Deep Learning Compensation of Rotation Errors During Navigation Assistance for People with Visual Impairments or Blindness

DRAGAN AHMETOVIC, SERGIO MASCETTI, and CRISTIAN BERNAREGGI,
Università degli Studi di Milano
JOÃO GUERREIRO, Carnegie Mellon University
URAN OH, Ewha Womans University
CHIEKO ASAKAWA, Carnegie Mellon University

Navigation assistive technologies are designed to support people with visual impairments during mobility. In particular, turn-by-turn navigation is commonly used to provide walk and turn instructions, without requiring any prior knowledge about the traversed environment. To ensure safe and reliable guidance, many research efforts focus on improving the localization accuracy of such instruments. However, even when the localization is accurate, imprecision in conveying guidance instructions to the user and in following the instructions can still lead to unrecoverable navigation errors. Even slight errors during rotations, amplified by the following frontal movement, can result in the user taking an incorrect and possibly dangerous path.

In this article, we analyze trajectories of indoor travels in four different environments, showing that rotation errors are frequent in state-of-art navigation assistance for people with visual impairments. Such errors, caused by the delay between the instruction to stop rotating and when the user actually stops, result in *over-rotation*. To compensate for over-rotation, we propose a technique to anticipate the stop instruction so that the user stops rotating closer to the target rotation. The technique predicts over-rotation using a deep learning model that takes into account the user's current rotation speed, duration, and angle; the model is trained with a dataset of rotations performed by blind individuals. By analyzing existing datasets, we show that our approach outperforms a naive baseline that predicts over-rotation with a fixed value. Experiments with 11 blind participants also show that the proposed compensation method results in lower rotation errors (18.8° on average) compared to the non-compensated approach adopted in state-of-the-art solutions (30.1°).

CCS Concepts: • **Human-centered computing → Empirical studies in accessibility**; • **Social and professional topics → Assistive technologies**; • **Information systems →** *Geographic information systems*;

Additional Key Words and Phrases: Turn-by-turn navigation, orientation & mobility, navigation assistance

ACM Reference format:
Dragan Ahmetovic, Sergio Mascetti, Cristian Bernareggi, João Guerreiro, Uran Oh, and Chieko Asakawa. 2019. Deep Learning Compensation of Rotation Errors During Navigation Assistance for People with Visual Impairments or Blindness. *ACM Trans. Access. Comput.* 12, 4, Article 19 (December 2019), 19 pages.
https://doi.org/10.1145/3349264

Authors' addresses: D. Ahmetovic, S. Mascetti, and C. Bernareggi, Università degli Studi di Milano, Via Celoria 18 - 20133 Milano, Italy; emails: {dragan.ahmetovic, sergio.mascetti, cristian.bernareggi}@unimi.it; J. Guerreiro and C. Asakawa, Carnegie Mellon University, 5000 Forbes Ave. Pittsburgh, PA, 15213, USA; emails: jpvguerreiro@gmail.com, chiekoa@cs.cmu.edu; U. Oh, Ewha Womans University, 52 Ewhayeodae-gil, Daehyeon-dong, Seodaemun-gu, Seoul, South Korea; email: uran.oh@ewha.ac.kr.

1936-7228/2019/12-ART19 $15.00
https://doi.org/10.1145/3349264

1 INTRODUCTION

Independent mobility is one of the greatest challenges for people with visual impairments or blindness (VIB). Among the many navigation assistive technologies proposed to address this challenge, turn-by-turn navigation is a common and particularly suitable approach because it provides step-by-step guidance toward a specified destination without requiring any prior knowledge about the traversed environment [13]. This approach involves three key steps: (1) definition of navigation routes in terms of sequences of straight paths and turns, (2) continuous tracking of the user's position and orientation within the route, and (3) notification of turn instructions upon reaching turning points within the navigation route.

In the case of sighted individuals, turn-by-turn guidance, such as Global Positioning System (GPS) navigation used while driving, relies on the user's sense of sight to visually confirm the goodness of the provided instructions and to correct for localization inaccuracies or disambiguate between adjacent paths. Instead, without the possibility to perform sight-based route confirmation and correction, navigation assistance for VIB people requires much more accurate guidance. Therefore, a great amount of research on turn-by-turn navigation assistance focuses on achieving as high localization accuracy as possible to guarantee persistent and reliable navigation assistance [42].

However, even when the localization accuracy is high, imprecision in following guidance instructions can equally impact the navigation outcome. In particular, our prior work [5] shows that even small errors while following rotation instructions can lead to a significant distance offset during the following frontal movement, which in turn can result in the user taking an incorrect path (Figure 1). Such errors are frequent and can cause delays in reaching the desired destination, make the user lose orientation and not being able to recover the navigation, or even lead toward potentially hazardous areas.

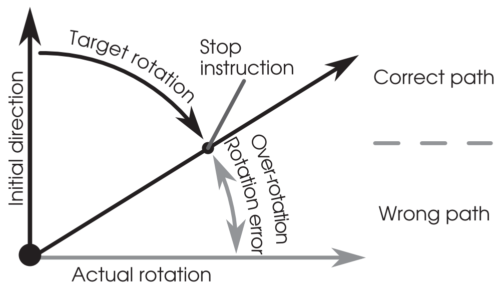
One of the main factors contributing to this kind of errors is that turn-by-turn navigation systems, particularly the considered system, use a single impulsive sound to instruct the user to stop rotating (henceforth called *stop instruction*). Using a short duration sound avoids prolonged occupation of the user's hearing channel. However, there is a temporal delay between the stop instruction and when the user actually stops rotating, which results in *over-rotation*. Considering that in the considered system the stop instruction is played only when the target rotation is reached, over-rotation results in rotation error (i.e., the angular distance between the target rotation and the actual one). This effect is presented in Figure 1(a).

1.1 Contributions

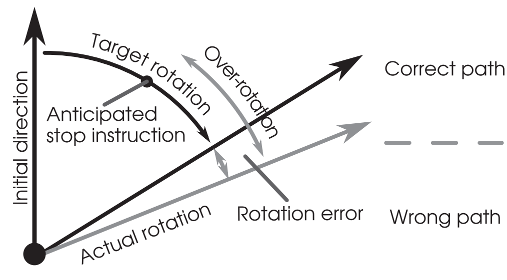
This article extends our prior work, presented at the 20th International ACM SIGACCESS Conference on Computers and Accessibility [5], with three additional key contributions:

(1) We confirm our prior findings on the magnitude of rotation errors and their impact on the navigation outcome during guidance. We also show that these findings generalize to very different environments and use cases. For this purpose, in addition to the data presented in our previous work [5], we analyze three additional datasets of navigation logs collected using NavCog [41, 54], a mobile turn-by-turn navigation assistant for people with VIB.

(2) We propose a novel technique that reduces rotation errors by compensating over-rotation with an anticipated stop instruction so that the user actually stops close to the target rotation. This is achieved by playing the stop instruction when the user's current rotation plus the predicted over-rotation equals the target rotation (see Figure 1(b)). We predict the over-rotation with a deep learning model, a multilayer perceptron (MLP) feed-forward



(a) Rotation without compensation



(b) Rotation with compensation

Fig. 1. The effect of over-rotation during rotation instruction with and without compensation.

neural network, using rotation speed, time, and angle as features. The model is trained with a dataset of rotations performed by blind participants.

(3) We evaluate the prediction model against a naive baseline approach, in which the predicted over-rotation is computed as the mean value from previous data. Using existing datasets of rotations performed by blind people, we show that the proposed approach outperforms the baseline in terms of prediction accuracy (i.e., the difference between the predicted over-rotation and the actual one). We also evaluate our technique in experiments conducted with 11 blind participants. Results show that our technique results in lower rotation errors, 18.8° on average, compared to 30.1° achieved without compensation (i.e., the state-of-the-art technique).

2 RELATED WORK

Orientation and way-finding are commonly achieved by visually exploring the surroundings in order to find reference points and make decisions on the path to follow [34]. For people with VIB, the same can be achieved using auditory, haptic, and vestibular sensing, along with their residual sight (if any) [18]. For example, blind people may identify nearby reference points or obstacles through haptic exploration using a white cane [32]. Thus, even though sighted people and people with VIB rely on different sensory abilities, their exploration model based on the inspection of the surroundings is very similar [55]. However, non-visual way-finding is less accurate and has a lower sensory range compared to visual exploration; hence, it is is also longer, more error prone, and cognitively demanding [60], particularly for long range exploration [62].

Thus, there are stark differences in how people with VIB and sighted people face navigation challenges in familiar and unfamiliar environments [64]. Indeed, since autonomous exploration of unfamiliar places is time consuming, cognitively demanding, and even possibly hazardous, people with VIB tend not to explore new routes for the first time unassisted [36]. Therefore, people with VIB generally learn new routes through orientation and mobility (O&M) training, assisted either by a sighted person [63] or by a visually impaired person who knows the route precisely [43]. Recently, several different assistive technologies have also been proposed for navigation assistance and route learning [11]. In particular, mobile assistive technologies have become a viable option for autonomous learning of new routes [47]. The adoption, frequency, and modality of usage of the available assistive technologies, however, are not homogeneous among people with VIB. Instead, these factors largely depend on the user characteristics and usage scenario [4, 28, 65].

2.1 Navigation Assistance

Thanks to the recent advances in the field of navigation assistive technologies, people with VIB can rely on several different solutions to plan and follow a route, or inspect their surroundings, both in indoor and outdoor environments [14]. Embedded navigation tools are a common approach to support non-visual way-finding [6, 16, 35, 50]. However, these solutions are built as proprietary devices, which can be cumbersome, expensive, or draw attention. These limitations result in a lower adoption [57] and higher discontinuance rate [51] of such assistive technologies. Conversely, smartphone applications [11] use widespread devices to support people with VIB in the process of planning ahead the navigation along a route [19] and during the navigation toward a target destination [54]. These include mainstream navigation assistance applications (e.g., Google Maps, Apple Maps) [47] and specialized software specifically designed to assist people with VIB in O&M [8, 29].

In particular, turn-by-turn navigation assistance has been proposed as an effective way of providing guidance to people with VIB [41, 54]. This approach translates a route into a graph of straight path segments and turning points, which are traversed sequentially. This paradigm is useful for people with VIB since contextual knowledge of the environment is not needed to be able to follow navigation instructions [13]. Furthermore, the provided instructions are provided one at a time and notified only when needed, so the user does not need to memorize the route or constantly track the navigation progress.

To support navigation assistance, several different localization approaches are used. For outdoor environments, GPS localization, also commonly used in vehicular navigators, typically is adopted [10]. In indoor environments where GPS signal is inaccurate or unavailable, localization can be achieved using computer vision–driven recognition of visual features [52], visual light communication [42], or radio signal mapping using WiFi [49] or Bluetooth beacons [33, 41]. In particular, Bluetooth beacon–based approaches achieve 1- to 2 m localization accuracy [41] using off-the-shelf equipment that can easily be installed [17]. Other approaches extract points of interest and contextual knowledge from geographical information services (GIS), such as shops or restaurants [29], or provide the user with sonified, verbal, or haptic guidance instructions toward a feature of interest [39, 44] or around an obstacle [48, 58].

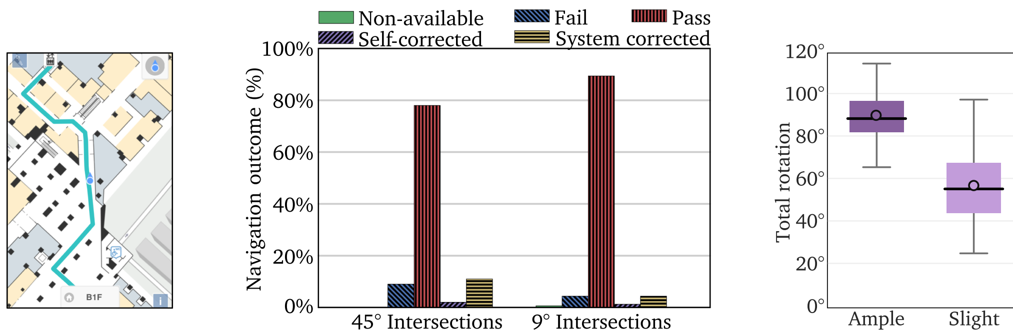
2.2 Navigation and Rotation Errors

During turn-by-turn navigation toward a target destination, sighted people can compensate for possible localization errors through visual exploration of the surroundings [67]. Conversely, errors during non-visual way-finding cannot be corrected by long-range inspection, and therefore the navigation outcome can be affected by imprecision in localization and navigation instructions. Such imprecision can be caused by environmental factors [30] or user movements such as veering [15, 24] and incorrect rotation [9, 26]. Although issues such as veering can be somewhat mitigated through O&M training [23], in practice they are a consistent source of navigation errors during turn-by-turn navigation assistance [20].

Similarly, our prior work shows that user imprecision during rotations is responsible for a significant number of navigation errors [5]. Indeed, previous research highlights that both sighted [53] and blind [37] people tend to misjudge rotations, and particularly that vestibular tracking of rotations is cognitively demanding [66], which facilitates the occurrence of rotation errors. Specifically, rotation errors in following instructions (e.g., speech or sonified messages) are affected both by how the instruction is decoded and by how the rotation is performed [9]. Such rotation errors are also related to the target turn angle [9]. Indeed, rotations at a 90° rate [2, 27] are easier to detect and track, and therefore are performed more accurately than other types of rotations.

2.3 Preliminary Work

Our prior work investigated the magnitude of the rotation errors and their effect on navigation assistance outcome in a real-world scenario [5]. For this purpose, we collected the navigation



(a) NavCog interface (b) Navigation outcomes at 45° and 90° intersections (c) Total rotation by angle

Fig. 2. Preliminary experiment.

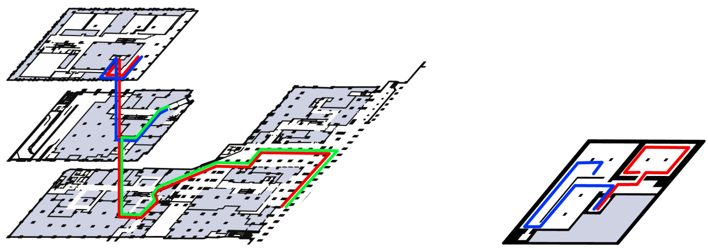
trajectories and videos of 11 participants with VIB during turn-by-turn guidance in a large shopping mall environment using NavCog navigation assistant [41, 54] (Figure 2(a)). NavCog is a turn-by-turn navigation system that relies on an installation of Bluetooth beacons and a map of their radio signal distribution in the environment to accurately localize and guide a user. Similarly to other turn-by-turn navigation software, NavCog instructs the user to start a turn with a simple verbal message (e.g., "Turn right"), then it uses an impulsive sound to notify the user to stop rotating when the target rotation is reached. Specifically, NavCog uses the default iOS "calendar alert chord" sound, a sequence of three notes:(F_4, C_5, and F_5) played at 0.1-second distances and decaying exponentially in about 1 second. During the rotations, the collected data included gyroscope measurements that were used to evaluate rotation errors. The analysis of the gyroscope data has shown that the participants have the tendency to over-rotate turns, which results in rotation error (see Figure 5(a) and Figure 5(b)). Furthermore, the rotation errors were found to be higher for slight turns (22.5° to 60°) than for ample turns (60° to 120°), as shown later in Figure 6(a). The analysis of recorded videos also highlighted that higher rotation errors during slight turns correlate to a greater occurrence of navigation issues at 45° intersections, which was also hinted to in prior works [54] (see Figure 2(b)). Instead, for ample turns, we noticed that the participants often rotated by 90° with remarkable precision, regardless of the target rotation (see Figure 2(c)[1]).

These findings suggest that simply notifying the user when the rotation should be stopped with a single impulsive sound is error prone, and therefore different interaction is required to convey accurate rotation instructions. Using sonification to provide continuous feedback while rotating can help the user slow the rotation when approaching the target direction [3]. Our prior studies confirm that such an approach can mitigate rotation errors compared to impulsive sound feedback but also results in slower rotations. Additionally, it is not clear how the increased auditory and cognitive load of such interaction may impact real-world navigation assistance and how the understanding of continuous sonification feedback is impacted by surrounding noise such as traffic. An alternative approach, which we investigate in this work, consists of compensating for the rotation errors by preempting the stop instruction. This way, the interaction remains unchanged without increasing the rotation time or the auditory and cognitive loads of the feedback.

3 EVALUATION OF ROTATION ERRORS ACROSS DIFFERENT ENVIRONMENTS

Our initial study [5] explored the impact of rotation errors on the outcome of turn-by-turn navigation assistance in a large-scale environment: a shopping mall during working hours [54]. To

[1]Boxplots range from the first to third quartile, whiskers are min. and max., and line and circle markers are median and mean.



(a) Mall floorplan (b) Museum floorplan

Fig. 3. Mall and museum testing environments (routes as colored lines).

ensure that our findings generalize to diverse scenarios, we repeat the analysis on three additional datasets of navigation trajectories, collected using NavCog in three different environments: a museum [7], an airport [20], and a hotel during a conference for VIB people [46]. These environments were selected because they present diverse and unique navigation challenges. For example, a museum requires precise short-range navigation, whereas an airport presents ample and long-range routes.

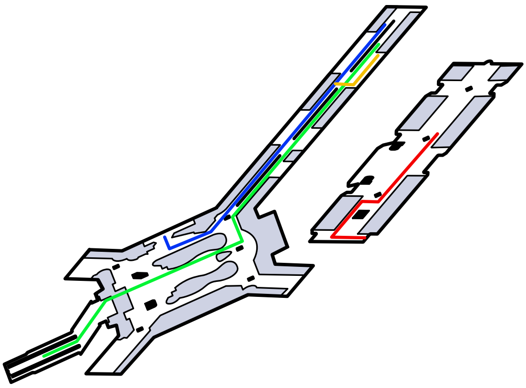
3.1 Datasets Description

We briefly describe the four datasets used. Detailed information is reported later in Figure 5(b).

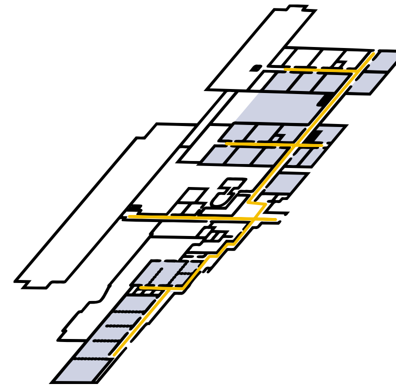
3.1.1 Mall Dataset. This dataset was collected in situ at the Coredo shopping mall in Tokyo during working hours. We installed the NavCog system on five floors of three buildings and a public basement area. For the experiment, we selected three routes, spanning across three floors, of the total length of 400m and containing 26 turns (Figure 3(a) and see Table 5(b)). The study was conducted with 11 participants with VIB [54] and consisted of three navigation tasks with the NavCog system on the defined routes. During the tasks, the system collected usage logs, including trajectories and rotations performed. The experimenters followed the participants to ensure their safety and collect navigation videos.

3.1.2 Museum Dataset. We collected this dataset during a user study at the Andy Warhol Museum in Pittsburgh [7]. We deployed the NavCog system on one floor of the museum and extended the app to support an independent museum experience combining both accurate navigation in the narrow museum area and appreciation of surrounding artworks. We created two routes with a combined length of 190m and containing 13 turns (see Figure 3(b) and Table 5(b)). In the Navigation mode, the app provides turn-by-turn guidance and informs the user about nearby artworks (e.g., "Campbell Soup Can is on your right"). In the Art Appreciation mode, the app plays audio content describing the artworks. The users can change between Navigation and Art Appreciation modes by turning toward artworks or a route, respectively. We performed a study with nine legally blind participants who followed two interest-based routes for museum visitors after a short practice route. Participants took approximately 1 hour to complete the experiment (including questionnaire and post-interview), including an average of 33 minutes navigating the two routes. Two participants performed the tasks with a guide dog, whereas the others used a white cane.

3.1.3 Airport Dataset. This dataset was collected during a user study at Pittsburgh International Airport (PIT). We deployed the NavCog system in the ticketing area, on the main terminal and one concourse, and extended the app to work on moving walkways and to include veering correction to help the users cope better with large open areas of the airport. In the study, 10 people with

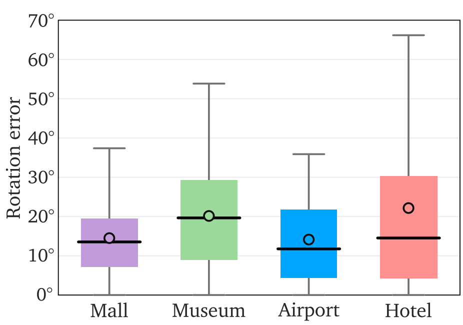


(a) Airport floorplan



(b) Hotel floorplan

Fig. 4. Airport and hotel testing environments (routes as colored lines).



(a) Rotation error for different datasets

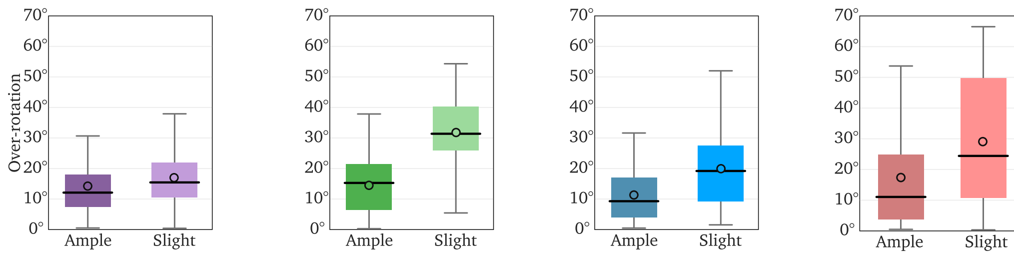
	Environment			
Info	**Mall**	**Museum**	**Airport**	**Hotel**
Participants	11	9	10	37
Avg. length	400m	190m	700m	108m
Rotations #	278	100	54	62
... slight	107	33	14	22
... ample	119	67	40	40
Avg. error	14.9°	20.7°	14.6°	22.1°
... slight	17.4°	32.1°	20.5°	29.0°
... ample	13.4°	15.1°	12.5°	18.4°

(b) Information about the collected datasets

Fig. 5. Information about the datasets and average rotation error.

visual impairments completed four navigation tasks on routes that are relevant to their air travel experience (e.g., navigating to the gate or from the gate to a restroom, restaurant, or store) [20]. The routes were approximately 700m long in total and contained 14 turns (Figure 4(a) and see Table 5(b)). Participants took approximately 2 hours to complete the experiment, including an average of 16 minutes actively navigating in the environment using NavCog. The dataset included four guide dog users, five white cane users, and one participant with low vision who did not use a navigation aid.

3.1.4 Hotel Dataset. This dataset was collected during a 4-day conference (October 26–29, 2017) held by the Pennsylvania Council for the Blind at a hotel in Pittsburgh [46, 54]. We deployed the NavCog system on the whole venue area, spanning the ground floor of the hotel, and announced to conference attendees that they could either download the navigation app from the iOS App Store on their personal devices or try it on testing mobile devices that we prepared for the occasion. The users were informed in-app and by organizers that the system would collect anonymous usage data, including user trajectories and gyroscope measurements. The users were allowed to navigate the venue freely by themselves, so there were no specific testing routes (see Figure 4(b) and Table 5(b)). No demographic data about the users was collected, and no in-person training was provided to the participants.



(a) Mall dataset (b) Museum dataset (c) Airport dataset (d) Hotel dataset

Fig. 6. Rotation error for ample and slight turns for the original and new datasets used.

3.2 Data Analysis and Results

Figure 5 shows the presence of significant rotation errors for all of the tested environments, confirming our prior findings [5]. As in our previous work, we also verify the presence of higher rotation errors for slight turns (22.5° to 60°) compared to ample ones (60° to 120°). These results are reported in Figure 6.

3.2.1 Mall Dataset. Of the 286 turns performed, 8 turns were not completed due to navigation errors and therefore were not included in the dataset. Among the 278 rotations in the dataset, 107 are defined as slight and 119 as ample, corresponding to 45° and 90° turning points, respectively. On average, the participants rotated 14.9° ± 9.9°[2] more than instructed, which resulted in 34 *incorrect* turns (12% of the total). Of those, 4 (10%) were labeled as *self -corrected* and 16 (47%) as *system corrected*. These errors did not prevent the users from reaching the target, but they still caused delays and fatigue. However, 42% of the incorrect turns (14 turns, 5% of the total) required stopping the navigation to avoid danger or because participants could not recover their orientation. These turns were labeled as *failed*. The distribution of correct and incorrect turns between 90° and 45° intersections was significantly different based on a Pearson's chi squared test ($\chi^2 = 56.06$, $p < .01$). Specifically, for 90° intersections, 93% of turns were correct and 7% were incorrect (0.5% failed). Instead, for 45° intersections, 78% were correct and 22% were incorrect (9% failed). The cause for this dissimilarity was found to be the rotation error difference between slight (17.4° ± 4.0°) and ample turns (13.4° ± 3.7°), which was statistically significant ($U = 9501.0$, $p < .01$). Additionally, we found a significant ($U = 7698.0$, $p < .05$) reduction of rotation errors for participants with high smartphone expertise (13.24° ± 9.62°) compared to others (15.65° ± 9.97°). Age also significantly influenced rotation accuracy ($U = 6305.0$, $p < .01$). Indeed, for participants younger than 45 years, the rotation error was 17.70° ($SD = 10.84°$) and 12.25° ($SD = 8.26°$) for others. Follow-up analysis reveals that participants older than 45 years were more cautious when starting to turn and more reactive to stop instructions.

3.2.2 Museum Dataset. This dataset includes a total of 100 turns; this number varied among participants due to the nature of the second task, in which the users were instructed to turn only toward the artworks they were interested in. On average, the participants rotated 20.7° ± 13.1° more than instructed. Such rotation difference did not prevent users from reaching the destination, but it caused either veering when navigating or a slight mismatch when facing the artworks to listen to the audio content. Among the registered rotations, 33 were slight turns and registered a rotation error of 32.1° ± 11.6°. Instead, ample turns were 67, with an error of 15.1° ± 9.7°. A Mann-Whitney U test revealed that the difference is statistically significant ($U = 294.0$, $p < .001$),

[2]We will use *Mean* ± *Standard Deviation* notation.

suggesting greater difficulty in estimating the correct orientation when performing a slight turn than an ample turn. Age and blindness onset information were also collected but did not seem to impact rotation errors.

3.2.3 Airport Dataset. This dataset includes 54 rotations, among which 40 are ample and 14 are slight turns. On average, the participants rotated 14.6° ± 12.4° more than instructed. Such level of rotation error did not prevent the participants from reaching the destination, but it caused several navigation errors. In particular, due to the presence of long and wide corridors, small rotation errors resulted in veering toward wrong directions. This issue was often corrected by the system through "bear right/left" messages or re-routing. Average rotation error was 20.5° ± 14.2° for slight turns and 12.5° ± 11.1° for ample ones. Again, the difference is statistically significant ($U =$ 176.0, $p < .05$), confirming a greater difficulty to estimate the correct turning angle for slight turns. We also explore the effect of participants' age and navigation tool used (guide dog or white cane) on rotation errors. Unlike the Mall dataset, there were no significant differences in rotation errors between participants older than 45 years of age and others. Instead, regarding the navigation tool used, participants supported by a guide dog had lower rotation errors (13.0° ± 14.3°) than those using a white cane (16.26° ± 9.9°). The difference was found to be statistically significant ($U = 256.0$, $p < .03$). This may suggest that in an environment composed of long and wide corridors, white cane users have little reference points during rotations, whereas guide dogs provide assistance in selecting the correct direction while turning.

3.2.4 Hotel Dataset. A total of 280 in-the-wild travel logs were collected from 37 NavCog users. Since the conference area spanned several nearby rooms on the same floor, the routes were often short and simple. Indeed, on average, the traveled distance was 108.0m ± 71.0 and the duration of the trips was 182.2 ± 131.0seconds. In total, 62 turns were logged, with 1.67 turns per user on average. The rotation error was 22.14° ± 21.72° on average. Among the registered turns, 22 were slight turns, whereas the other 40 were ample turns. The average rotation error for slight turns was 29.0° ± 21.3°, which was greater than the average rotation error for ample turns (18.4° ± 21.3°). A Mann-Whitney U test also revealed in this case that the difference is statistically significant ($U =$ 297.0, $p = .018$).

4 PREDICTIVE COMPENSATION OF OVER-ROTATIONS

The analysis of the three new datasets confirms consistent over-rotations that translate to high rotation errors during turn-by-turn navigation assistance. This shows that our prior findings generalize to diverse use cases and motivates the need for a compensation mechanism to offset over-rotation and mitigate the resulting rotation errors. A simple way to achieve this would be to anticipate all turns by a fixed amount. However, given the high over-rotation variability among turns of different magnitude, a compensation by a fixed value is not ideal. Indeed, our prior finding that slight turns result in greater over-rotations compared to ample turns is also verified in all three datasets.

Thus, we propose a technique to dynamically predict the over-rotation and preempt the stop instruction to reduce rotation errors. At every moment during the user's rotation, we predict what would be the over-rotation if the stop instruction were played at the user's current rotation angle. If the current rotation angle plus the predicted over-rotation equals the target rotation angle, the stop instruction is immediately played. We model the prediction problem as a regression that estimates the over-rotation performed by the user after the stop instruction, and we design an MLP feedforward neural network suitable for this task.

4.1 Dataset and Feature Selection

The datasets collected in previous works were useful for motivating our research and for evaluating the magnitude of the rotation errors in different environments. However, they include a limited number of data samples, which are unequally distributed (i.e., mostly 90° turns, as shown in Table 5(b)). Therefore, those datasets are not suitable to model the whole range of possible rotation angles. Additionally, they have been acquired in the wild and potentially include noisy or mislabeled data. For example, for the Hotel dataset, there have been occasions where the users did not follow the instructions correctly because they stopped to chat with other conference participants.

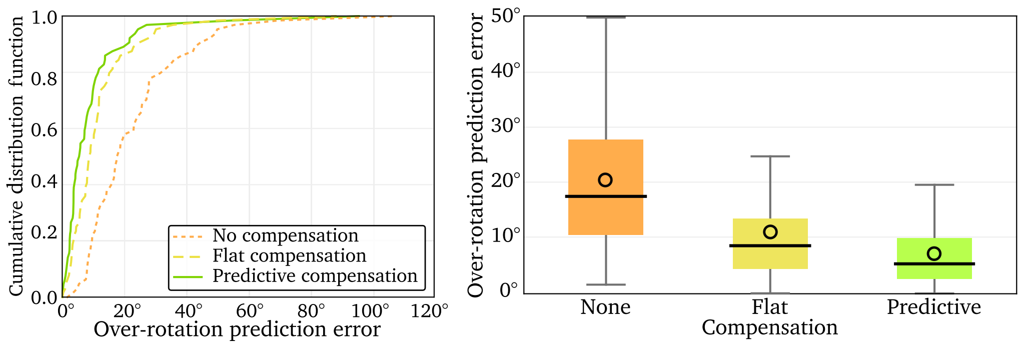
To guarantee the reliability of the data used to train our regressor, we use a dataset of rotations performed by people with VIB in a controlled setting [3]. This dataset contains rotation recordings with several different interaction approaches. However, we consider only data in which the rotation and stop instructions had the same interaction paradigm as in NavCog [54]. During the recordings, participants seated on a swiveling chair with the smartphone positioned on the chair handle to minimize the effect of involuntary movements on the rotation measurements. The dataset collects 324 rotations from 18 participants with VIB, divided in three rotation angles, both clockwise and counter-clockwise, with three repetitions for each angle. It includes gyroscope data measuring the angular velocity of rotations at a rate of 100Hz, with timestamps for each data sample. We divide the dataset in 80% training and 20% testing data (259 and 65 samples, respectively).

We identify four features that we select for our prediction task:

- *Rotation angle* is the angle rotated between the turn instruction and the stop instruction for each training sample. During usage, this corresponds to the user's current rotation angle, since we predict what the over-rotation would be if the stop instruction were provided at the current moment. Previous analysis of ample and slight turns shows that this metric impacts the over-rotation (see Figure 6). Spearman's rank correlation coefficient analysis shows significant correlation ($p < .001$) of weak magnitude ($\rho = -0.26$) between rotation angle and over-rotation, which further motivates the inclusion of this metric as a feature.
- *Rotation duration* is the time between the turn instruction and the stop instruction for each training sample. During usage, this corresponds to the current time since the turn instruction. Correlation analysis in this case reports a strong correlation with the over-rotation measurement ($p < .001$, $\rho = -0.53$).
- *Average rotation velocity* is the average angular speed during rotation for each training sample. During usage, this corresponds to average angular speed since the turn instruction. We find a significant correlation with over-rotation ($p < .01$) of moderate intensity ($\rho = 0.47$).
- *Maximum rotation velocity* is the maximum angular speed during rotation for each training sample. During usage, this corresponds to maximum angular speed since the turn instruction. We detect a Spearman's rank correlation coefficient of ($\rho = 0.43$), indicating moderate correlation with the over-rotation, which also is significant ($p < .01$).

4.2 MLP Architecture and Evaluation

An MLP neural network is composed of one input layer of the size N equal to the number of features and an output layer of the size M equal to the number of outputs of the network. In our case, we have $N = 4$ features, described earlier, and $M = 1$ output, which is the predicted over-rotation. When defining an MLP network, there are two key hyperparameters to select: the number of hidden layers that compose the neural network and the number of neurons for each hidden layer [59]. Although there have been diverse proposals and best practices [56], there is no consensus on the method to adopt for this purpose, and the effectiveness of different choices varies based on the problem and the data [61].



(a) Cumulative distribution of prediction errors (b) Boxplot of prediction errors by compensation approach

Fig. 7. Over-rotation prediction.

For the architecture of our regression MLP network, we consider two main criteria: (1) To avoid overfitting of the neural network, we set the upper bound of the number of neurons as $L - 1$, where $L = 259$ is the number of training samples [25], and (2) we adopt a pyramid-based architecture in which layers decrease in the number of neurons the closer they are to the output layer, which was shown to provide robust results for regression tasks [40]. We define an initial network with a single hidden layer having the same number of neurons as the input layer $N = 4$. We then increase the number of hidden layers, each time adding a new one on top, having four times the neurons of the previous layer. This is repeated until the network stops improving and while criterion (1) is still valid. The resulting network has three hidden layers, with 4, 16, and 64 neurons.

We implement our network as a model for the Keras [22] Python library. We then train the model with our dataset, and evaluate its over-rotation prediction accuracy with respect to the non-compensated approach and a naive baseline—a flat compensation that simply computes the prediction as the average over-rotation on the training data. We recall that 80% of the dataset was used for training and 20% for the evaluation. The evaluation metric is defined as over-rotation prediction error, and it is computed as the difference between the prediction and the actual over-rotation. In case of the non-compensated approach, we consider the prediction to be 0°. Therefore, the prediction error will be equal to the over-rotation ($20.65° \pm 20.79°$). In the case of flat compensation, the average over-rotation on the training dataset is 20.17°, which is used as a fixed-value prediction. The resulting over-rotation prediction error is $12.03° \pm 16.96°$. For our predictive compensation approach, the prediction error is $9.3° \pm 16.42°$ ($9.93° \pm 16.64°$ for the two-layer version and $10.13° \pm 17.14°$ for the single-layer version), which improves over the non-compensated approach by 11.35° and over the naive compensation approach by 2.73° on average (Figure 7).

5 USER STUDY

We evaluated our approach in a controlled user study with 11 blind participants. We implemented a test mobile app to conduct rotation experiments with and without compensation, measuring rotation error using gyroscope data. Since we wanted to evaluate rotation errors due to human imprecision and our system's capability to compensate for such errors, a controlled testing environment was used to minimize external sources of noise and distraction such as traffic or other people's voices. These factors instead depend on the environment context and could be learned from context-specific datasets, which is left as future work.

Table 1. Participants' Demographic Information

ID	Sex	Age	Visual Impairment		Experience with			
			Condition	Since	Music	Smartphone	NAT	Which Ones?
P1	M	23	Totally blind	5	High	Med.	High	Ariadne GPS, iMove
P2	M	34	Totally blind	Birth	None	Low	Med.	–
P3	M	44	Legally blind	10	Low	Med.	High	Google Maps, Navigon, Maps
P4	F	32	Legally blind	14	High	Low	Med.	–
P5	F	29	Totally blind	Birth	Low	Low	Low	–
P6	F	27	Totally blind	Birth	Med.	Med.	High	Maps, Kapten
P7	M	39	Totally blind	Birth	None	Low	Low	–
P8	M	38	Totally blind	3	High	High	High	Maps, Kapten, Google Maps
P9	M	24	Legally blind	10	Low	Low	Med.	–
P10	M	48	Totally blind	Birth	High	Med.	Med.	Tried Ultra Cane
P11	F	37	Totally blind	5	High	Low	High	–

5.1 Participants

We recruited 11 blind participants (seven male, four female) through word of mouth. Among them, 9 are totally blind and 2 are considered legally blind, having minimal residual sight. Their age ranged between 23 and 48 years (34.1 ± 8.0). Five were blind since birth, whereas others became blind between 3 and 14 years of age. High-level musical expertise was reported by 5 participants, whereas one reported medium expertise, three low, and two none. Six participants reported low expertise with smartphones, 4 considered their expertise medium, and 1 participant stated to have high smartphone expertise. In terms of expertise with navigation assistive tools (NATs), 5 participants considered their expertise high, and reported prior usage of tools such as Ariadne GPS [10], iMove [29], and Kapten GPS [31]. Participants' data is reported in Table 1.

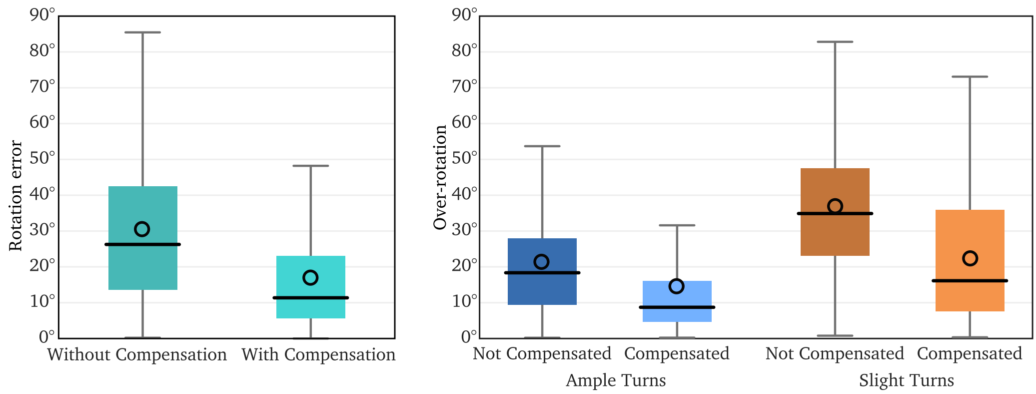
5.2 Apparatus

For the study, we adapted an Android app used in our previous work [3] to evaluate rotation accuracy with different sonification approaches. We defined two test modalities: without and with predictive compensation. In terms of interaction, there is no difference between the two. In both cases, the approach mimics the one used in the NavCog app [54]: the application presents a verbal turn instruction (e.g., "turn left") and, once the user has rotated sufficiently, it plays a single stop instruction to signal to the user to stop turning. Without compensation, the stop instruction is played when the current rotation angle is equal to or greater than the target rotation angle (see Figure 1(a)). With compensation, the system computes the over-rotation prediction for every gyroscope reading in real time. The stop instruction is played when the predicted over-rotation, added to the current rotation angle, is equal to or greater than the target rotation angle (see Figure 1(b)).

We used the DeepLearning4J [12] Java library to import and run our MLP Keras model for predictive compensation. A Google Pixel 2 smartphone was used, with over-rotation prediction computation executed for every gyroscope reading at a rate of 100Hz. Gyroscope readings were also used to measure the rotation ground truth during the experiments, since the error introduced by these sensors can be considered negligible [38]. The experiments were performed in controlled and silent environments, and no additional instruments (e.g., earphones or white cane) were used.

5.3 Procedure

At the beginning of the study, we provided each participant with a brief description of the experiment and its goals. Afterward, an experimenter collected participants' demographic information



(a) Overall rotation error

(b) Rotation error by slight and ample turns

Fig. 8. User rotation errors with and without predictive compensation.

and explained how to perform the rotation tasks. The participants were then led to the testing environment and provided with a smartphone device running the test app. They were instructed to stand straight, hold the device in front of the body, and follow the turn instructions, keeping the direction of the device consistent with the direction of the body. We also advised them to turn as they usually do during mobility, without specifically prioritizing speed or precision, and to stop when the stop instruction was provided. The participant could then double tap on the screen of the device to initiate a rotation task and could terminate a task and start the next one with the same interaction. The tasks followed in sequence until completion. A first sequence contained 4 training tasks to ensure that the participant correctly comprehended the procedure. Afterward, the main sequence of 48 tasks was performed. Since the tasks were very brief (about 10 seconds per task), all of them were conducted in one sequence. In total, the experiment took between 10 and 15 minutes for each participant.

5.4 Results

As shown in Figure 8(a), the tasks performed with compensation had lower rotation errors (18.8 ± 19.6) than those without (30.1 ± 20.8). The difference is statistically significant based on a Mann-Whitney U test ($U = 22637.0$, $p < .001$). The analysis of rotation errors divided by the turn type shows that there remains a difference in rotation error between slight and ample turns, even for the tasks performed with compensation. However, such difference is smaller for compensated than for uncompensated tasks (see Figure 8(b)). Specifically, for uncompensated tasks, the rotation error is 37.4° ± 20.5° and 22.7° ± 18.4° for slight and ample turns, respectively, resulting in a difference of 14.7° on average. Instead, for compensated tasks, the rotation error is 22.5° ± 19.6° for slight turns and 15.1° ± 19.0° for ample turns, totaling a lower difference of 7.4°. These findings not only confirm that our approach is effective in reducing rotation errors, but they also show that it is capable of adapting to over-rotation changes based on the rotated angle.

As hypothesized, there was no increase in time required to perform compensated rotations. Instead, compensated tasks were actually 8.8% faster to perform (3.1 ± 1.4 seconds) than others (3.4 ± 1.4 seconds). The difference was found to be statistically significant ($U = 32758.5$, $p < .001$). This result holds for both slight (from 3.3 ± 2.3 seconds to 3.0 ± 1.7 seconds) and ample turns (from 4.8 ± 2.3 seconds to 4.3 ± 2.0 seconds). The measurements exclude the overhead due to spoken feedback, for which a fixed time of 1.4seconds is allotted. This result is in contrast to the approaches that aim to reduce rotation errors using continuous sonification techniques [3], which instead cause the interaction to be slower. Regarding participant characteristics, we investigate the effect of musical, smartphone, and NAT expertise, as well as participants' age and visual impairment onset on the rotation errors. We discover that visual impairment onset is the only factor with statistically significant differences ($U = 4067$, $p < .001$). For participants visually impaired since birth, the average rotation error ($22.8° \pm 20.5°$) was lower than for others ($45.3° \pm 30.8°$). This factor persists even after compensation ($U = 4067$, $p < .001$), where participants with visual impairment onset at birth had a lower rotation error ($14.5° \pm 18.8°$) than the other participants ($26.5° \pm 25.8°$).

6 DISCUSSION AND LIMITATIONS

We discuss our main contributions: the analysis of rotations in real-world experiments and our technique for reducing rotation errors. We also compare our approach to different interaction approaches.

6.1 Rotation Analysis

Our previous work [5] shows that, despite accurate localization, NAT users incur problems because that they cannot precisely follow the provided rotation instructions. These results were obtained from real-world data collected in a specific environment: a shopping mall. It is worth wondering whether the same problems occur in different environments. Indeed, in an environment with different architectural characteristics (e.g., narrower corridors), users might be able to rely on their O&M abilities to compensate for the rotation errors. Our study shows that similar rotation errors actually occur across different environments, including those with open spaces and ample corridors (shopping mall, airport) and those with narrower corridors (hotel).

We highlight another phenomenon common to all environments: slight turns are more error prone than ample ones. This was hinted in our previous work and other contributions for specific environments [5, 21, 30]. We confirm this finding, providing even clearer results, on the observations from other three environments with different characteristics. Instead, the effect of participants' characteristics on rotation errors is not consistent for all datasets. For example, participants older than 45 years and those with high smartphone and NAT expertise had lower rotation errors in the mall environment but not in others. Conversely, although the blindness onset age did not have an effect in the mall environment, in our user study participants who were visually impaired at birth had lower rotation errors both with and without compensation. Similarly, being assisted by a guide dog reduced rotation errors in the wide airport environment but not in the museum. These results indicate that participant characteristics may have interaction effects with other factors such as environment characteristics and therefore need to be analyzed jointly, with more participants.

Another factor that was not investigated in this work is the impact of O&M training on rotation errors. Prior work shows that higher O&M expertise can significantly reduce veering [23], and that additional training to avoid veering can have lasting effects on participants. In our case, all participants had medium to high O&M expertise, and the experiment had a short duration. Therefore, we could not verify whether the same results apply to rotation errors.

6.2 Over-Rotation Compensation

The proposed approach was shown to effectively reduce rotation errors. In our experiment, such errors were reduced from 30.1° to 18.8°. However, even with over-rotation compensation, rotation errors remain present in our results, and their magnitude, measured with actual users, is higher than the average prediction error estimated on previous data (compare Figures 7(b) and 8(a)). This suggests that the trained over-rotation prediction model does not completely adapt to our experimental setting. This can be explained by the following two considerations.

First, our prediction model was trained with data collected in a different experimental setting than the one used to verify the model. Indeed, the data used for training, collected with participants sitting on a swiveling chair, shows smaller over-rotation ($20.65° \pm 20.79°$) than the data used for testing, which was collected with users standing (30.1 ± 20.8). An approach to better adapt the over-rotation prediction model to the experimental setting could be to retrain the model with data collected in similar conditions, or to use additional data from diverse datasets to better generalize the model to diverse conditions. Nonetheless, even with the great difference between the training and the testing data, and particularly with a much more challenging testing data, our model is still able to greatly reduce rotation errors and perform better than the baseline approach.

Second, the model was trained with data from a small number of individuals, different from the ones it was evaluated with. With a small number of participants, personal characteristics may have a significant impact on data variability and therefore on the predictive power of the model trained with such data. Indeed, our model, trained with data generated by a small number of participants, did not perfectly adapt to experiments conducted by a diverse set of participants. To achieve a more robust prediction, one possible approach is to create a training dataset containing data from a higher number of diverse participants. Alternatively, model personalization [45] could be used to better adapt to specific participants' characteristics.

6.3 Different Interaction Approaches

The technique proposed to reduce rotation errors uses a very simple form of sonification: a single impulsive sound. Clearly, continuous sonification techniques are possible [3] and were shown to be capable of reducing rotation errors at the cost of additional rotation time. Instead, our approach reduces both rotation error and rotation time, which could be better suited to time-critical tasks and situations in which audio feedback must be kept at a minimum. However, preliminary analysis conducted in Ahmetovic et al. [3] also highlights that a single impulsive sound is less pleasant for participants than continuous sonification techniques. Thus, we cannot conclude that a single impulsive sound approach is better than the one based on continuous sonifications, nor vice versa. Additional analysis is needed to evaluate the cognitive load associated with different interaction techniques, particularly in presence of external noise [1]. Results presented here can also contribute to improve the approach based on continuous sonification. Indeed, it has been shown that the best results are achieved by pairing a continuous sonification approach with a single impulsive sound as reinforcement feedback [3]. Thus, being able to anticipate the impulsive sound to compensate for over-rotation can be effective also in the case of continuous sonification approaches.

Although achieving better rotation accuracy is the main goal of the proposed technique, we also observe that the act of rotating is intrinsically subject to approximation, so NATs should be designed to take inaccurate rotations into consideration. One example is the veering correction adopted in the airport environment [20], which limits the error caused by the frontal movement after a wrong rotation. This sort of compensation mechanism occurs after the rotation and is complementary to the technique proposed in this article, which aims at limiting the rotation error itself.

7 CONCLUSION

We present a thorough investigation of rotation errors during turn-by-turn guidance for people with VIB on four datasets collected in diverse environments. We show the presence of consistent turn errors caused by over-rotation following the stop instruction and propose a method for limiting such errors by predicting and compensating for the over-rotation. Our approach is trained and validated on existing data, and it is evaluated in a user study with 11 participants with VIB. Although our approach successfully reduces rotation errors, some of the error remains.

As future work, we will explore whether training our model with datasets similar to the testing condition can achieve even better results. We will also investigate personalization and active learning to better adapt over-rotation prediction to user characteristics. In particular, active learning can automatically improve future predictions: each time our technique predicts over-rotation and plays the stop instruction, it can measure the actual over-rotation and use it to retrain future predictions. Similarly, we will consider environmental and contextual characteristics as features to better adapt to other use cases. Another possibility is to integrate our approach with interaction techniques that use continuous sonification to reduce rotation errors. We will evaluate our approach and compare it to continuous sonification techniques both quantitatively in terms of rotation time and error and qualitatively in terms of perceived cognitive load and user preferences. We will also investigate how the O&M expertise level impacts rotation errors and whether specific training can help offset such errors. Finally, we will investigate how diverse techniques perform in different environment soundscapes, such as the presence of background traffic noise or voices.

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*. ACM, New York, NY, 21.

[2] Eugene Abravanel and Herbert Gingold. 1977. Perceiving and representing orientation: Effects of the spatial framework. *Merrill-Palmer Quarterly of Behavior and Development* 23, 4 (1977), 265–278.

[3] Dragan Ahmetovic, Federico Avanzini, Adriano Baratè, Cristian Bernareggi, Gabriele Galimberti, Luca A. Ludovico, Sergio Mascetti, and Giorgio Presti. 2019. Sonification of rotation instructions to support navigation of people with visual impairment. In *Proceedings of the International Conference on Pervasive Computing and Communications*. IEEE, Los Alamitos, CA.

[4] Dragan Ahmetovic, João Guerreiro, Eshed Ohn-Bar, Kris M. Kitani, and Chieko Asakawa. 2019. Impact of expertise on interaction preferences for navigation assistance of visually impaired individuals. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*.

[5] 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 ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY.

[6] Ilias Apostolopoulos, Navid Fallah, Eelke Folmer, and Kostas E. Bekris. 2014. Integrated online localization and navigation for people with visual impairments using smart phones. *ACM Transactions on Interactive Intelligent Systems* 3, 4 (2014), Article 21.

[7] Saki Asakawa, João Guerreiro, Daisuke Sato, Hironobu Takagi, Dragan Ahmetovic, Desi Gonzalez, Kris M. Kitani, and Chieko Asakawa. 2019. An independent and interactive museum experience for blind people. In *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility*.

[8] Hsuan-Eng Chen, Yi-Ying Lin, Chien-Hsing Chen, and I-Fang Wang. 2015. BlindNavi: A navigation app for the visually impaired smartphone user. In *Proceedings of CHI'15 Extended Abstracts on Human Factors in Computing Systems (CHI EA'15)*. ACM, New York, NY, 19–24.

[9] E. R. Chrastil and W. H. Warren. 2017. Rotational error in path integration: Encoding and execution errors in angle reproduction. *Experimental Brain Research* 235, 6 (2017), 1885–1897.

[10] L. Ciaffoni. 2012. Ariadne GPS. Retrieved November 26, 2019 from https://www.ariadnegps.eu/tickets.

[11] 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. *Journal on Multimodal User Interfaces* 9, 4 (2015), 275–286.

[12] DL4J. 2017. Deep Learning for Java Library. Retrieved November 26, 2019 from https://deeplearning4j.org/.

[13] Elliot P. Fenech, Frank A. Drews, and Jonathan Z. Bakdash. 2010. The effects of acoustic turn-by-turn navigation on wayfinding. In *Human Factors and Ergonomics Society*. Sage.

[14] Hugo Fernandes, Paulo Costa, Vitor Filipe, Hugo Paredes, and João Barroso. 2019. A review of assistive spatial orientation and navigation technologies for the visually impaired. *Universal Access in the Information Society* 18, 1 (2019), 155–168.

[15] Alexander Fiannaca, Ilias Apostolopoulous, and Eelke Folmer. 2014. Headlock: A wearable navigation aid that helps blind cane users traverse large open spaces. In *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility.* ACM, New York, NY.

[16] Aura Ganz, James Schafer, Siddhesh Gandhi, Elaine Puleo, Carole Wilson, and Meg Robertson. 2012. PERCEPT indoor navigation system for the blind and visually impaired: Architecture and experimentation. *International Journal of Telemedicine and Applications* 2012 (2012), Article 894869, 12 pages.

[17] 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. *Proceedings of the ACM Conference on Interactive, Mobile, Wearable, and Ubiquitous Technologies.*

[18] Reginald G. Golledge, Roberta L. Klatzky, and Jack M. Loomis. 1996. Cognitive mapping and wayfinding by adults without vision. In *The Construction of Cognitive Maps*. Springer.

[19] João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. 2017. Virtual navigation for blind people: Building sequential representations of the real-world. In *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility.* ACM, New York, NY.

[20] 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 Conference on Human Factors in Computing Systems.*

[21] 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 Web for All Conference.* ACM, New York, NY.

[22] Antonio Gulli and Sujit Pal. 2017. *Deep Learning with Keras*. Packt Publishing Ltd.

[23] David Guth. 2007. Why does training reduce blind pedestrians veering. In *Blindness and Brain Plasticity in Navigation and Object Perception*, J. J. Reiser, D. H. Ashmead, F. Ebner, and A. L. Carn (Eds.). Lawrence Erlbaum Associates, New York, NY, 353–366.

[24] David Guth and Robert LaDuke. 1994. The veering tendency of blind pedestrians: An analysis of the problem and literature review. *Journal of Visual Impairment and Blindness* 88 (1994), 391–400.

[25] S.-C. Huang and Y.-F. Huang. 1991. Bounds on the number of hidden neurons in multilayer perceptrons. *IEEE Transactions on Neural Networks* 2, 1 (1991), 47–55.

[26] I. Israël, D. Sievering, and E. Koenig. 1995. Self-rotation estimate about the vertical axis. *Acta oto-laryngologica* 115, 1 (1995), 3–8.

[27] R. Jürgens, T. Boss, and W. Becker. 1999. Estimation of self-turning in the dark: Comparison between active and passive rotation. *Experimental Brain Research* 128, 4 (1999), 491–504.

[28] 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, New York, NY.

[29] Hernisa Kacorri, Sergio Mascetti, Andrea Gerino, Ahmetovic Dragan, 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 Transactions on Accessible Computing* 11, 1 (2018), Article 5.

[30] 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 2018 CHI Conference on Human Factors in Computing Systems (CHI'18).* ACM, New York, NY.

[31] Kapten. 2017. Kapten GPS. Retrieved November 26, 2019 from http://www.kapsys.com/.

[32] Dae Shik Kim and Robert Wall Emerson. 2014. Effect of cane technique on obstacle detection with the long cane. *Journal of Visual Impairment & Blindness* 108, 4 (2014), 335–340.

[33] Jee-Eun Kim, Masahiro Bessho, Shinsuke Kobayashi, Noboru Koshizuka, and Ken Sakamura. 2016. Navigating visually impaired travelers in a large train station using smartphone and Bluetooth Low Energy. In *Proceedings of the Symposium on Applied Computing*. ACM, New York, NY.

[34] Robert M. Kitchin. 1994. Cognitive maps: What are they and why study them? *Journal of Environmental Psychology* 14, 1 (1994), 1–19.

[35] Gordon E. Legge, Paul J. Beckmann, Bosco S. Tjan, Gary Havey, Kevin Kramer, David Rolkosky, Rachel Gage, Muzi Chen, Sravan Puchakayala, and Aravindhan Rangarajan. 2013. Indoor navigation by people with visual impairment using a digital sign system. *PLoS One* 8, 10 (2013), e76783.

[36] Roberto Manduchi and Sri Kurniawan. 2011. Mobility-related accidents experienced by people with visual impairment. *AER Journal: Research and Practice in Visual Impairment and Blindness* 4, 2 (2011), 1–11.

[37] V. V. Marlinsky. 1999. Vestibular and vestibulo-proprioceptive perception of motion in the horizontal plane in blindfolded man–II. Estimations of rotations about the earth-vertical axis. *Neuroscience* 90, 2 (1999), 395–401.

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

[39] 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. *International Journal of Human-Computer Studies* 85 (2016), 16–26.

[40] T. Masters. 1993. *Practical Neural Network Recipies in C++*. Morgan Kaufmann.

[41] Masayuki Murata, Dragan Ahmetovic, Daisuke Sato, Hironobu Takagi, Kris M. Kitani, and Chieko Asakawa. 2019. Smartphone-based localization for blind navigation in building-scale indoor environments. *Pervasive and Mobile Computing* 57 (2019), 14–32.

[42] Madoka Nakajima and Shinichiro Haruyama. 2012. Indoor navigation system for visually impaired people using visible light communication and compensated geomagnetic sensing. In *Proceedings of the 2012 1st IEEE International Conference on Communications in China*. IEEE, Los Alamitos, CA.

[43] Hugo Nicolau, Joaquim Jorge, and Tiago Guerreiro. 2009. Blobby: How to guide a blind person. In *Proceedings of CHI'09 Extended Abstracts on Human Factors in Computing Systems (CHI EA'09)*. ACM, New York, NY.

[44] Uran Oh, Shaun K. Kane, and Leah Findlater. 2013. Follow that sound: Using sonification and corrective verbal feedback to teach touchscreen gestures. In *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY.

[45] 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 International Conference on Intelligent User Interfaces*. ACM, New York, NY.

[46] PCB. 2017. PCB Conference Schedule: Pennsylvania Council of the Blind. Retrieved November 25, 2019 from http://pcb1.org/2017-pcb-conference-schedule/.

[47] Marko Periša, Ivan Cvitić, and Rosana Elizabeta Sente. 2017. Comparative analysis of mobile phone application solutions accessibility for informing visually impaired persons in traffic environment. In *Proceedings of the 25th International Symposium on Electronics in Traffic (ISEP'17)*.

[48] Giorgio Presti, Dragan Ahmetovic, Mattia Ducci, Cristian Bernareggi, Luca Ludovico, Adriano Baratè, Federico Avanzini, and Sergio Mascetti. 2019. WatchOut: Obstacle sonification for people with visual impairment or blindness. In *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'19)*. ACM, New York, NY.

[49] Jyri Rajamäki, Petri Viinikainen, Julius Tuomisto, Thomas Sederholm, and Miika Säämänen. 2007. LaureaPOP indoor navigation service for the visually impaired in a WLAN environment. In *Proceedings of the International Conference on Electronics, Hardware, Wireless, and Optical Communications*.

[50] Timothy H. Riehle, Shane M. Anderson, Patrick A. Lichter, William E. Whalen, and Nicholas A. Giudice. 2013. Indoor inertial waypoint navigation for the blind. In *Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, Los Alamitos, CA.

[51] Marti L. Riemer-Reiss and Robbyn R. Wacker. 2000. Factors associated with assistive technology discontinuance among individuals with disabilities. *Journal of Rehabilitation* 66, 3 (2000), 44–50.

[52] Alejandro Rituerto, Giovanni Fusco, and James M. Coughlan. 2016. Towards a sign-based indoor navigation system for people with visual impairments. In *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY.

[53] Edward K. Sadalla and Daniel R. Montello. 1989. Remembering changes in direction. *Environment and Behavior* 21, 3 (1989), 346–363.

[54] 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. *Transactions on Accessible Computing* 12, 3 (2019), Article 14.

[55] Barry M. Seemungal, Stefan Glasauer, Michael A. Gresty, and Adolfo M. Bronstein. 2007. Vestibular perception and navigation in the congenitally blind. *Journal of Neurophysiology* 97, 6 (2007), 4341–4356.

[56] K. Gnana Sheela and Subramaniam N. Deepa. 2013. Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering* (2013), Article 425740, 11 pages.

[57] Kristen Shinohara and Jacob O. Wobbrock. 2011. In the shadow of misperception: Assistive technology use and social interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*. ACM, New York, NY.

[58] Shankar Sivan and Gopu Darsan. 2016. Computer vision based assistive technology for blind and visually impaired people. In *Proceedings of the International Conference on Computing Communication and Networking Technologies*. ACM, New York, NY.

[59] D. Stathakis. 2009. How many hidden layers and nodes? *International Journal of Remote Sensing* 30, 8 (2009), 2133–2147.

[60] Simon Ungar. 2000. Cognitive mapping without visual experience. *Cognitive Mapping: Past, Present, and Future*, S. Freundschuh (Ed.). National University of Ireland.

[61] Tijana Vujicic, Tripo Matijevic, Jelena Ljucovic, Adis Balota, and Zoran Sevarac. 2016. Comparative analysis of methods for determining number of hidden neurons in artificial neural network. In *Proceedings of the Central European Conference on Information and Intelligent Systems.*

[62] Richard L. Welsh. 1980. Psychosocial dimensions. *Foundations of Orientation and Mobility*, R. L. Welsh and B. B. Blasch (Eds.). American Foundation for the Blind, New York, NY, 225–264.

[63] William R. Wiener, Richard L. Welsh, and Bruce B. Blasch. 2010. *Foundations of Orientation and Mobility.* American Foundation for the Blind, New York, NY.

[64] Michele A. Williams, Caroline Galbraith, Shaun K. Kane, and Amy Hurst. 2014. Just let the cane hit it: How the blind and sighted see navigation differently. In *Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility.* ACM, New York, NY.

[65] 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 ACM SIGACCESS Conference on Computers and Accessibility.* ACM, New York, NY.

[66] Lucy Yardley and Marsha Higgins. 1998. Spatial updating during rotation: The role of vestibular information and mental activity. *Journal of Vestibular Research* 8, 6 (1998), 435–442.

[67] G. L. Zacharias and L. R. Young. 1981. Influence of combined visual and vestibular cues on human perception and control of horizontal rotation. *Experimental Brain Research* 41, 2 (1981), 159–171.

Received April 2019; revised August 2019; accepted October 2019