Leveraging Augmented Reality to Create Orientation and Mobility Apps for People with Visual Disabilities

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

ABSTRACT
The introduction of augmented reality technology to iOS and Android enables, for the first time, mainstream smartphones to estimate their own motion in 3D space with high accuracy. For assistive technology researchers, this development presents an enormous opportunity. In this spirit, we present our work leveraging these technologies to create two smartphone apps to empower people with visual disabilities to more easily perform orientation and mobility tasks. Our first app provides automatic navigation guidance when backtracking along a route. Our second app utilizes crowdsourcing to provide automatic guidance to objects of interest. We present the design of these apps along with a usability study. Our work expands the capabilities of assistive technology for orientation and mobility that can be distributed cheaply at massive scale. We conclude by discussing the opportunities, pitfalls, and limitations of augmented reality for creating orientation and mobility apps.

ACM Classification Keywords
K.4.2. Assistive technology for persons with disabilities: Orientation and mobility tools for persons with visual disabilities

Author Keywords
orientation and mobility; augmented reality; assistive tech

INTRODUCTION
For people who are blind or visually impaired (B/VI), improvements in orientation and mobility (O&M) have been shown to increase economic opportunity as well as psychological well-being. For instance, while only 30% of working-age Americans who are B/VI are employed [2, 25], individuals with better O&M skills fare better at finding employment [11, 10, 26, 32]. Due to the importance of O&M, there is a long history of assistive technology designed to bolster these skills [3, 7]. Despite considerable effort, historically, few of these technologies have achieved much impact beyond the lab [36]. This lack of impact has been driven, largely, by the fact that the technologies were either expensive, unreliable, cumbersome, did not provide significant benefits over simpler solutions, or were hard to distribute at mass scale.

ASSETS'18, October 22–24, 2018, Galway, Ireland

ISBN 123-4567-24-567/08/06. . . $15.00
DOI: http://dx.doi.org/10.475/123_4

Major exceptions to this disappointing track record are GPS navigation apps for smartphones. Since these apps are mainstream technologies that happen to be universally accessible — as opposed to special purpose assistive technologies — they are highly robust, powerful, and extremely useful to people who are B/VI. Further, the fact that a majority of people who are B/VI own smartphones [30] makes them distributable at either no additional cost (e.g., Google Maps) or at modest cost (e.g., BlindSquare which is designed for people who are B/VI and costs about $30 dollars).

While being incredibly useful, GPS apps for O&M have major shortcomings. Most notably, GPSes in modern smartphones are only accurate to about $5m$ under open sky (and are even worse in challenging environments such as cities) and do not work indoors. Researchers are working to build systems that overcome these challenges (see *Related Work*).

Recently, smartphone manufacturers have introduced augmented reality modules (AR), which support high accuracy 3D-tracking. While the primary purpose of these modules is to enable AR applications — whereby virtual and real content are mingled, e.g., by overlaying virtual characters on a smartphone's camera feed — these modules have the potential to be repurposed to create assistive technology for O&M that is highly robust, accurate, usable, widely deployable, and free.

With the significant potential of smartphone-based AR technology come critical research questions. Are the motion estimates provided by AR robust enough to use for O&M, and, if so, which O&M tasks might be facilitated? In this document we begin to provide answers to these questions. Specifically, we present our work utilizing user-centered design to leverage the AR modules in modern smartphones to create assistive technologies to aid users who are B/VI with O&M tasks.

Our first app, *Clew*, enables users to backtrack along previously traveled routes. *Clew* is designed to alleviate various pain points experienced by non-visual travelers (e.g., finding one's way independently after being led to a location by a sighted guide). The second app, *ViewShare*, utilizes crowdsourcing to enable a user to find objects in cluttered environments. In the remainder of the paper we present related work on O&M assistive technology, discuss the algorithms that underlie AR technology on modern smartphones, present the design of *Clew* and *ViewShare*, provide preliminary usability data for *Clew*, and finally conclude with a discussion of future challenges and promising directions for smartphone-based AR technology for people who are B/VI.

RELATED WORK

A number of researchers have worked to overcome the limitations (e.g., less-than-ideal accuracy, lack of availability indoors) of GPS technology for assisting people who are B/VI with O&M. Researchers have pursued roughly two approaches. The first is to utilize crowdsourcing, whereby people who are B/VI connect over the internet with a sighted person for real-time assistance. Examples of this approach include the VizWiz project [4], BeMyEyes [21], and Aira [9] (both BeMyEyes and Aira support a video chat interface). A second approach combines rudimentary motion estimates derived from inertial sensors (gyroscopes and accelerometers) with detection of fixed environmental infrastructure (e.g., Bluetooth beacons, Wifi access points). For instance, [17, 16, 18] developed a system for navigation using smartphone-detectable RFID tags. Dias and her collaborators utilized WiFi fingerprinting and dead-reckoning for indoor navigation [13], and similar systems based on Bluetooth beacons have also been developed [23, 1]. Others have explored the use of robots, with sophisticated sensors, to act as guides for people who are blind [31].

Another area of research is the development of solutions to help people who are B/VI find objects that were misplaced, moved by a third party, or whose locations were never known. While the task of object finding doesn't fall under the umbrella of O&M, here we stretch the terminology to encompass it by noting that both O&M and object finding require a significant degree of spatial awareness. As in the case of navigation, crowdsourcing approaches have shown promise for object finding [5]. Other approaches use automatic object detection and tracking to navigate to objects [33, 24, 35].

AUGMENTED REALITY

Recently, both Apple and Google have released support for highly sophisticated smartphone-based AR experiences, whereby virtual and real content are combined. For instance, an app might show a virtual cat projected into a real world scene. As the user moves, the phone senses the user's motion and renders the cat at an appropriate distance and angle, providing the illusion that the cat exists in the physical world.

These AR systems utilize 3D motion-processing algorithms that are vastly more accurate than the inertial-based systems in previous O&M apps. The high accuracy of these systems is driven by two key trends: the development of sophisticated algorithms for visual-inertial odometry (VIO) [28, 27, 6, 15] (which combine optical tracking with inertial sensing to perform motion estimation) and the development of special purpose hardware that allows these computationally intensive algorithms to run with minimal heat generation and power consumption. It is the combination of high accuracy and low power consumption that has the potential to unlock many new potential O&M apps for people who are B/VI.

Algorithms for Visual Inertial Odometry

While a full explanation of VIO [19] is beyond the scope of this paper, it helps to have a conceptual understanding of VIO. VIO algorithms are designed for either the monocular (single camera) or stereo setting. Since the monocular setting is the one applicable to mass-market smartphones, here we use the term VIO to refer to monocular VIO specifically.

VIO algorithms utilize sensor fusion to blend optical and inertial motion estimates. Optical motion estimates are made by tracking salient visual features — e.g., corners or other highly textured portions of an image — through multiple video frames. Utilizing the mathematics of perspective geometry, one can estimate the rotation and translation of the phone [20]. Importantly, the accuracy of these estimates is dependent on tracking a large number of visual features that should, ideally, correspond to points at a range of depths from the camera.

Further complicating matters, the translation estimated using optical tracking is only determined up to an arbitrary scale factor. This indeterminacy arises due to the fact that the depths of the tracked visual features are unknown [20]. For example, given an estimate of the translation of the phone, it is possible that the phone moved twice as far and the depths of the tracked points were twice as great. The shortcomings of optical tracking, scale-indeterminacy and inaccurate performance in feature-poor environments, can be overcome by the fusion of inertial data (gyroscopes and accelerometers). Gyroscopes, which provide accurate estimates of angular velocity, can refine estimates of rotation and accelerometer data can be integrated over time to obtain estimates of linear velocity, overcoming the scale-indeterminacy problem.

VIO in Mass-Market Smartphones

Both Apple and Google have released AR modules based on VIO. While, the details of their VIO algorithms are not publicly available, there are distinctions between these frameworks that researchers should keep in mind.

Google Tango

Release in late 2014 by Google's ATAP (Advanced Technology and Projects) division, the Tango utilizes a wide-angle lens and a global image shutter to enable accurate visual-feature tracking. The platform also includes a PrimeSense depth-sensing camera. Two commercial smartphones have been released based on the Tango platform: the Lenovo Phab2 Pro and the Asus Zenfone AR. While the tracking capabilities of Tango devices are superior to both ARKit and ARCore (discussed next), the reliance on special-purpose hardware has severely limited the adoption of the technology. As a result, Google suspended the project in early 2018 [12].

ARKit and ARCore

Apple's ARKit [22] and Google's ARCore [29], both released in 2017, do not require special purpose hardware. Since these platforms utilize conventional cameras, the richness of visual features available for tracking is not as great as with Tango, and consequently, their motion estimates are less accurate. Further, since neither of these platforms have depth sensing cameras, the availability of 3D information is limited to objects with special structure (e.g., planar horizontal and vertical surfaces). Despite their drawbacks, these frameworks can run on a wider share of phones than Google Tango, however, given the large preference for iOS among people who are B/VI [30], ARKit is the primary platform of interest for researchers seeking to develop O&M apps for people who are B/VI.

DESIGN PROCESS AND INSIGHTS

We employed semi-structured interviews, participatory design approaches [8, 34], and digital prototype testing over a period of a month with members of the B/VI community. We engaged deeply with one co-designer named Joe (name changed for anonymous review), a college student who has no functional vision due to Retinitis Pigmentosa (a degenerative vision disorder that leads to the breakdown of cells in the retina); he is a non-visual traveler, mostly navigating with the help of his guide dog. We also learned from an O&M trainer for primary and middle school students who shared stories drawn from his own experiences. Furthermore, two members of our research team, who are themselves visually impaired, contributed both to the design and implementation of the apps and leveraged their personal experiences and knowledge of the B/VI community to inform our design process. Our research team identified a number of pain points and accessibility barriers through involving people who work with or identify as B/VI people throughout aspects of our design and evaluation.

Navigating newly traveled routes towards previously visited locations was difficult. Joe expressed that finding his seat (e.g., in a classroom) after going to the restroom was challenging, particularly when traveling without his dog. This story resonated for one of our visually impaired team members who found it difficult to find his seat on a dimly lit airplane while returning from the restroom. These insights motivated the creation of our app *Clew* to support precise indoor navigation for finding the way back in unfamiliar environments.

Finding objects in unfamiliar environments was difficult. While Joe mentioned he had no difficulty in finding objects in spaces he has control to organize himself, he admitted that locating objects of interest in unfamiliar places or when others disturb his organization can be very challenging. Additionally, finding precise locations in outdoor environments (e.g., the correct part of a train platform or button to activate a crosswalk) is very difficult. This set of insights helped support the creation of our app *ViewShare* for object finding.

Over the course of our multi-week co-design with Joe, he tried various prototypes as a means to explore potential concepts for assistive apps. Based on his feedback, we developed a mostly complete version of *Clew*, which he was able to test and help us to improve. While informed by his insights, Joe did not test *ViewShare* as it was developed after he had returned to college.

APP 1: CLEW

We present *Clew*, an iPhone app based on ARKit (link to Clew on the app store redacted). Clew provides turn-by-turn directions to guide users who are B/VI along previously traveled routes. The app has two primary use cases. First, when a user is led somewhere by a sighted guide, they may wish to return to their previous location at a later time — without wanting to be guided back. Secondly, sometimes it is easier for someone who is B/VI to navigate from a location than back to it. For instance, it is easier for someone who is B/VI to leave a conference room than it is to find their way back to their particular seat. For both of these uses cases, *Clew* provides high-accuracy, easy-to-follow, navigational guidance to enable people who are B/VI to travel independently indoors, without the need for modifications to the environment (e.g., the introduction of beacons or special signage).

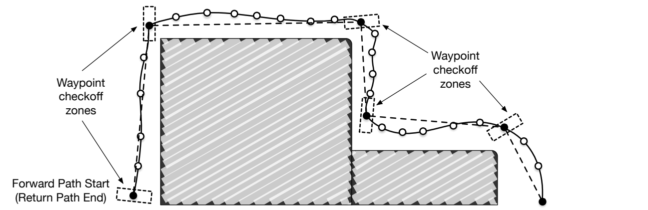

Figure 1. A topdown view of a path generated by the *Clew*. The figure shows the raw path (solid line) and breadcrumbs (white circles with black outlines). The Ramer-Douglas-Peucker reduces the breadcrumbs to a more manageable number of waypoints (solid black circles). The resultant path consists of straight lines connecting these waypoints in reverse order. Waypoint checkoff zones are shown as dashed boxes).

Clew works by dropping a trail of virtual breadcrumbs (representing timestamped 3D positions), which can be followed in reverse order at a later time. The phase of laying down these breadcrumbs is called *path recording* mode. In *path recording* mode, the user holds their smartphone with the screen facing towards them and the camera facing roughly parallel to the ground. The user then travels to a new location (either via their traditional O&M process or via assistance from a sighted guide) and stops the recording. Figure 1 shows a sample path along with virtual breadcrumbs.

When the user has recorded their path, they can either pause (described later) or navigate back along their route. When navigating back, the trail of breadcrumbs is processed by the Douglas-Ramer-Peucker (DPR) algorithm [14] for path simplification. This algorithm winnows down the path by removing sequences of breadcrumbs that are well represented by a straight line. Figure 1 shows the breadcrumbs selected by the DPR algorithm, called *waypoints*, and the resultant piecewise straight navigation route obtained by connecting the waypoints. In *navigation mode*, the app synthesizes directions to the next waypoint using one of three mechanisms: (1) speech (e.g., "continue straight for 10 feet"), (2) haptic or (3) audible feedback when the phone is pointing towards the next waypoint. When using (2) or (3), the low latency of the update of the phone's position allows the user to sweep their phone back and forth until they sense a haptic or auditory cue, providing an accurate sense of the direction to the next waypoint.

As the user navigates, the app continuously checks to see if the user has reached the next waypoint. We define the condition of "reached the next waypoint" as the user entering a *waypoint checkoff zone* (see Figure 1). Instead of making these checkoff zones spherical, we made them rectangular prisms with sides perpendicular to the direction of travel longer than for dimensions parallel to the direction of travel. This choice of shape enables the user to deviate laterally from the intended path without missing a waypoint. The app also announces flights of stairs by detecting if the vertical angle of the segment connecting two waypoints exceeds a threshold.

Length	*Stairs*	*Visuals*	*Success*	*Mean Error*	*Median Error*
21.3m	no	good	100%	0.15m	0.15m
40.2m	yes	mixed	100%	0.99m	0.91m
11.9m	no	poor	100%	0.91m	0.61m
22.9m	no	good	100%	1.45m	1.37m
47.2m	yes	mixed	100%	1.45m	1.37m
43.4m	yes	good	75%	2.03m	1.83m

Table 1. Usability results for Clew. The *Visuals* column indicates the availability of trackable visual features, with "mixed" indicating that a route had segments with both good and poor quality visual features.

Pause feature

While an ARKit tracking session is running, the user must not switch to a different app, lock the phone, or occlude the camera. This can be a hindrance when the user wants to wait a significant amount of time before navigating back, when the user would like to use another app, or when the user needs both of their hands to perform some task. ARKit does support a pause feature for a tracking session, however, when the session is resumed the phone's position will remain the same as before it was paused. Given this limitation, we designed a pause feature that prompts the user to place their phone in a position and orientation (collectively called a "pose") that is easy to return to when the user wants to resume. It is not crucial that the user be able to perfectly recreate this pose. Specifically, the most important aspect of the pose to recreate is its yaw. The reason for this is that the phone's accelerometer can easily determine the other two degrees of freedom of rotation (as these are not perpendicular to gravity) and deviations in the phone's position cause a fixed amount of error, whereas an error in the phone's yaw will be magnified the farther the user travels. Holding the phone's screen flat against features such as walls or doors works well for this purpose.

Novelty in Relationship to Previous Work

Our work on the design of *Clew* represents the first research into automatically guiding a user back along a route, without the need for special modifications to the environment.

Usability Testing

As a preliminary usability test, four college students used Clew to navigate six routes. For each route, the participant was led by one of the authors — mimicking the role of a sighted guide — and then navigated back using Clew. Participants were instructed to follow the guidance of the app, avoiding the use of their memory to correct for the app's errors. In order to become accustomed to the app, for the first two routes the participants were sighted for the duration of the trial. A trial was unsuccessful if the user couldn't follow the app's guidance (e.g., the app's estimation of a waypoint's location was inside a wall due to a failure of the ARKit's tracking algorithm). For successful trials (the user navigated to all waypoints), we recorded the distance from the user's final position to the ground truth position. A summary of each of the routes and the results are shown in Table 1. These results demonstrate the robustness of Clew for navigating fairly complex routes — even ones that involve stairs — and through places with poor visual features. The current study, while promising, is limited by its use of participants with occluded vision, rather than people who are B/VI. While our co-designer who is blind tested Clew extensively, we are currently working on a full evaluation of Clew with multiple users who are B/VI.

Future Work for Clew

Firstly, In order to make *Clew* successful, we must continue to engage deeply with users to understand how they experience the current version of the app and improve it based on their feedback. Secondly, we are working to improve the ability of the app to detect and communicate tracking failures. Specifically, when ARKit throws an error, it is straightforward to communicate this information to the user, however, if tracking drifts or becomes inaccurate without throwing an error, detection is much more difficult. Thirdly, we are working to enable the sharing of previously traveled routes with other users. This feature will work similarly to the pause feature in that the user will place their phone in a known starting pose (e.g., against a door as they enter a building), allowing their phone to register its current pose with the starting pose of a saved route.

APP 2: VIEWSHARE

Finding objects in unfamiliar environments can be challenging for people who are B/VI. Inspired by previous work [5] that utilizes a crowdsourcing model, whereby sighted online volunteers label an object of interest and sonic cues automatically guide a user to an object, we developed the *ViewShare* app.

ViewShare provides high precision, automatic guidance to objects of interest. To begin, the user, who is B/VI, launches the ViewShare app, starting an ARKit tracking session. When the user wants to find an object, they announce its name, the phone recognizes their voice command, and a localization job is created. This job is assigned to multiple sighted volunteers, who are sent push notifications on their smartphones. On the B/VI user's device, every two seconds a snapshot of their environment is captured and added to the job. Once a volunteer clicks on the push notification they are shown the snapshots (see Figure 2). After the volunteer locates the object, they tap its position on the screen. The object location in the image, represented as a 2D pixel coordinate, is relayed back to the B/VI user's app, which then attempts to convert this coordinate into a 3D position (see *2D to 3D Transformation*). If a 3D location can be determined, the user is provided with automatic guidance to the object in the form of computer-generated speech and haptic feedback.

Object Localization Feedback

ViewShare communicates location information using haptic and speech feedback. Specifically, whenever the phone is oriented towards a located object the phone will vibrate subtly and the distance and name of the object will be announced. To facilitate finding objects in 3D, the app provides two localization modes. In *2D feedback mode* the distance and orientation to an object are determined by projecting 3D spatial information into the floor plane. This mode is useful for navigating to objects that are far away (since the height of the object is irrelevant until the user gets closer). In *3D feedback mode* distances and orientations to objects are determined using the

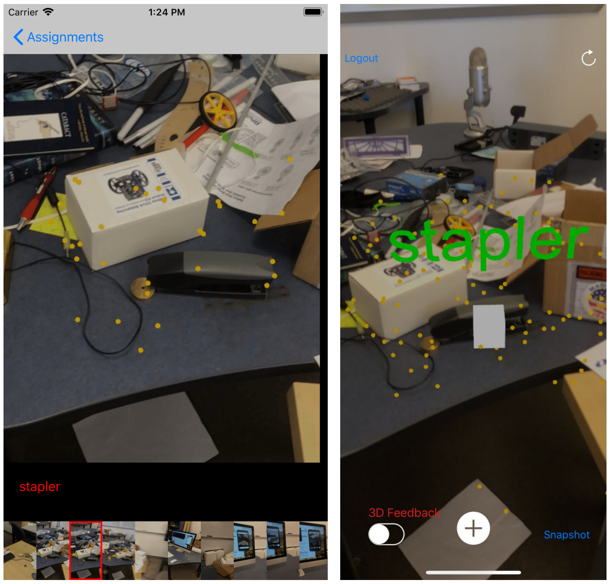

Figure 2. Left: ViewShare's interface for the crowdworker. The crowdworker is instructed to locate the stapler in the images collected from the B/VI user's phone. Right: the interface for the B/VI user showing the location of the stapler as indicated by the crowdworker (the visualization is useful primarily for users with low vision and as a debugging tool).

unmodified 3D information. This mode is useful for finding objects that are nearby.

2D to 3D Transformation

In order to map a 2D pixel coordinate to a 3D point in space, we utilize ARKit's plane fitting feature. We first test whether the ray that connects the optical center of the camera with the location of the pixel coordinate of the located object intersects a tracked 3D plane (ARKit is capable of finding both vertical and horizontal planes in 3D). If the ray intersects a plane, we mark the object as localized and compute its 3D position using standard geometric formulas. If the ray *does not* intersect a known plane (either because the object is not located on a plane or because the plane has not been found by ARKit) we return an error condition. In this situation, the volunteer is instructed to click on the object from a second viewpoint. Once the 2D pixel coordinate of the object is provided in two views, we use triangulation to compute its 3D position.

Novelty in Relationship to Previous Work

Our work on *ViewShare* represents the first development of an app designed specifically for object finding that leverages the built-in motion sensing capabilities to both localize and provide guidance to objects in 3D. Previous approaches either use a more brittle, 2D, approach to object localization, or require imprecise, mentally taxing, and high latency verbal communication to relay location information.

Future Work for ViewShare

While we designed this app primarily for finding objects, it can be utilized for other tasks. One possibility is finding things that are not objects (e.g., the precise location of a bus stop or a doorway in outdoor settings where GPS is not sufficiently accurate). A second possibility is automatically providing the location of important features in the environment (e.g., doors, windows, chairs) so that the user can make a mental map of the space around them. In addition to these new use cases, another area of future exploration is quality control. Aggregating judgments from multiple workers is an active area of research and these techniques can be utilized to filter inaccurate localizations from the crowd.

DISCUSSION AND FUTURE DIRECTIONS

Robustness of AR motion estimates for precise O&M

Our study demonstrates that for navigation routes of $\sim 45m$ the motion estimates of ARKit are sufficiently accurate. Additionally, for object finding within a room or on a tabletop we have found, anecdotally, that ARKit's performance is sufficient.

As the complexity of the environment increases other sources of information (e.g., the detection of landmarks such as optical tags or Bluetooth beacons) could reduce tracking drift, providing spatial anchors. Alternatively, crowdsourcing could correct drift in the phone's motion estimates (e.g., a crowdworker might identify a landmark at two points in time).

Usability Concerns of O&M technology using AR

Since AR relies on VIO, which leverages optical tracking, apps based on AR require the camera to be unoccluded. Currently, our apps assume that the user holds their phone in one hand while holding their cane in the other. Future research should consider whether a method of handsfree operation can be developed that allows the user to continue to use their phone as an interface. This could work by using an attachment (e.g., a neck lanyard) for holding the phone coupled with a speech-based user interface. We are currently exploring this idea.

For apps that require the user to hold the phone in their hand, recall that VIO algorithms work best when they are able to detect a large number of visual features at a range of depths. This condition is best achieved when the user holds their phone upright with the camera facing approximately parallel to the ground. We found that Joe had difficulty maintaining the phone in this configuration, perhaps due to the lack of visual feedback about the phone's orientation. As a result, tracking performance suffered. Developers of AR-powered assistive apps should consider adding feedback mechanisms to help the user to maintain their phone in an optimal orientation.

CONCLUSION

We have presented two smartphone apps that allow people who are B/VI to perform difficult O&M tasks with their smartphones. Our apps are among the first mass-distributable apps for navigation and object finding in unmodified indoor environments. While the results of our usability study are promising, more work, particularly in concert with users who are B/VI, is required to further refine the developed apps.

Additionally, we have outlined several promising areas of opportunity for the development of AR apps for O&M and provided a discussion of the usability factors of this technology. Through the development of new algorithms, co-design with users, and the improvement of AR technology itself, researchers can hopefully continue to leverage AR to create impactful smartphone technology for people who are B/VI.

REFERENCES

1. Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris M Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: a navigational cognitive assistant for the blind.. In *MobileHCI*. 90–99.

2. American Federation for the Blind. 2017. Interpreting Bureau of Labor Statistics Employment Data. (January 2017). http://www.afb.org/info/blindness-statistics/interpreting-bls-employment-data/24

3. J Malvern Benjamin. 1973. The new C-5 laser cane for the blind. In *Proc. Carnahan Conf. on Electronic Prosthetics*. 77–82.

4. Jeffrey P Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, and others. 2010a. VizWiz: nearly real-time answers to visual questions. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. ACM, 333–342.

5. Jeffrey P Bigham, Chandrika Jayant, Andrew Miller, Brandyn White, and Tom Yeh. 2010b. VizWiz:: LocateIt-enabling blind people to locate objects in their environment. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE, 65–72.

6. Michael Bloesch, Sammy Omari, Marco Hutter, and Roland Siegwart. 2015. Robust visual inertial odometry using a direct EKF-based approach. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 298–304.

7. Johann Borenstein and Iwan Ulrich. 1997. The guidecane-a computerized travel aid for the active guidance of blind pedestrians. In *Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on*, Vol. 2. IEEE, 1283–1288.

8. Christian Bühler. 2001. Empowered participation of users with disabilities in R&D projects. *International Journal of Human-Computer Studies* 55, 4 (2001), 645–659.

9. Aira Tech Corp. 2018. aira: your life, your schedule, right now. https://aira.io/. (2018).

10. Adele Crudden and Lynn W McBroom. 1999. Barriers to employment: A survey of employed persons who are visually impaired. *Journal of Visual Impairment and Blindness* 93 (1999), 341–350.

11. Adele Crudden, Lynn W McBroom, Amy L Skinner, and J Elton Moore. 1998. Comprehensive Examination of Barriers to Employment among Persons Who Are Blind or Visually Impaired. *Mississippi State: Rehabilitation Research and Training Center on Blindness and Low Vision, University of Mississippi.* (1998).

12. Tech Crunch. 2017. Android Guys. http://www.androidguys.com/2017/08/30/google-rebrands-tango-as-arcore/. (August 2017).

13. M Bernardine Dias . 2014. *NavPal: Technology Solutions for Enhancing Urban Navigation*. Technical Report CMU-RI-TR-21. Robotics Institute, Pittsburgh, PA.

14. David H Douglas and Thomas K Peucker. 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization* 10, 2 (1973), 112–122.

15. Christian Forster, Matia Pizzoli, and Davide Scaramuzza. 2014. SVO: Fast semi-direct monocular visual odometry. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 15–22.

16. Aura Ganz, Siddhesh Rajan Gandhi, James Schafer, Tushar Singh, Elaine Puleo, Gary Mullett, and Carole Wilson. 2011. PERCEPT: Indoor navigation for the blind and visually impaired. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 856–859.

17. Aura Ganz, James M Schafer, Yang Tao, Larry Haile, Charlene Sanderson, Carole Wilson, and Meg Robertson. 2015. PERCEPT based interactive wayfinding for visually impaired users in subways. *Journal on Technology & Persons with Disabilities* 3, 22 (2015).

18. Aura Ganz, James M Schafer, Yang Tao, Carole Wilson, and Meg Robertson. 2014. PERCEPT-II: Smartphone based indoor navigation system for the blind. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 3662–3665.

19. Jianjun Gui, Dongbing Gu, Sen Wang, and Huosheng Hu. 2015. A review of visual inertial odometry from filtering and optimisation perspectives. *Advanced Robotics* 29, 20 (2015), 1289–1301.

20. R. I. Hartley and A. Zisserman. 2004. *Multiple View Geometry in Computer Vision* (second ed.). Cambridge University Press, ISBN: 0521540518.

21. Bill Holton. 2015. A Review of the Be My Eyes Remote Sighted Helper App for Apple iOS. *Access World Magazine* 16, 2 (2015).

22. Apple Inc. 2018. ARKit Apple Developer. https://developer.apple.com/arkit/. (2018).

23. Tatsuya Ishihara, Jayakorn Vongkulbhisal, Kris M Kitani, and Chieko Asakawa. 2017. Beacon-Guided Structure from Motion for Smartphone-Based Navigation. In *Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on*. IEEE, 769–777.

24. Rabia Jafri, Syed Abid Ali, Hamid R Arabnia, and Shameem Fatima. 2014. Computer vision-based object recognition for the visually impaired in an indoors environment: a survey. *The Visual Computer* 30, 11 (2014), 1197–1222.

25. Corinne Kirchner, Emilie Schmeidler, and Alexander Todorov. 1999. Looking at Employment through a Lifespan Telescope: Age, Health, and Employment Status of People with Serious Visual Impairment. *Mississippi State, MS: Rehabilitation Research and Training Center on Blindness and Low Vision.* (1999).

26. Robin Leonard, Tana D'Allura, and Amy Horowitz. 1999. Factors associated with employment among persons who have a vision impairment: A follow-up of vocational placement referrals. *Journal of Vocational Rehabilitation* 12, 1 (1999), 33–43.

27. Stefan Leutenegger, Simon Lynen, Michael Bosse, Roland Siegwart, and Paul Furgale. 2015. Keyframe-based visual–inertial odometry using nonlinear optimization. *The International Journal of Robotics Research* 34, 3 (2015), 314–334.

28. Mingyang Li and Anastasios I Mourikis. 2013. High-precision, consistent EKF-based visual-inertial odometry. *The International Journal of Robotics Research* 32, 6 (2013), 690–711.

29. Google LLC. 2018. ARCore Overview. https://developers.google.com/ar/discover/. (2018).

30. John Morris and James Mueller. 2014. Blind and deaf consumer preferences for android and iOS smartphones. In *Inclusive designing*. Springer, 69–79.

31. Amal Nanavati, Xiang Zhi Tan, and Aaron Steinfeld. 2018. Coupled Indoor Navigation for People Who Are Blind. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*. ACM, New York, NY, USA, 201–202. DOI: http://dx.doi.org/10.1145/3173386.3176976

32. Bonnie O'Day. 1999. Employment Barriers for People with Visual Impairments. *Journal of Visual Impairment & Blindness* 93, 10 (1999).

33. Boris Schauerte, Manel Martinez, Angela Constantinescu, and Rainer Stiefelhagen. 2012. An assistive vision system for the blind that helps find lost things. In *International Conference on Computers for Handicapped Persons*. Springer, 566–572.

34. Douglas Schuler and Aki Namioka. 1993. *Participatory design: Principles and practices*. CRC Press.

35. Kaveri Thakoor, Nii Mante, Carey Zhang, Christian Siagian, James Weiland, Laurent Itti, and Gérard Medioni. 2014. A system for assisting the visually impaired in localization and grasp of desired objects. In *European Conference on Computer Vision*. Springer, 643–657.

36. William R Wiener, Richard L Welsh, and Bruce B Blasch. 2010. *Foundations of orientation and mobility*. Vol. 1. American Foundation for the Blind.