# Virtual Navigation for Blind People: Building Sequential Representations of the Real-World

João Guerreiro[1], Dragan Ahmetovic[1], Kris M. Kitani[1], Chieko Asakawa[1,2]
[1]Carnegie Mellon University, [2]IBM Research - Tokyo
{jpvguerreiro, dragan1}@cmu.edu, {kkitani, chiekoa}@cs.cmu.edu

## ABSTRACT

When preparing to visit new locations, sighted people often look at maps to build an *a priori* mental representation of the environment as a sequence of step-by-step actions and points of interest (POIs), e.g., turn right after the coffee shop. Based on this observation, we would like to understand if building the same type of sequential representation, prior to navigating in a new location, is helpful for people with visual impairments (VI). In particular, our goal is to understand how the simultaneous interplay between turn-by-turn navigation instructions and the relevant POIs in the route can aid the creation of a memorable sequential representation of the world. To this end, we present two smartphone-based virtual navigation interfaces: VirtualLeap, which allows the user to jump through a sequence of street intersection labels, turn-by-turn instructions and POIs along the route; and VirtualWalk, which simulates variable speed step-by-step walking using audio effects, whilst conveying similar route information. In a user study with 14 VI participants, most were able to create and maintain an accurate mental representation of both the sequential structure of the route and the approximate locations of the POIs. While both virtual navigation modalities resulted in similar spatial understanding, results suggests that each method is useful in different interaction contexts.

## CCS Concepts

•**Human-centered computing** → **Accessibility;** *Interaction techniques;*

## Keywords

Assistive Technology; Blind Navigation; Cognitive Mapping; Virtual Navigation; Orientation and Mobility.

## 1. INTRODUCTION

Before visiting a new location, sighted people often explore maps to understand the sequence of actions they need to follow to reach a destination. Both printed and digital maps visually convey information about POIs, which are used to support the construction of a person's mental representation of a route (e.g., turn left **after the post office**). This sequential representation of information is also very common when people provide verbal descriptions to guide someone to a destination. The reduction of a complex physical 3D space into a compact sequential representation is useful because it helps us to store navigational information efficiently. Previous research shows that the same sequential route-based instructions are also used by visually impaired (VI) people when building a mental model of the environment [21, 32]. However, extrapolating a sequential model of the environment requires a way to simultaneously access both the information about the route and POIs along it.

Our goal is to understand how we can provide simultaneous access to turn-by-turn navigation instructions and the POIs in the route to aid the creation of a memorable sequential representation of the world. Moreover we investigate how to use currently available technologies to provide the same kind of integrated knowledge non-visually.

We present a smartphone-based virtual navigation system that supports turn-by-turn navigation and announces POIs at their exact location and orientation along the route (e.g. "Starbucks is on your right"). We describe the design process of the virtual navigation app, including two navigation modalities to simulate route navigation. 1) VirtualLeap allows the user to jump through a sequence of intersections, turn-by-turn instructions and POIs of the route. 2) VirtualWalk simulates step-by-step walking at variable speeds using audio effects and also conveys a sequence of turn-by-turn instructions and POI information.

We performed a user study with 14 participants, using route reconstruction and real-world exposure to both assess users' ability to build sequential representations of the environment and understand the relative benefits and disadvantages of VirtualLeap and VirtualWalk. We gathered user feedback using a think-aloud protocol while using the full system and a semi-structured interview afterwards to inquire about the perceived usefulness, acceptance, and requirements of smartphone-based virtual navigation.

Our results show that the majority of the participants were able to build an accurate sequential representation of the route structure, while a subset was able to locate (all or most of) the POIs in the respective street block. Conversely, participants had difficulty estimating the length of the street blocks. The performance of the two interaction modalities was very similar and participants foresee benefits



*ASSETS '17, October 29-November 1, 2017, Baltimore, MD, USA*



ISBN 978-1-4503-4926-0/17/10...$15.00

DOI: https://doi.org/10.1145/3132525.3132545

in combining them depending on their goal. *VirtualLeap* was perceived as a faster way to get an overview of the route and relevant POIs (and explore their attributes). In contrast, they preferred *VirtualWalk* for building an accurate sequential representation of the environment.

# 2. RELATED WORK

VI people's Orientation and Mobility (O&M) depends on non-visual sensing for cognitive mapping of the environment [33]. A cognitive map is a comprehensive spatial representation, which contains information about routes, connections, directions, distances and landmarks [22]. Assistive technologies aim to improve O&M of VI people by supplementing their sensing capabilities or by providing accessible sources of spatial information.

## 2.1 In-Situ Navigation and Exploration

Assistive navigation caters to the users' need to move independently and to acquire knowledge about the surroundings [3]. In order to support navigation, several solutions convey non-visual turn-by-turn instructions similarly to mainstream apps such as *Google Maps*. These solutions provide a functional understanding of the spatial structure of the environment as a sequence of distances and turns (e.g. [2, 18]). For instance, NavCog [2] is an indoor navigation assistant that relies on Bluetooth beacons installed in the environment to provide accurate localization and turn-by-turn instructions. NAVIG [13] uses GNSS satellite data for localization and camera for near visual sensing, enabling landmark exploration using sonified audio cues.

Other systems and research projects focus on exploring the environment, either by augmenting the users' sensory range through visual detection of features of interest (e.g. crosswalks [20]), or by providing information about nearby landmarks and POIs (e.g. shops), encouraging serendipitous discovery of relevant locations (e.g. [1, 11, 34]). Blum et al. [4] augment this approach by rendering nearby POI through spatialized audio. These approaches provide radial orientation and euclidean distance to surrounding POIs, which conveys the relative disposition of the elements, but no knowledge about the structure of the environment (e.g., buildings and streets between the user and the POI).

These systems are very useful to support VI people's O&M. However, increasing the amount of information conveyed while navigating can be overwhelming to users [26]. For this reason, we combine both turn-by-turn and POI information beforehand, in a controlled environment, in order to increase users' confidence and knowledge about the environment.

## 2.2 Map Exploration and Virtual Simulation

Apart from real-world exposure, spatial knowledge can also be acquired indirectly [22], for instance through language, maps or virtual simulation. Tactile maps and 3-D models provide spatial information through haptic exploration [33]. However, they require special printers, time and effort to design and print, and they have a fixed resolution which makes it difficult to present detailed and multi-scale information [30]. Approaches relying on touchscreen interactive maps try to overcome these limitations, enabling the user to explore the screen while receiving auditory (and sometimes tactile) feedback. Interesting solutions include the use of sonification to convey geometric information [31]; two-handed map exploration on tabletops with simultaneous audio feedback [10]; touchscreens with raised-line overlays [5]; and software overlays that ease the access to spatially structured elements in the map [12]. However, in order to provide detailed spatial information, these approaches often require large touchscreens, which are less widespread and mobile, and therefore more difficult to use on a daily basis.

Other approaches rely on virtual navigation through egocentric exploration, usually relying on 3D audio to increase VI people's spatial knowledge [7, 28]. However, these approaches require quality audio equipment, silence, and are often restricted to specific scenarios that require building comprehensive, specialized virtual environments.

In addition to exploration during mobility, the smartphone app Blindsquare [1] allows users to simulate a static real-world location and examine nearby POIs. SpaceSense [35] extends this approach by enabling virtual route navigation from the simulated location to a selected POI, through sequential turn-by-turn instructions. After each instruction, a vibrotactile matrix addon on the back of the phone provides feedback indicating the direction of and distance to the destination, and to bookmarked POIs near the route. Both BlindSquare and SpaceSense convey the POIs location relative to the user's current location regardless of the user's route (e.g. the POI could be in an adjacent/parallel street). In contrast, we integrate the POIs as elements in the sequential representation of the route with the goal to create a memorable representation of the real-world. This difference, together with SpaceSense need for special hardware and BlindSquare not supporting route navigation makes it difficult to compare them with our two modalities.

# 3. ENABLING VIRTUAL NAVIGATION

Our virtual navigation system is built on top of NavCog, an open-source turn-by-turn smartphone navigation assistant for VI people [2]. We extend both NavCog's map server, which stores information about the environment, and its iPhone app in order to simulate the user's movement and convey guidance information.

## 3.1 Populate the Map

The NavCog system relies on a web-based editor to manually insert environment maps and POI. A map consists of a graph-based structure where the edges are walkable areas (corridors or streets), and the nodes are decision points such as intersections or orientation changes. A major concern for the simulation of real-world routes was the ability to automate the process of adding such information for large scale areas without human intervention. The map structure is extracted from the OpenStreetMap (OSM) street layout to automatically generate route edges and nodes. To populate the map with POIs, we combine information from Yelp and FourSquare POI retrieval geo-location services.

We extract POIs from a desired area along with several attributes: name, location, category, description, address, average rating, opening hours, price range, FourSquare tips and a Yelp review. We consider the functional position of each POI in the route, instead of relying on their euclidean distance and radial orientation, often used by existing assistive navigation apps [1, 11]. To link a POI with its respective street, we project the POI coordinates (from OSM, Google Maps, Yelp or FourSquare) on the closest edge having the same street name. When virtually navigating that street, the user is informed of the POI when passing by it.

## 3.2 Virtual Navigation Interface

The virtual navigation interface was adapted from the most recent version of NavCog [2] through an iterative design process. A preliminary study with three blind subjects aimed at understanding the perceived usability of our initial implementation and how could we improve its functionality in order to better support virtual navigation. We describe the resulting application, which implements two navigation simulation modalities that integrate turn-by-turn instructions and the exploration of POIs and their attributes.

### *3.2.1 Turn-by-Turn Instructions*

NavCog turn-by-turn instructions are grounded on previous research on how to guide VI people (e.g. [24]) and can be divided in three types of messages: distance announcements, action instructions and POI descriptions [2]. Virtual navigation does not use explicit distance announcements as presented in NavCog, since (implicit or explicit) distance information is conveyed by the simulation modalities. It does however provide action instructions in the proximity of decision points (e.g., intersections). When reaching an intersection, the system announces its name, for instance "*Fifth Avenue with Sixth Street*", the action required (e.g. "turn right"), and (after the turn) the distance to the next intersection/turn. When approaching a POI, the system announces its name and its relative position depending on the user's virtual orientation (e.g. "*an obstacle is in front of you*" or "*Starbucks is on your right*"). All interaction with the system in Virtual NavCog relies on gestures.

### *3.2.2 Virtual Navigation Modalities*

We implemented two different modalities for navigation simulation: **VirtualWalk** and **VirtualLeap**. VirtualWalk is based on literature that suggests that virtual exploration may be used to build a mental representation [16] of an environment. In this method, we try to mimic real world navigation by enabling the user to walk and turn in the environment, while being alerted by the surrounding POIs and required actions. The VirtualLeap approach resembles a sequence of instructions similar to what is found in *Google Maps*, but also containing information about the POIs. Herein, the user is able to jump through the relevant points of the route (POIs and intersections) while being informed about the distance the user would need to travel.

**VirtualWalk**. This approach uses gyroscope data[1] to detect walking and turning gestures (See Figure 1). Tilting the phone down enables walking (Figure 1 a)), with a speed ranging between 0.3 and 0.7 seconds per step. The speed is based on the amount of tilt (using the gyroscope pitch value, with a window of approximately 25 degrees). These walking speeds were obtained empirically based on the feedback from the preliminary study, and answer to the users' need to navigate faster, but not too fast to make sure they can process all information. The tilt range is limited to avoid angles uncomfortable to users, and the system allows users to calibrate their resting pitch inclination.

When reaching a location that requires a turn, the user needs to rotate the phone (detected by the yaw value from the gyroscope) to the announced direction (Figure 1 b)), receiving a confirmation sound when the rotation reaches the required angle. This intends to mimic actual turning

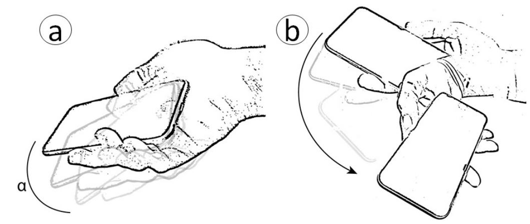


**Figure 1: VirtualWalk gestures. a) Walking speed is modified by tilting the phone. b) The user turns by rotating the phone in the required direction.**

that occurs in physical locomotion, frequently used in immersive Virtual Environments, which requires less cognitive processing to update the mental representation of the space, when comparing to imagined rotation/turning [29, 15]. After turning, users may rotate the phone back to its previous orientation or maintain its orientation.

Each virtual step is 0.7m long, corresponding to an average human step [23]. To increase the feeling of presence in the virtual environment [25] and to provide feedback on how much the user is walking, we provide auditory feedback of steps. The step length is fixed, but the speed depends on the tilting angle. When the user passes by a POI, it is announced together with its relative location. The user can tilt the phone up instead of down in order to walk backwards while keeping the same orientation. This is useful, for instance, to re-explore a previous POI.

The pilot study supported the addition of a command where users can go through a route at a faster speed, focusing only on the route segments and turns. A two-finger double tap starts an automatic walking mode (at the speed of 0.2 seconds per step), where the user gets auditory feedback about the steps and is required to turn when reaching an intersection but is not alerted about the POIs. The participants' comments are in line with the fact that knowing the environment is a dynamic process that is iteratively updated with supplemental information [8], as they claimed that they would like to learn the route first, and then learn information about the surrounding POIs.

**VirtualLeap**. This approach relies on swipe gestures (up and down) to move to the next or previous relevant points (intersection or POI) along the route. A *swipe up* gesture announces the distance to the next relevant point (e.g. "walked 50 meters/feet") followed by the respective instruction. For instance, if the user moves to a POI, the system announces its name and relative direction. If the user moves to an intersection, the system reads its name, the required turn, and the distance to the next turn/intersection. In case of a swipe down, the user moves back (e.g. "went back 50 meters/feet"), but the remaining instructions are the same. A two-finger swipe up/down moves the user to the next/previous intersection independently of the presence of POIs.

### *3.2.3 Explore POIs Information*

When next to a POI, users are able to get more information about it, by performing a *swipe right* gesture. The system starts by reading the POI category, and subsequent swipe right (or left) gestures go to the next (or previous) attribute, which can include a short description, average rating, opening hours, address, among others. This gestures

[1]pitch and yaw values from the *iOS Core Motion* framework

intended to ease and quicken the access to POI information, reducing the need to change screens and resume the navigation afterwards. In case there are no near POIs, the system alerts that the user is not next to a POI.

### 3.2.4 Additional Commands

Additional commands intend to help users orienting themselves in the route. A *double tap* repeats the current instruction. In case the user is not at an intersection nor next to a POI, it reads the name of the street and the distance to the next turn. The *shake gesture* is adapted from current navigation apps [1] and provides the name of and distance to the previous and next intersections. Finally, the user may go back to the starting point by performing a *triple tap*.

## 4. USER STUDY

The high level goal of our approach and system is to ease autonomous navigation by people with visual impairments. To accomplish that, we rely on the ability to simulate navigating a route using a smartphone to increase the knowledge of a particular location before physically visiting it.

The objective of this user study is to understand the following: 1) blind people's ability to build a mental representation of the sequential instructions and POIs of a route, after using virtual navigation; 2) the relative benefits and disadvantages of VirtualLeap and VirtualWalk modalities; 3) the perceived usefulness, acceptance and requirements of smartphone-based virtual navigation to learn a route while discovering relevant POIs. The study consisted of one session that took two to three hours (average around 2.5 hours), and can be divided in three parts: first, we use route reconstruction as the means to evaluate users' spatial representations of a simulated route; second, users are exposed to a real-world route after using the system; third, users interact with the full system using a think aloud protocol, followed by a semi-structured interview to gather feedback about the system design, usage scenarios and user preferences.

## 4.1 Participants

We recruited 14 participants (6 female). Their ages ranged from 41 to 75 (M=59.15, SD=11.92) years. Thirteen were legally blind and one had low-vision. Thirteen participants had a congenital or an early-onset visual impairment. Eleven participants own a smartphone (10 iPhone), and all but one walk alone in the street regularly. Besides the white cane or guide dog, nine participants use GPS-based navigation tools, such as *Google/Apple Maps* (5), BlindSquare (4), Nearby Explorer (2), Trekker Breeze (2) or Transit (2). Participants were compensated for their time ($25 per hour).

## 4.2 Apparatus

We used an iPhone 6 with iOS 10.1.1 and the virtual navigation application that was previously described. We had a set of LEGO blocks and respective base plates, which were used for route reconstruction (Figure 2). We video and audio recorded the whole experiment for further analysis.

## 4.3 Route Reconstruction Task

Map reconstruction is commonly used to assess route knowledge and the cognitive map of VI people [19, 14, 27, 28, 35]. In this phase, users were able to simulate a route with virtual navigation and build the same route afterwards using LEGO blocks (as in [19, 28]). Users performed two tasks,

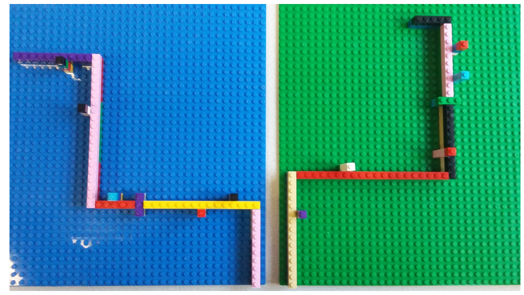

**Figure 2: The LEGO blocks and baseplate with a correct representation of the two routes used in the route reconstruction task.**

one with each method (*VirtualWalk* and *VirtualLeap*), and with similar routes. The order of the two methods and the two routes was counterbalanced among participants.

### 4.3.1 Method

**Conditions.** The *VirtualWalk* and *VirtualLeap* conditions work as described in the previous section. When using one condition, the commands in the other condition are disabled. Moreover, the function to get more information about the POIs (swipe right and left) is disabled.

**Routes.** Both routes (Figure 2) have the same length of 130 meters and the same number of intersections (4) and POIs (5). Three of those intersections require a turn. Route complexity is similar to the routes used in *SpaceSense* [35]. Our larger number of POIs reflect our main goal to use virtual navigation to learn the sequential steps of a route, which includes the POIs themselves. The POIs names are based on the most well-known food chains (via lists available in *forbes.com* and *businessinsider.com*), in order to reduce the cognitive load required to memorize the POIs.

**LEGO blocks.** Participants were required to assemble the LEGO blocks on top of a base plate (Figure 2). We provided four different LEGO lengths to represent the relative differences among the street blocks in the route (as in [35]). Their lengths can be measured in terms of the number of *studs* (the "bumps") on top of the LEGO block: 20, 15, 10 and 5 (each *stud* corresponding to two meters). Participants were informed that they could compose the route by using a single LEGO block for each street block [35]. The relative orientation of the LEGOs indicate the turns. Additional LEGO types include intersections that do not require a turn and the POIs (taller LEGOs with a single stud).

### 4.3.2 Procedure

After assigning the method for the current condition, the researcher explained all commands to the participants. Participants were able to practice for approximately 10 minutes, while the researcher provided assistance and clarifications if needed. Afterwards, the researcher explained the task, including the need to navigate a route and then build the same route using LEGO blocks. The participant was asked to explore the route until she feels comfortable and ready to reconstruct it. When the participant did not name all POIs, the researcher asks for their names. If at least one name is missing, the researcher informs the name of the remaining POIs and asks if the participant knows to what LEGO element in their route they correspond.

#### 4.3.3 Design and Analysis

We used a within-subjects design where participants performed both conditions (VirtualWalk and VirtualLeap) (in a counter-balanced order), one with each route. We excluded from this analysis one participant with low vision, since she could see, with some effort, the route on the screen. In order to evaluate participants' mental representation of the route, we adapted metrics used in two previous studies [27, 35]:

**NumberElementsError**: The number of missing or unnecessary street blocks and intersections used to recreate the route. Both routes have five street blocks and four intersections, resulting in nine elements.

**FormElementsError**: The Levenshtein distance [17] between the correct route and the participant's route composition. This corresponds to the minimum number of operations (deletions, insertions and substitutions) to correct the participant's route and targets the form of the route and not the size (of each block).

**PlacementErrors**: The number of street blocks used with an incorrect length.

In addition, we extended the analysis to include metrics about the POIs and their locations, and also to assess if participants were able to understand the relative lengths of adjacent streets.

**RelativeLengthAccuracy**: The accuracy of the perceived relation (longer, shorter or same) between the lengths of adjacent street blocks.

**POIsOrderingError**: The Damerau-Levenshtein distance [9] between the correct POI order and the users' ordering. It corresponds to the minimum number of operations (deletions, insertions, substitutions, transpositions between adjacent POIs) required to correct the participant's ordering.

**POIsInBlockError**: The number of POIs that are not in the correct street block or are missing.

**POIsOnSideError**: The number of POIs that are not in the correct side of the street.

**POIsDistanceError**: The average distance from the POIs correct location to where the user placed them. The position of the placed POI is re-calculated so it is represented in proportion to the correct length of the block. The distance is then normalized (0 to 1) depending on the block size.

Additionally, we measured participants' **completion times** to explore the route with the application. Shapiro-Wilkinson tests were applied to all dependent variables to check for normality. We ran Paired Samples T-Test to compare both approaches when the variables have a normal distribution, and the Wilcoxon Signed-Rank test otherwise.

### 4.4 Real-World Exposure Task

Measuring the accuracy of participants' route reconstruction is important to understand the spatial knowledge provided by virtual navigation, but its utility is better assessed in the real world [14]. However, it is very difficult to guarantee a fair quantitative comparison between different methods or alternatives, because external factors such as auditory cues or number of people may impact (positively or negatively) the users' ability to navigate in the real-world [14].

Nevertheless, we wanted to observe how virtual navigation in general, and both navigation methods in particular, could help to build a sequential representation of the environment that can be leveraged when physically navigating the previously simulated route. On one hand, we wanted to focus on the mental representation and not on navigation challenges caused by external factors. In order to accomplish that, instead of freely navigating the route, participants were asked to guide the researcher to the destination using the required turns and POIs as reference. On the other hand, we still wanted to take advantage of the real world sensory cues that help blind people localizing themselves. With this task (guiding the researcher), participants can still make use of such cues to provide more accurate instructions. For instance, if the user perceives she is next to a landmark, such as the elevator due to its particular sound, she can better position herself in her cognitive map.

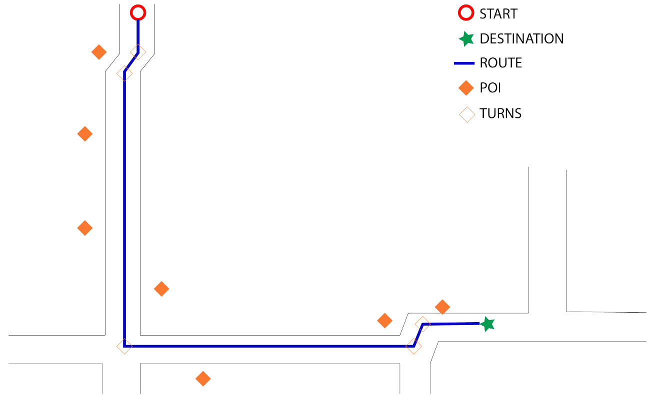


**Figure 3: The route used in the real-world exposure task, located at Carnegie Mellon University.**

#### 4.4.1 Method

**Condition.** Participants used the last method they tried in the previous phase. There was no additional practice period, nor additional commands were added.

**Route.** The route is in the University Center at Carnegie Mellon University and had FedEx as the destination. It has six segments, five turns and seven POIs besides the destination (Figure 3). The POIs include (in the following order): a book store, an exit, an art store, a post office, the career and professional development center, a package pick-up and a lounge with seats and tables.

#### 4.4.2 Procedure

Participants were instructed to explore the route for a maximum of seven minutes (based on the pilot study). They were able to finish their current trial until reaching the destination. They were also informed that afterwards they would need to guide the researcher to the destination using the turns and landmarks as reference. After simulating the route, the participant and the researcher walked to the respective building. It took seven to twelve minutes to get to the starting point.

When reaching the starting point, the researcher asked the participant to hold his elbow and guide him through the same route explored before. The researcher would not provide any new information about the environment, and would only comment about landmarks referred by the participant. For instance, if the participant comments that they *"should be passing by the Art Store"*, possible answers are: "*I do not see the Art Store*", *"We already passed the Art Store"*, "*I can see it, but we're not next to it*" or provide confirmation ("*yes, we are*". Moreover, before starting, the researcher in-

**Table 1: The description on how we assign accuracy values based on participants' instructions.**

| Accuracy | Value | Description |
|---|---|---|
| High | 1 | Distance lower than 5 steps |
| Medium | 0.75 | Distance between 5 and 10 steps |
| Low | 0.5 | Distance higher than 10 steps or vague description (e.g. "don't know the distance, but we should take the next left") |
| Unnamed | 0.25 | Acknowledge a POI, but does not recall the name |
| None | 0 | Turn error; POI not mentioned; POI in different segment |

formed that: 1) he would warn the user when they are about to reach a wall (or obstacles such as chairs and tables) and ask what to do; 2) participants could ask to come back to a known landmark/POI, in case they get lost; 3) the researcher could ask for an estimate on when they are about to reach a landmark or turn mentioned by the participant; this intends to understand participants' distance perception when they refer to elements that are far from their current position (e.g. "we should turn left after the *post office*", when the user is at the beginning of that segment).

### 4.4.3 Design and Analysis

We used a between-subjects design where participants used one of the navigation methods (VirtualWalk or VirtualLeap) to simulate the route. We collected the users' completion time and interaction behaviours while using the application. Moreover, in order to assess participants' performance guiding the researcher, we observed and coded the videos to account the following:

**Reached Destination**. It refers to participants ability to reach FedEx.

**Route Errors**. It refers to the number of wrong turns made by the user.

**Missing or Erroneous POIs**. POIs not referred by the participant or referred as belonging to a different street segment.

**Instruction Accuracy**. Accuracy of participants' POI / turn estimation from 0 (error/missing information) to 1 (high accuracy). We defined the maximum value for medium accuracy based on the GPS 4.9 meters accuracy (according to the United States government [2]). Since participants guided the researcher at a slow pace, with shorter steps than a regular step, we estimate this value to correspond approximately to 10 steps. We coded high accuracy as half of this value, meaning less than 5 steps distance. Table 1 describes the different accuracy levels.

**Route accuracy**. The average of the path instructions (five required turns) accuracy.

**POI accuracy**. The average of the POI instructions (eight, including the destination) accuracy. We also calculated the accuracy including only POIs that were referred by the participants.

We ran Independent Samples T-Test to compare both groups when the variables have a normal distribution, and the Mann-Whitney U test otherwise.

[2]https://www.gps.gov/systems/gps/performance/accuracy/

## 4.5 Overall System Exposure and Feedback

After completing the previous tasks, participants were able to explore a popular area in Pittsburgh (the Strip District). They were able to use the full system to navigate in the route while exploring the several POIs existent therein, as well as their attributes with the *swipe-right/left* gestures. First, participants were instructed to explore the route using a think-aloud protocol, where they would verbalize their thoughts about their experience with the system and this particular route. Afterwards, we performed a questionnaire and a semi-structured interview. Besides targeting feedback about the system itself, we intended to understand what the users were learning about the route and the perceived usefulness and acceptance of a system that enables both to simulate a route and to explore and discover relevant POIs along such route. Moreover, we wanted to collect information about additional requirements for such a system.

# 5. RESULTS

## 5.1 Route Reconstruction Task

Figure 4 presents the results concerning task completion time and all route and POI accuracy metrics.

### 5.1.1 Simulation Completion Time

Since in *VirtualLeap* users are able to move faster through the relevant elements of the route, we were expecting users to spend less time exploring the route. However, users took an average of 539.85 (SD=233.92) seconds exploring the route with *VirtualLeap* and 412.54 (SD=188.05) seconds with *VirtualWalk*. A *paired t-test* has shown significant differences between the two conditions ($t(12) = 2.932$, $p<0.05$), which have a mean difference superior to two minutes.

### 5.1.2 Route Accuracy

Although Figure 4 shows a slight advantage for the VirtualWalk method, differences were not statistically significant. Metrics that concern the route structure independently of the length of the street blocks show that most participants were able to accurately learn the route sequence with both methods. For instance, 10 out 13 participants reconstructed the route structure correctly (*FormElementsError* equals zero) with the VirtualWalk method, while 8 accomplished the same with VirtualLeap. Errors in route reconstruction were mostly caused by adding or omitting a street block.

Although able to grasp the sequential steps of the route, participants had more difficulty estimating the lengths of street blocks, as shown by participants' *PlacementErrors*. *RelativeLengthAccuracy* shows a better understanding of the relative lengths of adjacent street blocks, but still revealing a poor accuracy consistent with prior research (e.g. [35, 16]).

### 5.1.3 POI Accuracy

Participants placed in their LEGO route an average of 4.54 and 4.69 POIs with VirtualLeap and VirtualWalk, respectively (from a total of 5). Most reconstructed routes included all 5 (17 out of 26) or 4 (8) POIs. Again, the differences between VirtualLeap and VirtualWalk were not statistically significant for any of the POI accuracy metrics. The *POIsInBlockError* metric shows that users were able to add information about the POIs location to their mental representation of the route. In particular, five participants placed all five POIs in the respective street block

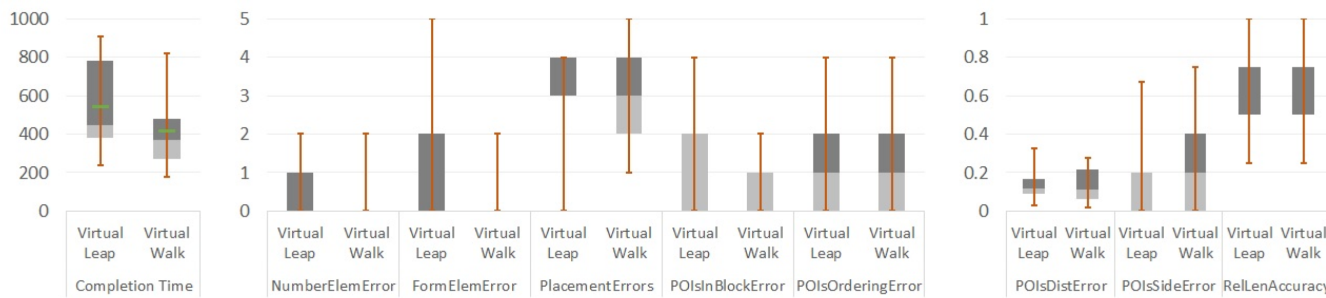


**Figure 4: Performance in route reconstruction. BoxPlots show median, IQR, min and max. Green marker indicates the mean for normally distributed variables. Plots without boxes represent measures equal to zero.**

in each of the conditions. Moreover, all participants placed at least three POIs in their correct block with the VirtualWalk condition, and eleven participants with VirtualLeap. While participants made a few errors concerning the side of the POIs in the street, they were able to have an approximate idea of the POI location in the respective block (most participants had an average POIsDistanceError below 0.20). *POIsOrderingError* adds up to the *POIsInBlockError* the POIs that are located in the same block but had their order swapped, as well as the POIs that were placed in the correct block but the participants did not recall their names.

#### 5.1.4 Interaction Patterns

The higher completion times for the VirtualLeap method can be explained by participants exploring the route more times (M=9.15, SD=5.16) with that method than with VirtualWalk (M=5.46, SD=3.15). Moreover, participants in both VirtualLeap (7) and VirtualWalk (5) tried, at some point, to explore the route independently of the POIs. Several VirtualWalk users tried to build a better mental map of the route by rotating their body when making a turn (3 users) or by keeping the phone oriented in that direction (6). However, none of the interaction patterns identified caused significant differences in the users' performance.

### 5.2 Real-World Exposure Task

#### 5.2.1 Virtual Navigation

In both methods, the interaction patterns were very similar to the ones described in the previous task. For instance, VirtualLeap users took an average of 540.6 seconds and explored the route 9.2 times, while VirtualWalk users took an average of 411.6 seconds and explored the route 5.5 times. However, there were no significant differences in user performance depending on the interaction patterns they used.

#### 5.2.2 Real World Navigation

All participants were able to guide the researcher to the destination. Figure 5 depicts the results for all metrics concerning the real world navigation task. The comparison between VirtualLeap and VirtualWalk showed no significant differences between methods, for all metrics.

Overall, participants took an average of 254.23 seconds to reach the destination and made an average of 0.77 errors. Seven participants went through the route without making any error. From the other six participants, four made one (or more) errors at a single point in the route. They often chose to return to a known POI (always right before a turn), being then able to reach the destination. The mean route accuracy (0.82) reflects participants' ability to estimate the distance to the next turn. In general, participants were able to distinguish short and long distances, even when providing less accurate information. The POIs, mainly the ones right before a turn, were key to provide accurate instructions. For instance, seven participants gave the instruction to "*turn left right after passing by the package pickup*", while eight were aware of the approximate distance to turn left after passing by the post office. When providing less accurate information, participants were also able to benefit from the environment itself (e.g. by reaching a wall or a table and chairs) to understand their next steps. For instance, three participants were aware that they should have the lounge on their left, so when they reached the tables and chairs, they immediately inferred they should turn right.

POI accuracy was much lower (0.48) than route accuracy. However, it is important to mention that participants were not instructed to mention all POIs, but to guide the researcher by providing important information about the environment. For that reason, participants focused more on the POIs that preceded a turn (bookstore, post office and package pickup), resulting in an average of 3.38 of missing or erroneous POIs. If we consider the *referred POI accuracy*, which disregards the POIs that were not mentioned by the user, the mean accuracy increases to 0.82.

Our main goal was to assess the ability of virtual navigation to convey an understanding of the sequential steps of the route, and all participants were able to reach the destination. While seven were able to guide the researcher without making any error, three of them had a combined route and referred POI accuracy above 90%, where two of them referred all POIs.

### 5.3 Overall System Exposure and Feedback

The participants' comments under the Think-Aloud protocol and during the semi structured interviews were recorded and transcribed for further analysis. The researcher that conducted the study compiled a pre-set of codes based on the questions and notes taken during the interviews. Then, two researchers coded, independently, one interview and discussed their differences until reaching an agreement on the codes and their interpretation. They coded an additional interview and due to the high inter-coder agreement (Cohen's kappa = 0.94) one researcher coded the other interviews [6].

The main themes arising from the interviews were: the comments related to the exploration of a known area, po-

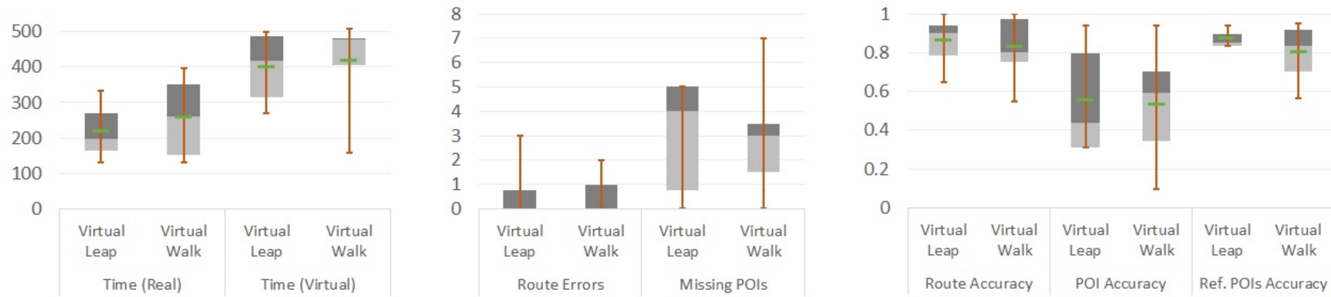


**Figure 5: Performance in the real-world exposure task. BoxPlots show median, IQR, min and max values. Variables with normal distribution also show a marker indication the mean.**

tential contexts for virtual navigation, the main advantages and disadvantages of VirtualLeap and VirtualWalk, potential improvements to such an application, and the originality or distinguishing features of this application.

**Exploring a Known Area**. All participants were very positive about being able to explore POIs in a popular route and wanted both to increase their knowledge about known POIs and to discover places they were unaware about. One participant stated: "*I live here for 19 years, and I've been to the Strip District, but I didn't know a lot of these places even existed. So you get all this good information*".

**Contexts for Virtual Navigation**. Among scenarios where participants would use such an application, all but one referred that it would be useful both in known and unknown places, usually mentioning it in general terms. A more concrete example to explore known places was their own neighborhood, while unknown places can include, for instance shopping or touristic activities. One participant commented that he could not say "*how many people say, where would you like to eat? And you don't know anything! And it would be neat to pull out an application like this and find something close, and the ratings, all these things.. is really neat*". Another participant referred that "having this knowledge could balance the power in the relationship", by being able to suggest places that are on a particular route.

**Modalities Pros and Cons**. Participants' comments regarding their favorite functions were scattered across the system functionality. For instance, 3 participants referred that they liked to have the ability to control the speed in VirtualWalk. Eight participants were able to see advantages in both modalities, supporting that both are useful, each in their own context. Ten participants referred to VirtualWalk as the way to convey a better spatial understanding of the route, while four participants pointed the slower exploration as its disadvantage. VirtualLeap on the other side, is able to provide a faster overview of a route and is more efficient when their goal is focused on browsing POIs. One participant stated: "*[VirtualWalk] (...) I think negative, it's gonna take you longer to get the overall information [but] you might have more intuition of how long it takes to get from one restaurant or POI to the next.*"

**Future Improvements**. All but two participants referred to navigational transit information as the main kind of information to add to the system. This includes bus stops, more information about intersections and crossings and sidewalk information including temporary construction. Additionally, one participant referred that he would like to have "*an auditory recording without even opening the app (...) sometimes me and my wife go on cruises, (...) we may want to put together a couple of recordings of virtual streets (...) and just have a virtual with the streets and the POIs.*"

**Originality**. All participants were enthusiastic about being able to explore a location before-hand, for instance by stating that it is the "*best virtual way to explore the route (..) you can actually navigate!*". A distinguishing factor of our application was the ability of simulate the navigation and at the same time grasp the knowledge surrounding POIs and their exact location within the route. One participant stated: "not only do you get directions, but you get your surrounding businesses (...) and what they do and what they are and how they are rated... is really neat. You can get that stuff with Google if you look for the place itself, but this turn-by-turn addition is really nice."

# 6. DISCUSSION

Building on the results obtained, we discuss our main findings and contributions:

**Virtual navigation enabled blind users to build accurate sequential representations of the environment**. The route reconstruction task showed that most participants were able to build a mental representation of the sequence of steps in a route. Although we believe that the knowledge about the POIs can be leveraged when navigating the real world, we anticipated that the ability to discover POIs along a route, could also increase the cognitive load. Still, most participants were able to recognize the different street blocks and direction of turns, while placing at least three POIs in their respective street block. Five participants, in each of the conditions, were able to recall the whole route structure and at the same time all POIs, placed in the correct block and sequence. This ability to build a mental representation in a sequential structure [21, 32] was also evident when participants guided the researcher when exposed to the real-world environment. All participants were able to reach the destination, where seven out of thirteen were able to do that without making any error. In this task, the POIs were essential landmarks for providing accurate instructions, but also to recover from errors. Participants that were not sure about their location often chose a landmark they knew in order to restart the navigation and were then able to succeed and reach the destination.

**VirtualLeap and VirtualWalk were equally able to support the creation of a mental representation of the environment, but users' preference depends**

**on their goal**. All metrics related to the route and POI knowledge provided no significant difference between the two methods. However, participants took significantly more time to use VirtualLeap than VirtualWalk, suggesting that the first may require a greater memory effort to obtain a similar cognitive mapping accuracy. While most participants would like to have the ability to use both methods, their comments point towards using VirtualWalk when the goal is to have an accurate cognitive map of the route, while using VirtualLeap when the goal is to have a quicker overview or to focus on the discovery and exploration of POIs.

**Accurate estimation of length is a challenge.** The main challenge to build an accurate representation of the route was the street length estimations. This result is not unexpected, in particular for the *VirtualLeap* modality, as previous research based on verbal descriptions (or navigation instructions) reported the difficulty to estimate distances [35]. However, we were expecting that VirtualWalk would provide a better understanding of the relative lengths of each street block. Although some participants referred that it gave "*a better perspective of length and the time it takes*", overall that was not reflected in participants' performance. Although guiding the researcher in the real world did not require a fully accurate sense of the distances, participants where generally aware of long and short distances; and they were very accurate when basing their instructions on landmarks that preceded a turn. Two participants stood out as they were able to recall all POIs and made very accurate estimations in the real-world task, but were also able to reconstruct the LEGO route (with the two modalities) without making any error regarding the sequential ordering of the route. Although these were the two youngest participants (41 and 42 years old), we could not find any effect of age in route nor POI metrics.

**Virtual navigation had high perceived usefulness and acceptance**. The main advantage of our approach was the ability to simulate and control the navigation, while at the same time being guided and informed by relevant information in the route. This was perceived as beneficial both in terms of building a mental representation of a route and as a way to discover relevant POIs in a route, which was referred as a boost in independence and knowledge that can be leveraged in their interaction with sighted peers.

# 7. CONCLUSIONS

Prior exploration of an environment by VI people can increase their spatial knowledge without the risks and cognitive load of direct exploration [5]. In order to overcome users' absent/fragmented knowledge about the real-world, we hypothesized if sequential representations of the environment containing information about turn-by-turn instructions and POIs could be used to increase route knowledge. We built a smartphone app with two virtual navigation methods and conducted a user study to assess how could they support sequential representations of the real-world.

Results have shown no significant differences between VirtualLeap and VirtualWalk, but have shown VI people's ability to build an accurate sequential representation of the route structure, complemented with POIs along the route. Most participants were also able to locate the POIs in their respective street block (at an approximate location). Moreover, in a subsequent exposure to the real world, participants were able to use the POIs as landmarks as they guided the researcher to the destination. While providing accurate sequential information, most participants were not able to retain the relative lengths of the streets blocks, which suggests the need for complementary approaches that enable learning other specific aspects of a route independently.

Subjective feedback also placed the ability to simulate the navigation while being informed about the POIs along the street as the main distinguishing aspect of our approach. Participants highlighted its ability to provide a spatial understanding of the route not only to increase the person's confidence, but also for trip planning based on the POIs explored. The most important feature to increase the usefulness of virtual navigation is the inclusion of information related to transit or navigation, such as more detail about the shape of intersections and traffic signs (including audible pedestrian signals), information about sidewalks, crosswalks and bus stops. In order to integrate this information, future work should try to understand how to convey larger amounts of information without overwhelming users. Such knowledge and the interaction with the system in a simulation mode, may help understanding how to convey timely and adequate information in the real world as well.

# 8. ACKNOWLEDGMENTS

This work was sponsored in part by Shimizu Corp, JST CREST grant (JPMJCR14E1) and NSF NRI grant (1637927).

# 9. REFERENCES

[1] Blindsquare ios application. http://blindsquare.com/. Accessed: 2017-05-04.

[2] D. Ahmetovic, C. Gleason, C. Ruan, K. Kitani, H. Takagi, and C. Asakawa. Navcog: a navigational cognitive assistant for the blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 90–99. ACM, 2016.

[3] N. Banovic, R. L. Franz, K. N. Truong, J. Mankoff, and A. K. Dey. Uncovering information needs for independent spatial learning for users who are visually impaired. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, page 24. ACM, 2013.

[4] J. R. Blum, M. Bouchard, and J. R. Cooperstock. What's around me? spatialized audio augmented reality for blind users with a smartphone. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pages 49–62. Springer, 2011.

[5] A. M. Brock, P. Truillet, B. Oriola, D. Picard, and C. Jouffrais. Interactivity improves usability of geographic maps for visually impaired people. *Human–Computer Interaction*, 30(2):156–194, 2015.

[6] J. L. Campbell, C. Quincy, J. Osserman, and O. K. Pedersen. Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42(3):294–320, 2013.

[7] E. C. Connors, E. R. Chrastil, J. Sánchez, and L. B. Merabet. Virtual environments for the transfer of navigation skills in the blind: a comparison of directed instruction vs. video game based learning approaches. *Frontiers in human neuroscience*, 8:223, 2014.

[8] R. G. Golledge. *Wayfinding behavior: Cognitive mapping and other spatial processes*. JHU press, 1999.

[9] W. H. Gomaa and A. A. Fahmy. A survey of text similarity approaches. *International Journal of Computer Applications*, 68(13), 2013.

[10] T. Guerreiro, K. Montague, J. Guerreiro, R. Nunes, H. Nicolau, and D. J. Gonçalves. Blind people interacting with large touch surfaces: Strategies for one-handed and two-handed exploration. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, pages 25–34. ACM, 2015.

[11] H. Kacorri, S. Mascetti, A. Gerino, D. Ahmetovic, H. Takagi, and C. Asakawa. Supporting orientation of people with visual impairment: Analysis of large scale usage data. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 151–159. ACM, 2016.

[12] S. K. Kane, M. R. Morris, A. Z. Perkins, D. Wigdor, R. E. Ladner, and J. O. Wobbrock. Access overlays: improving non-visual access to large touch screens for blind users. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 273–282. ACM, 2011.

[13] B. F. Katz, S. Kammoun, G. Parseihian, O. Gutierrez, A. Brilhault, M. Auvray, P. Truillet, M. Denis, S. Thorpe, and C. Jouffrais. Navig: augmented reality guidance system for the visually impaired. *Virtual Reality*, 16(4):253–269, 2012.

[14] R. Kitchin and R. Jacobson. Techniques to collect and analyze the cognitive map knowledge of persons with visual impairment or blindness: issues of validity. *Journal of Visual Impairment and Blindness*, 91(4):360–376, 1997.

[15] R. L. Klatzky, J. M. Loomis, A. C. Beall, S. S. Chance, and R. G. Golledge. Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological science*, 9(4):293–298, 1998.

[16] R. L. Klatzky, J. M. Loomis, R. G. Golledge, J. G. Cicinelli, S. Doherty, and J. W. Pellegrino. Acquisition of route and survey knowledge in the absence of vision. *Journal of motor behavior*, 22(1):19–43, 1990.

[17] V. I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady*, volume 10, pages 707–710, 1966.

[18] J. M. Loomis, R. G. Golledge, and R. L. Klatzky. Navigation system for the blind: Auditory display modes and guidance. *Presence: Teleoperators and Virtual Environments*, 7(2):193–203, 1998.

[19] M. Lumbreras and J. Sánchez. Interactive 3d sound hyperstories for blind children. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 318–325. ACM, 1999.

[20] S. Mascetti, D. Ahmetovic, A. Gerino, and C. Bernareggi. Zebrarecognizer: Pedestrian crossing recognition for people with visual impairment or blindness. *Pattern Recognition*, 2016.

[21] S. Millar. *Understanding and representing space: Theory and evidence from studies with blind and sighted children.* Clarendon Press/Oxford University Press, 1994.

[22] D. R. Montello. Spatial cognition. In N. J. Smelser and B. Baltes, editors, *International Encyclopedia of the Social and Behavioral Sciences*, pages 7–14771. 2001.

[23] C. Morency, M. Trépanier, and M. Demers. Walking to transit: an unexpected source of physical activity. *Transport Policy*, 18(6):800–806, 2011.

[24] H. Nicolau, J. Jorge, and T. Guerreiro. Blobby: how to guide a blind person. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pages 3601–3606. ACM, 2009.

[25] R. Nordahl, S. Serafin, N. C. Nilsson, and L. Turchet. Enhancing realism in virtual environments by simulating the audio-haptic sensation of walking on ground surfaces. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, pages 73–74. IEEE, 2012.

[26] S. A. Panëels, A. Olmos, J. R. Blum, and J. R. Cooperstock. 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*, pages 2107–2116. ACM, 2013.

[27] R. Passini, G. Proulx, and C. Rainville. The spatio-cognitive abilities of the visually impaired population. *Environment and Behavior*, 22(1):91–118, 1990.

[28] L. Picinali, A. Afonso, M. Denis, and B. F. Katz. Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge. *International Journal of Human-Computer Studies*, 72(4):393–407, 2014.

[29] J. J. Rieser. Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(6):1157, 1989.

[30] J. Rowell and S. Ungar. Feeling our way: tactile map user requirements-a survey. In *International Cartographic Conference, La Coruna*, 2005.

[31] J. Su, A. Rosenzweig, A. Goel, E. de Lara, and K. N. Truong. Timbremap: enabling the visually-impaired to use maps on touch-enabled devices. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, pages 17–26. ACM, 2010.

[32] C. Thinus-Blanc and F. Gaunet. Representation of space in blind persons: vision as a spatial sense? *Psychological bulletin*, 121(1):20, 1997.

[33] W. R. Wiener, R. L. Welsh, and B. B. Blasch. *Foundations of orientation and mobility*, volume 1. American Foundation for the Blind, 2010.

[34] R. Yang, S. Park, S. R. Mishra, Z. Hong, C. Newsom, H. Joo, E. Hofer, and M. W. Newman. Supporting spatial awareness and independent wayfinding for pedestrians with visual impairments. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, pages 27–34. ACM, 2011.

[35] K. Yatani, N. Banovic, and K. Truong. Spacesense: representing geographical information to visually impaired people using spatial tactile feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 415–424. ACM, 2012.