participants were not confused. Participants’ suggestions for improving photos instead included wanting more photos per breadcrumb (P7), and picking photos from important locations like the car (P10) or turns (P8).

Verbal. Given the simplicity of the verbal instructions (floor and quadrant of garage) which necessitated a search of the entire quadrant to find the car, we were surprised by its popularity. Five participants ranked it as most useful (Figure 5), and it tied for lowest frustration level, effort, and mental demand (Figure 4). Several participants’ comments help clarify the popularity of the verbal condition, including “specific accurate and could take the elevator” (P8), “familiar, not a lot of unknowns,” (P10) and “just need verbal description to be effective.” (P6). However, other participants were less enthusiastic. P9, who ranked Verbal lowest, said “I was less confident that I was going to find it than in other tasks” and P3, who ranked Verbal third, said “[he] found that it was more mentally demanding because [he had] to actually think about whether to take stairs or elevator, etc.” (rather than having the device tell him).

In summary, participants generally preferred SF+Compass and Verbal overall, followed by SF+Photos, then Steps/Floors. This

ordering also agrees with post-study ratings for each condition, where users generally perceived SF+Compass (median=1.5) and Verbal (2.5) to be more helpful than both SF+Photos (3.5) and Steps/Floors (4), rated on a scale from (1 = “Very Helpful” to 7 = “Not Helpful”). Looking across environmental conditions, we found participants actively used GPS data to find cars when it was available, and that the task of car-finding was made noticeably more difficult in the absence of such data, but was still possible.

6.2 Quantitative Comparison of Conditions

As a loose measure of how well participants performed the car finding task, we calculated for each parking spot the *step difference*, defined as the increase in the number of steps taken by participants over the ground truth trail, and the *time difference*, defined as the increase in the time taken by participants over the ground truth trail. It is important to note that these are rough measures; participants did not necessarily follow the same trail, have the same stride length or walking speed, and external factors such as waiting for elevators or using stairs rather than elevators also affected these results. Nonetheless, comparing the time and step differences between conditions gives some insight into the relative challenges of each condition. Because participants’ subjective feedback suggested that they relied heavily on GPS data when it was available, we focused on the data collected from the task trails in G_{NoGPS} to gain better insight into the relative value of the different presentation conditions in supporting activity-based navigation when GPS was not available.

Table 1 shows the median step and time differences across the presentation conditions in G_{NoGPS} . Ground truth trail lengths for these four parking spots ranged from 134 steps to 283 steps. Of our ten participants, we had data collection problems with the step count and timing data for P4 and P6 and a hardware problem prevented capturing durations for P5. Occasions when participants’ steps and time were *less* than the ground truth, resulting in negative differences, are included in the median calculations, although this happened only twice for steps and once for time, out of a total of 34 traces.

The Verbal condition had the lowest median time and step differences, pointing to the fact that participants took relatively direct and efficient paths in the Verbal condition, which may explain why half the participants ranked the Verbal condition as “most preferred” (Figure 5). Among the three GUI-based conditions, SF+Compass had a median time difference roughly one minute faster than the other two. Interestingly, the additional time it took users to find cars using SF+Photos could not be explained by users taking longer paths, since the step differentials with SF+Photos were generally on par with, or even lower than, the SF+Compass condition; rather, we believe participants lost time due to studying the photos and mentally reconstructing the scene. Not surprisingly, Steps/Floors had the highest median step and time differences, highlighting that users took more circuitous and less focused paths in this condition. Thus, for Greenfield presentations in absence of GPS, these numbers suggest that

Table 1. Median differentials of participant data collected versus our ground truth data in G_{NoGPS} . Number of data samples in subscripts, data collection problems prevented capturing some of the step and time differences for P4, P5, and P6.

	Verbal	SF+Compass	SF+Photos	Steps/Floors
Step Differences	56 ₍₈₎	67 ₍₉₎	58 ₍₉₎	159 ₍₈₎
Time Differences (seconds)	17 ₍₉₎	65 ₍₈₎	137 ₍₈₎	170 ₍₈₎

SF+Compass generally supported the most efficient navigation, followed by SF+Photos, then Steps/Floors, which is consistent with users' relative subjective experiences across the conditions.

6.3 When Participants had Trouble

Analogous to Jonsson's [16] exploration of breakdowns to understand how people make cognitive maps, we looked in detail at problems our participants had to help us understand challenges with activity-based navigation using a mobile device. Surprisingly, the 14 times we needed to prompt participants to get them back on track were distributed across the four presentation conditions—3 in Steps/Floors, 3 in SF+Compass, 4 in SF+Photos, and 4 in Verbal—and they occurred in both the G_{GPS} (6) and G_{NoGPS} (8) garages. Although we did find gender effects in the spatial orientation test we administered, with all but one female scoring lower than all the men, these effects did not seem to carry through to the need for prompts, which were nearly evenly distributed across men (6) and women (8). We identified the following reasons for prompts:

Wrong Floor (4): In G_{NoGPS} , which had an underground garage directly below the building, 4 participants required prompting about being on the wrong floor. The first trail that participants followed in this building included a 14-step walking breadcrumb between the stairs to the ground floor and those that continued down to the garage levels. This was technically correct as these are separate stairwells, albeit close to one another. P1 required a prompt for this and expressed an interest in feedback saying “wish [there was] something telling me [I'm going] in the wrong direction.” Three other participants also had trouble, but did not require prompts. Participants appeared to conceptually think of the two sets of stairs as continuous, and when they saw the walk breadcrumb they were not certain if they needed to walk somewhere else before continuing down. For example P8, who initially went to the wrong floor, wondered if 2 floors down meant he should count building floors or sections of stairs, as floors were separated by two sections of stairs connected by a switchback landing.

The trail for parking spot 8 caused the remaining 3 prompts related to floors. As can be seen in Figure 1b, in the Elevator breadcrumb, floor levels are presented as relative distances (e.g., 5 Floors Down). This was particularly problematic for trails in G_{NoGPS} , because the office floors are numbered 1 – 4 and underground parking levels are G1, G2, and G3. Thus for parking spot 8 where participants started at a 4th floor office, they had to determine that 5 floors down meant G2. Both P5 and P10 mistakenly thought the car should be on G1, and took 654 and 617 extra steps respectively. These were the only instances where the step difference was more than two standard deviations from the average. More generally, 6 participants talked through the “floor math” verbally and sometimes experienced difficulties that they self-corrected. For example, P2, who initially got off the elevator on the wrong floor in G_{NoGPS} , said “[I] couldn't do math on floors.” P8 also talked about how he “had to count relative floors in the elevator.” P5 said “if I'm on the 4th floor I think about 4 floors down as ground and I bet others have that problem too.” This particular issue may be less of a concern in cultures where the ground floor is “G” rather than “1”.

Wrong Place (3): Participants did not always try to follow the exact trail between an office and parking spot and instead picked

up the trail once they reached the garage or another landmark. Eight of ten participants did this at least once and most of the time resumed the trail without difficulty. However, we saw two cases where participants could not successfully resume the trail once they reached the G_{GPS} garage. Both participants made the incorrect assumption that there was only one set of stairs in G_{GPS} , looked for the car near the north stairs instead of the south stairs, and had to be prompted to reconsider their assumptions. These are good examples of the increased difficulty in following someone else's trail, since they would have known which stairs to take had it been their own trail. Lastly, P7 got off track in the G_{NoGPS} Steps/Floors condition and needed a prompt to look for the car in another direction.

Memory Failures (4): In the verbal condition, participants had to rely on their memory of the floor and quadrant information given at the beginning of the task, and four prompts were required to remind our participants of what they had been told.

Missed information in user interface (3): In three cases we pointed participants to information contained in the interface that they had not observed, reminding P8 and P10 in G_{GPS} of the map in the detail view and pointing P1 to additional breadcrumbs.

7. DISCUSSION

Overall, we have shown that activity-based navigation can help users successfully navigate using trails for a car finding task. While our results used hand-crafted trails, and we do not underestimate the technical challenges that remain, much of the necessary inference, such as step counting and floor detection, has already been demonstrated in the literature (e.g., [14, 22, 28]). While users preferred having absolute positioning information and maps when they were available, users were still able to find cars without resorting to exhaustive search, with as little information as a step count and floor number. In the rest of this section we discuss our key findings that have implications for future work in this area.

7.1 Trails Were Useful Without Being Precisely Followed

While Greenfield was designed to take users step-by-step along trails, 8 of our 10 participants chose to self-navigate at least once and refer to the device interface only for guidance when they had to make a choice, e.g., which floor to go to. For activity-based navigation, this motivates using absolute rather than relative measures where possible, e.g., “turn North” rather than “turn left”, since then users who pick up a trail mid-stream will not be disadvantaged by the lack of prior context. It also motivates simplifying the user interface for trails to facilitate quickly finding such decision points. This could be done by automatically chunking long trails and labeling them, e.g., grouping by “outside” and “inside” using GPS signal, or segmenting at elevator/stairs breadcrumbs since they are typically key navigation stages. If similar trails are recorded every day (as would be the case for car finding), another method for highlighting decision points would be to learn over time which breadcrumbs varied and which parts stayed relatively constant, and highlight the former.

While we did not explore automatic feedback in the Greenfield UI in the study (e.g., auto-advancing between breadcrumbs), users' self-navigation has implications pertaining to feedback for activity-based navigation. In particular, the interface should allow

for the fact that the user may deliberately choose not to precisely follow the path, but still provide feedback when possible. For example, a device could tell the user how many floors they have just descended without being concerned how they moved between floors. This might help address “floor math” issues participants experienced by highlighting errors in number of floors ascended or descended. We are particularly interested in exploring ways that feedback could be more interactive. For example, Greenfield could indicate the likelihood that a user is currently at a given breadcrumb using beacon-based location techniques [1], but also allow a user to self-declare where they are in the trail if they wish, facilitating joining trails partway.

Finally, our participants were familiar with the general layout of the buildings/garages used, but in other activity-based navigation scenarios where a user is navigating through unknown surroundings (e.g., finding a meeting room in an unfamiliar building), the trail may be more faithfully followed and automatic feedback with auto-advancing of the trail, which is familiar from car navigation systems, may be useful. For example, P8 said “it would be great if I didn’t have to count [steps] in my head,” and would have liked Greenfield to help him keep track of how far he had walked.

7.2 Breadcrumbs Interpreted at High Level

Our study showed that human interpretation of where breadcrumbs should start and end differs significantly from what naïve inference based on sensor data would produce. Participants routinely voiced that the step counts were inaccurate, even when the worn pedometers proved the numbers to be quite accurate. Typically this was due to participants failing to begin counting steps when the pedometer did—even when following the trail precisely. For example, participants would unconsciously omit the steps taken from the elevator to the door of the elevator vestibule, as they viewed the “elevator” breadcrumb as ending there. Similarly, participants were confused by a “walking” breadcrumb between two “stairs” breadcrumbs that was arguably correct since the two stairwells were not vertically aligned, but participants considered it logically as a continuous set of stairs that traversed multiple levels.

It may be difficult for inference algorithms to correct for this human interpretation, since some information may be missing, e.g., the location of the elevator vestibule door in the trail. However, another way to look at this is that reporting excessive precision can be counterproductive. For example, providing only N, S, E, W rather than precise degree values worked well during our study. In addition, hiding low-steps breadcrumbs and providing round values (e.g., to one significant figure) rather than precise step counts may avoid confusing users with too much detail. Finally, the popularity of the verbal condition indicates that users liked having a high-level summary. Simple summation of the breadcrumb information could be used to offer this in Greenfield, e.g. “3rd floor, approximately 100 steps NW from elevator.”

7.3 Relative Value of Sensed Information

In our study, we looked at five types of activity information: step counts, floor changes, compass directions, photographs, and absolute location from a GPS. We felt strongly that providing a minimum of floor and step information was important to avoid

exhaustive search strategies, and therefore focused on understanding the relative benefits of absolute position, compass and photos for helping participants make more efficient navigation choices. We found that absolute position information, both in the form of paths (GPS traces), and floor/quadrants locations (verbal), was the most effective and required the least amount of effort in helping users reach their goal. However, our study also showed that in situations where absolute location was not available, activity-based navigation using trails is a viable means of leading users to geographical locations.

The popularity of SF+Compass overall, and the fact that it benefited users over Steps/Floors in both garages, highlight that real-time directional data was valuable for users in finding the cars, regardless of the presence of GPS information. This motivates including barometric pressure sensors and magnetometers on mobile devices. It also motivates further research into localizing a device with respect to the person carrying it, building off previous work (e.g., [5, 12, 20, 25]), so that magnetometer data can be translated into movement direction. Similarly, work in accurately determining floor level irrespective of where a device is carried (e.g., hand, pocket, purse) is also motivated to ensure this information is available in practice.

While photos provided value for some participants, they were less valued than compass information in our study and increased the mental demand on participants. Offering more photos per breadcrumb as well as associating them with salient landmarks in the trail might improve their value to users. This would be possible in real deployments, as photos would be captured by the same device performing the activity inferencing. In different applications of activity-based navigation, photos may also prove more valuable than for our car finding setting, as photos often look similar in garages and therefore provide relatively little orientation information.

8. CONCLUDING REMARKS

We have explored the feasibility and usefulness of activity-based navigation on a mobile device through the lens of the application of car finding (which is useful when one loses one’s car, or has to find a car parked by someone else). We used hand-crafted *trails* of information (e.g., “walk 10 steps north”, “go up two floors using the stairs”) based on data that can be inferred from sensors: step counts, compass direction, floor changes (including the mode of change, stairs/elevator), photos and absolute positioning (when available).

In general, participants were quite successful using Greenfield. As P9 said “I did not have to remember specific details (which floor, how far from stairs, which aisle)—rescuing my own memory is highly beneficial since this is the type of detail I typically don’t pay that much attention to.” While our study had limitations including the lack of real-time feedback in following the trails, we found lessons for future activity-based navigation user interfaces. These lessons include the need to provide designs that allow users to join a trail at any point along the way, that users interpreted breadcrumb start and end points differently than the literal interpretations from sensed data (e.g., elevators “end” after the vestibule, staircases near each other are regarded as a single stairwell), that too much precision and detail hindered rather than helped users, and that providing summaries of trails would be beneficial.

Directions for future work include furthering the state of the art of inference algorithms to better support activity-based navigation, studying activity-based navigation in less familiar environments, and exploring how and what information could be presented to help users cope with and overcome inference errors which will be present in any real-world system.

9. ACKNOWLEDGMENTS

Our thanks to our study participants.

10. REFERENCES

- [1] Bahl, P. and Padmanabhan, V. N. RADAR: an in-building RF-based user location and tracking system. In Proc INFOCOM 2000, IEEE (2000) 775-784.
- [2] Bao, L. and Intille, S., Activity Recognition from User-Annotated Acceleration Data. In Proc. Pervasive 2004, Springer (2004) 1-17.
- [3] Beeharee, A., Steed, A., A natural wayfinding exploiting photos in pedestrian navigation systems, In Proc. MobileHCI'06, ACM (2006) 81-88.
- [4] Benford, S., Seagar, W., Flintham, M., Anastasi, R., Rowland, D., Humble, D., Stanton, D., Bowers, J., Tandavanitj, N., Adams, M., Farr, J. R., Oldroyd, A., and Sutton, J. The Error of our Ways: The experience of Self-Reported Position in a Location-Based Game. In Proc. UbiComp 2004, Springer (2004).
- [5] Blanke, U., and Schiele, B. Sensing Location in the Pocket. In Extended Abstracts of UbiComp 2008.
- [6] Car Spotter, appshopper.com/navigation/car-spotter
- [7] Consolvo, S., McDonald, D., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., & Landay, J.A. Activity Sensing in the Wild: A Field Trial of UbiFit Garden. In Proc. CHI 2008, ACM (2008) 1797-1806.
- [8] Downs, R.M, Stea, D., Cognitive Maps and spatial behavior: Process and Products. In Image and Environment. (pg. 8-26). Aldine (1997).
- [9] Frank, C., Bolliger, P., Mattern, F., and Kellerer, W. The Sensor Internet at Work: Locating Everyday Items Using Mobile Phones. Pervasive and Mobile Computing 4:3, June 2008, 421-447.
- [10] Froehlich, P, Obernberger, G., Simon, R., Reichl P., Exploring the Design Space of Smart Horizons, In Proc. Mobile HCI 2008, ACM (2008) 363-366.
- [11] GparkS, www.freewarepocketpc.net/ppc-tag-parking.html
- [12] Harrison, C. & Hudson S., Lightweight material detection for placement-aware mobile computing. In Proc UIST 2008. ACM (2008) 279-282.
- [13] Hile, H. Grzeszczuk, R., Liu, A., Vedantham, R., Kosčeka, J., Borriello, G., Landmark-Based Pedestrian Navigation with Enhanced Spatial Reasoning. In Proc. Pervasive 2009, Springer (2009), 59 – 76.
- [14] Huynh, T., Blanke, U. and Schiele, B. Scalable Recognition of Daily Activities from Wearable Sensors. In Proc. LOCA 2007, Springer (2007).
- [15] Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N. and Wood, K. SenseCam: a Retrospective Memory Aid. In Proc. UbiComp 2006, Springer (2006). 81 – 90.
- [16] Jonsson, E., Inner Navigation, Scribner (2002).
- [17] Kray, C., Elting, C., Laakso, K. and Coors, V. Presenting route instructions on mobile devices. Proc. IUI 2003. ACM Press (2003), 117-124.
- [18] Krüger, A., Butz, A., Müller, C., Stahl, C., Wasinger, R., Steinberg, K., Dirschl, A. The Connected Interface: Realizing a Personal Situated Navigation Service. In Proc. IUI 2004, 161-168.
- [19] Krumm, J., & Hinckley, K. The NearMe Wireless Proximity Server, In Proc. UbiComp 2004, Springer (2004) 283-300.
- [20] Kunze, K., Lukowicz, P., Junker, H., & Tröster, G. Where am I: Recognizing On-body Positions of Wearable Sensors, In Proc. LOCA 2005, Springer (2005) 264-275.
- [21] Lee, M., Dey, A., Providing Good Memory Cues for People with Episodic Memory Impairment, In Proc. Assets 2007, 131-148.
- [22] Lester, J., Choudhury, T., Borriello, G., A Practical Approach to Recognizing Physical Activities. In Proc. Pervasive 2006, Springer (2006) 1-16.
- [23] Liao, L., Patterson, D. J., Fox, D., and Kautz, H., Building Personal Maps from GPS Data. In Proc. MOO 2005.
- [24] Liu, A., Hile, H., Kautz, H., Borriello, G., Brown, P., Harniss, M., Johnson, K., Indoor Wayfinding: Developing a Functional Interface for Individuals with Cognitive Impairments. In Proc. ASSETS 2006, 95-102.
- [25] Mizell, D., Using Gravity to Estimate Accelerometer Orientation, International Symposium on Wearable Computers, White Plains, 2003.
- [26] NaviSeer, www.seertechnology.com/naviseer.html
- [27] Ofstad, A., Nicholas, E., Szcodronski, R., and Choudhury, R. R. AAMPL: accelerometer augmented mobile phone localization. In Proc. MELT Workshop. ACM (2008).
- [28] Ravi, N., Dandekar, N, Mysore, P., and Littman, M. Activity Recognition from Accelerometer Data, In Proc IAAI 2005, AAAI (2005) 1541—1546.
- [29] Robertson, P., Angermann, M, Krach, B., Simultaneous Localization and Mapping for Pedestrians using only Foot-Mounted Inertial Sensors. In Proc. UbiComp 2009, ACM (2009) 93-96.
- [30] Woodman, O., & Harle, R., Pedestrian Localisation for Indoor Environments, In Proc. UbiComp 2008, ACM (2008) 114-123.