The 21st International ACM SIGACCESS Conference on Computers and Accessibility

ASSETS 2019

Author Response

Title: Leveraging Augmented Reality to Create Apps for People with Visual Disabilities: A Case Study in Indoor Navigation
Authors: Chris Yoon, Ryan Louie, Jeremy Ryan, MinhKhang Vu, Hyegi Bang, William Derksen and Paul Ruvolo

Instructions

The author response (also known as rebuttal) period has begun. The reviews for your submission are displayed on this page. If you want to respond to the points raised in the reviews, you may do so below.

You are not obligated to respond to the reviews.

PLEASE NOTE that, although there are four text boxes for your rebuttal, all comments go to all of the reviewers. The word limit is 600 words across all of the boxes. You may place comments in each box or place all of your comments in the "General Response to Reviewers" box.

For reference, you may see the review form that reviewers used to evaluate your submission. If you do not see some of the filled-in fields in the reviews below, it means that they were intended to be seen only by the committee. See the review form HERE.

Review #1

Appropriateness (1-5): 5
Clarity (1-5): 4
Originality / Significance (1-5): 4
Soundness / Correctness (1-5): 4
Recommendation (1-5): 4
Reviewer's Confidence (1-5): 4
Submit as Poster/Demo?: Yes

Detailed Comments

This paper presents the user-centered design process and deployment of a prototype that employs smartphone-based augmented reality tools to support blind and visually impaired users in recording indoor routes and re-navigating those routes.

The paper has several strengths. It describes an important challenge that people with visual impairments face, which is learning a new environment and traveling back to a starting point within that unfamiliar environment. The prototype builds on ARKit to enable blind users to record a route, add landmarks, and use the recorded route to navigate a new environment (forward and backward). To my knowledge, the idea is novel, and the solution seems to be reasonably effective. I also appreciated that four blind people were involved in co-designing the system, which helped to identify the major usability challenges with the app, such as the difficulty of positioning the phone-based camera, and understanding the direction of the phone in relation to the user's body. Another strength that sets this paper apart from most ASSETS papers is that the prototype was deployed in the iOS app store and data on this large-scale deployment is included. While the prototype received a thumbs up rating for only 68% of routes, many users appear to have been pleased with the system. Moreover, the deployment data revealed practical issues that would not have been found in a lab-based study and yielded feedback from a user group (low vision users) that the research team had not anticipated.

I have some questions and suggestions about the paper:

1. From what I understand, the co-design process included 10 hours of participation from each of the co-designers. It would help to provide more detail in this section on what happened during that time and how the design changed (beyond the two insights listed). More information on the co-designers themselves would also be helpful to include at the start of this section (e.g., how long have they had visual disabilities? Did they have prior experience with camera-based app?) to help the readers understand their experiences with the app. Although some of this detail is mentioned within the findings section, an earlier table and description would be useful. Given that the paper is currently only 9 pages long, there is plenty of room to add more detail.

2. Quantitative evaluation of the ARKit's accuracy. Who are the "experimenters" who walked with the app? The co-designers or the researchers? Do they have visual disabilities? The results may differ substantially based on the experimenter's visual condition because of their ability to use the camera. Moreover, it seems that accuracy was only evaluated when navigating forward on the recorded route. What about backward? Also, as mentioned later, navigating a reloaded route is less accurate than navigating a newly recorded one because of the route alignment. So, when evaluating the accuracy, wouldn't it be useful assess how route alignment affects accuracy?

3. The results in Table 1 are interesting. I'd like to see more explanations about the characteristics of the different routes. For example, why does the 26m route have much higher error than the other routes? What's special about it?

4. The reporting is incomplete for the logistic regression analysis. Please provide a table (or some other form) that includes detail for the full the model (log odds ratios, p values, etc.). There are figures for 3 of the 5 independent variables, but these may be difficult for readers who are not used to interpreting

log odds. One suggestion is to put the log odds in the full table, then use the figures to show descriptive statistics such as mean and standard deviation for the different groupings of data. (Please also label what the error bars represent). Also regarding this logistic regression model, it wasn't clear why there were three categorical features, when these actually appear to be ordinal or scalar (if they hadn't been binned).

5. One aspect that's missing from this paper is Discussion beyond the "considerations" and "future work" points. For example, what are the main limitations of the work? How do these findings and the potential utility of this approach compare against other possible approaches (i.e., discussion that grounds the results in the research literature).

Overall, although this paper has limitations, my assessment is that it would be of value to the ASSETS community.

Review #2

Appropriateness (1-5):	5
Clarity (1-5):	4
Originality / Significance (1-5):	4
Soundness / Correctness (1-5):	4
Recommendation (1-5):	4
Reviewer's Confidence (1-5):	4
Submit as Poster/Demo?:	Yes

Detailed Comments

Summary: This paper presents an indoor navigation system leveraging augmented reality, where the novelty lies in not relying on assumptions about the environment (e.g., presence of signs, 90-degree turns). Because the paper is well written and includes long-term design with design partners and in the wild testing with international users, it makes it a strong paper for ASSETS.

Main review: This paper presents an indoor navigation system, Clew, the utilizes augmented reality (via ARKit) that allows people with visual impairments to create, save, amend, and replay routes forward and backward. This work provides a novel artifact contribution, with the ability to save routes and removing assumptions about the environment (e.g., signs, 90-degree turns). The system was motivated from initial interviews with people with visual impairments, along with three research team members who have personal experience. The app records paths by leaving breadcrumbs based on position and orientation, simplifies the paths by removing breadcrumbs (via Ramer-Douglas-Peucker algorithm), and then replays the paths in path navigation mode. There were multiple types of evaluations including design insights gathered from co-design partners, a proof of concept experiment to show the potential of Clew, and an "in the wild" evaluation using

data from international users using the app. The paper describes the effect of certain factors (e.g., phone tilt, route length) on user ratings.

The paper has several strengths. First, the system provides a strong technical contribution, furthering the capability of indoor navigation. Second, I appreciated that people with visual impairments were included throughout the design and evaluation process and were a part of the research team. I appreciated the depth of evaluations, from generating insights, to showing the apps potential, to showing how it performs in the wild. I felt that even with an approval rating of 68%, the authors were able to capture through user data an understanding behind the ratings. The paper was written clearly and was easy to follow, making the above strengths easier to understand.

One concern I have is with the lack of statistical analysis on the data presented in Figures 5-7 – the error bars (all of Figure 5, first 3 of Figures 6&7) very much overlap with each other, showing that the differences are not statistically significant. I would express this caution in the paper.

I have a couple of clarifying questions: When the user creates a starting location, with the phone camera pointing toward the floor, what if the user starts in the wrong position and the floor is patterned? Could that result in a false positive starting point and make the first direction inaccurate? What instructions are given to avoid this error?

To find the starting location, the phone is described to have the short edge flush with the wall such that the camera is pointing down and the screen is pointing up. Later, in the Longitudinal design it states that the app works best if the camera is pointing parallel to the ground (which is similar to Figure 4), where the phone's top edge is parallel to the ground and the camera is pointing in the same direction as the user. Is the phone camera only pointing toward the ground for landmark creation, but then pointing forward otherwise?

Despite my above questions, this paper has strengths with a fully functioning deployed system with several layers of evaluation. I recommend this paper as an accept.

Other thoughts: When first reading the paragraph just before "Registration through Physical Alignment," I initially thought of another limitation – that the person would not properly aim their phone. Given that was a recurring theme in the results, it may be worth listing that limitation here as well.

Review #3

Appropriateness (1-5): 5
Clarity (1-5): 4
Originality / Significance (1-5): 4
Soundness / Correctness (1-5): 4
Recommendation (1-5): 4
Reviewer's Confidence (1-5): 5

Submit as Poster/Demo?: Yes

Detailed Comments

The authors developed a navigation app "Clew" for the visually impaired which is characterized by the feature to remember routes by using visual inertial odometry (VIO) developed for augmented reality. The app was deployed through the app store and evaluated by the usual app evaluation framework (Thumbs up or down), and user log.

The app is novel given the features to save and load accurate VIO data for backtracking. The survey is thoroughly completed and comprehensive set of papers are referred. It may be better to cite other backtracking apps like this one: Shu, Yuanchao, et al. "Last-mile navigation using smartphones." Proceedings of the 21st Annual International Conference on Mobile Computing and Networking. ACM, 2015.

I agree that this approach is technically sound because of the high accuracy of "relative localization" of VIO technologies.

And also the fact that authors developed and distributed the app through AppStore and collected real-world data should be highly evaluated.

On the other hand, there are some issues I should mention.

The authors reported the results from open deployment, but the main results are effects of "errors" and "length of routes". These factors are obvious without deploying the system.

They summarized the takeaways as the section for "Summary of Consideration", but most of the items are general topics such as internationalization and localization. They should focus on items based on the use of VIO for blind and visually impaired.

One of the unanswered questions is the effectiveness of VIO technologies for the blind and visually impaired people since they can not visually check the appropriateness of captured video. That is why the evaluation of errors (Table 1) should be done by blind users, but there is no mention about the participants of the evaluation.

It is well known that the accuracy will decline because of motion blur according to the faster walking speeds. [18]

[18] Fusco, Giovanni, and James M. Coughlan. "Indoor Localization Using Computer Vision and Visual-Inertial Odometry." International Conference on Computers Helping People with Special Needs. Springer, Cham, 2018.

In the SUMMARY OF CONSIDERATIONS, it says "for navigation routes of ~ 61m the motion estimates of ARKit are accurate", but Figure 6 shows that the accuracy is significantly lower in routes with more than 33 m. This fact should be discussed.

Overall, the paper is well written as a VIO paper, but as an assistive technology paper, authors should consider the target user for their technical approaches. Please reconsider the effects of disability to VIO, and also technologies to compensate such difficulties.

Submit Response to Reviewers

Use the following boxes to enter your response to the reviews. Please limit the total amount of words in your comments to 600 words (longer responses will not be accepted by the system).

Response to Review #1:

Response to Review #2:

Response to Review #3:

General Response to Reviewers:

Submit

START Conference Manager (V2.61.0 - Rev. 5828)