5-2017

Follow-my-lead: Intuitive indoor path creation and navigation using see-through interactive videos

Quentin ROY
Singapore Management University, quentinroy@smu.edu.sg

Simon T. PERRAULT
Yale-NUS College, Singapore

Shengdong ZHAO
National University of Singapore

Richard DAVIS
Act8 Design Pte Ltd

Anuroop PATTENA VANIYAR
Singapore Management University, anurooppv@smu.edu.sg

Citation

ROY, Quentin; PERRAULT, Simon T.; ZHAO, Shengdong; DAVIS, Richard; PATTENA VANIYAR, Anuroop; VECHEV, Velko; LEE, Youngki; and MISRA, Archan. Follow-my-lead: Intuitive indoor path creation and navigation using see-through interactive videos. (2017). *CHI '17: Proceedings of ACM CHI Conference on Human Factors in Computing Systems, Denver, May 6-11*. 5703-5715.

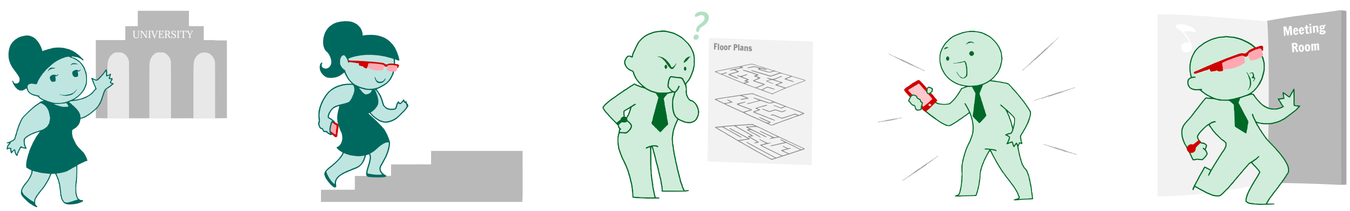
Available at: https://ink.library.smu.edu.sg/sis_research/3745

Follow-My-Lead: Intuitive Indoor Path Creation and Navigation Using Interactive Videos

Quentin Roy [1], **Simon T. Perrault** [2], **Shengdong Zhao** [3], **Richard C. Davis** [4], **Anuroop Pattena Vaniyar** [1], **Velko Vechev** [3,5], **Youngki Lee** [1], **Archan Misra** [1]

[1] Singapore Management University, Singapore [2] Yale-NUS College, Singapore
[3] NUS-HCI Lab, National University of Singapore, Singapore
[4] Act8 Design Pte. Ltd., Singapore,
[5] Chalmers University of Technology, Gothenburg, Sweden

quentin@quentinroy.fr, simon.perrault@yale-nus.edu.sg, zhaosd@comp.nus.edu.sg, rcdavis@act8design.com, anurooppv@smu.edu.sg, velko.vechev@gmail.com, youngkilee@smu.edu.sg, archanm@smu.edu.sg



Anna is organizing a talk on her university's campus. | As she arrives, she records the path **leading** to the room using *FML*. | Bob arrives later. First, he is disoriented. | Then, he realizes Anna sent him an *FML* path. | He **follows** it and reaches the room quickly and easily.

Figure 1. *Follow-My-Lead* (FML) scenario. Drawings ©Marie Dorso.

ABSTRACT
We present *Follow-My-Lead*, an alternative indoor navigation technique that uses visual information recorded on an actual navigation path as a navigational guide. Its design revealed a trade-off between the fidelity of information provided to users and their effort to acquire it. Our first experiment revealed that scrolling through a continuous image stream of the navigation path is highly informative, but it becomes tedious with constant use. Discrete image checkpoints require less effort, but can be confusing. A balance may be struck by adding fast video transitions between image checkpoints, but precise control is required to handle difficult situations. Authoring still image checkpoints is also difficult, and this inspired us to invent a new technique using video checkpoints. We conducted a second experiment on authoring and navigation performance and found video checkpoints plus fast video transitions to be better than both image checkpoints plus fast video transitions and traditional written instructions.

CHI 2017, May 06-11, 2017, Denver, CO, USA

DOI: http://dx.doi.org/10.1145/3025453.3025976

ACM Classification Keywords
H.5.2. User Interfaces (D.2.2, H.1.2, I.3.6): Interaction styles (e.g., commands, menus, forms, direct manipulation)

Author Keywords
Indoor Navigation; Mobile Computing; Smartglasses; Video; Wearable Computing; Leader-Follower; Egocentric Visual Navigation.

INTRODUCTION
Way-finding remains a complicated but essential task in many urban environments. Traditionally, users rely on maps and signs to orient themselves in unfamiliar outdoor environments, a task known to be difficult [16, 36]. Today, however, outdoor navigation has many viable commercial *You are Here* solutions, e.g. Google or Apple Maps, that offer fast and convenient guidance and have been universally adopted by mobile users [25].

In contrast, user-friendly indoor navigation continues to be a challenging problem, with commercial solutions being limited in scale and coverage. Google Indoor Maps, for example, currently covers about 10,000 locations, a tiny sliver of commercial buildings worldwide [54]. Indoor navigation poses two widely recognized challenges. First, GPS is less reliable indoors and localization requires additional infrastructure (e.g., WiFi APs or BLE beacons). Second, most buildings do not have up-to-date floorplans that are readily available [28].

Other challenges include the need for 3-D wayfinding [23] and the higher complexity of indoor vs. outdoor environments. Consequently, indoor navigation systems that show a user's location on a map are likely to be rare for some time.

We present *Follow-My-Lead* (*FML*), an alternative solution to indoor navigation. Unlike previous approaches, *FML* needs neither localization nor maps. Thus, it makes no assumptions about the user's environment or infrastructure support, does not require any particular data, and can be used anywhere. To meet this goal, *FML* relies on a leader-follower approach [43] and implements an *egocentric visual navigation* strategy [11, 21]. A leader walks through the entire path once and captures what followers should see along the route. Followers navigate by comparing the views presented by *FML* (Figure 2) to what they see in the real world. The system is fully functional on a mobile phone, but it has been designed to be used with a head-mounted camera and display (i.e. smart-glasses) controlled through a smartwatch or a smartphone in a cross-device fashion.

Though egocentric visual navigation has been investigated before, our work shows that the interaction design is not straightforward. On one extreme, scrolling through a continuous image stream – a technique we call *Continuous Flow* (*CF*) – allows followers to reach their destination with few errors, but it also requires continuous interaction and attention. This quickly becomes tedious and can make it difficult for followers to maintain proper situational awareness. At the opposite extreme, our results show that a succession of still images [5, 11, 21] – a technique we call *Image Checkpoints* (*IC*) – requires much less interaction and attention, but is hard to follow. To mitigate both drawbacks, we designed an intermediate solution that we call *Image Checkpoints + Fast Forwarding* (*IC+FF*). This third technique relies on image checkpoints, but illustrates transitions by playing the path video between checkpoints at high speed. This technique strikes a balance between the extremes, giving followers richer information while reducing their effort to acquire it. However, precise control remains required to deal with complicated segments.

Also, our investigations show that selecting the right set of checkpoints can be challenging for the leader, who must author this path. Consequently, we modified *Image Checkpoints + Fast Forwarding* to develop an alternative approach, called *Video Checkpoints + Fast Forwarding* (*VC+FF*). Using this technique, leaders have the ability to mark a segment of the path as a checkpoint. Video checkpoints ease the pressure on leaders to capture a single right image, and they turn out to be more helpful to followers as well.

To our knowledge, this is the first time that the development of egocentric visual paths has been studied from the perspective of both *authoring* and *following*. Our key contributions include:

1. An efficient indoor navigation system, based on egocentric visual guides, that works anywhere, requiring neither floor plans nor infrastructural support,
2. An exploration of four alternative egocentric visual navigation strategies and their trade-offs in terms of (1) fidelity of the leader's path to the follower's view, (2) effort required to follow paths, and (3) effort required to author paths.
3. Empirical evidence collected through two experiments: one with twelve participants and three tasks that shows the benefits of *IC+FF* for followers, and another with 18 participants and three tasks that shows the benefits of *VC+FF* for both leaders and followers.

RELATED WORK

Indoor Localization

In outdoor settings, localization is typically performed using either GPS [18] or via triangulation of cell tower signals [7]. Both approaches are highly unreliable indoors [50].

Strategically installing localization artifacts inside buildings is generally the most accurate strategy for indoor localization. Radio Frequency Identifier Description (RFID) [35, 53] and infrared [1, 49] tend to be used most often in the literature. Other systems make use of ultrasound [39, 50], RF beacons [26,39], Bluetooth beacons [24,47] or even visual markers (e.g. barcodes) to be scanned [2, 34, 44]. As these approaches rely on enhancing the environment with new equipment, installing and maintaining this equipment can be expensive, and each organization (e.g. museum, airport, or university) must make its own investment.

Another approach is to record and process the signal of wireless local area networks (WLAN) to determine user location [3, 17, 56], e.g. RADAR [3]. However, the precision can vary, in particular due to multipath issues and signal similarity (i.e., *fingerprinting ambiguity*) [2, 39, 54].

Visual localization approaches make use of computer vision to correlate images taken from the pedestrian's device with localized reference images from a database [54, 55, 57]. However, reference images can easily be missed on the way, particularly if the user is not invested in the reference matching task. Visual localization user interfaces may need to be adapted to encourage users to help the system [31].

Finally, dead-reckoning solutions incrementally calculate users' location by aggregating the data of movement sensors (e.g. accelerometers or pedometers) since the last known location [19, 28, 42, 43]. However, the sensor data is noisy and needs to be carefully processed [19]. Also, as the localization is an incremental process, errors accumulate over time (*drift* effect). As a result, these techniques often suffer from poor accuracy and landmark-based corrections are occasionally needed to recover from drift [48].

Despite more than a decade of research, indoor localization has still seen very limited deployment [54]. Our approach, based on the leader-follower paradigm, does not need underlying location support [4] and thus dispenses with the usual assumption that indoor localization will soon be widespread.

Indoor Navigation Guides

Maps are by far the most widespread outdoor navigation guides [1, 2, 25, 32], but they are sometimes difficult to read [16, 36], even when they indicate the user's location (*You-Are-Here* maps) [25]. Indoor navigation in single-floor buildings is often even more complicated [6,40], and the complexity grows when one has to move from floor to floor [23]. Moreover, floor plans are often not available: Google Indoor Maps, arguably the largest indoor maps database, only covers about 10,000 locations worldwide, which remains a "tiny fraction of millions of shopping malls, museums, and airports worldwide" [54], not to mention universities and office campuses.

Instructions, e.g. turn-by-turn instructions telling the user what he should see and do at each step [10], are another common navigation guide. Instructions are generally textual [5,20,21], pictographic [10,33] or audio [20,22]. Pure activity-based instructions (e.g. directions to take, distance to walk, stairs to climb, etc) [10,33] can sometime be automatically inferred by a computer from sensor data [10,41]. Instructions that include landmark clues (e.g. *Pass the door X23, cross the open-space*) [9,20] are preferred [30] but need to be created manually.

In addition, the use of vibrations to give instructions has been the focus of much research work. "Shoulder-tapping" approaches use the location of the vibration (e.g. left wrist) to induce the user to take a turn [8,45,46]. Pielot et al. instead used temporal patterns to indicate direction [38], which only requires one actuator. Vibrations are relatively unobtrusive compared to visual navigation: users remain free to concentrate their sight on their surroundings. However, vibrations have been found to be less-efficient than GPS-based personal navigation devices for pedestrians [37].

Egocentric visual navigation relies on giving users imagery of what they should be seeing along the way [5,11,12,17,21,51,55]. A major advantage of this approach is the visibility of landmarks [5,12], by far the most preferred type of navigation cues [30]. Many of these systems provide a sequence of images at strategic locations along the path [5,11,20]. Some systems adjust the visual perspective of landmark images automatically based on the orientation of users' mobile devices, but these approaches have been evaluated primarily outdoors and require manual curation of geotagged landmark image corpora [21]. Some use video instead of images [11,12], but since the video speed seldom matches users' preferred walking speed, many users view videos before starting to move. Wenig et al. addressed this by having users manually scroll through a stream of images displayed on a watch [51]. In studies, egocentric visual navigation has proven efficient for people with cognitive disabilities [11,27] and for aging populations [20]. Compared to maps without localization, egocentric visual navigation has been shown to be more useful, easier to use, less cognitively demanding, and less stressful for younger populations with no disability [55].

FOLLOW MY LEAD

The *Follow-My-Lead* (*FML*) project aims to develop a realistic indoor navigation tool that can be used today in any building. This implies that we cannot rely on new infrastructure (e.g. RF beacons) or new data (e.g. buildings maps, which are often unavailable).

Figure 2. *Follow-My-Lead*'s display with the *Video Checkpoints + Fast Forwarding* technique. The blue areas on the progress bar at the top represent checkpoints.

The scenario represented in Figure 1 illustrates a common situation in indoor navigation. Using *FML*, Anna can easily create a path to any meeting's location and distribute it to attendees, and Bob can find his way without additional assistance.

Approach

To meet these constraints, we opted for a leader-follower approach [43,57]: a *leader* creates a path and sends it to *followers* to help them reach the destination. We formulated two main design goals for *FML*:

1. A follower should be able to efficiently and easily follow a path.
2. A leader should be able to create a high-quality path quickly, ideally by walking the path only once.

Our work is inspired by previous work that leveraged images of landmarks to enhance navigation [5,11,21,55]. We pushed this concept to the extreme and rely exclusively on *egocentric visual navigation*: followers are provided with an interactive display of what they should be seeing all along the way. Hence, followers' navigation strategy is to follow the interactive display by comparing it to the real world. As our experiments will show, a careful balance of trade-offs in the design of this interactive display can make this approach easy and efficient for both leaders and followers alike.

Apparatus

We envision a future where wearable devices are widespread and interconnected, allowing cross-device interactions [13,14]. With this in mind, while our implementation is fully functional on a single smartphone, we designed both leader and follower interfaces to be viewed on smartglasses and controlled through smartphones or smartwatches. The camera of the smartglasses can accurately and effortlessly capture what leaders see along their path. In contrast, holding a smartphone at eye level while walking can be tiresome. For followers, manipulating a smartphone on the go usually means walking "head-down", which reduces awareness of their surroundings [8]. Instead, a head-mounted display enables them to rapidly switch their focus between the display and the real world without any head movement.

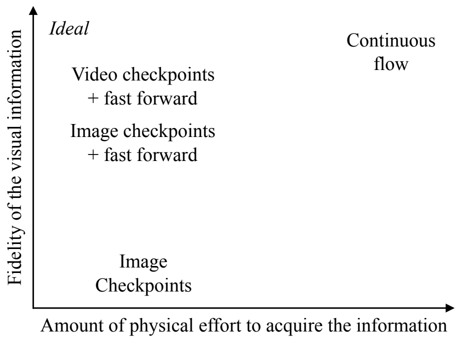


Figure 3. The fidelity / effort trade-off. Fidelity to the follower's view depends on the quantity of images shown, how informative those images are (assuming leaders select meaningful checkpoints), and when images appear. Physical effort is estimated by how much interaction a technique requires.

Design for Following: Fidelity versus Effort

Our early efforts focused on designing the information display for followers. Our prototypes and pilots revealed a relationship between two essential properties: the *fidelity* of the path information, and the *effort* needed to navigate through it. We say path information is high-fidelity if it is close to what a follower actually sees, and we say it is low-fidelity if visual imagery is sparse, poorly chosen, or shown at the wrong time. Similarly, navigation effort is low if it is automatic and high if it takes constant interaction. In the absence of automatic positioning, these properties are linked. For example, a continuous image stream of the leader's view provides high fidelity, but it requires high effort to synchronize the images with the follower's view. At the opposite extreme, a path with a few images may take little effort to step through, but it has such low fidelity to the follower's view that it may be hard to follow. To explore this trade-off, we designed our first two techniques at opposite ends of this design space (Figure 3): *Image Checkpoints* provide a low-fidelity path view and requires little user interaction, while *Continuous Flow* provides a full recording of the leader's view but requires continuous interaction while navigating. We then designed a third technique, *Image Checkpoints + Fast Forwarding*, as a compromise between the two.

All techniques are designed to be manipulable without the need to remove one's eyes from the route.

Image Checkpoints (IC)
Using image checkpoints to mark important steps on a route is a straightforward navigation solution proposed several times in the literature [5, 21, 27], though rarely as a self-sufficient navigation method [11]. Leaders can easily record checkpoints by taking snapshots as they walk. Following these checkpoints requires very little interaction: after reaching one checkpoint, followers switch to the next with a single action. However, the ease of finding the next checkpoint depends on how well the leader chose it. If the next checkpoint is not clearly visible from the current one, followers may get lost.

With our system, when a smartwatch is present, followers tap on the watch to go to the next checkpoint or press and hold for 330ms to go back to the previous one. When running on a smartphone, followers use the volume buttons (up = next, down = previous). This is similar to a technique from Chang et al. developed for cognitively impaired users, but we add simple controls to remove the need for localization technology [11].

Continuous Flow (CF)
At the opposite end of the spectrum, our second design displays the totality of the leader's view from beginning to end. Followers *scroll* through the entire image stream and can precisely adjust the progress of the video to their pace and exact location on the route.

Since recording the path requires no action other than walking the route, the quality of the path is less dependent on the leader. However, the information is raw and unfiltered, and some of it may be distracting (e.g. turning to greet someone) or uninformative (e.g. slow walks down long corridors). It is the follower's responsibility to find the most relevant segments, such as turns and intersections.

When using a smartwatch for input, followers scroll using circular gestures around the bezel of the watch, as in the Cyclostar approach [29]. When using a smartphone, followers scroll by dragging on the touch screen. (Note: While this paper was being written, Wenig et al. proposed a similar approach for smartwatches [52].)

Image Checkpoints + Fast Forwarding (IC+FF)
We designed *IC+FF* to be a compromise between the two other techniques. Like *IC*, it relies on a succession of checkpoints chosen by the leader. However, the transition between two checkpoints is shown as a high speed video. We empirically determined that playing the video at 15 times the original speed kept the transition short while ensuring that it was easily intelligible. With this design, followers receive all the information they received in *CF*, but with a greater focus on the parts that the leader believed to be most relevant. The fidelity of this display is only slightly lower than *CF*, while the ease of interaction for the follower is equivalent to *IC*. The controls for this design are also identical to *IC*.

IC+FF resembles another technique from Chang et al. [11], but our image flow is faster and we added controls. Our conceptual model is also slightly different: Chang et al. make use of series of video prompts while *IC+FF* is a single video of the path that is paused at strategic locations.

Display
In our experiments, we used the Epson Moverio BT-200 smart glasses, which puts the display at the center of the user's vision. Ideally, the display should appear at the bottom or the right [15]. However, most of today's mobile head mounted displays have a limited field of view and little or no control over the placement of the display. We expect the next generation to overcome this limitation. In case of discomfort, we allowed users to hide the display by tapping the watch with two fingers or or by pressing both of the phone's volume buttons simultaneously. However, our participants rarely used this feature in our experiments.

Design for authoring

Our later efforts focused on path creation. Our first experiment showed that the *IC+FF* technique was promising, but that choosing a good set of image checkpoints can be challenging. Choosing well is important, as previous work shows that poor choice of landmark images makes a path hard to follow [5, 21, 27]. This problem inspired a fourth design.

Video Checkpoints + Fast Forwarding (VC+FF)

With *VC+FF*, instead of taking image snapshots for checkpoints, leaders can record a whole segment of the path as a video checkpoint. This recording can include audio, giving the leader an opportunity to point out important details by speaking. As with *IC+FF*, the follower's display will show less relevant segments of the route as high-speed (15×) videos, while the video checkpoints will be shown at normal speed. While video checkpoints may take slightly longer to watch and might slow down the follower, they provide more information and reduce pressure on the leaders to choose the right image. For example, instead of wondering when to put a checkpoint to represent a turn, a leader may record the whole turning movement as a video checkpoint. To our best knowledge, there are no techniques similar to *VC+FF* in existing literature.

Authoring

To create a path, a leader need only record the route and indicate where checkpoints should be. We designed *FML*'s authoring tool so that both are captured in parallel: leaders mark checkpoints while they walk the route. Hence, no post-processing is required, creating a path is fast, and the authoring interface is minimal.

When using a smartphone, we again relied on volume buttons so the interface could easily be used eyes-free. In the case of *IC* and *IC+FF*, a checkpoint is recorded when a button is released. In the case of *VC+FF*, checkpoints are recorded while a button is being pressed. When using a smartwatch, in the case of *IC* and *IC+FF*, a checkpoint is recorded with a tap on the display. In the case of *VC+FF*, a first tap starts the recording of a video checkpoint, and a second tap ends it.

EXPERIMENT 1: FOLLOWERS' CONTROL STRATEGY

In our first experiment, we explored the trade-off between fidelity to the follower's view and the effort needed to explore the path (see Figure 3). We compared the *Continuous Flow* (*CF*) and *Image Checkpoints* (*IC*) techniques, located at two different extremes of the scale, and the *Image Checkpoints + Fast Forwarding* (*IC+FF*) technique, an in-between design.

We hypothesized that *CF* would allow users to follow paths with minimal cognitive load, but would quickly become uncomfortable due to the constant interaction required. We also hypothesized that *IC* would be cognitively demanding, because it only provides partial information, though it requires little interaction. We thought *IC+FF* would be a good compromise between these two techniques.

Path Creation

Path creation was a critical step in our experimental design. Each participant can follow a path only once. Thus, we had to create equivalent paths in terms of length and difficulty.

Positioning checkpoints was another critical aspect, as bad checkpoint positioning has an impact on path difficulty [5, 21, 27]. After multiple tests and pilots, we defined four following rules: (1) checkpoints should contain at least one salient feature (e.g. sign, window...); (2) it should always be possible to see the next checkpoint from the current checkpoint; (3) there should be one checkpoint before and after each turn, staircase or door; (4) there should be no more than 30 seconds between checkpoints.

We used pilots to assess the equivalence of the routes under each *Technique* and iterated to obtain a set of routes that met our criteria. The routes were created on the campus of the National University of Singapore (NUS). All were confined to pedestrian areas, mostly indoors, and approximately 250 meters long.

Participants

Twelve participants were recruited from Singapore Management University (SMU) students and staff (5 females) aged 18 to 30 years old ($M = 22.6$, $SD = 3.8$). They received the equivalent of 7.4 USD for their participation. They were all in good physical condition. The experiment was conducted on the campus of NUS, and all participants declared that it was their first time on campus.

Apparatus

The experimental software was run on Epson Moverio BT-200 smart glasses and a Huawei Watch. The watch was worn on the non-dominant wrist and participants used its touch screen to input commands to the system. Additionally, we used a Samsung Galaxy S7 Edge to connect the glasses and the watch together.

Task

Participants were asked to use our system to follow three different paths using the three different techniques. They were instructed to balance speed and accuracy to reach the destination as quickly as they could while still taking as few wrong turns as possible. They could walk rapidly but were not allowed to run (for safety reasons).

Procedure

Participants began the experiment by filling out a demographic questionnaire. We then described the whole experiment and led them to a training path, where participants tried each of the three techniques on different ∼100 meter segments. After training, participants completed the three experimental paths, each with a specific technique. Finally, they filled out a post-experimental questionnaire where they gave subjective feedback.

The experiment was run during week days, and paths were not blocked. Crowd size was random and usually varied from 5 to 50 persons on the way.

Design

We used a within-subjects design with one independent variable: *Technique {CF, IC, IC+FF}*. Our design was fully counterbalanced: three paths were created, and each technique

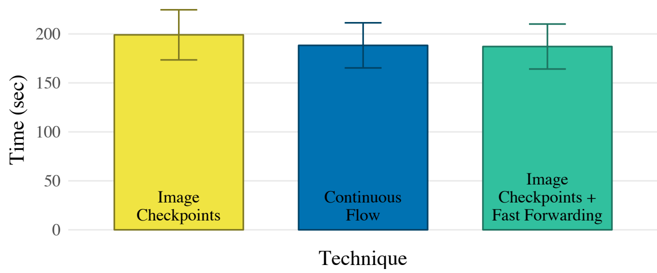


Figure 4. Total path time. Error bars show .95 confidence intervals.

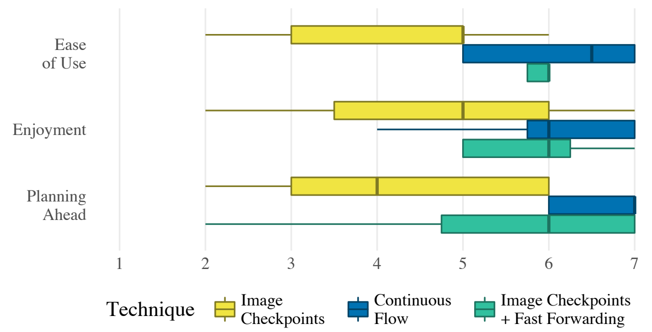


Figure 5. Subjective Feedback by Technique. Each box shows the 25th and 75th percentiles with the median as a bold line. The whiskers show min and max values.

was used on each path four times. We measured time and errors as dependent variables. Time was measured from when a participant started a path to when they reached the destination. Errors were counted if the user went the wrong way for more than two seconds. In case of such errors, participants were asked to correct their path.

Each participant took approximately 1 hour 30 minutes to complete the experiment. Participants were allowed to take breaks between each trial. We recorded 12 participants × 3 techniques = 36 trials.

Results

All measurements were taken manually by an experimenter following the participant. We recorded the total time to reach the destination, the number of hesitations and the number of errors (i.e. when a participant took a wrong direction). Note: for ANOVA, we applied Greenhouse Geisser correction when the sphericity assumption was violated.

Homogeneity of the Paths

Since our experiment required it, we checked to ensure that our three paths were equivalent in terms of length and completion time. The average path completion times were 192s, 195s, and 186s. A one-way ANOVA did not show any significant effect of *Path* on completion time ($p = .17$), suggesting that the three paths we created for this experiment were roughly equivalent.

Total time

Total path time results are shown in Figure 4. The average time to complete a path was 191 seconds, which suggests an average walking speed of 4.7 km/h. A one-way ANOVA showed a significant main effect of the *Technique* on Total Time ($F_{1.62,17.92} = 5.39, p = .019$). *IC+FF* was the fastest technique ($M = 187$s), closely followed by *CF* ($M = 188$ s) and *IC* ($M = 199$ s). Pairwise t-tests with Bonferroni corrections showed that *IC+FF* was significantly faster than *IC* ($p < .05$). No other significant differences were found.

Errors

As mentioned earlier, errors were counted when participants went the wrong way for more than two seconds, after which they were asked to correct their path. Overall, participants performed an average of 0.47 errors/trial, with 0.33 errors/trial for *IC+FF*, 0.42 for *CF* and 0.66 for *IC*. A Cochran Q test did not show any effect of *Technique* on the number of errors ($p > .05$).

Subjective Feedback

We asked our participants to rate their experience with each technique, using a 7-point Likert scale on (a) how easy they found the technique to use, (b) how much they enjoyed using it and (c) if they were able to plan ahead when using it. We performed statistical analysis on the results for each question using Friedman tests and pairwise Wilcoxon signed ranks tests with Bonferroni correction as a post-hoc test. The general feedback for each affirmation was very positive, with no technique scoring lower than 4.17 on average (see Figure 5).

Ease of Use. We observed a significant effect of *Technique* on participants' rating ($\chi^2(2) = 17.59, p < .001$). Post-hoc comparisons also showed a significant difference between *IC* ($M = 4.17/7$) and the two other techniques ($p < .01$ in both cases, $M = 6.17/7$ for *CF* and $5.75/7$ for *IC+FF*).

Enjoyment. Overall, participants enjoyed the three techniques ($M = 5.38/7$), and we observed a significant main effect of Technique on Enjoyment ($\chi^2(2) = 7.75, p = .02$). *CF* received the highest rating ($M = 6/7$), followed by *IC+FF* ($M = 5.58/7$) and *IC* ($M = 4.58/7$). *CF* received a significantly better rating than *IC* ($p < .05$).

Planning Ahead. On average, participants declared that they were able to plan ahead using our techniques ($M = 5.3/7$). *Technique* had an effect on the rating ($\chi^2(2) = 17.6, p < .0001$). *CF* received the best ($M = 6.25/7$), followed by *IC+FF* ($M = 5.5/7$). Both scored significantly better than *IC* ($M = 4.17/7$, both $p < .01$).

Feedback and Discussion

Based on objective and subjective measures, the *IC* technique appears weaker than the other two: it is slower and participants rated it lower. Participants reported that the technique suffers from a "*lack of details*" (P8) and that one "*can get lost easily as it doesn't show you how to get to the next checkpoint*" (P5). This validates our initial hypothesis, and we chose not to pursue this technique any further.

Our followers were overall faster with *IC+FF* and *CF*, and the subjective feedback does not significantly differ between the two. Thus, both seem viable.

In the light of the Effort/Information trade-off, the limited interaction using *IC+FF* was appreciated. P1 stated that *IC+FF* "*was much simpler for me and convenient as I do not have to continuously scroll on my smart watch (a simple tap would bring me a step forward) and it shows me a path to walk instead of just screenshots like the IC technique*". P4 also mentioned that it was "*a good balance between [IC] and [CF]*". However, participants reported that one "*might still get lost if the path is complicated*" (P5), e.g. when "*there are a lot of turns*" (P9). About *CF*, participants liked the "*possibility to look for details*" (P11) and that they were able to "*rewind the path and find [their] way easily*" when lost (P3). However, 7 out of 12 agreed that having "*to keep scrolling on the watch can be quite inconvenient*" (as stated by P5). It is likely that this issue will get worse if users' hands are full, e.g. with a bag. P9 also noted that it "*might not be that great of an idea especially when one is climbing down or up the stairs*" and P6 complained "*CF requires constant attention. I can't look at the road if I am staring at the screen all the time*". We also noted that one participant ran into a bench while using this technique. This may reveal safety issues with *CF*.

In summary, *CF*, the technique with highest fidelity to the follower's view, performed particularly well. But high fidelity is not the only important concern: participants were also concerned about the need for constant interaction with *CF*. *IC+FF* performed as well as *CF* but required much less interaction, which our participants appreciated. Thus, *IC+FF* looks like a good compromise in terms of Effort and Fidelity. However, *IC+FF* does lack *CF*'s precise control and the ability to slowly rewind, e.g. when a checkpoint is not well positioned. Hence, the best design seems to be a combination of techniques. *IC+FF* is a good choice for the common case, when less interaction is desirable, and *CF* is a good choice when precise control is required.

Finally, note that we focused on followers and carefully crafted the paths ourselves. Authoring for *CF* is trivial (one need only walk the route), while our experience and previous work indicate that choosing the right checkpoints for *IC+FF* is both difficult and critical for path quality [5,21,27]. Thus, our next experiment studied path authoring to assess the performance of techniques like *IC+FF* in real-life scenarios.

EXPERIMENT 2: PATH AUTHORING

In the first experiment, our followers performed well with *Image Checkpoints + Fast Forwarding* (*IC+FF*). However, we carefully crafted the path ourselves — with multiple iterations — to make them as close to optimal as possible. In a real-life scenario, leaders should be able to create high-quality paths quickly and painlessly, ideally just by walking the route one time. Hence, we designed a second experiment to investigate leaders' ability to author paths.

Our first experiment also taught us that choosing a good set of checkpoints could be challenging for an author. This problem inspired a fourth design, *Video Checkpoints + Fast Forwarding* (*VC+FF*), that allows authors to record a whole segment of the path as a video checkpoint instead of a single image.

For this second experiment, we wanted to compare the authoring experience with *Image Checkpoints + Fast Forwarding* (*IC+FF*) and *Video Checkpoints + Fast Forwarding* (*VC+FF*). To better validate our results, we also decided to include a baseline that does not rely on any technology: *Written Instructions* (*WI*). The discussion for our first experiment explains that *IC+FF* and *VC+FF* should ideally be combined with *Continuous Flow* (*CF*) to allow precise control during difficult situations. However, we did not make this combination in experiment 2, because it would benefit both *IC+FF* and *VC+FF* equally and would increase variance in time measurements (making analysis more difficult). Also, we were interested in the performance of these techniques in isolation. One may notice, though, that this put *IC+FF* and *VC+FF* at a slight disadvantage compared to *WI*.

To record a path, leaders need only indicate where checkpoints should be as they walk. Our authoring tools for *IC+FF* and *VC+FF* made use of physical volume buttons. Image checkpoints were captured when a button was released (*IC+FF*) and video checkpoints were recorded while a button was being pressed (*VC+FF*). Video checkpoints also captured audio, enabling authors to give special instructions.

For this experiment, we chose three new routes. As in experiment 1, all were at NUS, indoors, and approximately 250 meters long. As before, we took steps to ensure that the three routes had similar complexity.

Participants

We recruited two groups of participants, nine leaders and nine followers. For the leaders, we recruited nine students and staff from NUS (3 females), aged 25 to 44 years old ($M = 30, SD = 6.12$). All leaders were highly familiar with the environment. For the followers, we recruited nine students and staff from SMU (5 females), aged 25 to 33 years old ($M = 27.22, SD = 2.49$). None of the followers were familiar with the environment. All participants received the equivalent of 7.4 USD reward for their participation and were in good physical condition. The best performing leader and follower received the equivalent of 35.1 USD each as an extra incentive.

Apparatus

FML Trials. The apparatus in this experiment was similar to experiment 1. Both followers and leaders used the Epson Moverio BT-200 smartglasses as a display. Leaders' input device was an Epson trackpad with volume controls connected by wire to the smartglasses. (This device is equivalent to a smartphone for this study's purposes.) Followers wore a Huawei smartwatch on their non-dominant wrist and used this for input. Followers also carried a Samsung Galaxy S7 Edge in their pocket that managed communication between the watch and the glasses.

WI Trials. The apparatus was a clipboard with a sheet of paper. Leaders also had a pen to write with.

Tasks

Our experiment took place in two phases. In phase one, each leader authored three paths, one with each of the three different techniques. In phase 2, each follower tried to follow the three paths created by a single leader.

Each leader trial began with the experimenter taking the leader to the beginning of a route. The leader would then walk through the entire route once and was asked to remember it as accurately as possible. After reaching the end, the leader was asked to take the same route backwards to reach the starting point. After this, leaders recorded a path for one of the techniques. In all but one trial, leaders completed their paths in a single pass. In one trial, a leader went the wrong way and was allowed to restart. In *WI* trials, leaders used a clipboard to write instructions on paper along the way.

Followers had to follow the paths while trying to balance speed and accuracy, as in experiment 1.

Procedure
Participants began the experiment by filling out a demographic questionnaire. We then described the whole experiment to them and led them to a training route. During the training route, participants (both leaders and followers) tried each of the three techniques on a 100-meter segment. Leaders were given a chance to look at the *IC+FF* and *VC+FF* training paths that they authored. To save time, leaders did not actually write out full *WI* training paths, but they were shown example instructions written by two previous leaders. We also discouraged leaders from drawing maps in *WI* trials, because attempts to do this produced confusing results in pilot tests.

After training, participants authored or followed all three paths, each with a different technique. Finally, participants answered a post-experimental questionnaire in which they gave subjective feedback on all techniques. The experiment was run on week days and crowd size varied similarly to experiment 1.

Design
A within-subject design was used with one independent variable: *Technique {VC+FF, IC+FF, WI}*. We counterbalanced with a Latin Square, ensuring that each technique was used on each route three times for both groups.

Leaders. We measured authoring time as a dependent variable. Authoring time was the total time required for the leader to walk through the entire path while setting up checkpoints or writing instructions. As subjective feedback, we also measured the ease of use, enjoyment, and path quality (how easily they thought others could follow their path). Leaders performed the experiment in around 1 hour and 30 minutes and were able to take breaks between each trial.

Followers. We measured task time and errors as dependent variables. Task time was the total time required for the participant to reach the destination while following the entirety of the path. For this experiment, we counted an error if the user went the wrong way and was not able to recover from this error in 40 seconds. After 40 seconds, the experimenter would help the participant back to the path. As subjective feedback, we also measured ease of use, enjoyment and ability to plan ahead. Followers performed the experiment in around 1 hour and 10 minutes and were able to take breaks between each trial.

The total number of trials was 54 = (9 leaders + 9 followers) $\times$ 3 techniques.

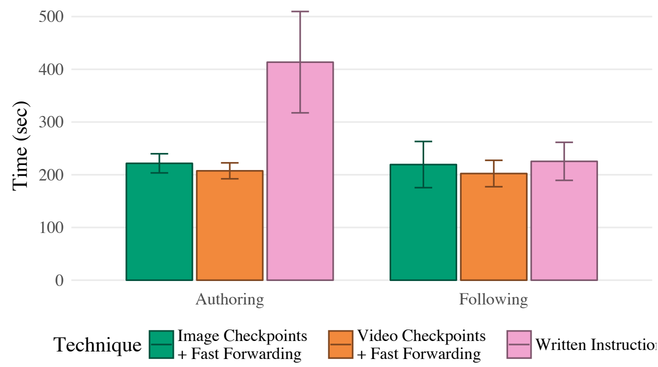


Figure 6. Authoring and Following time.

Results
We used ANOVA coupled with pairwise t-tests and Holm-Bonferroni correction to analyze time, a Cochran Q test to analyze errors, and Friedman's test with pairwise Wilcoxon signed ranks tests and Bonferroni correction for subjective feedback.

Leaders
Our main dependent variable for the leader phase was the time to author a path. The quality of the path itself is hard to evaluate but is reflected in the performance of the follower.

Authoring Time. A one-way ANOVA showed a significant main effect of Technique on authoring time ($F_{1.03,8.25} = 24.79, p < .001$, see Figure 6). Our participants were faster at creating paths with *VC+FF* ($M = 207$ s) than with *IC+FF* ($M = 221$ s) and *WI* ($M = 413$ s). Post-hoc comparisons showed significant differences between the three techniques (all three p<.05). The average walking speed of our leaders was 4.32 km/h for *VC+FF*, 4.05 km/h for *IC+FF* and 2.16 km/h for *WI*.

Subjective Feedback. Both *IC+FF* and *VC+FF* got good ratings (5.56/7 average). *WI* on the other hand received slightly negative feedback rating between 3/7 and 4.88/7 (Figure 7). We observed an effect of Technique on ease of use ($\chi^2(2) = 12.06, p < .01$), with *IC+FF* ($M = 5.88/7$) and *VC+FF* ($M = 6.67/7$) rating significantly better (both $p < .05$) than *WI* ($M = 3/7$). We found the same effect ($\chi^2(2) = 13.86, p < .001$) for enjoyment, with *WI* ($M = 3.56/7$) getting a significantly lower result than *IC+FF* ($M = 5.88/7$) and *VC+FF* ($M = 6.44/7$) (both $p < .05$). The trend continued with path quality ($\chi^2(2) = 15.77, p < .001$), with leaders giving a neutral rating to *WI* ($M = 4/7$), significantly lower than *IC+FF* ($M = 5.56/7$, $p < 0.5$) and *VC+FF* ($6.33/7$, $p < 0.5$).

Audio. All nine leaders used audio to give general directions (e.g. "turn left") or to highlight specific landmarks (e.g. "find the yellow trash bin").

Followers
Trial completion time was strongly dependent on the quality of the leader's path. Also note that we helped participants in the case of hard to recover errors, which lowers the time for conditions in which these errors occurred.

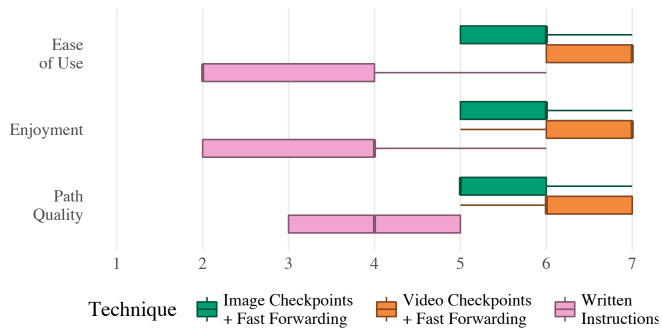


Figure 7. Leaders' Subjective Feedback.

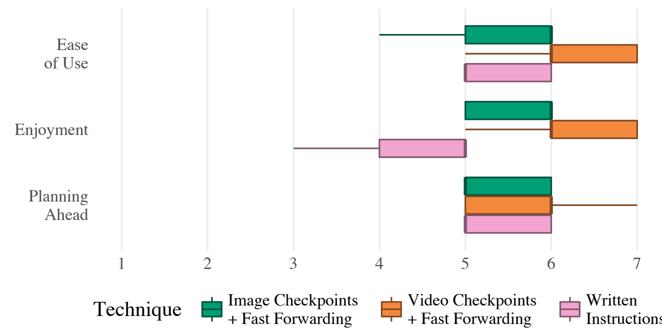


Figure 8. Followers' Subjective Feedback.

Time. The average time to complete a path was 215 s (Figure 6). An ANOVA did not show any significant main effect of Technique on time ($p = .61$). It took less time for our participants when they were using *VC+FF* ($M = 202$ s), followed by *IC+FF* ($M = 219$ s) and *WI* ($M = 225$ s), which suggests average walking speeds of 4.44, 4.11, and 3.99 km/h, respectively.

Errors. We counted a total of 5 hard to recover errors. Four were with *WI* and one was with *IC+FF*. For *WI*, errors would happen if the instructions were ambiguous or there was a missing step. The *IC+FF* error occurred outdoors when a leader's path showed a very fast turn with no checkpoint before or during the turn. A Cochran Q test did not show any significant main effect for *Technique* ($p = .71$).

Subjective feedback. Feedback from participants was positive overall (Figure 8). We observed no significant main effect of Technique on ease of use ($M = 5.3/7, p = .06$) but *VC+FF* was slightly higher than *IC+FF* and *WI*. For ability to plan ahead, the ratings are close (4.67 to 5/7, $M = 4.81/7$) without significance ($p = .41$). However, the technique had a significant effect on enjoyment ($\chi^2(2) = 14.21, p < .0001$). Our participants enjoyed using *VC+FF* ($M = 6.22/7$) significantly more than *IC+FF* ($M = 5.67, p < .001$) and *WI* ($M = 4.11/7, p < .001$). We also found a significant difference between *IC+FF* and *WI* ($p < .05$).

Audio. We also asked our participants to rate the following affirmation on a 5-point Likert scale: *I think I would have been able to find my way without audio.* The average score was 3.88/5 with only one participant giving a 2 (Somewhat Disagree) which suggests that our followers were confident that they could have used *VC+FF* even without audio. Audio was still appreciated, as P7 stated "I like audio [annotations] because it's almost as if I had someone walking with me".

Discussion

In this second experiment, our primary purpose was to investigate path authoring. Overall, creating a path with *VC+FF* is 6.3% faster than with *IC+FF* and 49.8% faster than with *WI*. Our subjective feedback showed that using *WI* made participants feel less confident in the quality of their paths, and they found it harder and less enjoyable to use. It is worth mentioning that all of our leaders were unfamiliar with the *VC+FF* and *IC+FF* techniques, and we believe that they could produce better paths with more training. However, the results are very encouraging and show that paths can be created quickly and accurately, even with minimal training.

Followers were also quite fast and usually accurate while following paths, as their average walking speed with *VC+FF* and *IC+FF* is respectively 4.44 and 4.11 km/h, slightly faster than *WI* (3.99 km/h). Our leaders' lack of confidence in their *WI* paths did not show itself in the followers' time, where we observed a mere 20-second difference (not significant). However, error rates did show problems with *WI*: 4/9 of our participants got lost when using *WI* and were unable to recover without external help. Another 1/9 got lost using *IC+FF* due to a lack of suitable checkpoints and low brightness in the video.

Overall, authoring times were shorter with *VC+FF*. The improvement over *IC+FF* is low (6.3%) but significant. In general, *VC+FF* also received slightly better subjective feedback. We believe that this difference is due to three factors. First, creating an image checkpoint requires more thought, since the choice of landmark and frame has to be carefully considered. Second, video checkpoints allow leaders to show the complete movement they are performing (e.g. a turn) instead of having to select a small set of relevant frames, as in *IC+FF*. A final advantage of *VC+FF* is the inclusion of audio commentaries, which were used by all 9 leaders. We believe this gave them more confidence, allowing them to go slightly faster. Interestingly, while audio was appreciated by followers, they were confident that they could have followed the paths without it.

DISCUSSION OF FOLLOW-MY-LEAD

Fidelity/Effort Trade-off for Followers

The different techniques we designed using the *Follow-My-Lead* approach represent different choices for the Fidelity/Effort trade-off. The *Continuous Flow* (*CF*) technique shows the most detailed information about the route at the cost of constant manual scrolling through the video. At the other end of the spectrum, *Image Checkpoints* (*IC*) requires minimal interaction effort at the cost of significantly reduced fidelity, showing important landmarks only. The results of our first experiment show that it was important for our participants to have access to most of this information. However, it was inconvenient and potentially dangerous to manually scroll through the whole interactive video using *CF*. As such, we believe that fast-forwarding–based techniques such as *Image Checkpoints + Fast Forwarding* (*IC+FF*) or *Video Checkpoints +*

Fast Forwarding (*VC+FF*) strike a better balance. However, followers may still occasionally need more details; therefore, complementing *IC+FF* and *VC+FF* with *CF*-like interaction is useful to handle difficult situations.

Authoring Paths

We designed *Follow-My-Lead* to allow quick path authoring with little or no practice. Our first experiment taught us that choosing the right path checkpoints is not as straightforward as it seems. This inspired us to invent the *Video Checkpoints + Fast Forwarding* (*VC+FF*) technique.

Our second experiment showed that leaders were able to create paths with both *IC+FF* and *VC+FF* just by walking the route once. However, it was faster, easier and more enjoyable with *VC+FF*. Compared to *WI*, our leaders were approximately twice as fast with *VC+FF* and *IC+FF*. The paths authored with *VC+FF* were easy to follow: followers using *VC+FF* paths quickly recovered from any errors. *IC+FF* achieved a good performance, but one follower was confused when a leader failed to put a checkpoint at a critical turn. These events confirm our hypothesis that the creation of discrete checkpoints can be complex and requires more training than video checkpoints.

Audio annotations were used by all nine of our leaders, making them more confident about the quality of their *VC+FF* paths. Interestingly though, our followers were confident (3.88/5) that they would have been able to follow the paths without these annotations. Audio can help to give additional contextual information, such as the floor to select in an elevator, the fact that a door needs an access card to open, or anything that may be hard to capture with a camera, e.g., pressing a small switch that is barely visible.

Improvements for Video Checkpoints + Fast Forwarding

Our authoring tool currently allows video checkpoints to be created only while walking a route, but leaders might go the wrong way while authoring a path. Allowing the leader to reshoot the last few seconds of a path would be an interesting improvement.

Limitations

While our experiments help to validate the *FML* prototype, we recognize that real-world usage scenarios raise additional issues that need further investigation. *FML* is susceptible to changes in the environment — e.g., storefronts seen in the leader's path may have changed since it was recorded and may no longer match the follower's view. This problem has bedeviled a wide variety of indoor localization technologies. We believe that this problem is more acute for other methods, such as WiFi-based localization, where radio signals fluctuate on timescales of seconds. In contrast, the majority of checkpoints used in our studies referred to *structural* characteristics (e.g., stairwells and turns) that are very likely to be stable.

Finally, while authoring paths with *Follow-My-Lead* requires a leader to walk through a route, some leaders may wish to provide path instructions from a remote location, such as a home or office. In such situations, leaders tend to write instructions from memory. However, our pilot studies indicate that such attempts are prone to oversights and imprecisions, even for someone who is highly familiar with a route.

CONCLUSION

We present *Follow-My-Lead*, a system for creating and following paths based on interactive videos captured by a leader. *Follow-My-Lead* is ideal for indoor wayfinding, because it needs neither localization technology nor floor plans.

In our first experiment, we investigated the Fidelity/Effort trade-off. We showed that our *Image Checkpoints + Fast Forwarding* technique strikes a better balance between fidelity to the follower's view and the interaction effort required to maintain this view. However, precise control of an interactive video – as enabled by the *Continuous Flow* technique – may still be helpful as a fall-back technique when leaders choose checkpoints poorly.

In our second experiment, we compared *Video Checkpoints + Fast Forwarding*, *Image Checkpoints + Fast Forwarding* and *Written Instructions*. We found out that authoring for *Video Checkpoints + Fast Forwarding* is faster and easier than *Image Checkpoints*, and both are much faster than creating *Written Instructions*. In the second part of the study, participants were able to follow these paths easily. While we did not find significant differences in terms following time, our followers made less hard-to-recover errors when using *Video Checkpoints + Fast Forwarding*. Participants also enjoyed *Video Checkpoints + Fast Forwarding* more than *Written Instructions*.

For future work, we would like to further investigate the relative usage of *Continuous Flow* and *Video Checkpoints + Fast Forwarding* when both are available to followers. We would also like to use devices' sensor data to automatically create checkpoints during the authoring process.

ACKNOWLEDGEMENTS

This work was partially supported by the Singapore Ministry of Education Academic Research Fund Tier 2 under research grant MOE2014-T2-1063 and National University of Singapore Academic Research Fund Tier 1 under research grant 251RES1617. All findings and recommendations are those of the authors and do not necessarily reflect the views of the granting agencies or the universities.

REFERENCES

1. Gregory D. Abowd, Christopher G. Atkeson, Jason Hong, Sue Long, Rob Kooper, and Mike Pinkerton. 1997. Cyberguide: A Mobile Context-aware Tour Guide. *Wireless Networks* 3, 5 (oct 1997), 421–433. DOI: http://dx.doi.org/10.1023/A:1019194325861

2. Masatoshi Arikawa, Shin'ichi Konomi, and Keisuke Ohnishi. 2007. Navitime: Supporting Pedestrian Navigation in the Real World. *IEEE Pervasive Computing* 6, 3 (jul 2007), 21–29. DOI: http://dx.doi.org/10.1109/MPRV.2007.61

3. Paramvir Bahl and V.N. Padmanabhan. 2000. RADAR: an in-building RF-based user location and tracking system. In *Proceedings IEEE INFOCOM 2000.*

Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064), Vol. 2. Ieee, IEEE, 775–784. DOI: http://dx.doi.org/10.1109/INFCOM.2000.832252

4. Karolina Baras, A. Moreira, and F. Meneses. 2010. Navigation based on symbolic space models. In *2010 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2010 - Conference Proceedings*. IEEE, IEEE, 1–5. DOI: http://dx.doi.org/10.1109/IPIN.2010.5646810

5. Ashweeni Kumar Beeharee and Anthony Steed. 2006. A Natural Wayfinding Exploiting Photos in Pedestrian Navigation Systems. In *Proceedings of the 8th Conference on Human-computer Interaction with Mobile Devices and Services (MobileHCI '06)*. ACM, New York, NY, USA, 81–88. DOI: http://dx.doi.org/10.1145/1152215.1152233

6. Gordon Best. 1970. Direction-finding in large buildings. *Architectural Psychology. London: RIBA Publications* (1970), 72–91.

7. Daniel G Borkowski, Hingsum F Fung, Hadi F Habal, Kenneth Chao, Sheng-roan Kai, and D Packard I I Robert. 1996. Cellular network-based location system. (1996).

8. Sal Bosman, Bas Groenendaal, J. W. Findlater, Thomas Visser, Mark de Graaf, and Panos Markopoulos. 2003. GentleGuide: An Exploration of Haptic Output for Indoors Pedestrian Guidance. In *International Conference on Mobile Human-Computer Interaction (Mobile HCI)*. Springer, 358–362. DOI: http://dx.doi.org/10.1007/978-3-540-45233-1_28

9. Nicholas A. Bradley and Mark D. Dunlop. 2002. *Understanding Contextual Interactions to Design Navigational Context-Aware Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 349–353. DOI: http://dx.doi.org/10.1007/3-540-45756-9_37

10. A J Bernheim Brush, Amy K Karlson, James Scott, Raman Sarin, Andy Jacobs, Barry Bond, Oscar Murillo, Galen Hunt, Mike Sinclair, Kerry Hammil, and Steven Levi. 2010. User Experiences with Activity-based Navigation on Mobile Devices. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. ACM, New York, NY, USA, 73–82. DOI:http://dx.doi.org/10.1145/1851600.1851616

11. Yao-Jen Chang and Tsen-Yung Wang. 2010. Comparing Picture and Video Prompting in Autonomous Indoor Wayfinding for Individuals with Cognitive Impairments. *Personal Ubiquitous Computing* 14, 8 (dec 2010), 737–747. DOI: http://dx.doi.org/10.1007/s00779-010-0285-9

12. Billy Chen, Boris Neubert, Eyal Ofek, Oliver Deussen, and Michael F Cohen. 2009. Integrated Videos and Maps for Driving Directions. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology (UIST '09)*. ACM, New York, NY, USA, 223–232. DOI: http://dx.doi.org/10.1145/1622176.1622218

13. Xiang 'Anthony' Chen, Tovi Grossman, Daniel J Wigdor, and George Fitzmaurice. 2014. Duet: Exploring Joint Interactions on a Smart Phone and a Smart Watch. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems (CHI '14)*. ACM Press, New York, New York, USA, 159–168. DOI: http://dx.doi.org/10.1145/2556288.2556955

14. Pei-Yu (Peggy) Chi and Yang Li. 2015. Weave: Scripting Cross-Device Wearable Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3923–3932. DOI: http://dx.doi.org/10.1145/2702123.2702451

15. Soon Hau Chua, Simon T Perrault, Denys J C Matthies, and Shengdong Zhao. 2016. Positioning Glass: Investigating Display Positions of Monocular Optical See-Through Head-Mounted Display. In *Proceedings of the Fourth International Symposium on Chinese CHI (ChineseCHI2016)*. ACM, New York, NY, USA, 1:1–1:6. DOI:http://dx.doi.org/10.1145/2948708.2948713

16. Roger M Downs and David Stea. 1973. *Image and environment: Cognitive mapping and spatial behavior*. Transaction Publishers.

17. Niklas Elmqvist, Morten Fjeld, David Axblom, Jonas Claesson, Joseph Hagberg, Daniel Segerdahl, Yan Tai So, Anders Svensson, Mattias Thorén, and Martin Wiklander. 2007. 3DVN: A Mixed Reality Platform for Mobile Navigation Assistance. In *ACM CHI2007 Workshop on Mobile Spatial Interaction*. 1–4.

18. Ivan Alexander Getting. 1993. The Global Positioning System. *IEEE spectrum* 30, 12 (1993), 36–38.

19. Andrew R. Golding and Neal Lesh. 1999. Indoor navigation using a diverse set of cheap, wearable sensors. In *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*. 29–36. DOI: http://dx.doi.org/10.1109/ISWC.1999.806640

20. Joy Goodman, Stephen Brewster, and Phil Gray. 2005. How can we best use landmarks to support older people in navigation? *Behaviour & Information Technology* 24, 1 (jan 2005), 3–20. DOI: http://dx.doi.org/10.1080/01449290512331319021

21. Harlan Hile, Radek Grzeszczuk, Alan Liu, Ramakrishna Vedantham, Jana Košecka, and Gaetano Borriello. 2009. Landmark-Based Pedestrian Navigation with Enhanced Spatial Reasoning. In *Proceedings of the 7th International Conference on Pervasive Computing (Pervasive '09)*. Springer-Verlag, Berlin, Heidelberg, 59–76. DOI: http://dx.doi.org/10.1007/978-3-642-01516-8_6

22. Simon Holland, David R Morse, and Henrik Gedenryd. 2002. AudioGPS: Spatial Audio Navigation with a Minimal Attention Interface. *Personal Ubiquitous Comput.* 6, 4 (jan 2002), 253–259. DOI: http://dx.doi.org/10.1007/s007790200025

23. Christoph Hölscher, Tobias Meilinger, Georg Vrachliotis, Martin Brösamle, and Markus Knauff. 2006. Up the down staircase: Wayfinding strategies in multi-level buildings. *Journal of Environmental Psychology* 26, 4 (dec 2006), 284–299. DOI: http://dx.doi.org/10.1016/j.jenvp.2006.09.002

24. Haosheng Huang, Georg Gartner, Manuela Schmidt, and Yan Li. 2009. Smart environment for ubiquitous indoor navigation. In *New Trends in Information and Service Science, 2009. NISS'09. International Conference on*. IEEE, 176–180.

25. Toru Ishikawa, Hiromichi Fujiwara, Osamu Imai, and Atsuyuki Okabe. 2008. Wayfinding with a GPS-based mobile navigation system: A comparison with maps and direct experience. *Journal of Environmental Psychology* 28, 1 (mar 2008), 74–82. DOI: http://dx.doi.org/10.1016/j.jenvp.2007.09.002

26. Anthony LaMarca, Yatin Chawathe, Sunny Consolvo, Jeffrey Hightower, Ian Smith, James Scott, Timothy Sohn, James Howard, Jeff Hughes, Fred Potter, Jason Tabert, Pauline Powledge, Gaetano Borriello, and Bill Schilit. 2005. Place Lab: Device Positioning Using Radio Beacons in the Wild. In *Proceedings of the Third International Conference on Pervasive Computing (PERVASIVE'05)*. Springer-Verlag, Berlin, Heidelberg, 116–133. DOI:http://dx.doi.org/10.1007/11428572_8

27. Alan L. Liu, Harlan Hile, Henry Kautz, Gaetano Borriello, Pat A. Brown, Mark Harniss, and Kurt Johnson. 2006. Indoor wayfinding: Developing a Functional Interface for Individuals with Cognitive Impairments. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06)*. ACM Press, New York, New York, USA, 95. DOI: http://dx.doi.org/10.1145/1168987.1169005

28. Markus Löchtefeld, Sven Gehring, Johannes Schöning, and Antonio Krüger. 2010. PINwI: Pedestrian Indoor Navigation Without Infrastructure. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI '10)*. ACM, New York, NY, USA, 731–734. DOI: http://dx.doi.org/10.1145/1868914.1869016

29. Sylvain Malacria, Eric Lecolinet, and Yves Guiard. 2010. Clutch-free panning and integrated pan-zoom control on touch-sensitive surfaces. In *Proceedings of the 28th international conference on Human factors in computing systems (CHI '10)*. ACM Press, New York, New York, USA, 2615. DOI: http://dx.doi.org/10.1145/1753326.1753724

30. Andrew J May, Tracy Ross, Steven H Bayer, and Mikko J Tarkiainen. 2003. Pedestrian Navigation Aids: Information Requirements and Design Implications. *Personal Ubiquitous Comput.* 7, 6 (dec 2003), 331–338. DOI:http://dx.doi.org/10.1007/s00779-003-0248-5

31. Andreas Möller, Matthias Kranz, Stefan Diewald, Luis Roalter, Robert Huitl, Tobias Stockinger, Marion Koelle, and Patrick A Lindemann. 2014. Experimental Evaluation of User Interfaces for Visual Indoor Navigation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3607–3616. DOI: http://dx.doi.org/10.1145/2556288.2557003

32. Daniel R Montello. 2005. *Navigation*. The Cambridge handbook of visuospatial thinking, New York. 257–294 pages.

33. Alessandro Mulloni, Hartmut Seichter, and Dieter Schmalstieg. 2011. Handheld Augmented Reality Indoor Navigation with Activity-based Instructions. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 211–220. DOI: http://dx.doi.org/10.1145/2037373.2037406

34. Alessandro Mulloni, Daniel Wagner, Istvan Barakonyi, and Dieter Schmalstieg. 2009. Indoor Positioning and Navigation with Camera Phones. *IEEE Pervasive Computing* 8, 2 (apr 2009), 22–31. DOI: http://dx.doi.org/10.1109/MPRV.2009.30

35. Lionel M Ni, Yunhao Liu, Yiu Cho Lau, and Abhishek P Patil. 2004. LANDMARC: Indoor Location Sensing Using Active RFID. *Wirel. Netw.* 10, 6 (nov 2004), 701–710. DOI: http://dx.doi.org/10.1023/B:WINE.0000044029.06344.dd

36. Michael J O'Neill. 1991. Effects of signage and floor plan configuration on wayfinding accuracy. *Environment and Behavior* 23, 5 (1991), 553–574.

37. Martin Pielot and Susanne Boll. 2010. Tactile Wayfinder: Comparison of Tactile Waypoint Navigation with Commercial Pedestrian Navigation Systems. In *Proceedings of the 8th International Conference on Pervasive Computing (Pervasive'10)*. Springer-Verlag, Berlin, Heidelberg, 76–93. DOI: http://dx.doi.org/10.1007/978-3-642-12654-3_5

38. Martin Pielot, Benjamin Poppinga, Wilko Heuten, and Susanne Boll. 2011. A Tactile Compass for Eyes-free Pedestrian Navigation. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-computer Interaction - Volume Part II (INTERACT'11)*. Springer-Verlag, Berlin, Heidelberg, 640–656. http://dl.acm.org/citation.cfm?id=2042118.2042179

39. Nissanka B Priyantha, Anit Chakraborty, and Hari Balakrishnan. 2000. The Cricket Location-support System. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom '00)*. ACM, New York, NY, USA, 32–43. DOI: http://dx.doi.org/10.1145/345910.345917

40. Martin Raubal and Max J Egenhofer. 1998. Comparing the complexity of wayfinding tasks in built environments. *Environment and Planning B: Planning and Design* 25, 6 (1998), 895–913. DOI: http://dx.doi.org/10.1068/b250895

41. Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L Littman. 2005. Activity Recognition from Accelerometer Data. In *Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence - Volume 3 (IAAI'05)*. AAAI Press, 1541–1546. http://dl.acm.org/citation.cfm?id=1620092.1620107

42. Günther Retscher and Michael Thienelt. 2004. NAVIO: A Navigation and Guidance Service for Pedestrians. *Journal of Global Positioning Systems* 3, 1&2 (dec 2004), 208–217.

43. Yuanchao Shu, Kang G Shin, Tian He, and Jiming Chen. 2015. Last-Mile Navigation Using Smartphones. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, New York, NY, USA, 512–524. DOI: http://dx.doi.org/10.1145/2789168.2790099

44. Asim Smailagic and Richard Martin. 1997. Metronaut: a wearable computer with sensing and global communication capabilities. In *Digest of Papers. First International Symposium on Wearable Computers*. IEEE Comput. Soc, 116–122. DOI: http://dx.doi.org/10.1109/ISWC.1997.629927

45. Koji Tsukada and Michiaki Yasumura. 2004. ActiveBelt: Belt-Type Wearable Tactile Display for Directional Navigation. In *UbiComp 2004: Ubiquitous Computing*. Springer, 384–399. DOI: http://dx.doi.org/10.1007/978-3-540-30119-6_23

46. Jan van Erp, Hendrik van Veen, Chris Jansen, and Trevor Dobbins. 2005. Waypoint navigation with a vibrotactile waist belt. *ACM Transactions on Applied Perception* 2, 2 (apr 2005), 106–117. DOI: http://dx.doi.org/10.1145/1060581.1060585

47. Philipp Wacker, Kerstin Kreutz, Florian Heller, and Jan Borchers. 2016. Maps and Location: Acceptance of Modern Interaction Techniques for Audio Guides. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1067–1071. DOI: http://dx.doi.org/10.1145/2858036.2858189

48. He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, and Romit Roy Choudhury. 2012. No Need to War-drive: Unsupervised Indoor Localization. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services (MobiSys '12)*. ACM, New York, NY, USA, 197–210. DOI: http://dx.doi.org/10.1145/2307636.2307655

49. Roy Want, Andy Hopper, Veronica Falcão, and Jonathan Gibbons. 1992. The Active Badge Location System. *ACM Transactions on Information Systems* 10, 1 (jan 1992), 91–102. DOI: http://dx.doi.org/10.1145/128756.128759

50. Andy Ward, Alan Jones, and Andy Hopper. 1997. A new location technique for the active office. *IEEE Personal Communications* 4, 5 (oct 1997), 42–47. DOI: http://dx.doi.org/10.1109/98.626982

51. Dirk Wenig, Johannes Schöning, Brent Hecht, and Rainer Malaka. 2015. StripeMaps: Improving Map-based Pedestrian Navigation for Smartwatches. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 52–62. DOI: http://dx.doi.org/10.1145/2785830.2785862

52. Dirk Wenig, Alexander Steenbergen, Johannes Schöning, Brent Hecht, and Rainer Malaka. 2016. ScrollingHome: Bringing Image-based Indoor Navigation to Smartwatches. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '16)*. ACM, New York, NY, USA, 400–406. DOI: http://dx.doi.org/10.1145/2935334.2935373

53. Scooter Willis and Sumi Helal. 2005. RFID Information Grid for Blind Navigation and Wayfinding. In *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*. IEEE, 34–37. DOI: http://dx.doi.org/10.1109/ISWC.2005.46

54. Han Xu, Zheng Yang, Zimu Zhou, Longfei Shangguan, Ke Yi, and Yunhao Liu. 2015. Enhancing Wifi-based Localization with Visual Clues. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 963–974. DOI: http://dx.doi.org/10.1145/2750858.2807516

55. Qianli Xu, Liyuan Li, Joo Hwee Lim, Cheston Yin Chet Tan, Michal Mukawa, and Gang Wang. 2014. A Wearable Virtual Guide for Context-aware Cognitive Indoor Navigation. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 111–120. DOI: http://dx.doi.org/10.1145/2628363.2628390

56. Zheng Yang, Chenshu Wu, and Yunhao Liu. 2012. Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking (Mobicom '12)*. ACM, New York, NY, USA, 269–280. DOI: http://dx.doi.org/10.1145/2348543.2348578

57. Yuanqing Zheng, Guobin Shen, Liqun Li, Chunshui Zhao, Mo Li, and Feng Zhao. 2014. Travi-Navi: Self-deployable Indoor Navigation System. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking (MobiCom '14)*. ACM, New York, NY, USA, 471–482. DOI: http://dx.doi.org/10.1145/2639108.2639124