ScrollingHome: Bringing Image-based Indoor Navigation to Smartwatches

Dirk Wenig
Digital Media Lab
TZI, University of Bremen
dwenig@tzi.de

Alexander Steenbergen
University of Bremen
alexandersteenbergen@gmail.com

Brent Hecht
Dept. of Comp. Sci. and Engineering
University of Minnesota
bhecht@cs.umn.edu

Johannes Schöning
Expertise Centre for Digital Media
Hasselt University – tUL – iMinds
johannes.schoening@uhasselt.be

Rainer Malaka
Digital Media Lab
TZI, University of Bremen
malaka@tzi.de

ABSTRACT
Providing pedestrian navigation instructions on small screens is a challenging task due to limited screen space. As image-based approaches for navigation have been successfully proven to outperform map-based navigation on mobile devices, we propose to bring image-based navigation to smartwatches. We contribute a straightforward pipeline to easily create image-based indoor navigation instructions that allow users to freely navigate in indoor environments without any localization infrastructure and with minimal user input on the smartwatch. In a user study, we show that our approach outperforms the current state-of-the-art application in terms of task completion time, perceived task load and perceived usability. In addition, we did not find an indication that there is a need to provide explicit directional instructions for image-based navigation on small screens.

AUTHOR KEYSWORdS
smartwatches; cartography; mobile maps; pedestrian navigation; stripe maps

ACM CLASSIFICATION KEYSWORdS
H.5.2 [Information Interfaces and Presentation]: User Interfaces — input devices and strategies, interaction styles

INTRODUCTION
Designing interactive applications for smartwatches is challenging as one of the most prominent problems in mobile HCI, the “fat finger problem” [20], heavily impacts interaction with small screens. As such, researchers have explored additional input modalities and techniques [18,26], as well as developed new visualization approaches [23]. Additionally, while the problem of adapting interfaces originally developed for larger-screen displays to smaller devices like mobile phones is well-known and well-studied [16,25], smartwatches are not just downscaled mobile phones: techniques that work on mobile phones do not necessarily work on smartwatches (e.g., text input using software keyboards).

Pedestrian navigation using maps is an important smartwatch use case that faces obstacles due to both of these challenges. While map apps are included by default on both the Apple Watch and Android Wear platforms, these are lightweight adaptations of their corresponding smartphone apps. Recent work [23] has discussed the limitations associated with using these mobile phone-like map apps for navigation tasks.

The StripeMaps project [23] addresses this problem by adapting the mobile web design technique of linearization for displaying maps on the small screens of smartwatches. The StripeMaps team demonstrated that StripeMaps outperforms both traditional mobile map interfaces and turn-by-turn directions for indoor navigation using smartwatches.

Figure 1: The ScrollingHome application concept; the user scrolls through a series of images showing the route.
However, despite the success of StripeMaps, it did not leverage one important potential opportunity: StripeMaps relies on traditional cartographic representations, but there is evidence that image-based navigation approaches can have benefits with small displays [3, 6, 8, 9]. The goal of this paper is to combine StripeMaps with these image-based approaches and evaluate whether doing so outperforms the StripeMaps state-of-the-art.

The contribution of this paper is thus threefold. First, we present the ScrollingHome app (see Figure 1). By building on the StripeMaps principle [23], ScrollingHome allows users to navigate indoors using image-based approaches without any specialized positioning infrastructure: the user enters a building through a specific entry (e.g., the main entrance), selects a starting point as well as a destination point and then uses ScrollingHome to navigate through the building. We also present an accompanying straightforward pipeline to easily create image-based navigation instructions. Secondly, we investigate the effectiveness and usability of image-based navigation on smartwatch screens and show that ScrollingHome outperforms the current state-of-the-art app, StripeMaps, in terms of navigation task completion time, perceived mental workload as well as perceived usability. Thirdly, we did not find any indication that there is a need to provide further direction information for image-based navigation on small screen devices as is typically done by in research and practice.

RELATED WORK

Work on mobile device-based navigation dates back almost two decades. Interestingly, the DeepMap system [14], one of the earliest mobile guides, presented animated route information on a wrist-mounted display. Other early projects used maps on pre-smartphone mobile devices. The Cyberguide [1] presented a schematic map, which was automatically updated based on the user’s position. The goal of the GUIDE [5] project was to build a mobile tourist guide for the city of Lancaster. It allowed the user to switch between an overview map and a map of the local area. For an overview of early map-based mobile guides see the survey of Baus et al. [2]. Huang and Gartner [11] provide an overview focused on indoor navigation systems.

Early work in the field of image-based navigation demonstrated the supporting value of photographs augmenting auditory instruction and maps [6] or combined with textual instructions and maps [3]. Goodman et al. [8, 9] used photographs in a navigation system for older adults. In a field experiment, the participants preferred to use audio instructions and images [9]. In another experiment, landmarks pictured on images performed better than landmarks presented by speech [8].

Walther-Franks and Malaka [22] built an image-only navigation system for mobile devices and used photographs of decision points augmented with visual instructions to guide users incrementally from waypoint to waypoint without a map. Recently, Wither et al. [24] automatically created landmark-based navigation instructions by detecting salient landmarks in panoramic street imagery. Vaittinen et al. [21] described the user-centered design of an image-based navigation service for mobile devices and confirmed in a field trial that images help with detecting the destination. Because of this and the aforementioned advantages of using images for route information, some image-based navigation systems exist in industry as well; with the navigation feature of Google Street View, which uses 360° panoramic images, as the most successful.

Practitioners and researchers have also explored navigation on various wearable devices. Google’s Android Wear platform supports turn-by-turn navigation while Google Glass additionally supports visualizing route information on maps. Pielot et al. [17] explored the use of a vibrotactile belt to continuously indicate a destination’s direction relative to the user’s orientation. Schirmer et al. [19] presented a similar system, but integrated the vibration actuators into the user’s shoes. McGookin and Brewster [15] investigated undirected navigation for runners by designing a navigation system called RunNav, which could also be used on a smartwatch. RunNav does not provide explicit routes, but rather a high-level overview to inform runners of areas that are good and bad places to run. Kerber et al. [12] demonstrated how smartwatches could be used as magic lenses to browse through maps, but also showed that this technique does not outperform a “classical” UI. Wenig et al. [23] introduced StripeMaps, a cartographic approach for indoor navigation with smartwatches. StripeMaps transforms 2D maps with route information into a 1D stripe (see Figure 2). In a user study, StripeMaps outperformed both traditional mobile map interfaces and turn-by-turn instructions. This shows that users can benefit from alternative visualization approaches for route information when navigating with smartwatches.

Figure 2: StripeMaps converts a 2D map to a 1D stripe. The original path on the 2D map is shown on the mini-map. The cut (shown on the smartwatch) indicates the direction of the turn the user needs to make to navigate along the path (figure adapted from Wenig et al. [23]).
THE SCROLLINGHOME APPLICATION

The underlying concept of the ScrollingHome application is very simple: users scroll through a series of images showing their route. The images are extracted from videos taken from a first person view (see below for details). Like StripeMaps, users are able to scroll through these images by scrolling vertically on the display. Moving the finger upwards on the smartwatch shows previous images while moving downwards shows the following images that guide the users on their route through a building.

Controlling route information by linearizing user interaction and confining it to simple, single scrolling gestures have been proven successful with StripeMaps. ScrollingHome takes the same approach, but with image-based route information instead of the traditional cartographic representations by StripeMaps. As existing image-based navigation approaches for mobile devices have also included visual instructions (e.g., path visualizations or arrows) [22], instead of only presenting plain images on the smartwatch, we implemented a slightly modified second version of ScrollingHome that augments the images with a 3D path visualization of the route.

Implementation

Example indoor navigation videos were taken with a bridge camera (Panasonic DMC-FZ 200) by traversing different routes at a regular walking pace while holding the camera at an ‘eye-level’. No further equipment (e.g., light or stabilization) was used. Videos were taken in the evening hours to ensure that no other people were captured on the video to preserve their privacy and to not distract the user during the navigation task. Microsoft Hyperlapse† [13] was used to remove trembling and jerky motion as well as to smooth the video.

In a next step, for the extended version of ScrollingHome with path visualizations, Adobe After Effects and Cinema 4D were used to composite 3D elements over the 2D footage. The camera motion was extracted using the 3D camera tracker by Adobe After Effects and exported to Cinema 4D to create an animated camera. In the last manual step, in Cinema 4D the 3D path as well as virtual walls were added to occlude the path where needed in the video.

Finally, the videos were scaled down to 320×320 pixels to match the resolution of the smartwatch used for our study (see below) and turned into a sequence of JPEGs. This approach allows direct manipulation of the image sequence at various speeds in high quality – similar to the video browsing approach by Dragicevic et al. [7] – and for maps, similar to StripeMaps. For the user, the image sequence appears to be an interactive video. 15fps turned out to be sufficient for smooth slow scrolling and also well-suited suited for scrolling faster than real time.

different colored floors) and added color to some noteworthy features like grass in a patio to make sure that the stripes are as similar as possible to the maps stripes used in the evaluation by Wenig et al. [23]. In the second condition (ScroHo), the participants used the image-based navigation approach as described above. In the third condition (ScroHo+P), additional path visualizations providing route information were added to the video (see previous section).

During the navigation task, an experimenter followed the participants a few meters behind, measuring time and counting navigation errors. We counted one error when a participant took a wrong turn without noticing the error within five meters. After five meters, the experimenter told the participant of the error and sent her/him back on the correct direction. In addition to time and error, we measured the perceived usability using the System Usability Scale (SUS) [4] and the perceived task load using the NASA-TLX [10] for all conditions. The participants filled in both questionnaires after each navigation task. We encouraged the participants to think aloud when navigating and recorded noteworthy incidents in writing. After they completed all three navigation tasks, we conducted a semi-structured interview with each participant. Overall, the total time for participating in the study was around 45 minutes.

Results & Discussion
All participants were able to complete all navigation tasks.

Time and Error
On average, the participants needed about 187 seconds per task and made 0.59 errors per task (Figure 4 summarizes the results). The ScroHo condition performed best with 170 seconds on average (SD = 65 s) with an error rate of 0.22 (SEM = 0.17) and a maximum of 3 errors. The ScroHo+P condition performed second best with 171 second on average (SD = 44 s) and an error rate of 0.44 (SEM = 0.2) again with a maximum of 3 errors. In the SM condition, the participants took the most time (220 seconds on average, SD = 53 s) and also committed the most errors (1.11 on average, SEM = 0.31).

Regarding time, Kolmogorov-Smirnov tests showed that for the SM condition time is not normally distributed. A non-parametric Friedman’s two-way ANOVA by ranks for related samples revealed significant differences ($\chi^2 = 7, p = .02$). Bonferroni-corrected (significance level of .016) pairwise Wilcoxon tests confirmed significant differences between SM and ScroHo ($p = .011$) as well as between SM and ScroHo+P ($p = .012$). For errors, statistical analysis did not reveal any significant difference. Overall, while the participants navigated faster via ScrollingHome without and with additional path visualization than via StripeMaps, we did not find any indication that there is a need to provide direction information for image-based navigation on small screen.

Interestingly, only three test runs took more than 300 seconds (one for each condition): one participant (P9: male, 34 years old) was much faster in the SM condition than in both image-based conditions (SM = 185 s, ScroHo = 395 s, ScroHo+P = 307 s) while another one (P13: female, 25 years old) was much slower in the SM condition than in both image-based conditions (SM = 345 s, ScroHo = 158 s, ScroHo+P = 163 s). It seems that P9 struggled extensively with image-based navigation while P13 does the same with map-based navigation, suggesting that there may be important individual differences to consider here.

Questionnaires
The mean overall NASA-TLX values (normalized between 0 and 100) are low for both ScrollingHome conditions (ScroHo = 19.0, SD = 16.6 and ScroHo+P = 15.8, SD = 11.2) while the values for SM are higher (30.5, SD = 14.6). A low NASA task load index indicates a relative low subjective mental workload. A one-way ANOVA for repeated measures ($F_{2,34} = 10.8, p < .001$) and post-hoc Sidak-corrected pairwise analysis revealed significant differences between SM and ScroHo ($p = .027$) as well as between SM and ScroHo+P ($p < .001$).

Mean SUS scores can be considered good for both ScrollingHome conditions (ScroHo = 81.4, SD = 17.2 and ScroHo+P = 83.8, SD = 13.4) followed by SM (54.5, SD = 17.1). Again, a one-way ANOVA for repeated measures ($F_{2,34} = 28.7, p < .001$) and post-hoc Sidak-corrected pairwise analysis confirmed statistical differences between SM and ScroHo ($p < .001$) as well as between SM and ScroHo+P ($p < .001$). That means that ScrollingHome with and without direct navigation instructions outperformed StripeMaps regarding the perceived task load as well as regarding the perceived usability. Again, there were no significant differences between ScroHo and ScroHo+P.

Figure 4: Overview of the study results; mean time the participants needed to complete a single route, mean navigation errors and mean NASA-TLX overall values (from left to right).


**Observations and Interviews**

In general, the participants understood the ScrollingHome concept very quickly and were fluent in scrolling after a few seconds.

When taking a wrong turn, in most cases the participants noticed their mistake after a few meters by themselves. They scrolled back, compared the images on the smartwatch with the surroundings, and almost always were able to return to the correct route. Rarely, false positive errors occurred; the participants thought that they were wrong but they actually were traveling along the right path. In these cases, the participants stopped for a moment to compare the images on the smartwatch with the surroundings. After a few seconds, they realized that they were on the correct route and continued walking.

Regarding the scrolling behavior, some participants were constantly scrolling while they were walking and kept the images on the smartwatch exactly synchronized with their current position. Others used a slightly different approach: they scrolled through the images to the next turn first and then walked to that point without looking at the smartwatch. After taking the turn, they stopped and scrolled to the next turn again. In the end, the second approach turned out to be faster than the first one.

In the semi-structured interviews, ScrollingHome was described as a “very good idea” and “very intuitive”. One participant explained, that because of the images he only had to glance at the smartwatch out of the corner of his eye. However, feedback on the path augmentation was mixed, e.g. one participant criticized that it covers large parts of the image. We can refer to this as the “Fat Arrow Problem” for smartwatches in imaged-based navigation. The app itself was described as “respond[ing] very fast to user touch input” and the “interaction was very smooth”.

**CONCLUSION**

In this paper, we introduced novel methods for image-based navigation on smartwatches. The basic idea is to use a video-like series of images of a pre-recorded walk along a path that can be scrolled by the user as they navigate.

Our first prototype works on images automatically extracted from video only (ScroHo), while the second requires additional manual steps for pre-processing and image registration in order to overlay the images with additional graphical path information (ScroHo+P). We compared both prototypes with StripeMaps (SM), the state-of-the-art map-based approach for indoor navigation on smartwatches. Our user study revealed promising results and shows that image-based navigation can indeed outperform StripeMaps with respect to navigation time, perceived taskload, and other metrics. Interestingly, the ScroHo+P method, which requires substantially greater content generation effort than the ScroHo method, did not outperform ScroHo and thus we see no support for the need to expend that effort.

We also observed individual differences indicating that one single technique may not work best with all users. Additionally, more individual predispositions may lead to personal preferences and restrictions (e.g., visual impairments). However, our results suggest that the image-based approach is a more effective method for pedestrian navigation using smartwatches in general.

ScrollingHome also has benefits regarding the effort that is needed to create the content. Creating routes with the ScrollingHome approach is straightforward (only the ScroHo+P condition requires extra work). Creating good maps, which is needed for full 2D maps as well as for StripeMaps, requires an expert and is a time-consuming task. At the moment, we are working on a mobile app assisting the user to create all possible routes in a building.

ScrollingHome as well as StripeMaps both work without position information, making it suitable for indoor navigation. Compared to a traditional full 2D map, a limitation of both approaches are situations where the user has become lost (the basic orientation use case). Here, a 2D map allows the user to scroll in both dimensions along the route and to localize herself/himself again. However, apart from the very small number of navigation errors in the study with ScrollingHome, an advantage of image-based approaches is that the user likely notices her/his mistake very quickly because of all the details in the images, while corridors on floorplans (indoor maps) often look the same. This allows him/her to go back on the correct route.

**FUTURE WORK**

For future work, it will be important to identify means by which we can provide a tracking mechanism in order to give the user feedback in the case of errors. The additional integration of an orientation indicator might also be a useful addition. Furthermore, while the routes used in the study did include narrow floors as well as some halls, the routes did not include changing levels in multi-level buildings via stairs or elevators. Future work should investigate such situations as well as more special ones, e.g., mezzanines. Lastly, we recorded our videos especially for the purpose of this study. Future work should examine more scalable solutions, for instance leveraging the indoor imagery available in Google Street View.

**ACKNOWLEDGMENTS**

This work was funded in part by a Google Research Faculty Award and a 3M NTFA, as well as BOF R-6060, NSF IIS-1552955, and UID/EEA/50009/2013.

**REFERENCES**


