

Modeling Expertise in Assistive Navigation Interfaces for Blind People

Eshed Ohn-Bar¹, João Guerreiro¹, Dragan Ahmetovic¹, Kris M. Kitani¹, Chieko Asakawa^{1,2}

¹Carnegie Mellon University, ²IBM Research - Tokyo
{eshed1, jpvguerreiro, dragan1}@cmu.edu, {kkitani, chiekoa}@cs.cmu.edu

ABSTRACT

Evaluating the impact of expertise and route knowledge on task performance can guide the design of intelligent and adaptive navigation interfaces. Expertise has been relatively unexplored in the context of assistive indoor navigation interfaces for blind people. To quantify the complex relationship between the user's walking patterns, route learning, and adaptation to the interface, we conducted a study with 8 blind participants. The participants repeated a set of navigation tasks while using a smartphone-based turn-by-turn navigation guidance app. The results demonstrate the gradual evolution of user skill and knowledge throughout the route repetitions, significantly impacting the task completion time. In addition to the exploratory analysis, we take a step towards tailoring the navigation interface to the user's needs by proposing a personalized recurrent neural network-based behavior model for expertise level classification.

Author Keywords

blind navigation; knowledge; experience; interface adaptation; user modeling; personalization; recurrent neural networks.

INTRODUCTION

Navigation assistive technologies play an important role in increasing the independence of blind people. Current solutions for indoor navigation are now able to provide practical localization accuracy and guidance [2, 6, 7]. Although such interfaces are often static in the instructions they convey, the user's information needs depend on the user's knowledge of the environment and navigation techniques [3]. In order to accommodate these dynamic needs, we envision interfaces that are able to automatically identify the user's behavioral states and expertise level. By focusing on the evolution of user behavior as they gain task expertise, we aim to provide insights into this relatively unexplored dimension of user-system interactivity for navigation assistance for blind people.

Even considering accurate navigation assistance, mistakes are common when traveling unfamiliar routes [1]. The lack of knowledge and confidence may lead users to hesitate in



Figure 1. We model the evolution of user expertise throughout repetitions of a navigation task with a smartphone-based turn-by-turn navigation guidance interface. Given motion of blind users along a planned path (shown for one participant for trial 1, T1, and trial 4, T4), the model learns to categorize expert behavior. In the example, the left and immediate right turn in a portion is often missed in T1, but not in T4, due to better knowledge of the interface and route.

complex areas, walk slowly, miss a turn due to a late reaction or have longer error-recovery times (Fig. 1). Previous research has shown that route knowledge can also be acquired through maps or virtual navigation [5, 10, 20], but its utility is better assessed with real-world navigation tasks [14]. However, route learning and familiarization is a process that takes time and it is difficult to conduct long-term, longitudinal user studies to assess its evolution.

In general, as a user interacts with an interface repeatedly, we expect gradual adaptation and improvement to occur. Nonetheless, the extent of this phenomenon is not trivial, in particular for navigation tasks. For instance, guided navigation is known to negate route learning as people become reliant on the system [8], but indoor navigation systems for blind people face significant localization and safety challenges. An intelligent navigation interface for blind people must facilitate a certain level of situational awareness in order to avoid missing a turn, bumping into obstacles and pedestrians, or getting lost.

By employing an assistive navigation interface, our study emphasizes expertise in Human-Machine Interactivity settings, as opposed to general mobility. The analysis is relevant to scenarios where the cognitive load of the task is shared between the user and the system, and this is reflected in the way expertise evolves. For instance, users do not need to remember certain information, such as the direction of instructed turns, but instead they can attend to learning challenging route areas.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IUI'18, March 7–11, 2018, Tokyo, Japan

© 2018 ACM. ISBN 978-1-4503-4945-1/18/03...\$15.00

DOI: <https://doi.org/10.1145/3172944.3173008>

As a step towards assistive navigation interfaces that can accommodate each user's unique walking and learning style, this paper presents: (1) expertise evolution assessment - the results of a real-world repetition study with 8 blind participants aimed to quantify the gradual increase of task knowledge, and (2) expertise model - a deep learning-based system for modeling user expertise that can be used in adaptive user interfaces for identifying expert behavior and contextually assisting the user in real-time. We find that model personalization is essential as expertise manifests differently among users. Specifically, appropriately modeling personal expertise leads to an improvement in classification accuracy of up to 17%.

USER STUDY

Our first step was to assess how route exposure affects the performance, knowledge and behaviour of blind users while using a turn-by-turn navigation guidance smartphone app. To our knowledge, this is the first study that quantifies the evolution of expertise through real-world assisted navigation of blind people.

Participants

We recruited 8 blind participants (4m/4f), who use a white-cane as the primary navigation aid. Their ages range from 43 to 76 ($M=65.63$, $SD=10.61$) years old. Two participants do not own a smartphone and five had experience with navigation apps. The study took approximately two hours per participant, who were compensated for their time (\$25 per hour).

Real-World Navigation and Apparatus

We used NavCog3 [21], an open-source¹ indoor navigation system that provides turn-by-turn instructions to guide a user to a destination. When the user reaches a turning point, the system provides a verbal instruction (e.g., "turn left") and a short vibration and sound effect. When the user completes the turn, achieving the correct orientation, the vibration and sound are provided again together with the next instruction (e.g., "proceed 50 feet and turn right"). While navigating that segment, the system provides periodic information about the remaining distance (e.g., "30 feet"), and an "approaching" message right before the turn. Additionally, the system provides information about relevant landmarks and Points of Interest (POIs) when the user is within close proximity (e.g., "a restroom is on your right").

We instrumented three buildings in our university campus ($58,800m^2$) with *iBeacons* and used an *iPhone 7* with the navigation app, which logged all events during the navigation tasks. Participants used their free hand to hold the smartphone and used *AfterShokz* bone-conductive headphones. Video was recorded for visualization and analysis purposes from a first-participant and third-person (researcher) perspective, using two *GoPRO* cameras.

Method and Procedure

The user study consisted of the repetition of two routes with similar complexities (number of turns and POIs), which order was counter-balanced among users. *Route A* has a total length

of 500 feet, 8 turns and 13 landmarks/POIs, while *Route B* has a length of 250 feet, 7 turns and 10 landmarks/POIs. For their first route, participants walked the same route six times. It included three consecutive trials with the navigation app, a trial from memory (using only the white cane), another trial with the app, and a second trial from memory. The unassisted trials intended to make participants reason about the route and to gain insights on their ability to leverage the knowledge previously acquired. For their second route, participants walked the same route four consecutive times with the app. We avoided the walk from memory due to fatigue constraints. Immediately after each trial, participants were asked to describe the route (as in [4, 15]) by referring its turns, POIs and landmarks. Then, they walked with the researcher to the starting point through a different path. Participants took a longer break after finishing all trials of the first route, but only minor breaks (up to 1 minute) between same-route trials.

Design and Analysis

The analysis focuses on the comparison of the two extremes of route expertise, the first (T1) and last (T4) trials with the navigation app. To assess user performance we rely on task **Completion Time**. We did not focus on navigation errors as those depend mainly on the system accuracy. In contrast, completion time reflects the fast/slow recovery from errors. In order to assess participants' route knowledge with their verbal descriptions, we used and adapted metrics from prior research [10, 18, 24]: (1) **FormElementsError** reflects the number of wrong, missing or extra turns and is analyzed using the Levenshtein distance [16] between the correct and participant's description; (2) **POIsMentioned** refers to the number of POIs and landmarks referred by the participant; (3) **POIsCorrect-Location** refers to the POIsMentioned in the correct segment and side of the route. We ran *Shapiro-Wilkinson tests* to all dependent variables to check for normality, and used a *Paired-Samples t-test* when the variables have a normal distribution and a *Wilcoxon Signed Rank test* otherwise.

EXPLORATORY ANALYSIS

Performance and Route Knowledge

Fig. 2 shows a gradual increase in performance, depicted by a decrease in *Completion Time* over the four trials. Results show significant differences between the first and last trial for both routes, with an average decrease of 45.2 seconds in *Route A* and 56.5 in *Route B* ($p < 0.05$ in both routes).

Route description metrics show that participants were able to significantly improve their knowledge about the route structure (**FormElementError**) from the first to the last trial, in both routes ($p < 0.005$). Although participants were never able to report the full route structure, they were only able to report an average of 26% of the route after one trial, against 69% after the last trial. This can be explained by participants being "focused on the instructions to guarantee I [they] reach the destination". As they got more familiar with the route, they were able to retain more of its information.

The rate of **POIsMentioned** by the participants was very similar among trials ($p > 0.05$ in both routes). Participants ended up referring mostly to landmarks that had a direct impact on

¹HULOP: <http://github.com/hulop>

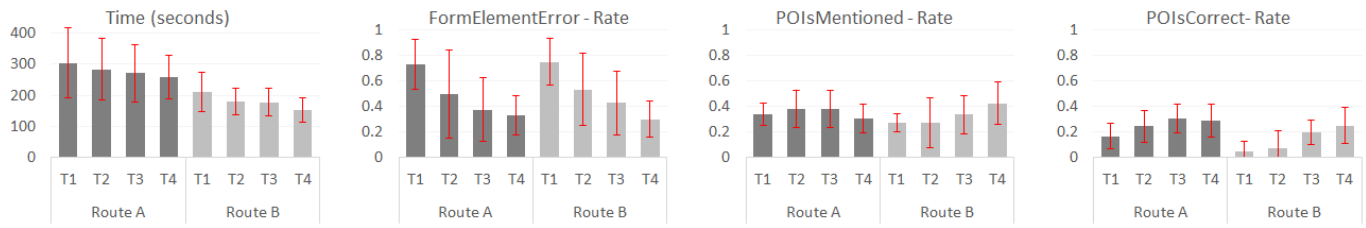


Figure 2. User performance metrics of completion time and route knowledge. Analysis is shown for each of the four trials and two routes in the study.

their navigation or landmarks/POIs close to the start or end of the route. For instance, the doors that users opened were mentioned in almost all trials, as well as floor changes next to a turning point. The rate of *POIsCorrect* increased slightly over time in *Route A* (n.s. $Z = -1.300, p = 0.194$), but significantly ($Z = -2.379, p < 0.05$) in *Route B*, showing a greater ability to incorporate the landmarks/POIs in the rest of the route.

The experimental procedure included walking the route from memory twice throughout the repetitions in order to further evaluate the evolving expertise. Although we found that most participants were not able to fully recall the routes, some participants were still able to complete the route when navigating from memory. One participant (out of four) completed *Route A* in their first unassisted trial, while three completed it on their second trial. All but one participant in *Route A* (and all in *Route B*) reached the last or second to last turn. In *Route B*, one and two participants completed the first and second unassisted trials, respectively.

Speed Analysis

The continuous location and orientation tracking provided by the phone and beacon signals enables us to analyze the impact of task knowledge on user walking style throughout the route. For instance, we can identify the frequency of user stopping or slowing down. Fig. 3 shows the walking speed distribution over all participants for each route and trial. Overall, average walking speeds increase over trials ($p < 0.001$), but more significantly on *Route A*, increasing from 0.54 meters per second (m/s) in T1 up to 0.64 m/s in T4. Since these speed values are computed over all time steps (including events of turning and stopping, resulting in low average speeds) and all participants, even small changes are significant. *Route B* has shorter segments and high density of turns and POIs, so users tend to walk slower. At times, we observed users walking more cautiously in this dense route in order to validate their route knowledge before continuing to move forward.

Throughout the repetitions, users were able to generally complete tasks faster by avoiding getting lost, not missing turns, and leveraging information about POIs and landmarks. Specifically, the number of hesitation events reported by the users decreased. Fig. 3(right) validates the reports, showing a reduction in the percentage of time users spent stopping or walking slowly (defined as any speed under 0.2 m/s). The observation that walking speed evolves in certain situations as users gain route navigation expertise motivated us to consider whether a sequence classification model can capture this subtle phenomenon. The model would have to account for our diverse group of participants, including their unique walking styles and mobility and orientation strategies.

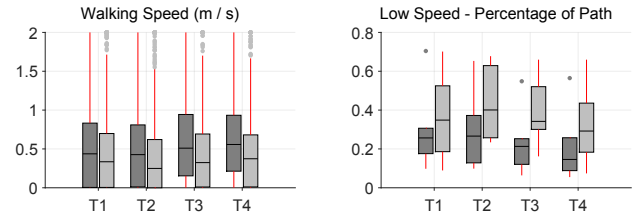


Figure 3. Box plots of walking speed (left) and percent of the path spent at low speed (right) over the four trials (color coding follows Fig. 2).

EXPERTISE CLASSIFICATION MODEL

With the overarching aim of developing navigation interfaces which can interpret the state of the user continuously, we propose a personalized behavior model for classifying route expertise.

Recurrent Neural Network Model

Given a sequence of observations of user walking and reaction patterns, we learn a model for user’s behavioral states and expertise level. Following state-of-the-art approaches for sequential behavior modeling [9, 17, 19, 22], a Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) hidden units [11] is trained. The model is trained for 100 epochs using Stochastic Gradient Descent (SGD) with momentum. In training, we use Dropout [23] and two 128-cell hidden layers.

Model Observations and Training

Our set of features is kept as general as possible to avoid overfitting to our specific settings, while reflecting our observations that user velocity changes throughout the trials. At each time step, the model observes sensor signals and events generated by the interface. The state of the user is parameterized by the tracked real-world position and orientation, $\mathbf{P} = [x, y, \theta]$. User movement is characterized by the linear velocity \mathbf{S} in m/s and the angular velocity \mathbf{A} in radians per second (r/s) components, derived from \mathbf{P} . The model also has access to the current instruction \mathbf{I} provided by the interface, encoded as a one-hot vector where each element in the vector corresponds to an instruction type. In the experiments, the instructions are clustered into four categories: approaching a turn, turn, proceed forward, and POIs/landmarks.

Modeling route expertise is formulated as classification between user behavior during T1, the initial navigation task, and T4, the last navigation task, for each route. The model generates route expert/novice labels, and these are compared with their trial number. To ensure the model generalizes well to different settings, each route is split into spatially disjoint sets used for cross validation. Over the entire collected dataset, sampling at 1 second intervals leads to a total of 5,147 samples.

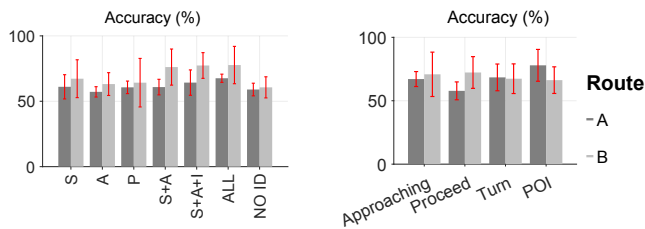


Figure 4. Behavior model performance for different feature combinations (left) and during different instruction events (right).

We found it crucial to explicitly personalize the expertise model. This is achieved by (1) adding the **person ID** as a one-hot vector to the feature space, and (2) modification of the training loss, so that in training time the 2-class log loss computed over the expertise label L_{expert} is combined with scalar multiple of the person ID classification log loss, $L = L_{expert} + \lambda L_{ID}$. The multi-task formulation can regularize training by encouraging the model to leverage person-specific cues [17, 12]. Adding the auxiliary training task will be shown useful for learning representations that can robustly generalize to unseen test cases.

Results and Analysis

Results for the behavior model in terms of normalized classification accuracy, computed as the average percentage of correctly classified instances for each class, are shown in Fig. 4. On the left, it studies expertise classification for different combinations of inputs to the model with the goal of analyzing the contribution of different cues for our modeling task.

A combination of all the proposed cues (**ALL**) is shown to produce the best classification results of 67.60% and 77.69% for *Route A* and *Route B*, respectively. We can observe how the velocity components $S + A$ alone account for a significant portion of the predictive power of the model. The model further benefits from incorporation of the instructional feature I , as it includes additional context regarding the evolution of expert behavior during specific navigation events. Interestingly, *Route B*, which contains a higher concentration of instructional and POI events, also provides more opportunities for the model to capture expert behavior (reflected by a higher classification accuracy). The position features provide a useful baseline, but result in only a small improvement when combined with the other features.

The person ID cue is shown to be essential for a flexible expertise model which can accommodate the large variations between T1 and T4 among participants. As shown in Fig. 4, without the explicit personalization in training (**NO ID**), the LSTM model performance deteriorates significantly, to 59.36% and 61.28% for *Route A* and *Route B*, respectively. This large drop in performance occurs with a model learned over the complete set of features (**ALL**). The need for personalization, in particular for our classification task, became clear during the user study. The large variability in mobility skills and learning styles makes expertise classification a challenging task; although the level of user task performance improved significantly between the initial and subsequent trials, the improvement and its rate was found to be highly personal. Therefore, it is not surprising that a non-personalized model is

struggling at the task of identifying how each user’s behavior changes along their own expertise spectrum.

A breakdown of the expertise classification results by different navigational instruction events is shown in Fig. 4. The figure analyzes user behavior changes in response to navigation instructions over the trials, as well as differences between the two routes. For instance, on *Route A*, the model is shown to perform best at POI/landmark-based instructions, such as doors or floor changes. Our data-driven model reveals how POI-related behavioral changes are a strong indicator for expert behavior, in particular on *Route A*. We can draw a similar conclusion for instructions of “approaching” a turn and “proceed” during navigation on the *Route B*. These results are in alignment with the findings in Fig. 2 and the unique properties of each route. On *Route A*, the common “proceed” instructions result in a more subtle difference between reactions of novice and experienced users. This is due to the several long forward segments throughout the route. Expertise classification of behavior during “turn” instructions is most consistent among the two routes when compared to other instructions. User behavior during turns is indeed more consistent with respect to the route properties, yet the model is still able to capture expertise-dependent behavior during turn events. This analysis demonstrates that our classification model is able to leverage several types of expertise-related cues specific to the route, the user, and the instructions provided.

CONCLUSIONS AND FUTURE WORK

This paper discussed challenges in deploying intelligent and adaptive assistive navigation interfaces for blind people. A real-world study with 8 blind participants explored how route expertise evolves with task repetition. Results show that both route knowledge and route completion time gradually improve. The results on timing and knowledge acquisition can be used to address some of the significant challenges faced by indoor navigation interfaces for blind people, including reduction in cognitive load and supporting the dynamic and personal needs of the user. The evolution of route expertise resulted in rich behavior characteristics which were captured with a personalized RNN model. The model was also used to gain insights into how user behavior changes with increased knowledge.

As future work, the exploratory analysis in this study can be complemented by the results of a real-world longitudinal study. The user performance findings can be used to compare with approaches that focus on other forms of route learning [10]. Although we kept our features as general as possible, additional cues, such as environmental context [13], may provide further classification performance improvements. The goal of our analysis was to compare two extremes of route expertise, yet we hope this study motivates future researchers to pursue additional general measures and models for user expertise level. Another next step would be to study how the interface should be modified to support each user and their identified expertise level.

ACKNOWLEDGMENTS

We thank the participants and the support of Shimizu Corporation, JST CREST (JPMJCR14E1) and NSF (1637927).

REFERENCES

1. Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: a navigational cognitive assistant for the blind. In *International Conference on Human-Computer Interaction with Mobile Devices and Services*.
2. Dragan Ahmetovic, Masayuki Murata, Cole Gleason, Erin Brady, Hironobu Takagi, Kris M. Kitani, and Chieko Asakawa. 2017. Achieving Practical and Accurate Indoor Navigation for People with Visual Impairments. In *Web for All*.
3. Nikola Banovic, Rachel L Franz, Khai N Truong, Jennifer Mankoff, and Anind K Dey. 2013. Uncovering information needs for independent spatial learning for users who are visually impaired. In *International ACM SIGACCESS Conference on Computers and Accessibility*.
4. Mark Blades, Reginald Golledge, R Jacobson, Robert Kitchin, and Yvonne Lippa. 2002. The effect of spatial tasks on visually impaired peoples' wayfinding abilities. *Journal of Visual Impairment & Blindness* 96, 06 (2002).
5. Julie Ducasse, Anke M Brock, and Christophe Jouffrais. 2018. Accessible interactive maps for visually impaired users. In *Mobility of Visually Impaired People*. Springer, 537–584.
6. Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2012. The user as a sensor: navigating users with visual impairments in indoor spaces using tactile landmarks. In *SIGCHI Conference on Human Factors in Computing Systems*.
7. Navid Fallah, Ilias Apostolopoulos, Kostas Bekris, and Eelke Folmer. 2013. Indoor Human Navigation Systems: A Survey. *Interacting with Computers* 25, 1 (2013), 21.
8. Elliot Fenech, Frank Drews, and Jonathan Z Bakdash. 2010. The Effects of Acoustic Turn-by-turn Navigation on Wayfinding. 54 (2010).
9. Srujana Gattupalli, Dylan Ebert, Michalis Papakostas, Fillia Makedon, and Vassilis Athitsos. 2017. CogniLearn: A Deep Learning-based Interface for Cognitive Behavior Assessment. In *International Conference on Intelligent User Interfaces*.
10. João Guerreiro, Dragan Ahmetovic, Kris M. Kitani, and Chieko Asakawa. 2017. Virtual Navigation for Blind People: Building Sequential Representations of the real-world. In *International ACM SIGACCESS Conference on Computers and Accessibility*.
11. Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation* 9, 8 (1997).
12. Hernisa Kacorri, Kris M. Kitani, Jeffrey P. Bigham, and Chieko Asakawa. 2017. People with Visual Impairment Training Personal Object Recognizers: Feasibility and Challenges. In *SIGCHI Conference on Human Factors in Computing Systems*.
13. 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 *SIGCHI Conference on Human Factors in Computing Systems*.
14. Rob Kitchin and RD Jacobson. 1997. 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 (1997), 360–376.
15. Orly Lahav and David Mioduser. 2008. Haptic-feedback support for cognitive mapping of unknown spaces by people who are blind. *International Journal of Human-Computer Studies* 66, 1 (2008), 23–35.
16. Vladimir I Levenshtein. 1966. Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet Physics Doklady*, Vol. 10. 707–710.
17. Zachary C Lipton, David C Kale, Charles Elkan, and Randall Wetzell. 2016. Learning to diagnose with LSTM recurrent neural networks. In *International Conference on Learning Representations*.
18. Romedi Passini, Guylene Proulx, and Constant Rainville. 1990. The spatio-cognitive abilities of the visually impaired population. *Environment and Behavior* 22, 1 (1990), 91–118.
19. Haşim Sak, Andrew Senior, and Françoise Beaufays. 2014. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In *Annual Conference of the International Speech Communication Association*.
20. Jaime Sánchez, Marcia de Borba Campos, Matías Espinoza, and Lotfi B Merabet. 2014. Audio haptic videogaming for developing wayfinding skills in learners who are blind. In *International Conference on Intelligent User Interfaces*.
21. Daisuke Sato, Uran Oh, Kakuya Naito, Hironobu Takagi, Kris M. Kitani, and Chieko Asakawa. 2017. NavCog3: An Evaluation of a Smartphone-Based Blind Indoor Navigation Assistant with Semantic Features in a Large-Scale Environment. In *International ACM SIGACCESS Conference on Computers and Accessibility*.
22. Harold Soh, Scott Sanner, Madeleine White, and Greg Jamieson. 2017. Deep Sequential Recommendation for Personalized Adaptive User Interfaces. In *International Conference on Intelligent User Interfaces*.
23. Nitish Srivastava, Geoffrey E Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *JMLR* 15, 1 (2014), 1929–1958.
24. Koji Yatani, Nikola Banovic, and Khai Truong. 2012. SpaceSense: representing geographical information to visually impaired people using spatial tactile feedback. In *SIGCHI Conference on Human Factors in Computing Systems*.