Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

ASSIST: Evaluating the Usability and Performance of an Indoor Navigation Assistant for Blind and Visually Impaired People

Vishnu Nair[a], Greg Olmschenk[a,b], William H. Seiple[c,d], Zhigang Zhu[a,b*]

[a]*Department of Computer Science, The City College of New York, NY, USA*

[b]*Department of Computer Science, CUNY Graduate Center, NY, USA*

[c]*Lighthouse Guild, New York, NY, USA;*

[d]*Department of Ophthalmology, NYU School of Medicine, New York, NY, USA*

Abstract. This paper describes the interface and testing of an indoor navigation app - ASSIST - that guides blind & visually impaired (BVI) individuals through an indoor environment with high accuracy while augmenting their understanding of the surrounding environment. ASSIST features personalized interfaces by considering the unique experiences that BVI individuals have in indoor wayfinding and offers multiple levels of multimodal feedback. After an overview of the technical approach and implementation of the first prototype of the ASSIST system, the results of two pilot studies performed with BVI individuals are presented – a performance study to collect data on mobility (walking speed, collisions, and navigation errors) while using the app, and a usability study to collect user evaluation data on the perceived helpfulness, safety, ease-of-use, and overall experience while using the app. Our studies show that ASSIST is useful in providing users with navigational guidance, improving their efficiency and (more significantly) their safety and accuracy in wayfinding indoors. Findings and user feedback from the studies confirm some of the previous results, while also providing some new insights into the creation of such an app, including the use of customized user interfaces and expanding the types of information provided.

Keywords: Indoor navigation, blind and visually impaired, mobile apps, human subject studies.

* Corresponding Author: Zhigang Zhu, Department of Computer Science, The City College of New York, 138th St and Convent Ave, New York, NY 10031. Email: zzhu@ccny.cuny.edu

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

Introduction

World Health Organization (2019) estimated that at least 2.2 billion people have a vision impairment or blindness globally. Although existing technologies (GPS) have been leveraged to provide outdoor navigation, there is a need for an assistive technology that aids these individuals in indoor navigation. Indoor navigation by blind and visually impaired (BVI) users requires information that is not accessible due to a lack of visual input. In sighted individuals, wayfinding in a novel environment is aided by landmarks that provide spatial location and/or task-relevant information. In such wayfinding, the goal is usually known; however, the spatial location of and path to reach this goal are unknown. As a sighted person walks through an environment to reach a goal, they search for visible landmarks to guide them to that goal. For individuals with vision loss, wayfinding based upon visual landmarks is impossible, and GPS is mostly unworkable inside of buildings due to its low accuracy indoors. In addition, BVI individuals lack the knowledge of their own position within a building's floor plan and relative to salient features (stairs, door, elevators, etc.), and obstacles (Remmen and Toft, 2015). As a result, navigating inside buildings and public spaces is an extreme challenge, and up to 70% of BVI individuals avoid indoor spaces and rely on assistance from a sighted guide when they are required to do so (Jeamwatthanachai et al., 2019; Miao et al., 2011).

Wold and Padøy (2016) identified the needs of BVIs when navigating indoors. These included: determining one's walking direction, knowing one's position within a building, obtaining route information, knowing building information, and being able to detect obstacles. Miao et al. (2011) identified several additional points of information that should be provided to users, namely, descriptions of functional waypoints and direction changes ("left/right/forward/back") in addition to information about the current position after entering a

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

new floor or area. Remmen and Toft (2015) also found that BVIs wished to be informed about locations of stairs, doors and elevators when navigating through complex buildings independently. These needs guided the design of our indoor navigation system.

From a general observation of the BVI community, we noted that the most popular technologies used are still long canes and guide dogs (Sato et al., 2019). From our studies and discussions with OM professionals and BVI users, this may be due to a lack of consideration of user's needs as well as low availability and production-readiness in new and upcoming technologies. We were unable to find any suitable *existing* commercial products for use in our navigation studies, which prompted us to develop our own testing system, **ASSIST** (an acronym for *Assistive Sensor Solutions for Independent and Safe Travel*). ASSIST is a mobile application ("app") with a server component that leverages Bluetooth Low Energy (BLE) beacons in conjunction with an augmented reality (AR) framework to provide a narrative of wayfinding instructions, much like having a sighted person's feedback during navigation. The highly precise positioning and navigation provided by ASSIST is not only important to safely navigate a BVI user through a cluttered indoor environment such as a transportation hub in New York City; it also enables the potential usages, such as guiding the user toward elevator buttons, door handles, or braille signs. As one example of a real-world need, in 2018, the Port Authority of NY&NJ issued a still open Request for Information (PANYNJ, 2018) to which we responded, exploring the possibility for different types of robotics and artificial intelligence (AI) technologies to meet their various needs, including customer service/wayfinding and traffic management.

ASSIST also utilizes environmental annotations to provide even further information on static characteristics of the user's current environment. These capabilities are combined and presented in a flexible and user-friendly app that can be operated using either touch or voice inputs.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

It can be configured as needed by varying the level of feedback, allowing for a customized experience for each user. In order to evaluate the usability and performance of using ASSIST (the implementation of which was detailed in a conference paper) (Nair et al., 2018b), we conducted two user-centric tests with BVI users and blindfolded-sighted users, recording their (objective) performance and (subjective) experience with the app. In this paper, we give a brief overview of the system's implementation for completeness and then focus on the user interface design and user evaluations. The main contributions of the paper include: (1) A user interface that provides options for varying levels of visual impairments and navigational ability. (2) A usability study that evaluated users' opinions of ASSIST and provided subjective measures of the app's usability. (3) A performance study that evaluated whether ASSIST improved a user's navigational performance and provided objective measures of the app's benefits. (4) An in-depth analysis of the studies with a presentation of findings that may be of use to researchers and developers in the space.

Related Work

Though accurate indoor positioning and navigation has been a popular area of research, it is still relatively underdeveloped compared to outdoor navigation systems (Real et al., 2019). Methods of indoor positioning have proposed the use of various technologies (Karkar & Al-Maadeed, 2018; Real et al. 2019), including but not limited to the use of cameras on smartphones or other mobile devices (Mulloni et al., 2009; Caraiman et al., 2017), passive RFID tags (Ganz et al., 2012; Chumkamon et al., 2008), NFC signals (Ozdenizci et al., 2011), inertial measurement unit (IMU) sensors (Ruiz et al., 2012; Sato et al., 2019) and Bluetooth Low Energy (BLE) beacons (Cheraghi et al., 2017; Sato et al., 2019; Murata et al., 2019).

Where passive RFID and NFC typically have significantly limited ranges (Ganz et al.,

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

2012) and are commonly limited to proximity detection, BLE beacon signals can be detected several meters or more away, allowing for localization based on signal strengths. Furthermore, BLE is compatible with most consumer devices reducing the need for users to own specialized equipment. However, localization using BLE may be inaccurate. Nair et al. (2018a) found that positions returned using BLE localization tended to be very noisy and unreliable. With users, Ahmetovic et al. (2016) tested NavCog with 6 participants and found that the turn-by-turn navigation paradigm was important. Among other findings, they also found that subjects tended to over-rotate when turning and that BLE localization accuracy was, at times, very low and contributed to missed turns for users. These findings suggest that more may be needed beyond BLE for highly accurate localization.

Google Tango (which uses a 3D sensor and computer vision) has also been of interest with the most relevant project being ISANA, a context-aware indoor navigation system implemented using Tango (Li et al., 2016); however, no user studies were performed. Kunhoth et al. (2019) provide a comparison of computer vision and BLE approaches. Although they did not explore a hybrid solution, the authors showed mixed benefits from both approaches. In another study, Nair et al. (2018b) proposed and tested a novel hybrid-based navigation system with BVI individuals as well as those with autism spectrum disorder and found that both groups positively rated the system, emphasizing the ability of a system to be personalized to each user's and group's unique needs. This study, though, focused on the technical approach to the system; thus, few details were provided on usability, and no performance study had been performed. However, the advantages of a hybrid approach to indoor navigation for those with visual impairments are worth exploring further, especially given the high accuracies that Nair et al. (2018a) found with this approach. For instance, augmented reality frameworks (such as Google Tango, its successor

ARCore, and its iOS counterpart ARKit) are natively compatible with the vast majority of modern smartphones. Furthermore, BLE has even wider compatibility. Thus, a hybrid system that combines both solutions (such as the one we test here) will allow for a universal solution that simply requires a recent smartphone and no additional hardware.

Beyond the technical details, several studies have also involved evaluations with users and identified needs for these groups. Abdolrahmani et al. (2017) present a study examining what kind of errors are acceptable to BVI users, and what kinds of errors are not acceptable. Such studies should help to guide what improvements need to be made in BVI assistive technologies. Nair et al. (2018a) found that BLE and Tango-based hybrid navigation required significantly fewer interventions when compared to pure BLE-based navigation. Their tests with 11 participants also found that users perceived the hybrid system as better. Sato et al. (2019) performed three studies (and held a focus group) with users using NavCog3. They came to several conclusions, noting that providing high accuracy is important (especially for finding small targets such as elevator buttons) and that personalizing the information provided is helpful in reducing the cognitive load induced in the user while walking. Users in the study also noted that they wished to use the system to complement their existing navigational aids and that it afforded them a sense of independence. Through their own studies, Yoon et al. (2019) recommend designing for multiple levels of vision and considering differences in spatial information processing among users. Ahmetovic et al. (2019) suggested that the need for a user to be assisted may decrease with prior knowledge and experience of the route and that changing cues depending on the mobility aid that the user uses (e.g., obstacles using a guide dog vs. cane) may be helpful. Ganz et al. (2012) tested the PERCEPT system with 24 BVI users; they found that users desired distances in steps, wanted instructions to be adjusted based on user preference, and solely wanted to use a smartphone (i.e., with no extra equipment). These findings suggest crucial considerations that must be taken into account when creating a navigation and wayfinding system for the blind and visually impaired. We considered these suggestions and findings, and present our new system in the following sections.

Sensory Components and System Architecture

This paper focuses on ASSIST's interface and human subject studies performed using the app. We have, thus, relegated a detailed overview of the system's technical components to Section S1 of the supplementary material.

In brief, ASSIST consists of two primary modules: location recognition via hybrid sensors, and map-based semantic recognition. These two modules interact with each other to provide a user with enough information to guide them successfully to their destination while augmenting their understanding of the environment around them. Note that the app does not intend to replace a BVI user's normal aids (e.g., white canes or guide dogs) for avoiding obstacles and finding doorways, heeding what previous studies have suggested (Sato et al., 2019). Rather, we simply aim to provide positional and situational information to enhance the user's travel experience.

ASSIST localizes mobile devices via a hybrid positioning method that utilizes BLE beacons for coarse localization in conjunction with an AR framework (in the prototype for this paper, Google Tango) for fine positioning. Note that, although Tango has been deprecated by Google, the underlying principles of 3D mapping and localization via device and pose estimation are applicable to other modern AR technologies. As such, our current work has focused on integrating ARCore on Android and ARKit on iOS for newer prototypes of the ASSIST app (Chang et al., 2020). Alongside these localization capabilities, ASSIST uses existing floor plans to mark the map with points of interest and perform related calculations (such as distance measurements). We also use these maps to annotate various static characteristics of the environment (e.g., doors and elevators), which are used to alert the user of these elements and incorporate them into navigation.

A client-server structure is used for ASSIST due to speed and scaling concerns (Figure 1). The client (the app on the phone) provides the user with a multimodal interface on top of the onboard localization program. The server component forms the system's core and contains all information that the app needs to operate properly. This allows our solution to scale-up to a large indoor facility such as the Port Authority Bus Terminal in NYC, a very complex transportation hub. Note that the app is able to download the required data when an Internet connection is available; thus, a connection is not necessarily needed when a user invokes navigation.

User-Centric Navigation Experience

ASSIST employs a user-centric navigation interface (Figure 2) by promoting a high level of configurability. Both the type (i.e., audio, visual, and vibrotactile) and level (e.g., information density and vibration intensity) of feedback can be adjusted to suit varying levels of disabilities.

Multimodal User Interface & Feedback

Previous work (Ahmetovic et al., 2019; Yoon et al., 2019) has demonstrated how varying levels of feedback benefit the users of BVI assistive systems, concluding that different users with varying levels of expertise benefit most from different levels of feedback. As such, our system provides options for varying levels for feedback. There are currently three options in the ASSIST mobile app: *minimal*, *medium*, and *maximal*. As the needs of BVI individuals vary over a spectrum, these options provide multiple densities of information to users. At one end of the spec-

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

trum is the *minimal* level, which utilizes the least reactive feedback, providing simple audio guidance and vibrotactile alerts. The *minimal* level provides simple audio guidance and vibrotactile alerts, which is intended to act as a basic option for everyone regardless of their disability status. In the *medium* level, we use both the visual cues and vibrations. (e.g., flashing and changing colors on the screen) and the "higher priority" event (e.g., the act of turning, or arriving at a destination) are signaled with increasing intensity. This is designed for low-vision individuals. The *maximal* option provides the densest amount of feedback and is designed for people who are totally blind. Feedback includes measures for course correction and guidance for ensuring that the user is facing the correct direction. In such a situation, ASSIST will pause the main navigation and, via audio cues, have the user slowly rotate, or move if needed, until they are correctly re-aligned with the path. Once this is achieved, the main navigation will resume. In our studies, since most users were totally blind, we gave them the highest level of feedback per their requests. The level to use for the tests was changed if the user requested it. Table 1 shows a feature comparison between the three modes.

Information Provided to Users

In its base form, the system provides turn-by-turn instructions much like those that would be provided by a sighted guide (e.g., *"In 50 feet, turn left," "Now turn left," "In 25 steps, you will arrive at your destination,"* and *"You have arrived at your destination: Cafeteria."*). These directives are repeated every 7.5 seconds to continually remind the user of their next step. If a user who requires it approaches within 10 feet (3 meters) of an obstacle or object of interest, they are informed of the type of object (*"You are approaching a security door."*). When they approach an elevator, they are instructed to call the elevator and go to a specified floor *("Now call the elevator and go down to the second floor."*). In its current form, ASSIST does not explicitly guide the user to the elevator buttons though future work could focus on facilitating this. (During tests, the authors pressed the buttons for the users.) Note that instructions can be communicated in imperial (feet), metric (meters), or general (steps) units; step units were provided following user feedback (Ganz et al., 2012). The true step size can also be customized to each user via a small program within the app that updates step size in advance.

Realignment of the user to the path is considered a major operation in that this event suspends navigation while the user turns. When a user deviates more than 35 degrees from the next navigational node, the app instructs the user to stop walking and begin rotating *("Stop and begin rotating to your right.")*. Every few seconds, this is repeated *("Continuing rotating to your right.")*. When the user is finally aligned, they are instructed to resume walking forward *("Now stop rotating and begin walking forward.")*.

Implementation and Configuration on the Modalities

We now describe specific implementation details of features across all three modalities. Note that specific parameters were picked as per the users' feedback and extensive self-testing.

Speech for audio cues was implemented using the phone's built-in text-to-speech functionality. Many of these instructions are very similar to those found in popular outdoor navigation apps (e.g., Google Maps). We believed in keeping the instructions as simple as possible and thus opted for an announcement scheme that users may already be familiar with (for example, through their use of these apps outside or while riding in a car).

Vibrotactile feedback was implemented to supplement the audio cues and was done so using the phone's built-in vibration functionality. There were several situations where vibration feedback would be issued. In *minimal* mode, no vibrotactile feedback is issued (as it is optimized for those who do not require any assistance beyond simple verbal instructions). In *medium* mode,

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

extra vibrotactile emissions are added in the form of a one-second burst before alerts (e.g., when approaching a door). *Maximal* mode does not use any special visual cues as it is meant for those who are blind. Instead, it relies on more detailed vibrotactile emissions. In addition to the pre-alert bursts found in *medium* mode, when a user is within 10 feet (3 meters) of a major navigational node, the phone begins vibrating in short 150ms bursts every 1.5s. When a user reaches the node and while the appropriate audio cue is being made, these bursts would get more frequent (one every 200ms) for a short amount of time until the user satisfies the instruction (e.g., turns).

Although not directly relevant for blind users, the app also includes a visual interface so that it is accessible to those who are sighted, including those with low vision (see Figure 2). The interface shows the current, upcoming instruction at the top alongside the distance to that instruction and an appropriate icon to symbolize the current (and next) instruction. The map takes up most of the screen, and the bottom panel shows the name of and the total distance remaining to the destination.

In *maximal* mode, veering is detected using Tango's built-in position and orientation functionalities. When the phone/user has pointed more than 35 degrees away from the next node on the route (as per Tango's readings), the app will register this as a "veer" and pause navigation in order to direct the user to rotate towards that node again. The app also provides redundancy in its cues to users by communicating information via multiple modalities (vision, sound, and haptics). In the app's current state, the user cannot customize (or filter out) individual modalities. However, such functionality may be useful in future iterations of the app. For example, a user could opt to only use haptic feedback if they are in a noisy environment or one where they must maintain silence. This adds an additional level of personalization, especially if multiple actua-

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

tors, sound sources, or screens are present.

User Evaluations

In order to evaluate the usability and acceptability of the ASSIST mobile app, we performed studies with blind & visually impaired users as well as (blindfolded) sighted users. Two types of tests were done: (1) A *usability study* was performed with BVI users to collect user evaluation data on the perceived helpfulness, safety, ease-of-use, and overall experience while using the app; and (2) a *performance study* was done with BVI and blindfolded-sighted users to collect data on mobility (walking speed, collisions, and navigation errors) while using the app. We used a Lenovo Phab 2 Pro (an Android smartphone with the Google Tango 3D sensor built-in). Users heard instructions through the phone's onboard speaker, and the phone's onboard vibration motors provided vibrotactile feedback. They did not use any other devices. These tests were performed across two floors of a six-story building in New York City. The studies were approved by the Institutional Review Board of the authors' institution.

Usability Study

Participants & Materials

The usability study was performed with BVI users to collect user evaluation data on the perceived helpfulness, safety, ease-of-use, and overall experience while using the app. A convenience sample of eleven (11) adults who were diagnosed as totally blind, legally blind, partially sighted, or low vision were offered participation in this study. Table 2 shows the participants' demographic data. We administered two surveys: a pre-experiment survey and a post-experiment survey. The pre-experiment survey included a demographic section, which asked the participants

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

to disclose their sex, age, and level of visual impairment. It also asked participants to rate their familiarity with smartphones as well as their overall difficulty (and strategies) in indoor navigation. The post-experiment survey assessed the perceived helpfulness, safety, ease of use, and overall experience of the navigation. The survey questions are listed in a table in the supplementary materials (Section S2), and many of the post-survey questions are based on a 5-point Likert scale (with 5 indicating strongly agree, 4 agree, 3 neutral, 2 disagree and 1 strongly disagree).

Procedure

All participants completed three walking paths, two of which included travel between floors. Figure 3 shows the paths used. All three paths ranged from 80 to 110 feet (24.3 to 33.5 meters) in total walkable distance and would each take around 5 minutes to complete (excluding time spent waiting for an elevator). The paths consisted of 3-7 turns (the exact number was dependent on the elevator taken), 1-3 doors, and 2 pillars in the immediate test area. During the tests, the authors opened locked doors and called the elevator for the user once they were within a few meters of each since, at the time of our studies, the app did not have functions to recognize door handles and elevator buttons. We chose these three paths in order to provide sufficient variety and opportunities for the user to acclimate to the app.

Starting from a pre-defined location in each of the three paths, we asked participants to use the app to navigate to a pre-determined destination. Five participants opted to start navigation using the voice assistant. All participants opted to use the maximal level of feedback. Subject used their habitual mobility aids during testing: 2 subjects used their guide dogs and 9 used their canes. The purpose of these navigation experiments was to provide users with a test of the system in a real-world environment.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

Quantitative and Qualitative Results

In the following, we will discuss the three aspects of the usability study: responses to the pre-test survey, responses to the post-test survey, and issues and limitations of the experiments.

Pre-test survey. According to the results of the pre-survey, 73% of the participants relied on others for assistance while navigating inside a building. The nature of this assistance varied widely form help pushing elevator buttons to leading the subject the entire way to their desired destination. A majority (73%) of the participants found navigation within a familiar environment easy or very easy (see Figure 5a), with subjects reporting that they used auditory cues, mental maps, and landmarks to find their way around. Within an unfamiliar environment, another majority (73%) of participants find navigation difficult or very difficult, with subjects mainly relying on other sighted people to assist them.

Post-test-survey. After testing, 10 participants agreed that the app was helpful (all except P3), 9 agreed that they could easily reach a destination with the app (all except P3 and P5), and all 11 agreed that using the app was easy (see Figure 5b). In addition, the voice features of the app were very well received. All 11 subjects found the voice feedback helpful, and all 5 subjects (P6, P7-P11) who used the voice assistant to initiate navigation also found that feature helpful as well.

During the tests, the app encountered some bugs; in these situations, we were forced to take some time to reset and reload the app. This prompted some participants to mark *down* the helpfulness of the app in the post-survey (P3 in particular reported this). Nevertheless, users had very positive impressions of the app. Many users appreciated the speech cues, saying that they were very clear and that the instructions were simple and obvious (P2, P4, P8 in particular). Users liked the vibration feedback, especially the uptick in the vibration frequency just before a turn

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

or other major directive (P10 in particular). They also praised the accuracy of the instructions in the context of their near-exact timings (for example, P8).

Issues and limitations. Users also gave suggestions to improve the application. The app encountered a few bugs during the testing process. One such bug occurred if Tango was unable to "see" the environment when a user inadvertently covered it with their finger. In these cases, navigation would freeze and would not update; in the post-survey, some users suggested alternative arrangements (such as using smart glasses or finding a way to attach the phone to the body) (P6). In another case, one user walked so fast that the app missed the target navigational node and did not update the instruction, effectively freezing navigation; here, the user suggested taking into account users' varied walking speeds to avoid this issue (P5). Others wanted greater interaction around elevators, including having the app help them find the buttons and even notify when the elevator doors open (P8). Finally, one user noted that some sort of functionality to alert them if they were about to bump into something would be helpful (P11).

Changes made. After these usability studies, we made several changes to the app and system prior to the performance studies. We added reminders about doors and other important environmental features (e.g., elevators). We fixed numerous bugs, including the delay between Tango position updates that would cause a missed instruction. In addition, we added a novel (non-controlled) condition to the performance study for observation purposes that included the use of wrist-wearable, proximity-based, vibrotactile devices for the purpose of providing greater awareness for the user (Molina et al., 2015).

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

Performance Study

Participants & Materials

In the performance study, data were collected on mobility (walking speed, collisions, and navigation errors) while BVI and blindfolded-sighted used the app. Six (6) BVI users participated in this study: 5 used a cane and 1 used a guide dog along with he app. Eleven (11) sighted control subjects participated, blindfolded and allowed to use a long cane after becoming accustomed to the cane. While using blindfolded users is not ideal, we intended for the study to show both similarities and differences between the two groups in using two aids simultaneously.

Procedure

Users were asked to repeatedly traverse a path that spanned across a single floor in three separate runs. The path (Figure 4) was 65 feet (~20 meters) long and consisted of a long corridor with three turns that took the user through three narrow doorways (the doors were propped open). The path, on average, took 1 to 2 minutes to traverse depending on the user's normal walking speed. The main study covered two conditions: (A) Baseline (navigation with the user's mobility aid and **no** other assistance including the app) and (B) ASSIST App (navigation with the user's preferred mobility aid **and** the ASSIST app). We also added a third novel condition (C) for the sake of completeness: "post-training" (navigation with the user's preferred mobility aid **and** the ASSIST app *after they acclimated themselves to the app through the prior condition #B*).

In condition A, users used only their habitual mobility aid to assist them. In this condition, one of the authors would verbally give complete navigation instructions to the user for their destination *before* they started their run; this simulated a situation in which the user asked someone around them for directions (Section S3 of the supplementary material presents what was said

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

to the participant in this condition). In condition B, users used the ASSIST app (instead of verbal pre-journey instructions) alongside their normal mobility aid. By condition C, all users had some "training" with the app and the path; thus, we allowed the user to choose which aids they wanted to use for this run. All of our blindfolded users chose to use both the app and the proximity-based, wrist-wearable vibrotactile devices (Molina et al., 2015) that would vibrate if the user veered too close to a wall or other obstacle, but not the cane. BVI users chose more complex options: one user chose to use her dog only (without the app), two chose to use their canes only (without the app), two chose to use the app plus their cane, and one chose to use his hand with the wrist-wearable vibrotactiles (but without a cane or the app). Here, we wanted to see the effect of familiarity on a user's navigational ability; we also wanted to see what users would use once they "mastered" a path/environment.

Conditions A and B were counterbalanced among users. Half of the participants performed condition A first followed by B. The other half performed condition B, followed by A. Counterbalancing was done to reduce any bias caused by a learning effect across the first two conditions; thus, we alternated between *using* and *not using* the app first. Furthermore, the path directions of tasks A and B were reversed; that is, if a user followed the path forward in the first condition, they would follow it backward in the next (but we made sure that subjects were unaware of this by bringing users through several corridors and doors before bringing them back to the new starting point). In both protocols, users performed the novel "post-training" condition (C) last.

The goal of the study was to concretely quantify navigation and walking performance with and without the app to see if there were noticeable improvements. We collected data on walking speed and navigational *events* (encounters), which comprise of (1) bumps into walls and

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

other obstacles, (2) wrong turns, and (3) needed interventions by the authors while using the app. We counted accidental bumps with walls, doors, and static obstacles such as pillars in an open space and tables on one side of the corridor. Sometimes, users intentionally used their canes or hands to touch the walls, which were not counted as events. Interventions meant some verbal or touch assistance when researchers felt there was imminent danger of bumping into an object, or when users lost their sense of direction. Identically to the usability study, we also performed a pre-survey and a post-survey. Events were recorded by hand; other raw data, including time spent on the path, was recorded using analytics functionality that we implemented into the app.

Quantitative Results

Basic statistics across groups and conditions can be seen in Table 3. Note that the number of events is the average among all participants in that group. A number under 1.0 indicates that many users did not have any errors or encounters. In summary, BVI users were, on average, much faster than blindfolded users across all three conditions. BVI users also averaged fewer total "events" per run across all three conditions, presumably due to existing experience with navigation without sight. Both groups' average time and number of events per run decreased when using the app versus those runs when the app was not used.

We performed statistical analyses (Wilcoxon signed-rank test and descriptive statistics) to determine the statistical significance of the time and event differences between using the app and *not* using the app; the raw results are listed in Tables 4-6. In brief, we found that:

1. The overall time difference between condition A (not using the app) and condition B (using the app) was *not* statistically significant (Wilcoxon signed-rank test, $p > 0.05$).
2. However, the difference in the number of "events" observed between the two conditions ***was statistically significant*** (Wilcoxon signed-rank test, $p < 0.05$).

Both conclusions hold when considering the data of all 17 subjects together as well as when considering BVI subjects and blindfolded subjects separately. We would like to note that, although performing statistical analysis on the results of 6 BVI users may not be very telling, we still include the results here for transparency. We also include some descriptive statistics for completeness.

The purpose of the novel condition C was to mimic real-world conditions, particularly with respect to the freedom of choice in mobility aids given to users. Although not controlled, we present the results of this condition here for completeness in addition to basic statistics in Table 3. Interestingly, BVI users noticed a slight (1.3 second) average decrease in the total run time after "training" and with their aid of choice, whereas blindfolded users experienced a very large (18 second) increase. Furthermore, BVI users had a relatively smaller (0.4 events on average) increase in the average number of events per run (proportionally, a two-fold increase). By contrast, blindfolded users had a *much* larger (1.5 events) increase (a four-fold increase). We believe that the relative stability of these numbers for BVI users is due, in part, to their existing experience with blind navigation, especially when compared to our sighted-yet-blindfolded users. Furthermore, BVI users chose a varied amount of aids that they were comfortable with for use in condition C. However, all of our blindfolded users unilaterally chose to use the app with some vibrotactiles (*without* the more "protective" cane) which is comparably much more difficult, especially for someone without existing experience in blind navigation.

Conclusion and Discussion

Here we would like to summarize the findings from these studies, with the limitations of our system and studies in mind, in order to provoke ideas for future directions of research, development, and studies.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

General Findings

The studies recruited 11 BVI users for the usability studies and 6 BVI and 11 blindfolded users for the performance study. The app was generally very well received by all subjects, and the performance study showed that the app reduced their navigation errors in a simple scenario (a long corridor). BVI subjects approved of the turn-by-turn voice feedback provided by the app and those who tried the voice assistant liked its simplicity. This confirms prior studies that app-based indoor navigation that uses a turn-by-turn paradigm is welcome in the BVI community (Ahmetovic et al., 2016; Nair et al., 2018a; Nair et al., 2018b).

Technologies versus Other Factors

From a general observation of the BVI community, we noted that the most popular technologies used are still long canes and guide dogs (Sato et al., 2019). From our studies and discussions with OM professionals and BVI users, this may be due to a lack of availability of and experience (training) with new technologies. We were unable to find any suitable *existing* commercial products for use in our navigation studies, which prompted us to develop our own. Even ASSIST, our own app, has several known bugs (which occurred both in and out of our studies) and is not fully functional as a product. A fully reliable app needs a fleshed-out product development cycle, which we cannot afford to do as academic researchers. Developing a real-time, reliable, low- or no-cost, user-centric app needs not only the appropriate technologies in research and development, but also related policies and new ADA compliance for buildings and facilities and market mechanisms to provide incentives to industry.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

Additional Findings to Guide Future Researchers and Developers

BVI Users vs. Blindfolded Users

Even though recruiting blindfolded users was not ideal for the performance study, we observed some similarities and obvious differences when comparing them with blind and visually impaired users. Interestingly, both groups experienced a statistically significant improvement in reducing the number of navigational events when using the app. We also observed that both groups tended to walk relatively faster with the app (despite counterbalancing). However, blind users were generally braver and walked faster, both with and without the app. Blindfolded users (perhaps objectively closer to those who are newly blind) were much more hesitant; this group, therefore, generally took a longer time to travel and experienced more events. This possibly indicates that the app should tailor its feedback and interfaces for BVI users at various stages of impairment and is a point of further study.

Is Accurate Guidance Needed?

A vision-based method (e.g., using Tango) is much more accurate than a beacon-based method. Nair, et al. (2018a) state that this is a difference of 6.5 feet (~2 meters) on average versus about an inch (~2.5 cm). The error of beacon solutions could easily go as large as more than 10 feet (~3 meters). This raises a question: *Do we need inch/centimeter-level precision?* The answer depends on both the task-at-hand and the approach we take. Our testbed lay in the very dense environment of New York City. Indeed, the available testable area for our studies of 1500 to 2000 square feet (about 140 to 185 square meters) is not very large and was unfamiliar to all of our users. This required much more precise turning and veering as it proved to be very easy for a user to bump into a wall or door. This is in stark contrast to the tasks of some previous works, which took place in large campus-like settings (Ahmetovic et al., 2016; Murata et al., 2019); however, other studies (such as Sato et al., 2019) have also indicated that high accuracy is important for some other tasks. One of the authors of this paper experienced this first-hand while accompanying a BVI user off an NYC Subway train. The user was familiar with the location; however, by making a left turn a few meters too early, he went to the end of the train platform and almost walked off it, instead of walking to the long ramp he was used to. Thus, we believe that an accurate system is needed, at least in a dense, metropolitan area like NYC. This confirmed some previous studies (Nair et al., 2018a; Sato et al., 2019). The high accuracy of the app also would enable the accurate localization of stairs, doors and elevator buttons when navigating through complex buildings independently if recognition functions are provided.

The Importance of User Feedback

Our studies have shown that the design and execution of usability studies is paramount to the successful development of such an app. We took a two-pronged approach that consisted of a usability and a performance study. Via this user-centered approach, we were able to understand the current experiences of BVI users in indoor navigation and were able to use that information to add new features and fixes to the performance study version of the app. Without their feedback, it would have been impossible to create an application that best serves the interests of visually impaired users while navigating indoors. Indeed, one subject noted that he “appreciated” the questions in the post-survey, and several subjects provided in-depth feedback and ideas that we took into consideration.

Acknowledgments

This research was supported by the U.S. Department of Homeland Security (DHS), administered by the Oak Ridge Institute for Science and Education (ORISE) under DOE contract #DE-AC05- 06OR23100 and #DE-SC0014664. This work is also supported by the U.S. National Science Foundation (NSF) through Awards #EFRI-1137172, #CBET-1160046, #CNS-1737533 and #IIP-1827505; the VentureWell (Award #10087-12); a Bentley-CUNY Collaborative Research Agreement 2017-2020; the U.S. Office of the Director of National Intelligence (ODNI) via Intelligence Community Center for Academic Excellence (IC CAE) at Rutgers (Awards #HHM402-19-1-0003 and #HHM402-18-1-0007); and NYSID via the CREATE Program. Special thanks go to all of our subjects for their participation and cooperation as well as for providing extremely helpful feedback in improving our system. Readers may view a demo video showing ASSIST's system components and human subjects testing through this link: https://youtu.be/Hq1EYS9Jncg.

References

Abdolrahmani, A., Easley, W., Williams, M., Branham, S., & Hurst, A. (2017). Embracing errors: Examining how context of use impacts blind individuals' acceptance of navigation aid errors. In Proc. 2017 CHI Conf. Human Factors in Computing System,4158-4169.

Ahmetovic, D., Guerreiro, J., Ohn-Bar, E., Kitani, K. M., & Asakawa, C. (2019). Impact of expertise on interaction preferences for navigation assistance of visually impaired individuals. In Proc. 16th Web for All: Personalization-Personalizing the Web.

Ahmetovic, D., Gleason, C., Ruan, C., Kitani, K., Takagi, H., & Asakawa, C. (2016). NavCog: a navigational cognitive assistant for the blind. Proc. 18th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '16.

Bohonos, S., Lee, A., Malik, A., Thai, C., & Manduchi, R. (2007). Universal real-time navigational assistance (URNA): an urban Bluetooth beacon for the blind. Proc.1st ACM SIGMOBILE Int. Workshop on Systems and Networking Support for Healthcare and Assisted Living Environments - HealthNet '07.

Caraiman, S., Morar, A., Owczarek, M., Burlacu, A., Rzeszotarski, D., Botezatu, N., Herghelegiu, P., Moldoveanu, F., Strumillo, P., & Moldoveanu, A. (2017). Computer Vision for the Visually Impaired: the Sound of Vision System. 2017 IEEE Int. Conf. on Computer Vision Workshops (ICCVW), 1480-1489.

Chang, Y.J., Chu, Y.Y., Chen, C.N., & Wang, T.Y. (2008). Mobile computing for indoor wayfinding based on Bluetooth sensors for individuals with cognitive impairments. 2008 3rd Int. Symposium on Wireless Pervasive Computing.

Chang, Y., Chen, J., Franklin, T., Zhang, L., Ruci, A., Tang, H. & Zhu, Z. (2020). Multimodal Information Integration for Indoor Navigation Using a Smartphone. IRI2020 - The 21st IEEE International Conference on Information Reuse and Integration for Data Science, August 11-13, 2020

Cheraghi, S. A., Sharma, A., Namboodiri, V., & Arsal, G. (2019a). SafeExit4AII: An Inclusive Indoor Emergency Evacuation System for People With Disabilities. In Proc. 16th Web For All: Personalization-Personalizing the Web.

Cheraghi, S. A., Almadan, A., & Namboodiri, V. (2019b). CityGuide: A Seamless Indoor-Outdoor Wayfinding System for People With Vision Impairments. In 21st Int. ACM SIGACCESS Conf. Computers and Accessibility, 542-544.

Cheraghi, S. A., Namboodiri, V., & Walker, L. (2017). GuideBeacon: Beacon-based indoor way-finding for the blind, visually impaired, and disoriented. In IEEE Int. Conf. Pervasive Computing and Communications (PerCom), 121-130.

Chumkamon, S., Tuvaphanthaphiphat, P., & Keeratiwintakorn, P. (2008). A blind navigation system using RFID for indoor environments. 2008 5th Int. Conf. on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology.

Ganz, A., Schafer, J., Gandhi, S., Puleo, E., Wilson, C., & Robertson, M. (2012). PERCEPT indoor navigation system for the blind and visually impaired: architecture and experimentation. International Journal of Telemedicine and Applications.

Jeamwatthanachai, W., Wald, M., & WillsIndoor, G. (2019). Navigation by blind people: Behaviors and challenges in unfamiliar spaces and buildings. British Journal of Visual Impairment. Vol. 37(2) 140–153

Karkar, A., & Al-Máadeed, S. (2018). Mobile Assistive Technologies for Visual Impaired Users: A Survey. 2018 Int. Conf. on Computer and Applications (ICCA), 427-433.

Kunhoth, J., Karkar, A., Al-Maadeed, S., & Al-Attiyah, A. (2019). Comparative analysis of computer-vision and BLE technology based indoor navigation systems for people with visual impairments. International Journal of Health Geographics, 18(1), 29.

Li, B., Munoz, P., Rong, X., Xiao, J., Tian, Y., & Arditi, A. (2016). ISANA: Wearable context-aware indoor assistive navigation with obstacle avoidance for the blind, in: Hua G., Jégou H. (eds) Computer Vision – ECCV 2016 Workshops. Lecture Notes in Computer Science. Springer. Vol. 9914, pp 448-462.

Miao M., Spindler M., & Weber G. (2011). Requirements of indoor navigation system from blind users. In: Holzinger A., Simonic KM. (eds) Information Quality in e-Health. USAB 2011. Lecture Notes in Computer Science, vol 7058, pp.673-679. Springer, Berlin, Heidelberg.

Molina, E., Khoo, W. L., Palmer, F., Ai, L., Ro, T., & Zhu, Z. (2015). Vista Wearable: Seeing through whole-body touch without contact. 12th IEEE International Conference on Ubiquitous Intelligence and Computing.

Mulloni, A., Wagner, D., Barakonyi, I., & Schmalstieg, D. (2009). Indoor positioning and navigation with camera phones. IEEE Pervasive Computing 8, 22–31.

Murata, M., Ahmetovic, D., Sato, D., Takagi, H., Kitani, K. M., & Asakawa, C. (2019). Smartphone-based localization for blind navigation in building-scale indoor environments. Pervasive and Mobile Computing, 57, 14-32.

Nair, V., Tsangouri, C., Xiao, B., Olmschenk, G., Seiple, W. H., & Zhu, Z. (2018a). A hybrid indoor positioning system for the blind and visually impaired using Bluetooth and Google Tango. Journal on Technology & Persons with Disabilities, Volume 6, 61–81. http://hdl.handle.net/10211.3/202986

Nair, V., Budhai, M, Olmschenk, G., Seiple, W. H., & Zhu, Z. (2018b). Personalized indoor navigation via multimodal sensors and high-level semantic information, Sixth International Workshop on Assistive Computer Vision and Robotics (ACVR).

Ozdenizci, B., Ok, K., Coskun, V., & Aydin, M.N. (2011). Development of an indoor navigation system using NFC technology. 2011 Fourth Int. Conf. on Information and Computing.

PANYNJ. (2018). Request for Information - Robots and Artificial Intelligence for The Port Authority of NY & NJ. RFI-52870, Good and Service Opportunities, The Port Authority of NY & NJ. https://www.panynj.gov/content/dam/port-authority/pdfs/bid-proposal/RFIDOC_52870.pdf

Real, S., & Araujo, A. (2019). Navigation Systems for the Blind and Visually Impaired: Past Work, Challenges, and Open Problems. Sensors, 19(15), 3404.

Remmen, B. M. and Toft, T. J. (2015). Blind Path. pages 1–36, Bachelor Thesis, the Faculty of Informatics and e-Learning, Sør-Trøndelag University College, May 2015.

Ruiz, A.R.J., Granja, F.S., Honorato, J.C.P., & Rosas, J.I.G. (2012). Accurate pedestrian indoor navigation by tightly coupling foot-mounted IMU and RFID measurements. IEEE Transactions on Instrumentation and Measurement 61, 178–189.

Sato, D., Oh, U., Guerreiro, J., Ahmetovic, D., Naito, K., Takagi, H., Kitani, K.M, & Asakawa, C. (2019). NavCog3 in the wild: Large-scale blind indoor navigation assistant with semantic features. ACM Trans. Accessible Computing (TACCESS), 12(3), 1-30.

Tang, H., Tsering, N., Hu, F., & Zhu, Z. (2016). Automatic pre-journey indoor map generation using AutoCAD floor plan. Journal on Technology & Persons with Disabilities, Volume 4, http://hdl.handle.net/10211.3/180125.

Wold, W. Z., & Padøy, S. H. (2016). Indoor Navigation for the Visually Impaired - A Systematic Literature Review. Specialization Project, Department of Computer and Information Science. Faculty of Information Technology, Mathematics and Electrical Engineering, Norwegian University of Science and Technology. https://folk.idi.ntnu.no/krogstie/project-reports/2016/padoy/FordypningsProsjekt.pdf

World Health Organization (2019). Blindness and vision impairment. URL: http://www.who.int/mediacentre/factsheets/fs282/en/

Yoon, C., Louie, R., Ryan, J., Vu, M., Bang, H., Derksen, W., & Ruvolo, P. (2019, October). Leveraging Augmented Reality to Create Apps for People with Visual Disabilities: A Case Study in Indoor Navigation. In 21st Int. ACM SIGACCESS Conf. on Computers and Accessibility, 210-221.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

Table 1. Feature comparison of user interface feedback categories. An ‘X’ indicates that the feature listed in the leftmost column is present in that specific feedback mode.

Interface Features	**Minimal**	**Medium**	**Maximal**
Speech announcements of all major instructions	X	X	X
Flashing of icons at major points		X	
Changing of colors on-screen		X	
Obstacle announcements		X	X
Haptic feedback (total)		X	X
Haptics: Single burst before major alerts/instructions		X	
Haptics: Continuous bursts before and at instruction point			X
Veering correction			X

Table 2. Participant demographics in the usability study.

Subject number	Visual impairment classification	Age group	Gender	App feedback mode preferred
1	Legal blindness	55+	Male	Maximal
2	Total blindness	35-44	Male	Maximal
3	Legal blindness	55+	Male	Maximal
4	Total blindness	55+	Male	Maximal
5	Total blindness	55+	Female	Maximal
6	Total blindness	45-54	Female	Maximal
7	Total blindness	25-34	Male	Maximal
8	Legal blindness	55+	Female	Maximal
9	Total blindness	55+	Male	Maximal
10	Total blindness	55+	Female	Maximal
11	Legal blindness	55+	Male	Maximal

Table 3. Basic statistics across all three conditions and both groups.

Condition	Blind and visually impaired		Blindfolded	
	Average time (s)	*Average events*	*Average time (s)*	*Average events*
A *(aid + no app)*	84.4	1.5	111.8	1.8
B *(aid + app)*	78.5	0.3	101.6	0.5
C *(after "training")*	77.2	0.7	119.6	2.0

Table 4. Results of Wilcoxon signed-rank test and descriptive statistics for condition A ("no app") and condition B ("with app") *for ALL subjects*. Assume 95% confidence interval. (M = mean time(sec), SD = standard deviation (sec), n = sample size, W = test statistics (Wilcoxon's W), p = probability value, Z = Z-value)

	No App M	No App SD	No App n	App M	App SD	App n	W	p	Z
time	102.1	33.8	17	93.5	34.3	17	47	0.163	-1.397
events	1.7	1.1	17	0.4	0.6	17	0	0.002	-3.059

Table 5. Results of Wilcoxon signed-rank test and descriptive statistics for condition A ("no app") and condition B ("with app") *for BVI subjects ONLY*. Assume 95% confidence interval. (Notations the same as Table 4.)

	No App M	No App SD	No App n	App M	App SD	App n	W	p	Z
time	84.4	25.1	6	78.6	21.6	6	6	0.345	-0.944
events	1.5	0.8	6	0.3	0.5	6	0	0.043	-2.023

Table 6. Results of Wilcoxon signed-rank test and descriptive statistics for condition A ("no app") and condition B ("with app") *for blindfolded subjects ONLY*. Assume 95% confidence interval. (Notations the same as Table 4.)

	No App M	No App SD	No App n	App M	App SD	App n	W	p	Z
time	111.8	34.9	11	101.6	38.0	11	21	0.286	-1.067
events	1.8	1.3	11	0.5	0.7	11	0	0.018	-2.366

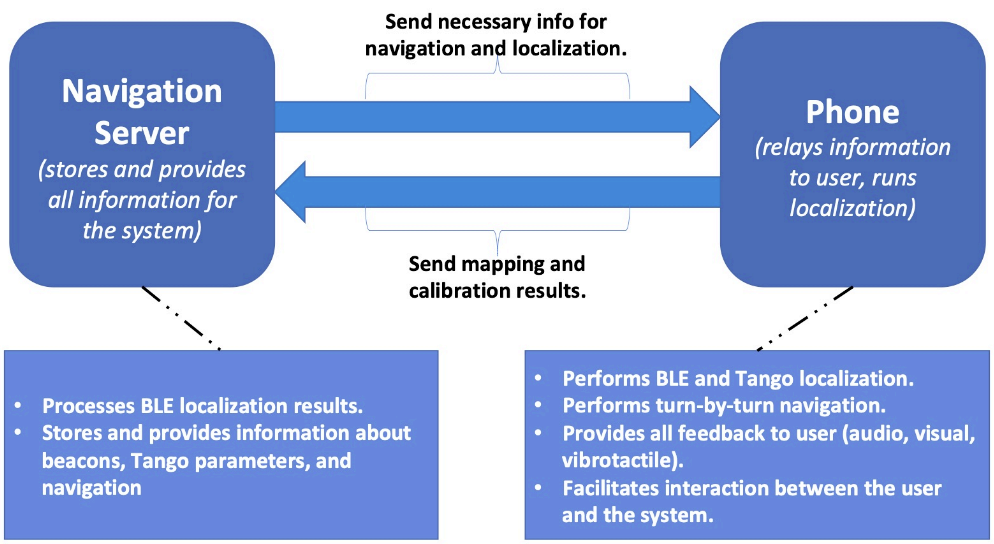


Figure 1 – Illustration of ASSIST's client-server structure.

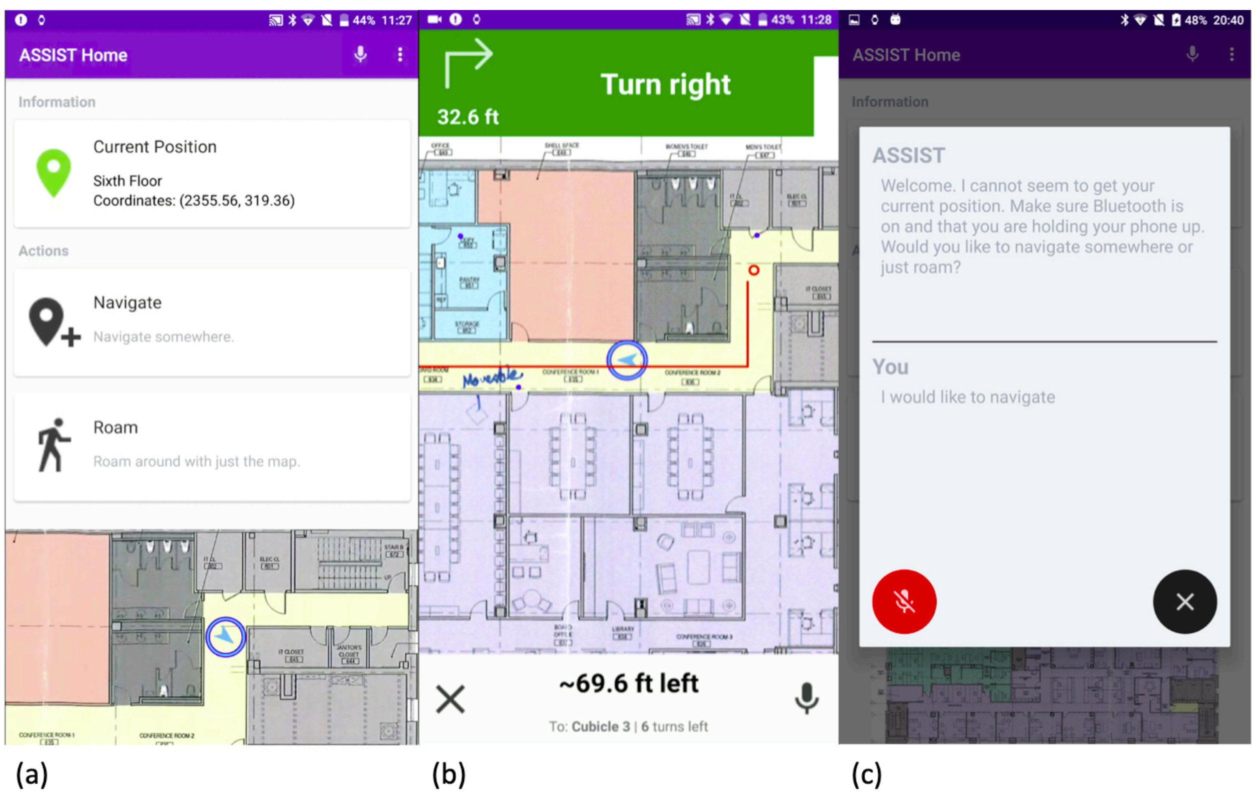


Figure 2 – Interface screens for ASSIST. From left to right: (a) home screen, (b) navigation interface, (c) voice engine interface.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

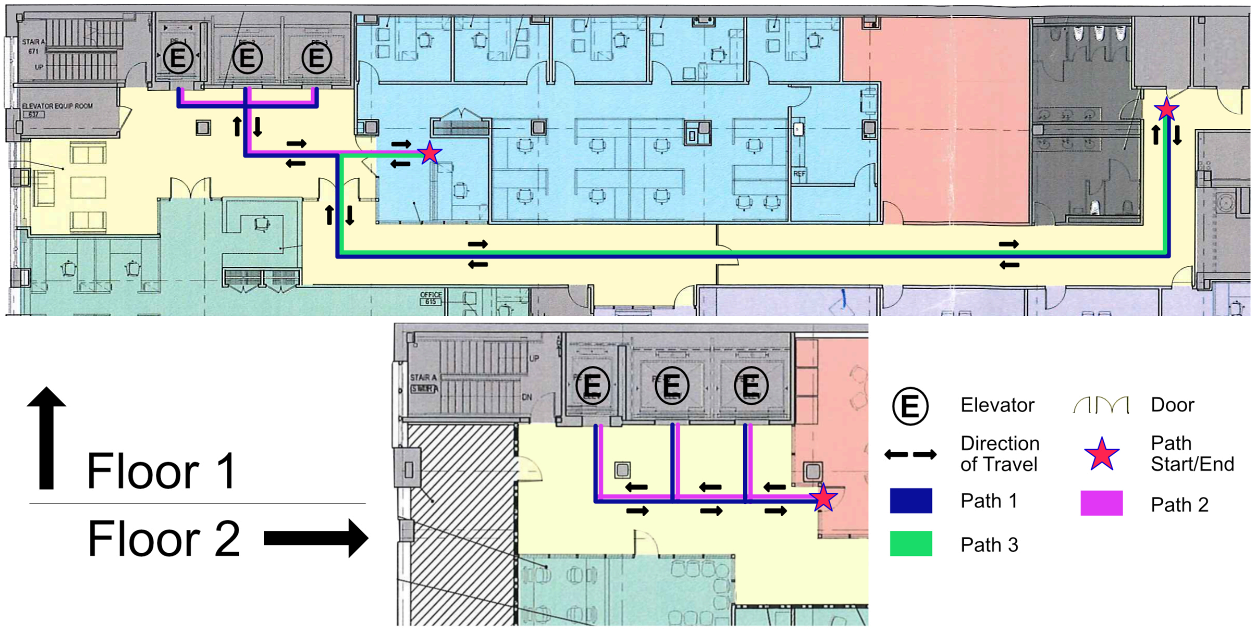


Figure 3 – Visualization of paths taken by participants during usability studies. Top panel map shows actual paths, path start/end points, directions of travel, and elevator and door locations for floor 1. The bottom-center panel map shows the same attributes on floor 2. A key explaining all symbols shown is presented in the bottom right of the figure.

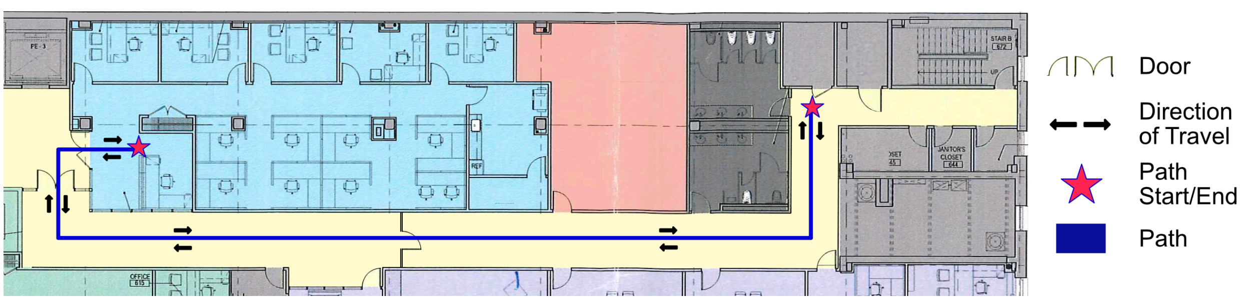
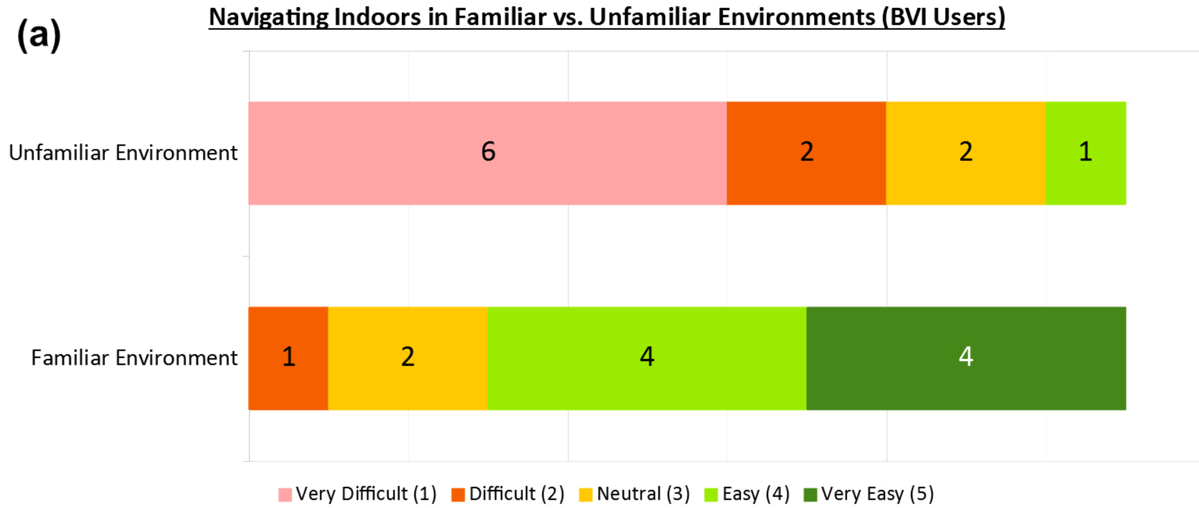


Figure 4 – Visualization of the path taken by participants during performance studies. The map shows the actual path, path start/end points, directions of travel, and door locations. A key explaining all symbols shown is presented on the right side of the figure.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553



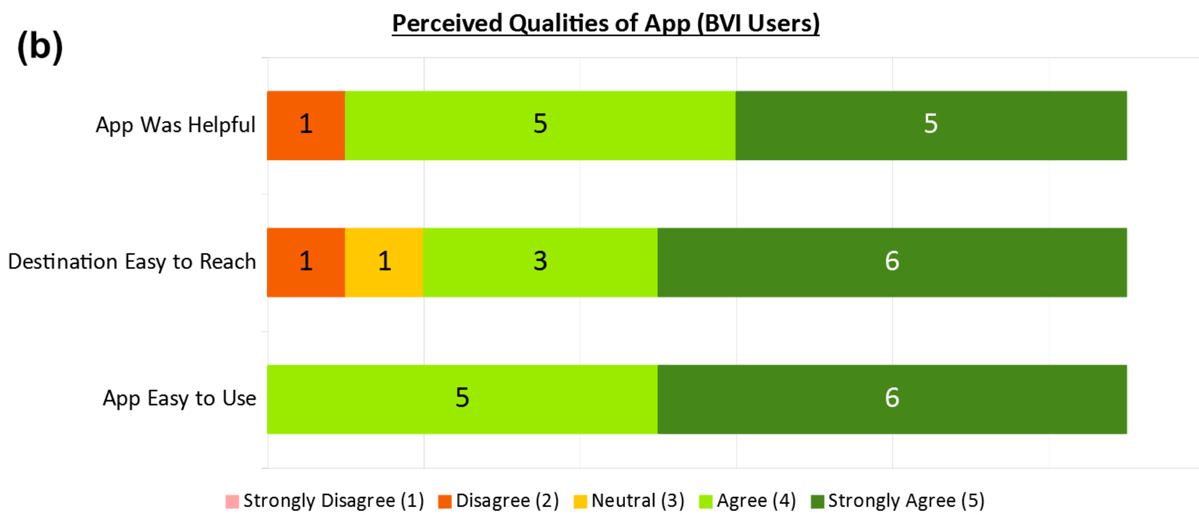


Figure 5 – Survey results. (a) Existing difficulties in navigating. (b) Perceived qualities of the app.

Supplementary Material 1 (S1): ASSIST sensory component overview

Location recognition via hybrid sensors

We explored many methods for positioning a user indoors via their mobile device; however, two methods were of particular interest to us: Bluetooth Low Energy (BLE) beacons and Google Tango. Two related works using BLE to provide turn-by-turn navigation include GuideBeacon (Cheraghi et al., 2017) and NavCog (Ahmetovic et al., 2016; Sato et al., 2019). GuideBeacon uses BLE for primary localization with the device's compass providing additional information to the wayfinding algorithm. NavCog is a smartphone-based "mobility aid" that provides turn-by-turn navigation and information about nearby points-of-interest. The original NavCog relied only on BLE; however, more recent versions include IMU measurements. A recent work by Murata et al. (2019) introduces and tests a BLE + IMU system similar to NavCog. Compared to these past works, this study notably integrates RGB-D information for precise localization. Another Bluetooth-based system proposed the use of beacons as part of a system to provide the visually impaired with information about the topology of an approaching urban intersection (Bohonos et al., 2007). BLE navigation methods have also been explored for specific BVI use cases, such as emergency evacuations (Cheraghi et al., 2019a) and the transition from indoor to outdoor navigation (Cheraghi et al., 2019b).

Google Tango, which uses a 3D sensor and computer vision, has also been of interest; the most relevant project is ISANA, a context-aware indoor navigation system implemented using Tango (Li et al., 2016). Nair et al. (2018a, 2018b) proposed a method of indoor localization that involves combining both BLE beacon localization and Google Tango map learning to create a highly accurate indoor positioning and navigation system. Kunhoth et al. (2019) provides a com-

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

parison of computer vision and BLE approaches. While they do not explore a hybrid solution, the authors show that mixed benefits from both approaches. This helps to motivate the possible advantages of a hybrid solution such as is presented here.

Note that, although Tango has been deprecated by Google, the underlying principles of real-time 3D mapping and localization are also applicable to emerging AR technologies. Yoon et al. (2019) explores using a pure AR approach to BVI navigation, demonstrating this as a potential replacement to the RGB-D Tango component of our hybrid system. Our current work has been focused on integrating Tango's successor, ARCore, into the system, and implementing an iOS version based on Apple's ARKit framework, which provides very similar real-time 3D mapping functionality. In this paper, we will use Tango as an example to describe the technology. However, the results of our human subject studies are largely agnostic to the specific technology, and thus, the observed results can be transferred to newer prototypes using ARCore and ARKit.

Despite BLE signals being easily attenuated and noisy, a prior study (Nair et al., 2018a) found that beacons were robust enough by themselves to approximate a user's position (i.e., determine a coarse location). Using it in a coarse localization step, the phone searches for all beacons it can detect in a one second interval. Of the beacons it detects in this interval, the three strongest beacons are taken and run against a pre-built database of "fingerprints" for all of the areas in question, which are created in a preceding offline mapping phase during which strategic positions are chosen to fingerprint. The region in which we installed beacons and tested are two floors of a six-story building, with an approximate total testing area of 1500 to 2000 square feet (about 140 to 185 square meters). The accuracy in using beacons for localization is about 6.5 feet (~2 meters) on average. The task of installing the beacons was trivial; in our experiments, we installed beacons above ceiling tiles (and, thus, out of view) at 10- to 15-foot intervals (about every 3 to 4.5 meters). Most of the time was spent registering the beacon signals at pre-defined locations with known coordinates, a task that could easily be done by non-experts. In our case, it took 5-10 seconds for each location, and overall, it took two days to build the fingerprint database for the two floors of a large building that served as our testbed.

Yet, some users, especially those who are BVI, require highly accurate (fine) positioning to avoid collisions with walls and other obstacles. Our early studies showed that it would be difficult for a user to navigate an unfamiliar and narrow space with relatively inaccurate and imprecise beacon localization. Thus, we investigated using Google Tango, which utilizes an RGB-D camera, that has been integrated into an Android device, to allow for device pose (orientation and position) estimation in a 3D environment (Li et al., 2016). Tango makes use of Area Description Files (ADFs), which are feature maps of an indoor environment. It can use its onboard sensors to determine a device's position within an ADF down to an inch (Nair, et al., 2018a). A single ADF may cover a single area of a floor, covering about 600 to 1000 sq. ft. (56 to 93 square meters) per ADF.

A main consideration for a hybrid solution is that it is not practical to use vision-based technologies for absolute localization in a large building using a mobile device since modeling and searching a large 3D model is computationally prohibitive. With a hybrid system, beacons are used to determine the approximate area that the user/device is located in. The area selected by BLE beacons is represented by a specific ADF that Google Tango uses to determine the user/device's exact position. Details of the hybrid localization algorithm and boundary-based ADF switching can be found in (Nair, et al., 2018b). Of note is the fact that building an ADF is a quick task, taking only a few minutes. Thus, it is reasonable to assume that a large building can also be mapped in a matter of just under a couple of days.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

Map-based semantic recognition and alerting of static environmental characteristics

Our system also utilizes pre-existing floor plans/maps of the area in question to mark the map on the interface, perform related calculations (such as distance measurements), and mark navigational nodes (along which a user will be guided) and checkpoints. We also use these maps by explicitly annotating them with various static characteristics of the environment (e.g., the locations of doorways and elevators) that are worth noting. We can then use these annotations to alert the user of these static elements and incorporate them into navigation. For example, when a visually impaired user approaches a door, the system can issue an advisory to the user, making them aware of the upcoming doorway. This concept is further prominently used in our system in the recognition of elevators, including the identification of the specific elevator that the user has entered and subsequent start of navigation from the front of this same elevator door on the destination floor. It should be noted that such an annotated map can be generated automatically, as shown in Tang et al. (2016).

System architecture and flow

In our system design, we considered scaling issues in, for example, creating a system for a large transportation hub (such as the Port Authority Bus Terminal in New York City) or a large campus. For such a large area, it is not possible to save all of the models onto a user's phone locally at once. We thus opted for a client-server architecture (Figure 1 in the main text). The client (the app on the phone) provides the user with a multimodal interface on top of the onboard localization program. The server component forms the system's core and contains all information that the app needs to operate properly, including but not limited to information about the map, Tango ADFs, coordinate transformations, and installed BLE beacon characteristics. Because of this, the phone relies on the server and connects with it as needed (both inside and outside of navigation) to exchange the necessary information.

This information exchange occurs very quickly, taking only about 100-300 ms (depending on the amount of data exchanged). These times are more than sufficient to provide a near real-time reactive experience for the user and were seen on a standard Internet connection of about 15 Mbps. Of course, the main risk is that an Internet connection might not be available. For this potential case, the required local models are downloaded in advance from the server when a connection is available. Thus, when a user invokes navigation, an Internet connection is not explicitly required.

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

Supplementary Material 2 (S2): Pre-survey and post-survey questions for usability study

Pre-survey	Post-survey
What is your age and gender? **(open-ended)**	How would you rate navigation *with* this app in comparison to navigation indoors *without* an app? **(5-point Likert scale)**
How would you classify your level of visual impairment? **(multiple choice)**	How would you rate the helpfulness of the voice feedback provided by the app? **(5-point Likert scale)**
Have you ever been on the 4th and/or 6th floor of this building before? **(multiple choice)**	Do you have any advice on how to improve the voice feedback? **(open-ended)**
How would you rate your familiarity with smartphones? **(5-point Likert scale)**	Please rate your level of agreement on the following statement: Using the app was easy. **(5-point Likert scale)**
How would you rate your level of familiarity with Android smartphones? **(5-point Likert scale)**	Please rate your level of agreement on the following statement: I could easily reach a destination with the app. **(5-point Likert scale)**
How frequently do you use your smartphone? **(5-point Likert scale)**	Complete the sentence: I found the navigation app... **(5-point Likert scale)**
How frequently do you use your smartphone for navigation? **(5-point Likert scale)**	How satisfied were you with the apps? **(5-point Likert scale)**
How would you describe the level of difficulty you experience in navigating indoors within a *familiar* building? **(5-point Likert scale)**	How would you rate the helpfulness of the voice assistant in initiating navigation? **(5-point Likert scale)**
How would you describe the level of difficulty you experience in navigating indoors within an *unfamiliar* building? **(5-point Likert scale)**	Do you have any feedback on how to improve the voice assistant? **(open-ended)**
How much do you rely on others for assistance? **(5-point Likert scale)**	How would you rate the helpfulness of the corrective turns? **(5-point Likert scale)**
How often do you get lost? **(5-point Likert scale)**	What changes (or additions) should we make to the app?
How often do you get disoriented? **(5-point Likert scale)**	What ideas (if any) do you have on how to better assist people in these situations? What can technology do to help you? **(open-ended)**
How often do you end at the wrong place? **(5-point Likert scale)**	What are other details can you share with us to help us understand the difficulties faced when navigating within a building? **(open-ended)**
How do you currently find your way inside a familiar building? **(open-ended)**	
How do you currently find your way inside an unfamiliar building? **(open-ended)**	
How do you handle situations which require you to take an elevator? **(open-ended)**	
Who do you tend to ask for help when you get lost? **(open-ended)**	
What level of help do they provide? **(open-ended)**	
How long does it take you before you are familiar/comfortable with finding your way around a building? **(open-ended)**	

Accepted Manuscript of an article published by Taylor & Francis in ASSISTIVE TECHNOLOGY, The Official Journal of RESNA, Accepted 08/04/2020
https://doi.org/10.1080/10400435.2020.1809553

Supplementary Material 3 (S3) – "Non-App" Condition Instructions (Performance Studies)

The following is a sample of the instructions given to participants in the performance study:

"You will begin by walking forward from here. After about 20 feet, you should turn right, after which, you will enter a long hallway that's about 60 feet long. Towards the middle of this hallway, you will encounter a door, so please watch out for that.

At the end of the hallway, you should turn right, whereupon you will encounter another door. Just a few feet past this door, you should turn right. You will need to go through one final door right after. Once you go through this final door, walk forward for a few feet and you will reach your destination."