NavCog3 in the Wild: Large-scale Blind Indoor Navigation Assistant with Semantic Features

DAISUKE SATO, IBM Research-Tokyo, Japan
URAN OH, JOÃO GUERREIRO, and DRAGAN AHMETOVIC, Carnegie Mellon University, USA
KAKUYA NAITO, Shimizu Corporation, Japan
HIRONOBU TAKAGI, IBM Research-Tokyo, Japan
KRIS M. KITANI, Carnegie Mellon University, USA
CHIEKO ASAKAWA, IBM Research, Carnegie Mellon University, USA

NavCog3 is a smartphone turn-by-turn navigation assistant system we developed specifically designed to enable independent navigation for people with visual impairments. Using off-the-shelf Bluetooth beacons installed in the surrounding environment and a commodity smartphone carried by the user, NavCog3 achieves unparalleled localization accuracy in real-world large-scale scenarios. By leveraging its accurate localization capabilities, NavCog3 guides the user through the environment and signals the presence of semantic features and points of interest in the vicinity (e.g., doorways, shops).

To assess the capability of NavCog3 to promote independent mobility of individuals with visual impairments, we deployed and evaluated the system in two challenging real-world scenarios. The first scenario demonstrated the scalability of the system, which was permanently installed in a five-story shopping mall spanning three buildings and a public underground area. During the study, 10 participants traversed three fixed routes, and 43 participants traversed free-choice routes across the environment. The second scenario validated the system's usability in the wild in a hotel complex temporarily equipped with NavCog3 during a conference for individuals with visual impairments. In the hotel, almost 14.2h of system usage data were collected from 37 unique users who performed 280 travels across the environment, for a total of 30,200m traversed.

CCS Concepts: • **Human-centered computing → Accessibility technologies**; • **Social and professional topics → People with disabilities**; • **Information systems** → *Location based services*;

Additional Key Words and Phrases: Indoor navigation, visual impairments, points of interest, voice interaction, user evaluation

This work was sponsored in part by JST CREST (Grant No. JPMJCR14E1) and NSF NRI (Grant No. 1637927).
Authors' addresses: D. Sato and H. Takagi, IBM Research-Tokyo, Hakozaki-cho 19-21, Nihombashi, Tokyo, 103-8510, Japan; emails: {dsato, takagih}@jp.ibm.com; U. Oh, J. Guerreiro, D. Ahmetovic, and K. M. Kitani, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA, 15213, USA; emails: {uranoh, jpvguerreiro, dragan1, kkitani}@cmu.edu; K. Naito, Shimizu Corporation, Kyobashi 2-16-1, Tokyo, 104-8370, Japan; email: k.naito@shimz.co.jp; C. Asakawa, IBM Research, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA, 15213, USA; email: chiekoa@us.ibm.com; Daisuke Sato's current affiliation is Carnegie Mellon University, USA; email: daisukes@cmu.edu; Uran Oh's current affiliation is Ewha Womans University, South Korea; email: uran.oh@ewha.ac.kr; João Guerreiro's current affiliation is Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento, Lisboa, Portugal; email: jpvguerreiro@gmail.com; Dragan Ahmetovic's current affiliation is Università degli Studi di Milano, Italy; email: dragan.ahmetovic@unimi.it.

1936-7228/2019/08-ART14 $15.00
https://doi.org/10.1145/3340319

ACM Reference format:
Daisuke Sato, Uran Oh, João Guerreiro, Dragan Ahmetovic, Kakuya Naito, Hironobu Takagi, Kris M. Kitani, and Chieko Asakawa. 2019. NavCog3 in the Wild: Large-scale Blind Indoor Navigation Assistant with Semantic Features. *ACM Trans. Access. Comput.* 12, 3, Article 14 (August 2019), 30 pages.
https://doi.org/10.1145/3340319

1 INTRODUCTION

The ability to travel independently is a fundamental need for everyone. However, independently traversing large and unfamiliar places, such as shopping malls, hotels, hospitals, and airports, is also one of the most significant challenges for people with visual impairments (PVI). While sighted persons can rely on visual information available in the environment, such as doorways, shops, and obstacles, PVI cannot use such semantic features to support their mobility in the real world. To address this issue, PVI learn to rely on non-visual exploration [13, 14, 27]. Moreover, in unfamiliar environments, PVI often have to request assistance from others. However, sighted assistants might not be readily available [72].

Many assistive navigation systems have been studied to facilitate independent mobility for PVI [23, 28]. Some are capable of providing accurate turn-by-turn guidance [73, 74]. However, most systems are tested in constrained laboratory scenarios without considering the challenges of navigation assistance in large-scale real-world environments [23]. Our prior research explored the technical challenges in providing accurate and practical localization and navigation assistance in wide and unconstrained real-world scenarios [62]. However, to the best of our knowledge, no studies reported users' experiences, behavior, and feedback while guided by an accurate turn-by-turn navigation assistant in real-world large-scale indoor environments.

This manuscript is an extended version of the research presented at ASSETS 2017 [76]. We extended this research involving two user studies with an additional third user study performed in-the-wild at a conference for PVI. We also provide detailed analysis of users' travels not only for the third study but also for the second study of the initial research. The two previous studies involved users navigating in a large shopping mall (approximately 21,000m^2) using NavCog3, an accurate turn-by-turn navigation assistant system we developed [62, 76]. These initial studies addressed the following research questions: (1) What level of accuracy can NavCog3 achieve during real-world turn-by-turn navigation, and is it sufficient for indoor navigation assistance for PVI? (2) What are the needs, preferences, and expected use cases of PVI for accessing semantic features during navigation? (3) To what extent can a navigation system be used to enhance orientation and mobility of PVI?

The first study was conducted in a controlled setting with ten PVI and consisted of three fixed route-navigation tasks. During this study, we measured localization accuracy while participants were performing the navigation tasks. The accuracy was 1.6m error on average, which is consistent with state-of-art smartphone-based indoor localization systems [59]. With this accuracy, 93.8% of the 260 turns made by participants were successful. We assessed the turn performance and found that 45° turns were more challenging to make than 90° ones, which was further explored in a follow-up study [6]. Findings from that first study confirmed that participants found semantic features to be useful for boosting their spatial awareness (e.g., obstacles), orientation, and mobility (e.g., using elevators). The second user study was conducted in the same venue with a more realistic scenario to reflect on PVI natural use of the system in a large-scale environment. We asked 43 PVI to use NavCog3 to reach any destination inside a shopping mall. Findings from the second study showed that NavCog3 allowed our participants to reach the shops of their choice with confidence.

The third study was conducted during a conference for PVI held in a hotel environment (approximately 11,000m^2). During this event, we carried out an ad hoc installation of NavCog3 in the conference venue. The conference participants, who were PVI, were allowed to navigate the venue by themselves without being accompanied by researchers. The attendees could freely use NavCog3 anytime and anywhere they wanted to go between their rooms, hotel facilities, and conference venues. After the conference, we examined the usage data of NavCog3, focusing on the following additional research questions: (4) What level of accuracy can be achieved in an ad hoc installation of NavCog3, and is it comparable to previous studies? (5) What are the use cases addressed in the wild, and what is the system performance during these use cases? (6) Are the user needs and preferences regarding semantic features similar across different environments? The results from the third study indicated that the participants could navigate independently in the hotel environment without asking anybody for help. These findings suggest that NavCog3 can increase users' confidence in walking independently and enable PVI to find and visit new places.

2 RELATED WORK

Navigating in an environment through only non-visual cues can be cognitively demanding [13]. Thus, PVI often receive orientation and mobility (O&M) training to improve their navigation skills [14]. O&M training enables safe and independent navigation on frequently traversed routes [11]. However, in new and unfamiliar environments, autonomous navigation is still unreliable and slow; therefore, it is generally avoided [53].

To support independent mobility of PVI, we consider prior research on assistive technology regarding: (1) localization and navigation assistance; (2) providing location-based semantic and contextual information; (3) interaction on-the-go. In addition to these research directions, we also review the literature on the NavCog system and how it evolved throughout the years.

2.1 Localization and Navigation Assistance

Mainstream GPS navigation tools (e.g., Google Maps [32]) and accessible GPS navigation assistants (such as Ariadne GPS [8]) are often used to support independent mobility for PVI [19]. However, GPS localization can degrade in construction-dense outdoor areas [80] and may even be unavailable in indoor environments [86]. Thus, researchers have studied alternative approaches to support indoor localization and navigation for PVI, such as using infrared [38], ultrasound [73], RFID [24], and ultra-wide band (UWB) [74] sensors. These approaches achieve reasonable accuracy but often require users to carry a proprietary dedicated device (i.e., a receiver).

Smartphone-based approaches can be used to provide localization and navigation capabilities without requiring a user to carry additional devices. Such approaches can provide guidance using Wi-Fi [17], Bluetooth Low Energy (BLE) beacons [51], inertial measurement units (IMUs) [25], videocameras [54], or a combination of sensors [26, 44, 55, 56]. In particular, Bluetooth-based solutions rely on commodity smartphones alone and off-the-shelf BLE beacons and have been used in large-scale installations supporting mobility for PVI [76, 77, 83]. Among these, NavCog3 [62, 76] supports reliable navigation assistance for PVI using built-in smartphone IMU sensors and BLE beacons and can estimate users' locations with high accuracy (1.65m on average).

2.2 Providing Location-based Semantic Information

Semantic features are useful for understanding the traversed environment and for confirming a person's location/orientation in it [49, 58, 64, 71]. Kammoun et al. [49] classify these semantic features into points of interest (POIs), such as buildings or shops, and landmarks, such as ground texture, or traffic lights for outdoor environments. However, detecting and recognizing semantic features present in the surroundings is difficult for PVI. There are a number of commercial mobile applications that provide semantic information about the environment [15, 46, 47, 60]. BlindSquare [15] and Soundscape [60], for instance, inform the user of nearby POIs based on the nearest BLE beacon or GPS localization. The available POIs are grouped into categories such as education, shopping, and transportation. Footnotes [31] is built on top of Soundscape and provides rich textual descriptions that are annotated by blind and sighted users and contains functional, visual, historical, and social descriptions.

Prior research explored how semantic features in the environment impact the navigation for PVI [20, 36, 49, 71]. Harper et al. [36] summarized problems of urban indoor navigation for PVI in terms of human-factors, socioeconomics, and systems and categorized POIs (mobility objects) into cues, obstacles, out-of-view items, and memories. Dias et al. [20] further assessed which features support indoor navigation and orientation of PVI through an interview with 20 participants. The majority of the participants reported that they use environmental cues (e.g., smells, sounds) or landmarks (e.g., doorways, elevators) to inform their decisions during navigation. Systems that provide location-based semantic information, however, provide no means to reach these features. For example, a POI on the other side of a block of buildings may be considered in proximity, but it requires the user to walk around the whole block. To the best of our knowledge, NavCog3 is unique in that it determines the position and orientation of these semantic features with respect to a route that navigates the user through the environment.

2.3 Interaction On-the-Go

Mobile-device usage was found to positively correlate to the independence of PVI in mobile contexts [50]. At the same time, using a mobile device during mobility can be cognitively demanding, since users have to concurrently pay attention to walking, interacting with the mobile device, and staying alert to ensure their safety [1, 57, 69, 85]. Moreover, since PVI often hold a cane or guide-dog leash during mobility, using a mobile device at the same time is even more challenging. To support mobile-phone usage in such use cases, one-handed or hands-free interaction paradigms have been investigated [65]. In particular, wearable technologies have been used to achieve hands-free navigation for PVI [75, 79, 81]. However, these paradigms often require custom interaction interfaces, which are not mass produced; therefore, they are not feasible to support a large population of visually impaired users.

In this regard, NavCog3 has been designed as a smartphone application, with gesture-based one-handed interaction in mind, to be readily accessible for PVI. To minimize the user's cognitive load, NavCog3 can also be accessed through a conversation-based interface. Such interfaces are widely used in many applications, such as those for elderly care [37] and knowledge workers [12]. Evaluating the conversation cognitive assistant interface is not within the scope of this article as it is in its early development phase. However, we expect that speech-based interaction will relieve the cognitive and physical load on users, as speech interaction is known to be faster and more immediate than manual input [10].

2.4 Background on NavCog

Our research goal is to increase the independence of PVI during mobility by providing turn-by-turn navigation assistance based on high localization accuracy using only off-the-shelf BLE beacons and commodity smartphones. We started the NavCog project in 2015. Since then, we have been studying and improving the system in two directions: (1) localization and navigation accuracy, and (2) interaction with users. We have also improved the system's robustness and made it easier to install, update, and maintain. As a part of this research, we published the key software components of the system as an opensource project called Human-scale Localization Platform (HULOP) [39]. The first iteration of the system used K-nearest neighbors (K-NN) regression to estimate the user's position based on BLE-beacon signal samples collected across the environment [2, 3]. The following improvement to the system introduced pedestrian dead reckoning (PDR) capabilities to stabilize the estimated localization [5]. Finally, the current iteration of the localization approach is based on a probabilistic model and includes sophisticated heuristics to enable accurate localization in large, real-world environments [62].

The initial interface of the system implements a simple navigation logic, assuming that the user always follows the route correctly. The interaction paradigm was improved in the latest iteration of the system with failure recovery and semantic features such as POIs and non-visual landmarks [76]. This version (NavCog3) was also deployed and evaluated at the Pittsburgh International Airport with slight adaptations to cope with the additional challenges of moving walkways and veering in large open spaces [34]. A version allowing a seamless interaction between turn-by-turn navigation and art appreciation using the users' body orientation was also recently piloted at the Andy Warhol Museum in Pittsburgh [9]. Recently, we have also investigated how the environment, user behavior, and expertise affect navigation and its outcome [35, 48, 66, 67]. We will adapt future iterations of the system interface based on the results of this investigation.

An orthogonal research direction explored ways to outsource system installation and maintenance procedures through physical crowdsourcing [30]. In particular, we investigated how the required workload varies based on installation characteristics (e.g., number of beacons and data samples) and how such characteristics impact the expected localization accuracy [5]. We then explored incentives for crowd-driven installations and validated their feasibility and the achieved accuracy [29].

We are currently exploring how to increase the knowledge of PVI in environments that are not equipped with the required infrastructure. For this purpose, we developed a smartphone-based virtual navigation app that enables users to build a sequential mental representation of routes before physically visiting them [33].

This manuscript extends our prior research [76], during which we conducted two user studies of NavCog3 involving 43 participants navigating in a large-scale real-world environment. For the extension study, we explored in-the-wild usage of NavCog3 during a temporary installation of the system at an annual conference for PVI.

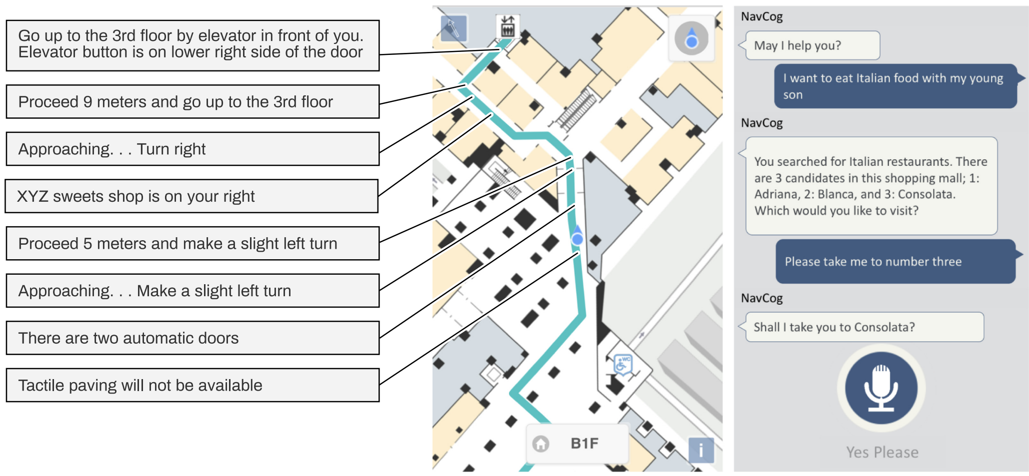
3 DESIGN OF NAVCOG3

As we previously reported [76], we designed and implemented NavCog3 as a smartphone-based indoor navigation system for PVI. It is readily available for any modern-smartphone user who wishes to receive navigational assistance without the need for an additional device. In this section, we describe the navigation instructions and modes of interactions. The details of our implementation can be found in Section 4.

3.1 Turn-By-Turn Instructions

NavCog3 provides turn-by-turn guidance using speech feedback to convey instructions and enable eyes-free interaction. The system also displays navigation information on the screen (e.g., a planned route, the current location) to support users who have residual sight and wish to receive visual feedback (see Figure 1(a)).

With the advantage of higher accuracy, NavCog3 can provide turn instructions in a timely manner so that PVI can easily make correct turns without visual aid. The system also provides "Approaching" notifications prior to the turning point to allow users to prepare. After a turn (or at the starting point), the system informs the user of the distance from the current position to the next turning point and the action that they need to take, such as turning or transitioning between floors (elevator, escalator, or stairs). For sufficiently long distances (≤15m), the system gives a



(a) Navigation screen and examples of announcements (b) Conversation

Fig. 1. Example screen-shots of NavCog3's interfaces.

verbal update on the remaining distance every 15m. At each turning point, the system provides verbal instructions to convey the angle of the turn: a *slight turn* ($22.5^\circ < \theta \leq 60^\circ$), *regular turn* ($60^\circ < \theta \leq 120^\circ$) and *big turn* ($120^\circ < \theta$). When a user reaches a turning point, a short vibration and short sound effect are provided simultaneously to instruct the user to start turning. Once the user reaches the correct heading, the feedback is provided again to instruct him/her to stop turning and continue walking.

To help the users get back on track when they veer from the route (by 6m or more) or head in a wrong direction (by 120° or more), our system provides failure-safe guidance that notifies the user immediately when an incorrect orientation is detected. For example, if a user heads in the opposite direction, the system says: "Turn around. You might be going in the wrong direction." If a faster route exists from the user's current location after deviating from the original route, then corrective guidance will recalculate the path and lead the user to the new route.

3.2 Nearby Landmarks and POIs

The system provides information about nearby landmarks and POIs so that a user can walk comfortably and confidently. Table 1 shows the types of semantic features supported by our system, which we previously categorized by summarizing previously examined features [20, 49] and features we found in our preliminary iterative on-site testing in the field with blind people and O&M experts [71]. We define landmarks as features that can provide physical or tactile cues to help users confirm their location such as tactile paving, doorways, and other building infrastructures (e.g., elevator, escalator, and stairs). POIs, such as shops and facilities (e.g., restroom), are defined as places that might interest users during navigation.

In general, information about POIs and landmarks is provided when a user is in their proximity. However, information about accessibility-related landmarks is already included in the navigation instructions, for example, "Proceed 14 meters on tactile paving and turn right at the end of the corridor."

3.3 Interaction

3.3.1 Search POI via Conversation. For unfamiliar environments, users may not know what places are available for them to navigate. To help users select their target destination, NavCog3

Table 1. Categories of Semantic Features for NavCog3

	Main	Sub	Information Provided
Landmark	Pathways	Floor	Tactile paving, ramp, step, slope, floor materials such as carpet and tiles
		Door	Types of doorways (e.g., automatic)
		Obstacles	Existence of object (e.g., trash cans)
	Floor Transition	Elevator	Call-button locations, control-panel location inside the elevator, Braille-button availability, audible-announcement availability, wheelchair accessibility
		Escalator	Correct standing side (left or right), directions to adjacent escalator(s)
		Stairs	Shape (e.g., straight, u-shape), number of steps and landings
POI	Shopping and Food	Shop	Type of shops, name
		Restaurant	Cuisine, name
	Facility/Utility	Restroom	Gender, wheelchair accessibility
		Other	ATM, information center

enables them to explore nearby POIs by asking the cognitive assistant via speech input. The assistant provides recommendations on the basis of user's request with descriptions.

As shown in Figure 1(b), a user can initiate a conversation with a request such as "I want to eat Italian food with my young son." The system then provides a set of recommendations on the basis of the search conditions extracted from the users' input; "Italian" and "kid friendly," in this case. If the user has a specific place in mind, then the system also enables the user to either manually search for the destination from a POI list or set the destination using speech input.

3.3.2 On-demand Instructions. NavCog3 supports on-demand instructions, which enables users to listen to the current instruction again upon request by simply tapping the screen or pressing a control button on a headset as needed. Example instances include confirming the next action, confirming the surrounding information (e.g., the position of elevator buttons), or checking the remaining distance to the next turning point. The system generates an appropriate instruction on the basis of the current navigation status and the user's location and heading.

4 IMPLEMENTATION

In this section, we present the main components of the NavCog3 system [40] (see Figure 2), which includes an iOS app distributed from Apple's App Store [63]. We also describe the details of the system's deployments in two large-scale environments: a shopping mall (two initial studies) and a hotel (extension study).

4.1 Location Engine

We now summarize the procedure to deploy the localization environment with our localization library [41]. Please refer to our previous study [62] for further details on the localization method. To achieve higher accuracy with the system, proper beacon placement and fine-grained fingerprinting of beacon signals are essential.

4.1.1 Beacon Placement. Since the localization using NavCog3 is estimated on the basis of the detected beacon signals and their received strength from a user's phone, we first need to place

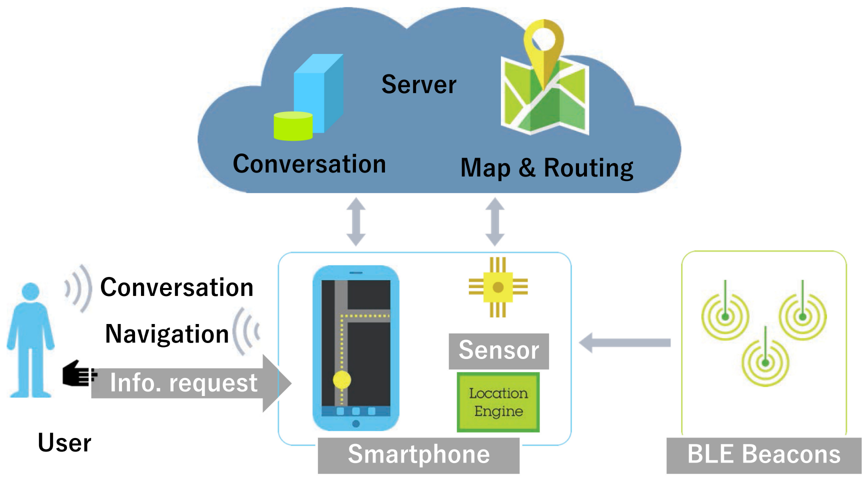


Fig. 2. System overview: User requests information or starts navigation through speech interaction using smartphone.

BLE beacons in the areas of the environment in which we wish to support navigation assistance. Based on our experience [5], the beacon-placement interval should be about 7–10m to achieve high accuracy. In addition, the ideal height of the beacon placement, to achieve good radio-wave-signal reachability, is between 2 and 3.5m. At such height, interference from passersby is weaker; therefore, signals are received with less noise. However, if the position is too high, the distance between the mobile device and beacons increases, which results in loss in accuracy, because the variance of the signal power increases with distance. Beacons can also be hidden (e.g., under ceiling tiles) if the covering material has little impact on signal power.

4.1.2 Fingerprint Collection. Once all the beacons are placed, we collect BLE-signal samples (fingerprints), which are measurements at a known location of signal strengths from all beacons within the range. Fine-grained fingerprinting is required for accurate localization. However, it is time-consuming to manually collect fingerprints point-by-point, and current alternatives do not guarantee high accuracy [45]. To address this issue, we developed a machine to assist us in collecting fingerprints. The machine is equipped with a LIDAR sensor that can scan the environment with lasers and obtain its coordinates in the environment with a centimeter-level error. This machine can reduce the fingerprint-collection time to 1/20 compared to point-by-point manual fingerprinting.

4.2 Map & Routing Server

To support indoor navigation for these environments, we first created indoor floor plans for the area from CAD drawings of the buildings. We then compiled the pedestrian network and POIs for the entire environments from scratch using our Web map editor [42].

The map data model is an extension of a spatial network specification for people with disabilities [21]. It converts a list of physical pedestrian maps and semantic features described in Section 3.2 into a spatial graph with semantic features. Each graph edge has attributes detailing path width, gradient, steps, stairs, elevators, escalators, and tactile paving.

The routing service computes the optimum route from the available routes given the current location and destination as well as users' mobility preferences. By default, it avoids escalators and

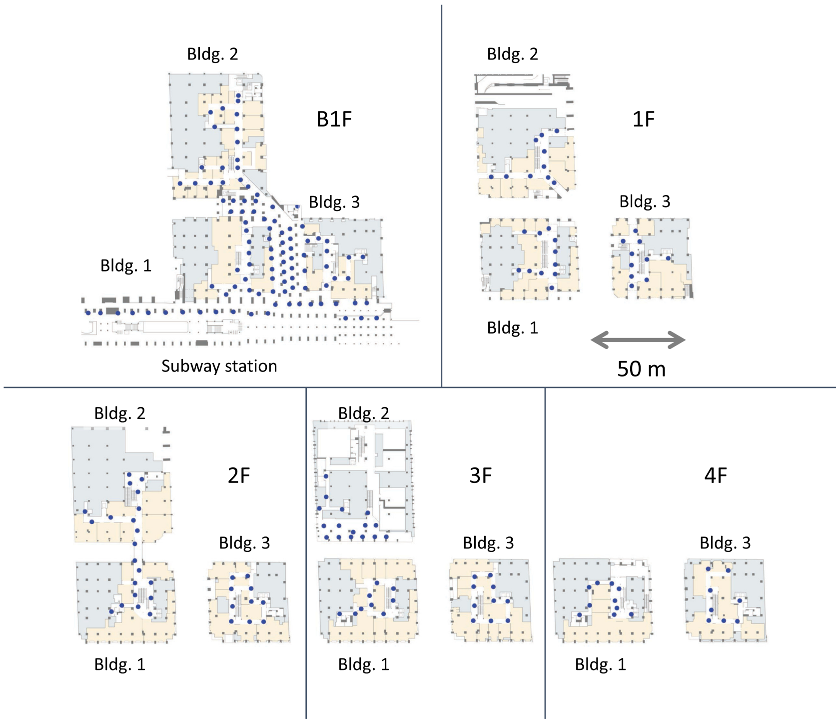


Fig. 3. Beacon-deployment locations in the shopping-mall environment. Blue dots indicate beacon locations.

includes routes with tactile paving for PVI. NavCog3 generates navigation instructions on the basis of these routes and semantic features along with them.

4.3 Conversation Server

A conversation system was developed as a Web service that combines two types of components: (1) a basic conversation script using the Watson Assistant API [82] for fixed common facilities and (2) a shop-recommendation engine based on the content of available shops (only for shopping malls), which can be adapted in case there are changes in the environment. Users' speech input is transcribed using the speech framework of iOS [78] then submitted to the server. The server runs both components simultaneously and returns the recommendation result if the confidence level of the result is greater than a certain value; otherwise, it returns the result of the conversation script. For shop recommendations, the system requires shop information in text format with annotations such as shop name, shop introduction, opening hours, and menu. The recommendation engine attempts to extract the search conditions from the query text to filter shops that meet the conditions and then match the text in the query with that in the shop information.

4.4 Large-Scale Deployments

The following sections describe the details of large-scale deployments with our two evaluation scenarios as examples: a shopping mall and a hotel.

4.4.1 Shopping Mall. The shopping-mall deployment involves three buildings that span about 21,000 m^2, and the connection to a subway station on the first basement level, as shown in Figure 3. We deployed approximately 220 battery-powered BLE beacons [7] inside the shopping mall, each

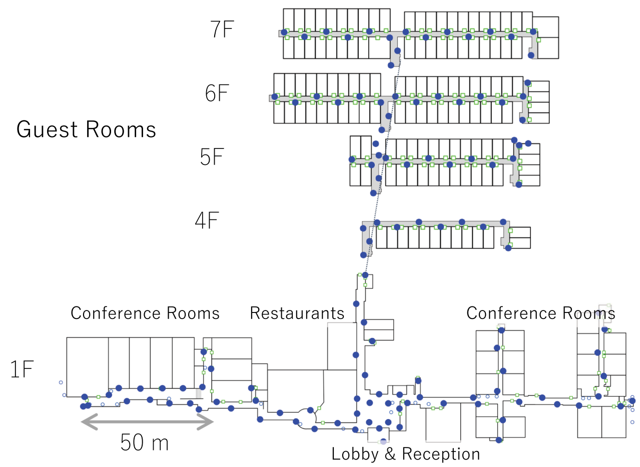


Fig. 4. Beacon-deployment locations in hotel environment. Blue dots indicate beacon locations.

costing about $20 and lasting for about one year. Most of the beacons were placed under ceiling tiles to hide them, while the rest were placed close to the floor if the ceiling was too high. The beacon locations are shown as blue dots in Figure 3. All beacons were deployed over two nights during closing hours. Due to ownership issues, we did not deploy beacons inside shop areas, although this would have been helpful to navigate users directly to the entrance of each shop. We also collected and annotated text information of the shops from their webpages to train the recommendation engine of the conversational assistant.

As we used a LIDAR sensor, we were able to collect fingerprints for creating the map of the entire area in 12h (3h for three buildings and corridors on the first basement level). This fingerprint collection was conducted during shop opening hours, because the signal strength from beacons can change depending on environmental changes over time, such as closed shutters. We also collected approximately 8,000 data points while walking in the environment to evaluate the localization error prior to the user study, and the result was 1.47m on average.[1]

4.4.2 Hotel. We used 130 battery-powered BLE beacons [18] for the hotel. Some of the beacons were deployed outside the hotel to support outside facilities such as guide-dog relief areas. The beacons locations are shown as blue dots in Figure 4. The first floor of the hotel consists of conference rooms, meeting rooms, restaurants, and facilities for their guests. All the other floors are for the guests, but we supported only accessible guest-room areas, where most of the conference participants were staying. We built a conversation script to deal with some use-cases of a typical conference at a hotel such as "I wanna see the exhibition X" or "I wanna go to my room." The recommendation engine was not used because the hotel has few dining options.

We collected about 1.5h of fingerprints with the LIDAR sensor. We also gathered approximately 1,500 data points to evaluate the localization error prior to the user study, which was 1.68m on average.

5 STUDY 1: FIXED-ROUTE AT SHOPPING MALL

To evaluate the localization accuracy of NavCog3 and collect subjective feedback on the usefulness of semantic features during navigation, we conducted a 90min single-session study with ten

[1]The result is slightly different from that in Reference [62], because data was added after the user study for further analysis of the localization engine.

Table 2. Participants' Demographic Information for Study 1

ID	Gender	Age	Visual Impairments	Mobility Aid	Smartphone ownership
P1	Female	42	20/2000 for both eyes	White cane	4 years
P2	Female	46	20/500 with right eye only	White cane	1.5 years
P3	Male	54	Totally blind	White cane	4 years
P4	Female	44	Totally blind	Guide dog	1 week
P5	Male	33	Totally blind	White cane	1 year
P6	Female	53	20/500 with left eye only	White cane	1 year
P7	Male	38	Totally blind	White cane	None
P8	Female	40	Totally blind	White cane	None
P9	Male	42	20/500 with left eye only	White cane	None
P10	Female	48	Totally blind	White cane	4 years

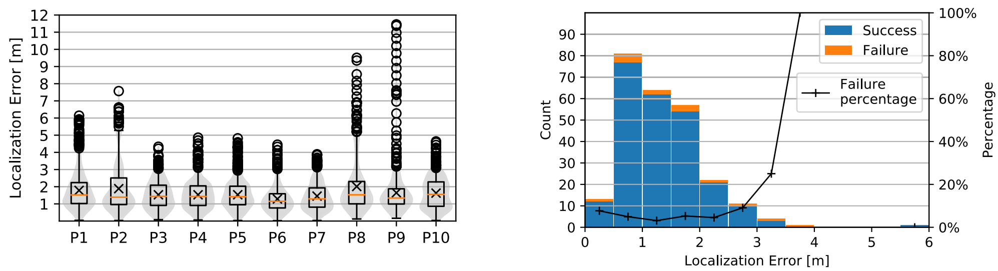
participants with visual impairments in the large shopping mall during opening hours. During the study, participants were asked to navigate three fixed routes (the results of this study are also reported in Reference [76]).

5.1 Method

5.1.1 Participants. We recruited ten PVI through a local Braille library (Table 2). Six were totally blind (P2 and P10 were visually impaired since birth), and the rest had low vision (at best 20/500). In terms of smartphone ownership, all but three participants (P7–P9) have been using smartphones for at least a year except for one (P4), who had been using a smartphone for a week. Four of the participants (P3, P5, P8, P10) reported that they have used a navigation system such as Google Maps before. All have participated in a study with the previous version of NavCog except three participants (P5, P6, P10). Participants were compensated ¥10,000 (approximately $90) for their time.

5.1.2 Apparatus. All participants used a revised version of the NavCog app during the study session. The app was running on an iPhone 6 smartphone, and bone conduction headphones were used to provide audio instructions while not impeding any environmental sounds, which was found to be important for safe navigation [1, 85]. All participants were asked to wear a waist bag with the phone attached to free their hands from holding the phone during navigation, which again is important when walking on a street, especially for PVI whose hands are often occupied holding a cane [1, 85]. All participants were provided with a remote control to interact with NavCog3 while the phone was secured in the bag. The app logged every event with a time stamp while the app was running (e.g., instructions), and an experimenter closely followed and videotaped all participants with a 360-degree camera.

5.1.3 Procedure. The session began with a background questionnaire. A short training session (5–10 min) was then provided to the participants who were trying NavCog3 for the first time until they were familiar with it, such as the navigation voice and types of information and instructions the system provides during navigation. Prior to the tasks, participants were asked to walk for 5–10m to calibrate their location and heading. They were then asked to navigate three different routes in the shopping mall. The order of the routes was the same for all participants. We instructed the participants to walk normally at their own pace, and the volume and speech rate were adjusted for each participant prior to the tasks. After the task was completed, we collected subjective responses on the usefulness of the semantic features we provided during navigation.



(a) Localization error in meters per participant across three routes (lower value indicates higher accuracy). Boxes indicate 25th, 50th, and 75th percentiles and 'x' denotes means. Outliers that do not fall in 95th percentile range are shown as circles.

(b) Distribution of localization error at turns (N = 254), where each turn is grouped on basis of its localization accuracy with interval of 0.5 m (e.g., [0.0 m, 0.5 m).

Fig. 5. Localization performance in Study 1.

5.1.4 Data Analysis. We visually annotated participants' actual location every second from the video data. For each of the 26 turns, we also annotated their turn performance (i.e., whether they successfully made a turn without the experimenter's help). The turns varied in terms of turning angle (16 regular, 10 slight turns), and width of a corridor after making a correct turn (M = 3.21m; SD = 1.37m; range 1.0–7.3m). We used paired t-test and Chi-square tests for independence testing throughout the analysis. We also collected audio recordings of participants' comments. The comments were transcribed, translated from Japanese to English, and analyzed on the basis of themes of interest following [16] (e.g., landmark, cognitive load).

5.2 Overall Localization Accuracy

The overall localization accuracy of NavCog3 was evaluated in terms of the Euclidian distance between a user's actual location and estimated location. We extracted 7,641 actual location points across the three routes from the participants' results. The average error rate was 1.65m (SD = 1.13m), which is comparable to the 1.47m accuracy measured offline prior to Study 1. It is also a better result than those from a former study (1.70m), which was conducted in a relatively small space (670m^2) [51].

The localization accuracy per participant is shown in Figure 5(a). Some participants had higher errors than others (P2, P8, and P9) due to issues with the localization engine. One issue was that the engine stopped working due to a problem during navigation while the user kept walking (P9). Another issue was that the heading estimation failed inside an elevator (P2 in route 2 and P8 in route 1).

5.3 Navigation Performance

All ten participants were able to complete the navigation tasks for all three routes. On average, the total task completion time was 990.5s (SD = 129.1s), and the total travel distance was 450.0m (SD = 32.4m) for all routes and participants.

We further examined the video data to analyze navigation errors, focusing on participants' turn performance. We labeled each turn performed by the participants as *success* if they were able to make the turn without requiring help from the experimenter, and *fail* otherwise. Of the total 260 turns, 221 turns were successful on the first attempt. Of the 39 missed turns, 22 were successful on the second attempt. An experimenter had to interrupt participants for the remaining 17 turns to prevent participants from entering shops (N = 8) and bumping into people (N = 2), because the participant seemed lost (N = 2) or due to system-related issues (N = 5). To show if our system can assist participants in navigation without requiring any help from others, we analyzed the success rate of the turns. The average rate per participant across all routes was 85.0% (SD = 10.6%) without any correction and 93.5% (SD = 5.8%) either with the help of the system's failure-safe guidance (N = 18) or by the participants themselves (N = 4). These results show that users were able to traverse the environment with very few navigation errors and that they (and the system) were able to recover from most errors that occurred. We further looked into factors that might have affected the turn-success rate. The results are presented below.

5.3.1 Localization Accuracy. With the hypothesis that the success rate would be lower as the localization accuracy decreases, we examined the distribution of localization accuracy with turn performance (*success* vs. *fail*). As shown in Figure 5(b), the localization error in distance does not seem to affect success rate if the localization error is less than 3m.

5.3.2 Width of Passage. We hypothesized that the turn success rate would be lower as the width of the turn gets narrower. While we were not able to find a correlation ($r = -0.016$, $p = 0.937$) between the success rate and width considering 25 turns (excluding one turn in the middle of the corridor), some participants commented that making a correct turn is more difficult for narrow widths.

For example, P3 mentioned that he would miss the timing of the turn if the corner was narrow. However, if he made too large of a turn into a wide passage, then he could correct himself later.

5.3.3 Turn Angle. Although the localization accuracy and average passage width were not significantly different between the types of turns (*regular* vs. *slight*), a Chi-square test for independence revealed that the turn types significantly impacted the turn success rates; $\chi^2_{(1)} = 6.245$; $p = 0.012$; $\phi = 0.155$. Of the 254 turns, 89.4% of 90-degree turns were successful while only 78.0% were successful for 45-degree turns, suggesting that 90-degree turns are easier to perform than 45-degree turns. For example, P4 stated that there were no problems when turning 90 degrees, but it was difficult to have her dog to make slight left turns. A follow-up analysis of the navigation trajectories computed with the system revealed a higher angular error in performing slight turns (22.5°–60°) than medium turns (60°–120°) [6].

5.4 Feedback on Semantic Features

Confirming prior findings that surrounding landmarks and POIs are useful for PVI [49], all participants reacted positively to the semantic features we provided during the navigation, especially for non-visual and silent landmarks such as elevators.

5.4.1 Tactile Paving and Obstacles. All participants considered tactile paving on the floor to be useful except for P4, a guide-dog user. She commented that since she walked with a dog, she did not care about tactile paving. The trend was similar for obstacles; eight participants reported that this information was useful for safety. Again, P4, who relies on her guide dog for avoiding obstacles, did not consider this information to be useful.

5.4.2 Elevators. We also provided information on the location of elevators as well as their button locations both inside and outside. While one participant (P9), who reported that he tends to rely on his sight whenever he can, did not consider this information to be useful, elevator-related information was appreciated by the rest of the participants. P8, for example, commented that, because elevators these days are mostly silent, she wishes to be notified when the door is open, since she does not know whether the elevator is open or closed unless she notices that people are getting on and off.

5.4.3 Points of Interest. All participants considered POI information to be useful. We noticed that some participants stated that POIs would increase the enjoyment of walking places ($N = 4$). P10, for example, stated that without NavCog3 she would have never walked into an interesting shop to buy something unless somehow encouraged to do so and that it was a pleasure to enjoy shopping from the information received. P7 also wished to use POIs information to enhance his spatial awareness. To be specific, he said that the information on nearby shops and vending-machine locations were informative as a clue for large crowds (e.g., people coming out from shops, standing in line, etc.) so he can be more cautious.

While POIs may increase the joy of walking around and improve spatial awareness, two participants specified that they would like to receive POI information only upon request or have two modes of navigation such as *exploration mode* where a user can receive detailed information about nearby POIs and *direction-only mode*, which does not provide any POI information.

5.4.4 Suggested Semantic Features. Besides the information about the existence and location of obstacles, four participants also wished to know the types of obstacles and the distance from them so that they can decide whether they should be alert. For example, P10 wished to receive more descriptive information, followed by an instruction that there was an obstacle on the right side; because she could not feel it, she was not sure if she had to walk carefully or differently from how she had been. However, P4, who travels with a guide dog, did not consider the distance information to be useful. Instead, she was more interested in learning about specific directions to the available target so that she can instruct her dog to walk towards it, such as a door or stairs.

The most commonly suggested landmarks by far were stairs and escalators, which participants considered being more accessible than elevators when transitioning between floors ($N = 7$). P8 stated that she did not like to use the elevator, because she prefers not to disturb other people. It would also be difficult to determine when to ride the elevator, when the door was open, and whether the floor buttons could be touched or not. She preferred not to use an elevator when alone, even if there was voice guidance, choosing to take the escalator when possible.

5.5 Overall Experience

The overall feedback was very positive for all participants. Most or all participants stated that the timing of the guidance was appropriate ($N = 9$) and the instructions were easy to understand ($N = 10$). In addition, nine participants reported that they would be able to walk alone in any unfamiliar place with NavCog3. Furthermore, all participants expressed their desire to use our system for navigating other places. We now summarize the factors that might have influenced the overall experience with NavCog3.

5.5.1 Spatial Mapping. As found in prior work [49], seven participants in our study reported that semantic features helped them understand the spatial layout of the environment as well as their orientation. For example, P3 mentioned that if he can recall the location from the name of a shop, it would be useful as a reference to other locations (e.g., "next to OO store"). P5 also commented, "While walking on tactile paving, [NavCog3] taught me the location of a stop and the name of a shop on the street. So, whether you are going out or returning, you can figure out where you are going by drawing a map inside your head."

5.5.2 Cognitive Load and Safety. While semantic features may help with the spatial understanding of the environment, seven participants reported that this information required extra cognitive load during navigation, which also introduced safety concerns. P6 mentioned, "If I become distracted by the system announcements, I cannot be aware of the surroundings. Therefore, I cannot avoid obstacles as I can usually do." P8 also reported that she only needs to listen to the guidance announcements with about 70–80% focus, so that she can concentrate on surrounding sounds and other people. However, P2 commented that she feels safer using NavCog3 as she felt like she can "see" the environment as she passes by semantic features using NavCog3. She said that the verbal description of the surroundings makes her feel like she can see the scenery, which reduces her fear while walking.

5.5.3 Hands-Free Interaction. As the need for hands-free interaction in mobile contexts was found to be important in prior studies for PVI [1, 85], half the participants expressed that they appreciate NavCog3 for enabling hands-free use. P2 said that it was easier to move, because she did not have to hold the phone in her hand and could concentrate on using her white cane. P9 specified that he preferred to put the phone in a pocket, especially when in elevators, as it is difficult to find and press the button while holding a phone. Interestingly, while they appreciated the hands-free feature, two participants mentioned that they would rather not use this feature, because it requires additional devices, a remote control and a headset in this case, other than the phone (P1), or to feel the haptic feedback better (P8).

6 STUDY 2: FREE ROUTE IN SHOPPING MALL

To investigate the usability of NavCog3 for supporting large-scale environments and to identify and confirm users' needs for indoor navigation reflecting their natural usages, we conducted a second study in which participants were asked to use NavCog3 for any destination of their choice in the large shopping mall during opening hours (in this article, we extend the results reported in [76] by adding usage log analysis).

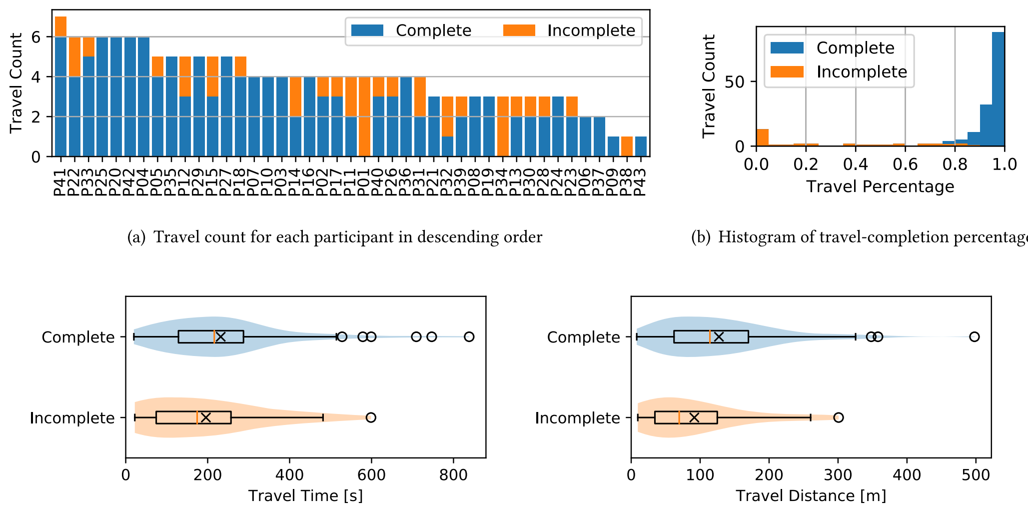
6.1 Method

6.1.1 Participants. Forty-three participants (22 males, 21 females) with visual impairments were recruited through a local Braille library for the study. The participants were almost equally split between people who were totally blind ($N = 21$) and people with low vision ($N = 22$). For the blind user group (**B**), 18 participants used a white cane and three had a guide dog. For the low-vision group (**LV**), 17 were cane walkers, two had a guide dog, and three did not use any mobility aid. We did not collect their age or smartphone ownership for this study.

6.1.2 Apparatus. The apparatus was very similar to that used in *Study 1.* All participants used the same version of NavCog3 and an iPhone 6. Again, bone conduction headphones were used to convey audio feedback. Unlike the participants in Study 1 who were asked to wear the phone in a waist bag, participants were allowed to carry the phone however they liked. For participants who wished to keep their phones in their pocket or a bag, we provided a remote control as an option. The app logged every event with a time stamp while the app was running. No video was recorded for any sessions in this study.

6.1.3 Procedure. We conducted three sessions in parallel, and an experimenter accompanied each participant to ensure their safety. The duration of a session was up to 60 minutes. After a brief explanation on how to use the app at the beginning, participants could freely select the destinations of their choice using the conversational interface. Unlike Study 1, no training was given to any participant. All participants were asked to navigate to at least three different destinations using NavCog3. At the end, we gave short questionnaires and conducted wrap-up interviews with the participants.

6.1.4 Data Analysis. We audio-recorded participants' comments, which were later transcribed and translated for further analysis. Comments were analyzed by themes of interest determined in Study 1 as well as new emergent themes, again following prior research [16]. We also collected



(a) Travel count for each participant in descending order
(b) Histogram of travel-completion percentage
(c) Distribution of travel duration (s)
(d) Distribution of travel distance (m)

Fig. 6. Usage statistics for Study 2. Box plots show 25th percentiles, medians and 75th percentiles. 'x' denotes means and 'o' denotes extremes.

usage logs from the app while participants were using it during the study, which includes the distance, duration, and selected destinations.

6.2 Overall Usage

Participants performed 170 travels in total. The total travelling time across participants was about 10.6h, where each participant traveled 224s on average ($SD = 147$) and the total travel distance was about 20,400m ($M = 120.0$, $SD = 82.5$ per user). As shown in Figure 6(a), all but five participants traveled a minimum of three routes, where each travelled four routes on average ($SD = 1.4$, range = 1 to 7).

They selected 45 unique destinations out of over 100 destinations. The destinations were diverse including shops ($N = 75$), restaurants ($N = 43$), public places such as information centers and metro stations ($N = 38$), and facilities such as restrooms and ATMs ($N = 14$).

As for further analysis regarding the travel distances, we grouped travels into two types based on the logs we collected from the app: *Complete* and *Incomplete*[2]. A travel was considered completed if the participant's last location was within 5 meters of the target destination, and incomplete otherwise.

Participants successfully completed 79.4% of the total travels (135 out of 170) where they traveled 127.3m ($SD = 83.1$) in 231.9s ($SD = 146.5$) on average. For the remaining 35 incomplete travels, participants traveled 91.6m ($SD = 74.4$) in 127.3s ($SD = 83.1$). They traveled 82.1% of their planned path on average; see Figure 6(b)) for the histogram of the traveled percentage for the planned path to the target destination.

To understand the reasons that led to incomplete travels, we investigated the impact of travel distance and time on travel completion. On the one hand, it is reasonable to think that incomplete routes result in shorter distances and times, as users did not complete the whole route. In our case, almost half of the incomplete travels ($N = 17$ out of 35) were canceled at the very beginning

[2]Note that logs contain estimated travel time and distance rather than exact values.

of travels (travel percentage ≤ 20%). On the other hand, longer routes contain more turns and instructions (increasing navigation difficulty), which may influence user performance (in errors and time) and lead them to quit. Further analysis is needed to understand the reasons that led participants to quit traveling (e.g., they could be trying out the app, be interrupted by others, or realizing they did not need further assistance, because they encountered a familiar place).

6.3 Feedback on Semantic Features

6.3.1 Landmarks of Interest. When asked about the perceived usefulness of landmark information (scaled from 1 to 5, where 5 is the best), the majority of participants showed a positive attitude by giving ratings higher than 3 ($N = 31$ out of 43), as confirmed in Study 1. Reflecting on Study 1's findings, some participants found some landmarks to be more or less useful than other landmarks. For example, a participant (B13) with a guide dog commented that the information on button location was helpful, since dogs cannot recognize buttons. However, LV5 did not consider the Braille availability of the buttons in the elevator to be useful, because the participant could not read Braille. The suggestions for other landmarks were more diverse. Some wished to know their current location and the location of chairs, stairs, moving objects, and crowds ($N = 2$ each). The most frequent requests were related to elevators ($N = 7$), which were found to be less preferred than stairs and escalators in Study 1. Again, participants wished to be informed of the current floor location ($N = 3$), arrival time ($N = 2$), heading direction (up/down), as well as which elevator has arrived if there are several ($N = 1$ each).

6.3.2 Points of Interest. In addition to seven participants commenting that they liked the nearby POI information as is, a greater number of participants ($N = 10$) expressed a different preference (e.g., level of details). Some wanted more compact announcements, especially when shops are on both sides ($N = 2$), while five participants wished to receive more detailed information on shops. For example, LV21 wanted to have a "window shopping mode" with more frequent updates on shops nearby. Other participants wished to turn POI information on and off ($N = 2$) or did not want any of this information ($N = 1$).

6.4 Overall Experience

6.4.1 Cognitive Load and Spatial Awareness. We asked participants if the amount of guidance the system provided was appropriate and did not cognitively overwhelm them during navigation. While most said the amount was appropriate, some participants showed concerns of losing spatial awareness of their surroundings ($N = 5$), which could lead to safety issues, as found in Study 1. For example, B11 commented that it was impossible to concentrate on the guidance and create a mental map at the same time and that it may be safer to cut down the amount of guidance so that one can also be more aware of their surroundings. One solution suggested by five participants was to provide a preview so that they could listen to the guidance prior to their itinerary and focus on their surroundings while walking.

6.4.2 Independent Navigation. As our primary goal is to support PVI to navigate independently in an unfamiliar environment, we asked participants if they think our app would allow them to walk alone in other places (i.e., *confidence to navigate by themselves*). Thirty-one participants were very positive about using the app elsewhere, not just in unfamiliar environments but also in familiar places at different times of the day (in daytime vs. at night). For example, B12 specified that "I would be able to navigate places without feeling anxious." Two participants also liked that NavCog3 would allow them to go out immediately without preparation (e.g., scheduling assistance). The places that participants wished to use NavCog3 for independent navigation included large facilities such as stations ($N = 6$), shopping malls, universities ($N = 3$ each), offices ($N = 2$), hospitals, airports, restrooms, and even at home (N = 1 each). While nine participants were neutral, three participants were more negative about using it in other places, usually due to particular requirements to the interface, such as: "I wanted to know where I am and the last navigation command", "I could not handle two-steps navigation announcement" (e.g., referring to instructions such as "Proceed 30 meters and turn right"), and "I prefer clock-wise directions rather than left or right." In addition, only one participant reported to be "disappointed with the accuracy" of the system, also causing a negative evaluation.

6.5 Other Suggestions for Improvement

Regarding distance, three participants mentioned that they would prefer the distance information to be described in a step count instead of in meters. For turns, four participants from Study 2 also commented that making slight turns was difficult, while two suggested using clock position. Interestingly, two participants suggested connecting NavCog3 with other services. B12, for example, wished that our app linked with other services such as Google Maps or shop websites.

7 STUDY 3: IN-THE-WILD USAGE DURING 4-DAY CONFERENCE FOR PVI

The shopping-mall studies showed NavCog3's ability to support indoor navigation assistance in a large-scale environment. Besides a fixed set of routes (*Study 1*), users were able to freely select a destination and navigate in the environment (*Study 2*). While *Study 2* presented much fewer constraints than *Study 1*, it was still a controlled environment with a constant presence of the researchers and a predetermined time for the tasks.

To fully understand the effectiveness of an indoor navigation system such as NavCog3, it is essential to study how PVI use it *in-the-wild*. Such analysis enables the capture of users' natural behaviors and performance while using the system by themselves, without the constraints that are found in the laboratory or controlled experiments [43, 61].

7.1 Method

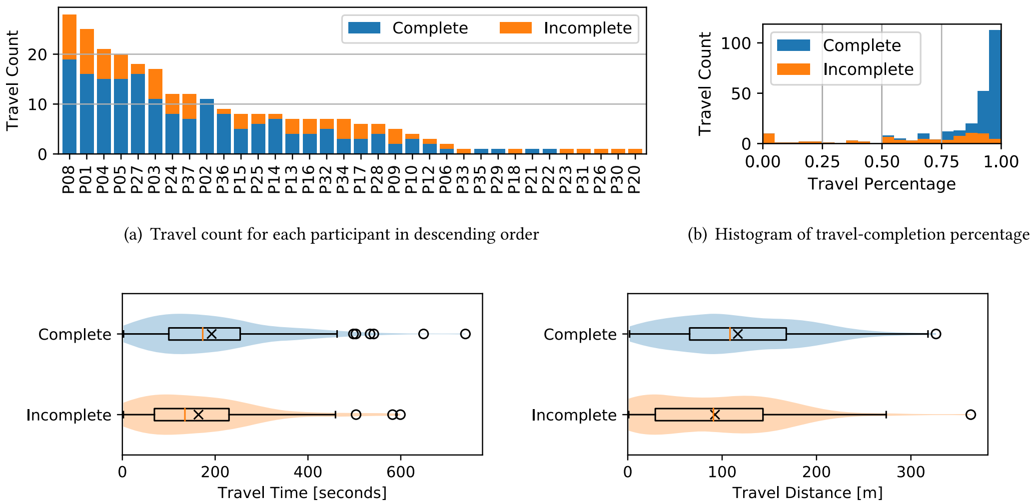
7.1.1 Data Collection. To study in-the-wild usage of NavCog3, we conducted an ad hoc installation of our system at the Hotel for a four-day conference (October 26–29, 2017) held by the Pennsylvania Council for the Blind [70]. By using our installation, the conference participants with visual impairments were allowed to navigate the venue by themselves, without being accompanied by researchers. NavCog3 was available in the *iOS App Store* for download and use by conference attendees. In-app instructions were provided to each user when they opened the app for the first time, but unlike the first two studies, we did not provide in-person training.

7.1.2 Data Analysis. For the usage analysis, we collected logs from the app during the conference. We also gathered overall experience with our system from users who agreed on providing their feedback. On the third day, we conducted a focus-group session with eight participants to collect subjective feedback after using NavCog3.

7.2 Overall Usage

During the conference, 37 unique users used the NavCog app and performed 280 travels in total for 34 different destinations. Their total traveling time across users was about 14.2h, where each user traveled 182.2s on average (SD = 131.0) and the total travel distance was about 30,200m (M = 108.0m, SD = 71.0 per user).

7.2.1 Usage per Travel. Of the 280 traveling records, users completed 179 travels, while 101 were incomplete. Like in Study 2, a travel was considered *complete* if the user ended his/her travel within 5m of the destination (see Figure 7(a)), and *incomplete* otherwise. We further analyzed incomplete



(a) Travel count for each participant in descending order

(b) Histogram of travel-completion percentage

(c) Distribution of travel duration (s)

(d) Distribution of travel distance (m)

Fig. 7. Usage statistics for Study 3. Box plots show 25th percentiles, medians, and 75th percentiles. “x” denotes means and “o” denotes extremes.

travels and found that 18 travels were canceled earlier (traveled less than 10% of the ideal travel distance), 17 travels were almost completed (traveled over 90% of the ideal travel distance), and 7 were stopped in front of the elevator connecting to guest rooms. Figure 7(b) shows the histogram of the travel percentage. The average travel percentage was 83.1%.

Users walked 116.7m (*SD* = 69.4) in 192.5s (*SD* = 133.0) on average for complete travels and 92.4m (*SD* = 71.3) in 164.1s (*SD* = 126.7) on average for incomplete travels (see Figures 7(c) and 7(d)). These results may be justified by users finishing their travels when they reached a familiar place closer to the destination or by stopping navigation either when interrupted by others or some navigation failure (which may justify the absence of differences in travel time). However, this requires further investigation to identify reasons for stopping navigation before reaching the destination.

7.2.2 Usage per Destination. We categorized the destinations into three types: conference rooms ($N = 129$), hotel facilities ($N = 112$), and guest rooms ($N = 39$). For the hotel facilities, the most popular destinations were restrooms ($N = 31$), followed by dog-relief areas ($N = 24$), restaurants, the lobby ($N = 15$ each), reception desks ($N = 10$), and the main entrance ($N = 6$). This can provide useful insight when installing a system when resources are limited (i.e., not being able to cover the entire venue) or when planning to provide higher accuracy for certain areas than others.

7.2.3 Usage per User. Each user travelled 7.6 routes on average (*SD* = 7.5, range = 1 to 28). As shown in Figure 7(a), 17 out of 37 users traveled less than or equal to 5 routes, 14 travelled between 6 and 15 routes, and the remaining 6 users traveled more than 15 routes. Based on the number of travels, we grouped users into *frequent* and *infrequent*. We defined a user as frequent if their number of travels was greater than 5 ($N = 20$ out of 37) and infrequent otherwise ($N = 17$). To check if there was any usage difference between the two groups, we compared their travel distances and times. The average traveled distance was 113.4m (*SD* = 71.1) for the frequent user group and 59.1m (*SD* = 44.3) for the infrequent group. This can be explained by the percentage of complete travels across users for each group; it was 67.1% for the frequent user group, while only 35.7% of total travels were completed for the infrequent user group.

To examine if there are more incomplete travels when starting to use our system, we also compared the usage between the first five travels of frequent users and all travels of infrequent users. We found that frequent users traveled significantly longer distances ($Z = 3.34$, $p < 0.001$), and times ($Z = 1.97$, $p = 0.049$). The average distance for the first five travels of the frequent user group was 108.6m ($SD = 75.1$) and the travel time was 183.0s ($SD = 149.2$), while the values for the infrequent user group were 59.1m ($SD = 44.3$) and 126.0s ($SD = 79.4$). A possible reason for this clear difference in usage behavior between frequent and infrequent users may be their individual characteristics such as familiarity with smartphones or navigation apps, as suggested in Reference [6]. However, further investigation would be needed to confirm such assumption.

7.3 Focus Group and Subjective Feedback

We conducted a focus-group session with eight attendees of the conference to collect qualitative feedback about their experience using our system during the conference. All but one participant (P1) owned an iPhone and used NavCog3 during the conference. P1 only used NavCog3 during a workshop session where people without an iPhone could learn and try the system. We also questioned other NavCog3 users on the last day of the conference about their experience with NavCog3. We transcribed the focus-group session and all feedback provided by users. We present the main findings of our thematic analysis below, which consists of four main themes: independent navigation, feedback on semantic features, cognitive load and spatial awareness, and errors and suggestions. We refer to the focus-group participants as P1–P8 and NavCog3 users questioned during the conference as *U1 and U2*.

7.3.1 Independent Navigation. All participants in the focus group provided very positive feedback about NavCog3, highlighting the independence, freedom, and confidence that they gained compared to their usual navigation in indoor environments. **Navigating independently** without asking for help was mentioned by most participants as the main advantage that NavCog brought to their experience during the conference, as mentioned by P2 and U1:

> **P2:** "*My experience with NavCog has been amazing, and my favorite part about it is the ability and confidence that I have to walk somewhere and if someone stops me and says, 'Do you need some help?', I can say 'No, I got it, thank you' and I'm on my way.*"
>
> **U1:** "*It's a great experience, it has given me more independence to get where I need to be without waiting for someone to come and assist me. It is just absolutely wonderful.*"

This ability to move independently without needing help from friends, family, or other people gave users a feeling of **freedom** as they were able to perform activities spontaneously, which is something they do not normally do. P3 and P4 exemplify this feeling:

> **P3:** "*[It] has been incredible for me, because it has been about 14 years since I lost my sight and quite honestly when going to a new hotel, and having that freedom to just walk around and get up spontaneously and go somewhere. I've missed that, and this has been the first time that I had that back... in quite some time, so I really appreciate it and want to see it in more places.*"
>
> **P4:** "*Tomorrow I plan to eat breakfast at the Bistro, so I know I can just get up and go there without asking anybody. It's really nice!*"

A major factor contributing to these feelings of freedom and independence is the **confidence** that users have that they will be able to get to the destination. While accuracy plays an important role, the feeling that the system supports error recovery or re-routing increases users' confidence that they will not get lost.

> **P1:** "*I had the phone in one hand, my cane in the other hand, I never walked with so much confidence in my life! I was even told to turn around when I did something wrong.*"

In contrast, a comment from U2 exemplifies the reduced freedom and independence to do such activities without navigation assistance, as it can result in getting lost or needing to require assistance from passersby:

> **U2:** "*For instance, this morning I went down to put the trash . . . I went down to the main lobby. And when I went back, I wasn't using NavCog, and I got disoriented, and I could not find the elevator. But had I had it ON, that wouldn't be a problem. So yeah . . . it's very useful and very helpful.*"

7.3.2 Feedback on Semantic Features. Confidence is often associated with the accuracy of the system but also to the semantic information about landmarks and POIs in the environment [68]. Several participants referred to the importance of the surrounding information to help them localize themselves and to be confident about their current position. For this purpose, tactile cues, such as floor changes, played a very important role and were appreciated by all participants of the focus group.

> **P1:** "*It said oh 'there's a sofa on your right.' I put my cane and sure enough, there's a sofa! The accuracy, and the feeling that someone was kind of . . . even though it's an inanimate object . . . but is with you; something is with you.*"
> **P2:** "*The other thing that I really liked was when it told you that the carpet was gonna change to tile, that is a great landmark*" [All other participants agreed]

Besides helping the participants localize themselves, the semantic information about the surrounding POIs was very important to increase the knowledge about the environment and give the users access to choices that are not always available given the absence of visual information. For instance, knowing where POIs such as restaurants or restrooms are located was found to be very relevant, but even simpler features such as couches or ATMs may also be useful, as mentioned by P3 and U2.

> **P3:** "*Yeah I liked that, knowing that the ATM is on your right or left . . . and all those landmarks.*"
> **U2:** "*We liked when it said there were couches on your left. Because sometimes you wanna go somewhere, but you wanna sit for a minute.*"

Some participants also referred to other scenarios that could benefit from additional information, such as providing the exact location of the important elements inside large restrooms or to receive additional information about a restaurant.

> **P3:** "*Going into the bathroom . . . would it be possible to include directions in there, like where are the stalls, where are the urinals? . . . And how much detail can you go with the landmarks, so for example, passing that ATM it would be great to know . . . and I know that it is accessible ATM, I believe . . . but things like that.*

And the restaurant on your left, would I be able to link to the restaurant at some point and get the menu?"

Moreover, a few participants stated that they would like to control the verbosity depending on their current knowledge of the environment, but others reported to prefer to listen to all information available even in places they are familiar with to confirm their location. In addition, U2 stated that users with different needs, such as guide dog or cane users, would probably require different information for navigation:

U2: "*But you should make a distinction . . . a difference between cane and guide-dog users. Because the cane users will want to know that the couch is on the left, but the dog will just go around it, so it is not a factor.*"

7.3.3 Cognitive Load and Spatial Awareness. Receiving continuous feedback was reported as very important to reduce users' cognitive load and stress levels compared to independent navigation without a navigation system. As part of orientation and mobility training and skills, PVI pay special attention to tactile and auditory cues to understand their current location and surroundings. By being informed about what is around them, users did not need to remember the information, but could still use it to confirm their location.

P6: "*My experience with NavCog is very positive and what I like about it is that it brings my stress level way down, because I don't have to remember as much. I can just set my destination, follow the directions and I'm gonna get where I wanna go and it will not gonna take me as much time.*"

Other important factors influencing not only the cognitive load but also users' awareness of the environment were the use of headphones and background noise. Almost all participants tried NavCog3 with a single earphone or the smartphone speaker. Such setting can affect users' spatial awareness, which is the reason bone conducting headsets are gaining more popularity among PVI, since they do not obstruct the user's ears [1]. However, even with such devices it can be challenging to listen to the instructions and be aware of the surroundings in areas with much noise, which usually happens in crowded environments.

P1: "*I noticed that in the areas where we were walking on tile . . . when there was echo or a lot of background noise . . . I didn't have a Bluetooth headset, but I didn't hear as well as I would like.*"

P5: "*With those [bone conducting headphones] it works really good . . . because that happened to me. The first time I really try to use it [NavCog] . . . they had some other convention right outside one of our sessions, and I walked out and no matter which way I went I was trying to listen . . . it was noisy, there were people everywhere and I walked into their meeting . . . plus you couldn't hear cause it was noisy and it was crazy and I thought, I need those bone conducting headphones.*"

7.3.4 Errors and Suggestions. Although user feedback was generally positive, being exposed to the app in-the-wild in an unconstrained scenario gave users the opportunity to further explore the app and experience different contexts, which would be difficult to find in a controlled laboratory environment. This resulted in a number of errors and suggestions to further improve the app.

Although it did not occur very often, a critical problem is regarding localization accuracy. Currently, localization accuracy can be affected by many factors, such as the user's speed. This happened mostly with guide-dog users who walk very fast, which may cause the instructions to be announced after the user has already passed by the corresponding point. Although the users ended up learning how to cope with this issue (by starting to prompt the dog to turn when they hear the *approaching* instructions), it is important to better adapt the timing of the instructions depending on the user's speed, as such coping mechanisms can also result in navigation errors [35].

> **P2:** "*The only issue that I experienced is that . . . I walk very very fast with my dog, so we sometimes would overshoot the turn before it told me to turn, so I had to make sure to adjust for that, but other than that it is a wonderful, wonderful app and I can't wait to see it in more public places.*"

Another factor that can affect localization accuracy is how users hold the phone, since it can impact the beacon-signal strength captured by the device. With our localization engine, it is assumed that users hold the smartphone with their hands, and for that reason having the phone in their pocket results in lower accuracy. Also, one user experienced lower localization accuracy when using an extended battery case for his smartphone but noticed a significant improvement without it.

> **P4:** "*One thing that we found out last night, by accident, was that some of us use battery cases for our phones and that obstructs the Bluetooth just a little bit. So it changes . . . I tried it without the case last night and I was pleasantly amazed.*"

Although holding the phone guarantees better accuracy for now, participants found it important to have other alternatives to keep one of their hands free (as the other is usually holding a white cane or guide-dog harness). Besides improving the accuracy independently of where the phone is kept, participants also discussed the option of keeping the smartphone in different locations (e.g., using a shoulder strap) or a Bluetooth receiver that could more consistently capture the beacon signals. A main requirement is that such a receiver should be small and possibly attached to something that they already use, such as the cane (P3), guide-dog harness (P2), or even combined with glasses or headphones (P3). However, participants were also concerned about the increased movement in such devices.

> **P3:** "*Something that would be right there on my hand or on my cane, I would be fine with that! Something I could place right on the handle. It's pointing in the direction I want to go. [P2: Or my dog's harness].* "
> **P5:** "*But the cane . . . this is Pittsburgh caning, I just . . . [P4 interrupts: sword-fighting!] That's exactly how I go, sword-fighting. I just go. That thing [referring to a receiver on her cane] would fly off. Defensive caning.*"
> **P3:** "*[After a reference to integrating a Bluetooth receiver with smart glasses] And with the temple of the glass with the bone conducting headsets . . . it's right there! [All agree]* "

Voice search was appreciated by most participants, as they were able to quickly set a destination and start the navigation. However, some participants referred to intelligibility difficulties, particularly in crowded, noisy environments. Moreover, even participants that often used the voice search, explored the list search before knowing the possible destinations.

> **U2:** "*NavCog did not seem to recognize my voice, but recognized a man's voice when he said it . . . but also it was noisy there.*"
> **P7:** "*I liked [that] with the voice recognition you don't have to hit the start button. [P2: The voice search was awesome].*"

A very problematic error was related to the loss of network connection, which happened in some areas of the hotel. Although this is a common situation, it is also true that losing access to the system can cause a temporary loss in mobility independence in unfamiliar environments when the user is relying exclusively on NavCog. Currently, NavCog only requires a network connection when starting the navigation (to calculate the route and use the voice search), which reduces the impact of this issue. Further options may involve storing specific routes in the device beforehand.

8 DISCUSSION

From our findings, we reflected on the implications for supporting large-scale indoor navigation for PVI.

8.1 Accuracy Requirements

Overall, most participants could complete their navigation tasks without any support both in the controlled and in-the-wild settings. The accuracy of NavCog3, evaluated before our three studies, was 1.47 and 1.68m in the shopping mall and in the hotel, respectively (see Section 4.4), which is comparable to one of the best accuracies for an indoor navigation system running on an off-the-shelf smartphone [51, 59]. The results indicated that the localization error during the user-study sessions was 1.65m on average, achieving almost the same performance as in the pre-evaluation in the shopping mall environment.

In the first study, we observed that the system could correct most of the errors by its fail-safe guidance system without improving accuracy (Section 5.3). These results indicate that our localization accuracy may be sufficient for supporting turn-by-turn navigation, as a well-designed fail-safe guidance system can complement the current accuracy limitations. However, it is still important to provide high accuracy for finding small targets such as an elevator button, doorknob, or water tap (referred to as the last meter problem [52]). Moreover, it is better to cap the maximum errors to prevent confusing users (Section 5.2), which may cause relatively higher cancellations of navigation in in-the-wild settings (Section 7.2). Still, the system can provide mobility independence of PVI with the current localization accuracy, including in ad hoc installations such as the one at the hotel.

8.2 Landmarks and Points of Interests

The majority of participants from all studies considered landmarks that our system provided to be useful (Sections 5.4, 6.3, and 7.3.2), confirming the importance of surrounding information for PVI to aid in their orientation and mobility [22, 49]. Information about elevator locations, for example, was appreciated, because they were silent, making them difficult to find. Information about floor-material changes, such as tile and carpet, was also useful to confirm their location and increase their confidence during navigation. Although our information was largely limited to constructed landmarks such as floor surface (ex. carpet, tile, and tactile paving) and doorways, some participants commented that non-tactile clues such as sounds from an escalator or changes in light were also useful, confirming that PVI use multiple sensory channels to understand the spatial layout of a physical environment [20]. In this regard, it would be useful to encode information on ambient sounds (e.g., escalator operating sounds) and other types of landmarks into the navigation instructions to inform users what clues to be aware of on the way.

Performing studies in two different environments allowed us to confirm that users' needs and preferences regarding semantic features were very similar across studies. Overall, participants wanted to be aware of POIs, such as shops, restaurants, or restrooms and of landmarks that can help them navigating in the environment or confirming their location.

However, we found that, in both environments, the landmarks and POIs that participants found *interesting* or *useful* varied depending on the user's mobility aid and objective (i.e., exploration vs. way-finding), as found in previous studies [4, 35, 84]. For example, the information about obstacles or tactile paving may be less appealing for a guide-dog user compared to white-cane users. Thus, the types of semantic features that a user wishes to be informed about should be personalized or customizable like in BlindSquare [15]. Personalization would also be helpful to reduce the amount of information that may result in cognitive overload while walking, which we will discuss further in the next section.

8.3 Augmenting Navigation Aids

Regardless of which primary mobility aid they use, all participants from the user studies wished to use our system as a complementary means of navigation assistance. In contrast, many participants complained about decreased attention to their surroundings due to the attention required for the navigation instructions (Sections 5.5, 6.4, 7.3.3.) For broad adoption, however, seamless augmentation of users' existing navigation skills would be important. To avoid disrupting users' attention, we should consider when designing a navigation assistant not to overwhelm the user's cognitive load, so that they can keep their routine during their way-finding. For example, the system can provide instructions when a user is standing still then offer as little guidance as possible when walking so as not to impede users' awareness of their surroundings. Moreover, if they can have a spatial layout of the environment in advance, they can focus on their safety, which is one of the major concerns for PVI [1, 85].

While indoor navigation for PVI is still not widespread, the studies presented in this article show that solutions such as NavCog3 can be deployed to support independent navigation for PVI. As new installations may follow (either for NavCog3 or similar systems), it is essential to take steps to promote awareness about these new technologies and educate PVI on how to interact with such systems and combine them with their primary navigation aids. During the development of NavCog3, we followed recommendations and had discussions with O&M trainers and local organizations for PVI. For our future work, we aim to explore how to integrate NavCog3 in current O&M training to teach users and prepare them to interact with navigation technologies. We also plan to investigate how virtual inspection methods [33, 87] can be used for this purpose and supplement current O&M training.

8.4 Support for Large-Scale Environments In-the-Wild

We were able to deploy an indoor navigation system in two different large-scale environments (Section 4.4) and demonstrated that the localization accuracy of NavCog3 is reliable in locations that combine various challenging aspects (e.g., open space, corridors with various widths, crowds, among others). Findings from *Study 2* and *Study 3* suggested that users were able to select diverse destinations and navigate to their desired POIs successfully (Sections 6.2 and 7.2). They also felt confident walking independently in unfamiliar environments (Sections 6.4 and 7.3.1).

Although our in-the-wild studies attempted to capture the natural usage of the system by allowing participants to navigate to any destination of their choice, a longitudinal field study may reveal additional implications for designing navigation systems for PVI. Based on the results and feedback of the third study, we are confident that NavCog3 can improve the quality of life of PVI by enabling independent mobility. Attendees' experience with the system was extremely positive, because their independence was not in the context of a user study, but of an activity that they were already joining. This allowed us to understand, in the context of a conference at a hotel, what were the main use cases for unconstrained navigation (in contrast to fixed tasks set up by researchers) using NavCog3. Besides learning that most attendees used NavCog3 multiple times during the conference, we also learned that they used it to navigate to many destinations, such as conference rooms, restrooms, dog-relief areas, restaurants, lobby, among others, while showing a consistent system performance. Overall, this gave them the sense of real mobility independence and of what they can find in other real-world scenarios and locations, enabling them to get anywhere in the environment without needing to ask for help.

While important, supporting indoor navigation in a large-scale environment is challenging. In addition to deployment costs, collecting landmark information for indoor areas is also time-consuming. We plan to investigate various approaches to help reduce installation costs such as physical crowdsourcing [30] or computer-based automation.

8.5 Limitations

We did not control participants' exposure to previous versions of our system, which might have affected navigation errors or subjective responses, although we did not find noticeable differences. Different results may have also been provided had we recruited a different population, especially for the findings related to guide-dog users, as the sample size was small. For future work, additional studies with a larger number of guide-dog users or other users with various needs (e.g., wheelchair users, foreigners) would be useful.

9 CONCLUSION

We designed and implemented NavCog3, a smartphone-based indoor navigation assistant for persons with visual impairments. Our system is characterized by turn-by-turn navigation with high localization accuracy and semantic features for spatial understanding. We deployed our system in a large shopping mall and evaluated it with over 50 participants to assess the feasibility of supporting indoor navigation for large-scale environments in the real world. Our findings suggest that NavCog3 can achieve high accuracy (1.65m error on average). Subjective responses confirmed that semantic features were perceived to be useful for building the spatial map of the environment as they navigate.

We also deployed the system in a hotel complex during a conference for people with visual impairments to validate the system's usability in the wild. The results indicate that with NavCog3, attendees were able to find and navigate to any destination of their choice without asking anybody for help. Our analysis and the feedback provided by the conference attendees suggest that our system can be adapted for use in other large-scale indoor environments and can provide opportunity for PVI to freely and confidently navigate in familiar or unfamiliar indoor environments by themselves.

REFERENCES

[1] Ali Abdolrahmani, Ravi Kuber, and Amy Hurst. 2016. An empirical investigation of the situationally induced impairments experienced by blind mobile device users. In *Proceedings of the 13th Web for All Conference (W4A'16)*. ACM, New York, NY. DOI: https://doi.org/10.1145/2899475.2899482

[2] Dragan Ahmetovic, Cole Gleason, Kris M. Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: Turn-by-turn smartphone navigation assistant for people with visual impairments or blindness. In *Proceedings of the 13th Web for All Conference (W4A'16)*. ACM, New York, NY. DOI: https://doi.org/10.1145/2899475.2899509

[3] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: A navigational cognitive assistant for the blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'16)*. ACM, New York, NY, 90–99. DOI: https://doi.org/10.1145/2935334.2935361

[4] Dragan Ahmetovic, João Guerreiro, Eshed Ohn-Bar, Kris Kitani, and Chieko Asakawa. 2019. Impact of expertise on interaction preferences for navigation assistance of visually impaired individuals. In *Proceedings of the 16th Web for All Conference (W4A'19)*.

[5] Dragan Ahmetovic, Masayuki Murata, Cole Gleason, Erin Brady, Hironobu Takagi, Kris Kitani, and Chieko Asakawa. 2017. Achieving practical and accurate indoor navigation for people with visual impairments. In *Proceedings of the 14th Web for All Conference on the Future of Accessible Work (W4A'17)*. ACM, New York, NY. DOI: https://doi.org/10.1145/3058555.3058560

[6] Dragan Ahmetovic, Uran Oh, Sergio Mascetti, and Chieko Asakawa. 2018. Turn right: Analysis of rotation errors in turn-by-turn navigation for individuals with visual impairments. In *Proceedings of the Conference on Computers and Accessibility*. ACM.

[7] Aplix. [n.d.]. MyBeacon general purpose type MB004 Ac—Aplix. Retrieved from http://business.aplix.co.jp/product/mybeacon/mb004ac/.

[8] Ariadne GPS. [n.d.]. Ariadne GPS—An innovative app for your mobility. Retrieved from http://www.ariadnegps.eu/.

[9] Saki Asakawa, João Guerreiro, Daisuke Sato, Hironobu Takagi, Dragan Ahmetovic, Desi Gonzalez, Kris Kitani, and Chieko Asakawa. 2019. An independent and interactive museum experience for blind people. In *Proceedings of the 16th Web for All Conference (W4A'19)*.

[10] Shiri Azenkot and Nicole B. Lee. 2013. Exploring the use of speech input by blind people on mobile devices. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'13)*. ACM, New York, NY. DOI: https://doi.org/10.1145/2513383.2513440

[11] J. M. Barlow, B. L. Bentzen, D. Sauerburger, and L. Franck. 2010. Teaching travel at complex intersections. *Found. Orient. Mobil.* 2 (2010), 352–419.

[12] Pauline M. Berry, Karen Myers, Tomás E. Uribe, and Neil Yorke-Smith. 2005. Task management under change and uncertainty. In *Proceedings of the Workshop on Constraint Solving under Change*.

[13] Alex Black, Jan E. Lovie-Kitchin, Russell L. Woods, Nicole Arnold, John Byrnes, and Jane Murrish. 1997. Mobility performance with retinitis pigmentosa. *Clin. Exper. Optom.* 80, 1 (1997), 1–12. DOI: https://doi.org/10.1111/j.1444-0938.1997.tb04841.x

[14] Bruce B. Blasch, William R. Wiener, and Richard L. Welsh. 1997. *Foundations of Orientation and Mobility*. American Foundation for the Blind.

[15] BlindSquare. [n.d.]. BlindSquare—Pioneering accessible navigation—indoors and outdoors. Retrieved from http://www.blindsquare.com/.

[16] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitat. Res. Psychol.* 3, 2 (2006), 77–101. DOI: https://doi.org/10.1191/1478088706qp063oa arXiv:https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa

[17] Yu-Chung Cheng, Yatin Chawathe, Anthony LaMarca, and John Krumm. 2005. Accuracy characterization for metropolitan-scale Wi-Fi localization. In *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services (MobiSys'05)*. ACM, New York, NY, 233–245. DOI: https://doi.org/10.1145/1067170.1067195

[18] Contact. [n.d.]. Double Battery Beacon. Retrieved from https://store.kontakt.io/our-products/30-double-battery-beacon.html.

[19] Ádám Csapó, György Wersényi, Hunor Nagy, and Tony Stockman. 2015. A survey of assistive technologies and applications for blind users on mobile platforms: A review and foundation for research. *J. Multimod. User Interfaces* 9, 4 (2015), 275–286.

[20] M. Bernardine Dias, Ermine A. Teves, George J. Zimmerman, Hend K. Gedawy, Sarah M. Belousov, and M. Bernardine Dias. 2015. Indoor navigation challenges for visually impaired people. In *Indoor Wayfinding and Navigation*. CRC Press, 141–164.

[21] Draft. 2017. Draft Development Specification for Spatial Network Model for Pedestrians. Ministry of Land, Infrastructure, Transport and Tourism, Japan. Retrieved from http://www.mlit.go.jp/common/001177505.pdf.

[22] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2012. The user as a sensor: Navigating users with visual impairments in indoor spaces using tactile landmarks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'12)*. ACM, New York, NY, 425–432. DOI: https://doi.org/10.1145/2207676.2207735

[23] Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2013. Indoor human navigation systems: A survey. *Interact. Comput.* 25, 1 (Jan. 2013), 21–33. DOI: https://doi.org/10.1093/iwc/iws010

[24] José Faria, Sérgio Lopes, Hugo Fernandes, Paulo Martins, and João Barroso. 2010. Electronic white cane for blind people navigation assistance. In *Proceedings of the World Automation Congress*. 1–7.

[25] German Flores and Roberto Manduchi. 2018. Easy return: An app for indoor backtracking assistance. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 17.

[26] Giovanni Fusco and James M. Coughlan. 2018. Indoor localization using computer vision and visual-inertial odometry. In *Proceedings of the International Conference on Computers Helping People with Special Needs*. Springer, 86–93.

[27] Nicholas A. Giudice. 2018. 15. Navigating without vision: Principles of blind spatial cognition. *Handbook Behav. Cogn. Geogr.* (2018). Edward Elgar Publishing, 260–290.

[28] Nicholas A. Giudice and Gordon E. Legge. 2008. *Blind Navigation and the Role of Technology.* John Wiley & Sons, Inc., 479–500. DOI: https://doi.org/10.1002/9780470379424.ch25
[29] Cole Gleason, Dragan Ahmetovic, Saiph Savage, Carlos Toxtli, Carl Posthuma, Chieko Asakawa, Kris M. Kitani, and Jeffrey P. Bigham. 2018. Crowdsourcing the installation and maintenance of indoor localization infrastructure to support blind navigation. *Proc. ACM Interact. Mob. Wearable Ubiq. Technol.* 2, 1, Article 9 (Mar. 2018). DOI: https://doi.org/10.1145/3191741
[30] Cole Gleason, Dragan Ahmetovic, Carlos Toxtli, Saiph Savage, Jeffrey P. Bigham, and Chieko Asakawa. 2017. LuzDeploy: A collective action system for installing navigation infrastructure for blind people. In *Proceedings of the 14th Web for All Conference on the Future of Accessible Work (W4A'17).* ACM, New York, NY. DOI: https://doi.org/10.1145/3058555.3058585
[31] Cole Gleason, Alexander J. Fiannaca, Melanie Kneisel, Edward Cutrell, and Meredith Ringel Morris. 2018. FootNotes: Geo-referenced audio annotations for nonvisual exploration. *Proc. ACM Interact. Mob. Wearable Ubiq. Technol.* 2, 3 (2018), 109.
[32] Google Maps. [n.d.]. Google Maps—See What We're up to. Retrieved from https://www.google.com/intl/en/maps/about/.
[33] João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. 2017. Virtual navigation for blind people: Building sequential representations of the real-world. In *The Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility.* ACM.
[34] João Guerreiro, Dragan Ahmetovic, Daisuke Sato, Kris Kitani, and Chieko Asakawa. 2019. Airport accessibility and navigation assistance for people with visual impairments. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI'19).*
[35] João Guerreiro, Eshed Ohn-Bar, Dragan Ahmetovic, Kris Kitani, and Chieko Asakawa. 2018. How context and user behavior affect indoor navigation assistance for blind people. In *Proceedings of the Internet of Accessible Things (W4A'18).* ACM, New York, NY. DOI: https://doi.org/10.1145/3192714.3192829
[36] Simon Harper, Stephen Pettitt, and Carole Goble. 2003. Sentinel: Towards an ambient mobility network. *Disabil. Rehab.* 25, 16 (2003), 940–948. DOI: https://doi.org/10.1080/0963828031000090533 arXiv:https://doi.org/10.1080/0963828031000090533 PMID: 12857586.
[37] Sumi Helal, Carlos Giraldo, Youssef Kaddoura, Choonhwa Lee, Hicham El Zabadani, and William Mann. 2003. Smart phone-based cognitive assistant. In *Proceedings of the 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications (UbiHealth'03).*
[38] Andreas Hub, Joachim Diepstraten, and Thomas Ertl. 2004. Design and development of an indoor navigation and object identification system for the blind. In *Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'04).* ACM, New York, NY, 147–152. DOI: https://doi.org/10.1145/1028630.1028657
[39] HULOP0. [n.d.]. Human-scale Localization Platform (HULOP) - GitHub. Retrieved from https://github.com/hulop.
[40] HULOP1. [n.d.]. hulop/NavCogIOSv3—GitHub, NavCog version 3 for iOS. Retrieved from https://github.com/hulop/NavCogIOSv3.
[41] HULOP2. [n.d.]. hulop/blelocpp—GitHub, BLE localization C++ library. Retrieved from https://github.com/hulop/blelocpp.
[42] HULOP3. [n.d.]. hulop/MapService—GitHub, map editing and routing server app. Retrieved from https://github.com/hulop/MapService.
[43] Amy Hurst, Jennifer Mankoff, and Scott E. Hudson. 2008. Understanding pointing problems in real world computing environments. In *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility.* ACM, 43–50.
[44] Tatsuya Ishihara, Jayakorn Vongkulbhisal, Kris M. Kitani, and Chieko Asakawa. 2017. Beacon-guided structure from motion for smartphone-based navigation. In *Proceedings of the Winter Conference on the Applications of Computer Vision.* IEEE.
[45] Suk Hoon Jung, Byeong-Cheol Moon, and Dongsoo Han. 2017. Performance evaluation of radio map construction methods for Wi-Fi positioning systems. *Trans. Intell. Transport. Sys.* 18, 4 (Apr. 2017), 880–889. DOI: https://doi.org/10.1109/TITS.2016.2594479
[46] Hernisa Kacorri, Sergio Mascetti, Andrea Gerino, Dragan Ahmetovic, Valeria Alampi, Hironobu Takagi, and Chieko Asakawa. 2018. Insights on assistive orientation and mobility of people with visual impairment based on large-scale longitudinal data. *ACM Trans. Access. Comput.* 11, 1, Article 5 (2018), 28 pages.
[47] Hernisa Kacorri, Sergio Mascetti, Andrea Gerino, Dragan Ahmetovic, Hironobu Takagi, and Chieko Asakawa. 2016. Supporting orientation of people with visual impairment: Analysis of large scale usage data. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility.* ACM.
[48] Hernisa Kacorri, Eshed Ohn-Bar, Kris M. Kitani, and Chieko Asakawa. 2018. Environmental factors in indoor navigation based on real-world trajectories of blind users. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI'18).* ACM, New York, NY. DOI: https://doi.org/10.1145/3173574.3173630

[49] Slim Kammoun, Florian Dramas, Bernard Oriolaand, and Christophe Jouffrais. 2010. Route selection algorithm for blind pedestrian. In *Proceedings of the International Conference on Control Automation and Systems (ICCAS'10)*. IEEE, 2223–2228.

[50] Shaun K. Kane, Chandrika Jayant, Jacob O. Wobbrock, and Richard E. Ladner. 2009. Freedom to Roam: A study of mobile device adoption and accessibility for people with visual and motor disabilities. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility (Assets'09)*. ACM, New York, NY, 115–122. DOI: https://doi.org/10.1145/1639642.1639663

[51] Dierna Giovanni Luca and Macha Alberto. 2016. Towards accurate indoor localization using iBeacons, fingerprinting and particle filtering. In *Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN)*.

[52] Roberto Manduchi and James M. Coughlan. 2014. The last meter: Blind visual guidance to a target. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*. ACM, New York, NY, 3113–3122. DOI: https://doi.org/10.1145/2556288.2557328

[53] Roberto Manduchi and Sri Kurniawan. 2011. Mobility-related accidents experienced by people with visual impairment. *AER J.: Res. Pract. Visual Impair. Blind.* 4, 2 (2011), 44–54.

[54] Roberto Manduchi, Sri Kurniawan, and Homayoun Bagherinia. 2010. Blind guidance using mobile computer vision: A usability study. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'10)*. ACM, New York, NY, 241–242. DOI: https://doi.org/10.1145/1878803.1878851

[55] Sergio Mascetti, Dragan Ahmetovic, Andrea Gerino, and Cristian Bernareggi. 2016. Zebrarecognizer: Pedestrian crossing recognition for people with visual impairment or blindness. *Pattern Recogn.* 60 (2016), 405–419.

[56] Sergio Mascetti, Dragan Ahmetovic, Andrea Gerino, Cristian Bernareggi, Mario Busso, and Alessandro Rizzi. 2016. Robust traffic lights detection on mobile devices for pedestrians with visual impairment. *Comput. Vision Image Understand.* 148 (2016), 123–135.

[57] Sergio Mascetti, Lorenzo Picinali, Andrea Gerino, Dragan Ahmetovic, and Cristian Bernareggi. 2016. Sonification of guidance data during road crossing for people with visual impairments or blindness. *Int. J. Hum.-Comput. Studies* 85 (2016), 16–26.

[58] Mei Miao, Martin Spindler, and Gerhard Weber. 2011. Requirements of indoor navigation system from blind users. In *Proceedings of the Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 673–679.

[59] Microsoft. [n.d.]. Microsoft Indoor Localization Competition—IPSN 2016. Retrieved from https://www.microsoft.com/en-us/research/event/microsoft-indoor-localization-competition-ipsn-2016/.

[60] Microsoft. [n.d.]. Microsoft Soundscape, a map delivered in 3D sound. Retrieved from https://www.microsoft.com/en-us/research/product/soundscape/.

[61] Kyle Montague, Hugo Nicolau, and Vicki L. Hanson. 2014. Motor-impaired touchscreen interactions in the wild. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 123–130.

[62] Masayuki Murata, Dragan Ahmetovic, Daisuke Sato, Hironobu Takagi, Kris M. Kitani, and Chieko Asakawa. 2018. Smartphone-based indoor localization for blind navigation across building complexes. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom'18)*. IEEE, 254–263.

[63] NavCog. [n.d.]. NavCog on the App Store. Retrieved from https://itunes.apple.com/app/navcog/id1042163426?mt=8.

[64] Hugo Nicolau, Joaquim Jorge, and Tiago Guerreiro. 2009. Blobby: How to guide a blind person. In *Proceedings of the CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 3601–3606.

[65] Uran Oh and Leah Findlater. 2014. Design of and subjective response to on-body input for people with visual impairments. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS'14)*. ACM, New York, NY, 115–122. DOI: https://doi.org/10.1145/2661334.2661376

[66] Eshed Ohn-Bar, João Guerreiro, Kris Kitani, and Chieko Asakawa. 2018. Variability in reactions to instructional guidance during smartphone-based assisted navigation of blind users. *Proc. ACM Interact. Mob. Wearable Ubiq. Technol.* 2, 3 (2018), 131.

[67] Eshed Ohn-Bar, João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. 2018. Modeling expertise in assistive navigation interfaces for blind people. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI'18)*. ACM, New York, NY, 403–407. DOI: https://doi.org/10.1145/3172944.3173008

[68] Rajchandar Padmanaban and Jakub Krukar. 2017. Increasing the density of local landmarks in wayfinding instructions for the visually impaired. In *Progress in Location-Based Services 2016*. Springer, 131–150.

[69] Sabrina A. Panëels, Adriana Olmos, Jeffrey R. Blum, and Jeremy R. Cooperstock. 2013. Listen to it yourself!: Evaluating usability of what's around me? For the blind. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2107–2116.

[70] PCB. 2017. PCB Conference Schedule—Pennsylvania Council of the Blind. Retrieved from http://pcb1.org/2017-pcb-conference-schedule/.

[71] J. Eduardo Pérez, Myriam Arrue, Masatomo Kobayashi, Hironobu Takagi, and Chieko Asakawa. 2017. Assessment of semantic taxonomies for blind indoor navigation based on a shopping center use case. In *Proceedings of the 14th*

Web for All Conference on the Future of Accessible Work (W4A'17). ACM, New York, NY. DOI: https://doi.org/10.1145/3058555.3058575

[72] Pablo-Alejandro Quinones, Tammy Greene, Rayoung Yang, and Mark Newman. 2011. Supporting visually impaired navigation: A needs-finding study. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems (CHI EA'11)*. ACM, New York, NY, 1645–1650. DOI: https://doi.org/10.1145/1979742.1979822

[73] Lisa Ran, Sumi Helal, and Steve Moore. 2004. Drishti: An integrated indoor/outdoor blind navigation system and service. In *Proceedings of the 2nd IEEE Annual Conference on Pervasive Computing and Communications (PerCom'04)*. IEEE, 23–30.

[74] Timothy H. Riehle, P. Lichter, and Nicholas A. Giudice. 2008. An indoor navigation system to support the visually impaired. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS'08)*. IEEE, 4435–4438.

[75] David A. Ross. 2001. Implementing assistive technology on wearable computers. *IEEE Intell. Syst.* 16, 3 (2001), 47–53.

[76] Daisuke Sato, Uran Oh, Kakuya Naito, Hironobu Takagi, Kris Kitani, and Chieko Asakawa. 2017. NavCog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'17)*. ACM, New York, NY, 270–279. DOI: https://doi.org/10.1145/3132525.3132535

[77] SFO. 2014. SFO app for visually impaired navigation by indoo.rs. Retrieved from https://indoo.rs/sfo/.

[78] Speech | Apple Developer Documentation [n.d.]. Speech | Apple Developer Documentation. Retrieved from https://developer.apple.com/documentation/speech.

[79] P. Strumillo, M. Bujacz, P. Baranski, P. Skulimowski, P. Korbel, M. Owczarek, K. Tomalczyk, A. Moldoveanu, and R. Unnthorsson. 2018. Different approaches to aiding blind persons in mobility and navigation in the "Naviton" and "Sound of Vision" projects. In *Mobility of Visually Impaired People*. Springer, 435–468.

[80] Frank van Diggelen and Per Enge. 2015. The world's first GPS MOOC and worldwide laboratory using smartphones. In *Proceedings of the 28th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS'15)*. 361–369.

[81] Ramiro Velázquez. 2010. *Wearable Assistive Devices for the Blind*. Springer, Berlin, 331–349. DOI: https://doi.org/10.1007/978-3-642-15687-8_17

[82] Watson Assistant. [n.d.]. Watson Assistant—Formerly Watson Conversation. Retrieved from https://www.ibm.com/watson/developercloud/conversation.html.

[83] Wayfinder. 2014. Wayfinder app helps the blind navigate the Tube. Retrieved from http://www.wired.co.uk/article/wayfindr-app.

[84] Michele A. Williams, Amy Hurst, and Shaun K. Kane. 2013. "Pray before you step out": Describing personal and situational blind navigation behaviors. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'13)*. ACM, New York, NY. DOI: https://doi.org/10.1145/2513383.2513449

[85] Hanlu Ye, Meethu Malu, Uran Oh, and Leah Findlater. 2014. Current and future mobile and wearable device use by people with visual impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*. ACM, New York, NY, 3123–3132. DOI: https://doi.org/10.1145/2556288.2557085

[86] Da Zhang, Feng Xia, Zhuo Yang, Lin Yao, and Wenhong Zhao. 2010. Localization technologies for indoor human tracking. In *Proceedings of the 5th International Conference on Future Information Technology (FutureTech'10)*. IEEE, 1–6.

[87] Yuhang Zhao, Cynthia L. Bennett, Hrvoje Benko, Edward Cutrell, Christian Holz, Meredith Ringel Morris, and Mike Sinclair. 2018. Enabling people with visual impairments to navigate virtual reality with a haptic and auditory cane simulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM.

Received May 2018; revised March 2019; accepted June 2019