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, here, 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 a crowdsourcing model to provide easy to follow, automatic guidance to objects of interest. We present information on 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 people 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 [3, 25] (compared with 65% of the general population), individuals with better O&M skills have a higher likelihood of being employed [11, 10, 26, 31]. Due to the high degree of importance of O&M, there is a long history of assistive technology designed to bolster these skills (e.g., [4, 8]). Despite considerable effort, few of these technologies have achieved significant impact beyond the research lab [34]. 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) are 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 a number of shortcomings. Most notably, GPSes in modern smartphones are only accurate to about $10m$ in ideal conditions (they are even worse in challenging environments such as cities) and do not work indoors. Due to these shortcomings, many researchers have been working to build systems that overcome these key challenges (see *Related Work* section).

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 world content are mingled, e.g., by overlaying images 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 comes critical 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 some of 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.

Need to sell harder in the intro once the story of the paper becomes more complete. For example, listing out the three most important contributions of this paper. Lay out the key differentiators with previous apps, e.g. no need to instrument the environment

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 on-demand assistance. Examples of this approach include the VizWiz project [5], 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 location 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] (similar systems based on low-energy Bluetooth beacons have also been develoepd [23, 1, 2]).

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 traditionally 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-based technologies have shown promise for object finding [6]. Other approaches use automatic object detection and tracking to both locate and navigate to objects [32, 24, 33].

figure out a place to integrate some of Aaron Steinfeld's work

might need to explicitly contrast with our

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, 7, 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 only one applicable to mass-market smartphones, here we use the term VIO to refer to monocular VIO.

VIO algorithms utilize sensor fusion to blend optical and inertial motion estimates. Optical motion estimates are made by tracking salient visual features (for instance, 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]. Of particular importance, 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. While some environments are intrinsically more difficult for optical tracking, we have observed that users may exacerbate this difficulty by holding their phones in a suboptimal orientation (e.g., with the camera facing the ground).

Further complicating matters, the phone's 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 (to some degree) 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. Since inertial sensors don't function optimally when subjected to fast rotations or high accelerations, users should hold their phones relatively stable for maximum accuracy.

VIO in Mass-Market Smartphones

Both Apple and Google have released AR modules based on VIO. Unfortunately, the precise details of the algorithms employed are not publicly available, however, there are important distinctions between these frameworks that assistive technology 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. Further, the platform 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 of the platform 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

Might be good to cite for a comparison of ARKit and ARCore https://medium.com/super-ventures-blog/how-is-arcore-better-than-arkit-5223e6b3e79d

Apple's ARKit [22] and Google's ARCore [29], both released in 2017, do not require special purpose cameras. 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. Further, given the high preference of iOS among people who are B/VI [30], ARKit is the primary platform of interest for researchers seeking to develop O&M apps based for people who are B/VI.

USER-CENTERED DESIGN CONSIDERATIONS

might want to put this later. The paper could run the risk of dragging at this point

Our lab has been developing O&M assistive technology for the last several years. We take a user-centered approach in which we deeply engage with people who are B/VI to understand their needs, wants, and values as well as any pain points and areas of opportunity that exist within their daily lives.

Based on background research, interviews, and observation we identified several areas of opportunity for the development of O&M apps for people who are B/VI. Specifically, two major areas of opportunity emerged: high precision indoor navigation, especially in unfamiliar environments, and, being able to browse and find objects, for instance in a grocery store.

Since AR relies on VIO, which leverages optical tracking, any app based on AR requires the user's camera to be unoccluded. Here, we assume that the user holds the phone in one hand while holding their long cane in the other. We are currently working on methods for enabling handsfree operation.

Participatory Design

Throughout our work we utilized a participatory approach to design. Firstly, we worked with a college student with Retinitis Pigmentosa (a vision disorder that leads to the loss of peripheral vision to the point where eventually little functional vision remains) over the course of five, three-hour co-design sessions to develop the basic concepts and test prototypes of our app *Clew*. Secondly, two of the authors of this paper are themselves visually impaired. In addition to their contributions to the design and implementation of the apps, their personal experiences helped guide our design process.

APP 1: CLEW

We present *Clew*, an iPhone app based on the 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. In 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).

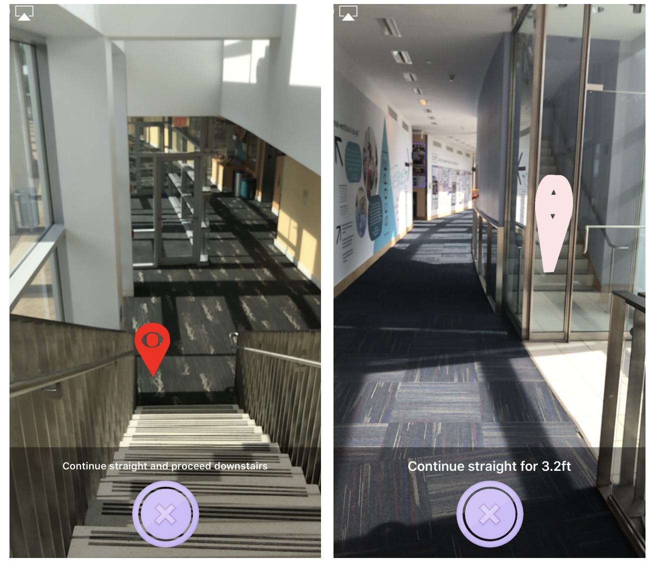

Figure 1. Two screenshots from our app "Clew" in navigation mode (where a user is retracing a previously traveled route). The text for the left image says "Continue straight and proceed downstairs" and the text on the right says "Continue straight for 3.2 feet."

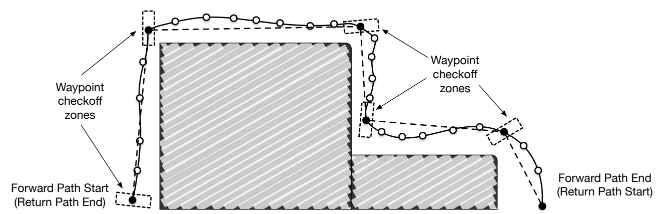

Figure 2. A topdown view of a path generated by the *Clew*. The figure shows the raw path (solid line) and breadcrumbs (white circles with black outlines). The Ramer-Douglas-Peucker reduces the breadcrumbs to a more manageable number of waypoints (solid black circles). The resultant path consists of 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 2 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 along the route, the trail of breadcrumbs is processed by the Douglas-Ramer-Peucker (DPR) algorithm [14] for path simplification. This algorithm winnows down the breadcrumb path by removing sequences of breadcrumbs that can be well=represented by a straight line. Figure 2 shows the breadcrumbs selected by the DPR algorithm (we call the selected breadcrumbs *way points*) and the resultant piecewise straight path obtained by navigating connected 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 feedback or (3) audible beeps when the phone is pointing towards the next waypoint. The low latency of the update of the iPhone's position when using mechanism (2) or (3), allows the user to rapidly sweep their phone back and forth until they feel a haptic feedback or hear the auditory cue, providing an accurate sense the direction to the next waypoint. Screenshots of the app in navigation mode are show in Figure 1.

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 2). Instead of making these checkoff zones spherical, we made them rectangular prisms with side lengths for dimensions perpendicular to the direction of travel greater 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.

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 between 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 that is easy to return to when the user wants to resume. It is not crucial that the user be able to perfectly recreate the position and orientation at pause time. 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 flat against features such as walls or doors work well for this purpose.

maybe add a figure showing what we mean by yaw

Usability testing

TBD.

Future Work on the Clew App

Firstly, In order to make *Clew* truly successful, we must engage deeply with users to understand how they experience the current version of the app in the app store and improve it based on their feedback. Secondly we are trying to improve the ability for the app to detect and communicate tracking failures. When ARKit throws an error, it is straightforward to communicate this information to the user, however, if the tracking drifts or becomes inaccurate without throwing an error, the user will be in a much more difficult situation. In future versions we would like to detect degraded tracking and alert the user. Thirdly, we are working to enable the sharing of previously traveled routes with other Clew 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 location with the starting location of a saved route.

APP 2: VIEWSHARE

jafri2014 (has good motivational material on object-finding)

Finding objects in unfamiliar environments can be challenging for people who are B/VI . Inspired by previous work [6] that utilizes a crowdsourcing model, whereby sighted online volunteers label an object of interest and sonic cues are then automatically generated to guide a user to an object, we developed the *ViewShare* app.

citation needed

ViewShare provides high-precision, automatic guidance to objects of interest. To begin this process, first, the user, who is B/VI, launches the ViewShare app, starting an ARKit tracking session. When user wants to find an object, they announce its name, the phone recognizes their voice command, and a localization job is created. This localization job is assigned to multiple sighted volunteers by sending push notifications to their smartphones. On the B/VI user's device, every two seconds a snapshot of their environment is captured and added to the localization job. Once a volunteer clicks on the relevant push notification they are shown the snapshots (see FIgure 3). Once the volunteer locates the object of interest, 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 section *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.

Flow chart of this?

Object Localization Feedback

The ViewShare app communicates location information using haptic and speech feedback. Specifically, whenever the phone is oriented towards a known object the final 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 unmodified 3D information. This mode is useful for finding objects that are close or small in terms of their vertical extant.

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 which connects the optical center of the camera with

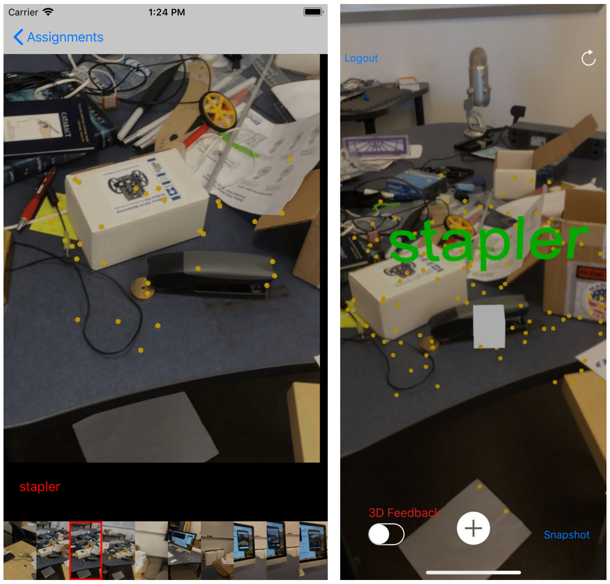

Figure 3. Left: ViewShare's interface for the crowdworker. The crowdworker is instructed to locate the stapler in the images collected from the B/VI user's phone. Right: the interface for the B/VI user showing the location of the stapler as indicated by the crowdworker (the visualization is useful primarily for users with low vision and as a debugging tool).

probably add a figure here

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. 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 can use triangulation to compute its 3D position.

Future Work on the ViewShare App

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 the issue of quality control. Aggregating judgments from multiple crowdworkers is an active area of research in the machine learning community and these techniques can be utilized to filter out inaccurate localizations from the crowd.

DISCUSSION AND FUTURE DIRECTIONS

The development of *Clew* and *ViewShare* provide useful information as to the potential of AR-based O&M apps. Our usability study and anecdotal observations demonstrate that for object finding — within a room or on a tabletop — and navigation along relatively short routes — Xm — the motion estimates of ARKit are sufficient for good performance. As the complexity of the environment increases (either because it is larger or has few easily trackable features) other sources of information (e.g., the detection of external landmarks such as special optical tags or Bluetooth Beacons) could be used to reduce tracking drift and provide anchors in the space. Alternatively, crowdsourcing could be used to correct drift in the phone's motion estimates (e.g., a crowdworker might identify a visible landmark at two points in time).

CONCLUSION

We have presented two smartphone apps that allow people who are B/VI to perform significant, new O&M tasks with their smartphones. Our apps are among the first mass-distributable apps or navigation and object finding in arbitrary indoor environments. While the initial results of our usability study are promising, more work, particularly in concert with our B/VI co-designers, is required to refine the developed apps to be maximally useful.

Additionally, we have outlined several promising areas of opportunity for the development of new AR apps for O&M and provided a discussion of the limitations of this technology. With the development of new algorithms, co-design with users who are B/VI, and the improvement of the AR technology from smartphone vendors, these limitations can hopefully be overcome to create impactful new assistive smartphone technology for people who are B/VI.

ACKNOWLEDGMENTS

Removed for anonymous review.

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. Dragan Ahmetovic, Masayuki Murata, Cole Gleason, Erin Brady, Hironobu Takagi, Kris Kitani, and Chieko Asakawa. 2017. Achieving Practical and Accurate Indoor Navigation for People with Visual Impairments. *Web for All* (2017).
3. 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`
4. J Malvern Benjamin. 1973. The new C-5 laser cane for the blind. In *Proc. Carnahan Conf. on Electronic Prosthetics*. 77–82.
5. 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.
6. 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.
7. 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.
8. 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.
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. Bonnie O'Day. 1999. Employment Barriers for People with Visual Impairments. *Journal of Visual Impairment & Blindness* 93, 10 (1999).

32. 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.

33. 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.

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