ASSIST: Assistive Sensor Solutions for Independent and Safe Travel of Blind and Visually Impaired People

Zhigang Zhu\textsuperscript{1,2*}, Vishnu Nair\textsuperscript{1}, Greg Olmschenk\textsuperscript{2}, William H. Seiple\textsuperscript{3,4}

1 Department of Computer Science, The City College of New York, NY, USA
2 Department of Computer Science, The CUNY Graduate Center, NY, USA
3 Lighthouse Guild, New York, NY, USA;
4 Department of Ophthalmology, NYU School of Medicine, New York, NY, USA
* zhhu@ccny.cuny.edu

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. 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.

1 Introduction
The World Health Organization (WHO, 2019) has estimated that at least 2.2 billion people have a vision impairment globally. Among those, there are over 285 million people with low vision and over 39 million people who are blind worldwide (WHO, 2012). In the U.S. alone, the population of blind or visually impaired (BVI) people has reached 6.6 million people and is expected to double by 2030 (Varma et al., 2016). Although existing technologies (such as GPS) have been leveraged to provide outdoor navigation, there is a need for an assistive technology that aids users in indoor navigation. The development of indoor navigation technologies is important to opening up new opportunities across a wide array of professional and personal contexts, thus significantly improving the quality-of-life of millions of BVI users.

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). These allow detection of proximal objects; however, they cannot support global tasks such as planning and executing navigation. Range-vibrotactile devices have been proposed for obstacle detection, such as Vista Wearable (Molina et al., 2015) and Enactive Torch (Froese et al., 2012). In fact, there are many obstacle detection/identification options, including ultrasonic detection "smart" canes (WeWalk, 2020), Microsoft Seeing AI (2020), and real-human interfaces (BeMyEyes, 2020). However, these are object detection solutions, and not navigation aids. From our studies and discussions with Orientation and Mobility (O&M) professionals and BVI users, there is a lack of consideration of users' needs as well as low availability and production-readiness in navigation guidance. 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) (Nair et al., 2018a; 2018b). 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 users with 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 BVI users through cluttered indoor environments such as transportation hubs; it also allows for the ability to guide users toward elevator buttons, door handles, or braille signs. For example, in 2018, the Port Authority of NY&NJ issued a still-open Request for Information (PANYNJ, 2018) to which we responded (Zhu et al., 2020), exploring the possibility of using different types of robotics and AI technologies to meet their various needs, including customer service, wayfinding, and traffic management.

ASSIST also utilizes environmental annotations to provide even more 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. 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, we conducted two user-centric tests with BVI users and blindfolded-sighted users, recording their (objective) performance and (subjective) experience with the app. This paper is an excerpt of our recently published work (Nair et al., 2020). Here, we give a brief overview of the system’s implementation and then summarize the user interface design and user evaluations, with some more discussion on various social issues, such as inclusiveness, ethics and privacy.
2. Related Work

Although outdoor navigation services (i.e., those that use GPS) can generally be considered mature, indoor navigation still needs a breakthrough (Real et al., 2019). Methods of indoor positioning have proposed the use of various technologies (Karkar & Al-Maadeed, 2018, Real & Araujo 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), NFC signals (Ozdenicci et al., 2011), inertial measurement unit (IMU) sensors (Sato et al., 2019), and Bluetooth Low Energy (BLE) beacons (Sato et al., 2019, Murata et al., 2019). Where passive RFID and NFC typically have significantly limited ranges (Ganz et al., 2012) and are thus limited to proximity detection, BLE beacon signals can be detected over several meters away, allowing for localization based on signal strengths. Google Tango (which uses a 3D sensor and computer vision) has also been of interest (Li et al., 2016); Kunhoth et al. (2019) provide a comparison of computer vision and BLE approaches. Beyond the technical details, several studies have also involved evaluations with users and identified needs for these groups. Abdelrahmani et al. (2017) present a study examining what kind of errors are acceptable to BVI users, and what kinds of errors are not acceptable. 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 of the user. 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. Ganz et al. (2012) tested their 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 considered when creating a navigation and wayfinding system for BVI users.

3. ASSIST Sensors and System

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 the findings of previous studies (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 cost-effective coarse localization, in conjunction with a vision-based AR framework (in our prototype, Google Tango) for fine positioning in real-time. Here, ASSIST uses the coarse location determined using the beacons to select a local 3D model to be used by Tango. 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 that can achieve very high localization accuracy. As such, our current work has focused on integrating ARCore on Android and ARKit on iOS for newer prototypes of the ASSIST app (Chen et al., 2019; 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). These floor plans are usually available for most buildings or can be easily built since only accurate measures of traversable paths and landmarks are needed. 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.

We use a client-server structure 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 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, complex indoor facility such as the Port Authority Bus Terminal in NYC (PANYNJ, 2018). Although the app tracks users’ locations using onboard sensors and cameras, it does not save any of the data out of respect for their privacy.

4. 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.
4.1. Multimodal user interface

There are currently three options for feedback 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 spectrum is the minimal level, which utilizes the least reactive feedback. The minimal level provides simple audio guidance and vibrotactile alerts and is intended to act as a basic option for everyone regardless of their disability status. In the medium level, we use both visual cues and vibrations. (e.g., flashing and changing colors on the screen); “higher priority” events (e.g., the act of turning, or arriving at a destination) are signaled with vibrations of increased intensity. This is designed for low-vision individuals. The maximal option provides the densest feedback and is designed for people who are totally blind. Feedback includes measures for course correction as well as guidance for ensuring that the user is facing the correct direction. In such a situation, ASSIST will pause the main navigation sequence 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 complete, navigation will resume.

4.2. 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 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.”). 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 of the user can also be customized to each user via a small program within the app that updates step size in advance.

5. User Evaluation

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. We used a Lenovo Phab 2 Pro (an Android smartphone with the Google Tango 3D sensor built-in). With non-stop use of the ASSIST app, the phone’s battery can last approximately 3 hours. If a user uses the app only when needed, the battery can last for about a day. 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. All users held the device in front of them; this was done out of user preference even though the device could have been affixed to their upper body. 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 City College of New York.

5.1. Usability study

Participants & materials. 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. A convenience sample of eleven adults (P1 to P11) who were diagnosed as totally blind, legally blind, partially sighted, or low vision were offered the opportunity to participate in this study. 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 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 user’s perceived helpfulness, safety, ease of use, and overall experience of the app while navigating.

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. Starting from a pre-defined location in each of the three

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

Figure 3. Paths taken by participants during usability studies.
paths, we asked participants to use the app to navigate to a pre-determined destination. Five participants opted to start navigation using the voice input assistant. All participants opted to use the maximal level of feedback. Subjects used their habitual mobility aids during testing: 2 subjects used their guide dogs and 9 used their canes.

Results. (1) Pre-experiment 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 from 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, 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. (2) Post-experiment 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 to use. 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, P8-P11) who used the voice assistant to initiate navigation also found that feature helpful as well.

5.2. Performance study

Participants & materials. In the performance study, data were collected on mobility (walking speed, collisions, and navigation errors) while BVI and blindfolded-sighted users used the app. Six BVI users participated in this study: 5 used a cane and 1 used a guide dog along with the app. Eleven sighted subjects participated, blindfolded and allowed to use a long cane after becoming accustomed to it.

Procedure. Users were asked to repeatedly traverse a path that spanned across a single floor in three separate runs. The path 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 approximately 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 preferred 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). The goal of the study was to concretely quantify navigation and walking performance with and without the app to determine if there were noticeable improvements. We collected data on walking speed and navigational events (encounters), which comprise of (1) bumps into walls and other obstacles, (2) wrong turns, and (3) required interventions by facilitators.

Table 1. Basic statistics of performance study

<table>
<thead>
<tr>
<th>Condition</th>
<th>BVI Users</th>
<th>Blindfolded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time (s)</td>
<td>Events</td>
</tr>
<tr>
<td>A (aid + no app)</td>
<td>84.4</td>
<td>1.5</td>
</tr>
<tr>
<td>B (aid + app)</td>
<td>78.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Results. Basic statistics across groups and conditions can be seen in Table 1; a more in-depth analysis can be found in Nair et al. (2020). 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.

5. Conclusions and Discussion

Here we would like to summarize the main findings from these studies, with the limitations of our system and studies in mind, in order to provoke ideas for future directions of research and development (more in Nair et al., 2020).

General findings. 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. BVI subjects approved of the turn-by-turn voice feedback of the app and those who tried the voice assistant liked its simplicity. This confirms prior studies that app-based turn-by-turn indoor navigation is welcome in the BVI community (Ahmetovic et al., 2016).

Findings to guide future work. A vision-based method (e.g., using Tango) is much more accurate than a beacon-based method, a difference of about an inch (~2.5 cm) versus 6.5 feet (~2 meters) (Nair et al., 2018a). This raises a question: Do we need 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 NYC. 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 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. Technologies versus other factors. A fully reliable app would require a complete software development product cycle, which is impractical for 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.

Acknowledgments

This work has been supported by VentureWell [#10087-12]; NY-SID [CREATE 2017-2018]; Bentley [CRA 2017-2020]; NSF [#CBET-1160046,#CNS-1737533,#EFRI-1137172,#IIP-1827505]; DHS [#DE-AC05-06OR23100, #DE-SC0014664]; ODNI [#HHM402-19-1-0003, #HHM402-18-1-0007]. Readers may view a demo video of ASSIST at: https://youtu.be/Hq1EY89Jneg.
References


[Sato et al., 2019] Sato, D., Oh, U., Guerreiro, J., Ahmetovic, D., Naito, K., Takagi, H., Kitani, K.M, & Aso-


