To be published in Joint International Conference on Digital Inclusion, Assistive Technology & Accessibility (ICCHP-AAATE) 2022

Non-Visual Access to an Interactive 3D Map

James M. Coughlan[1] [0000−0003−2775−4083], Brandon Biggs[1,2] [0000−0001−7380−1077] and Hui-ying Shen[1] [0000−0002−2195−7085]

[1] The Smith-Kettlewell Eye Research Institute, San Francisco, CA, USA
[2] Georgia Institute of Technology, Atlanta, GA, USA
`{coughlan, brandon.biggs, hshen}@ski.org`

Abstract. Maps are indispensable for helping people learn about unfamiliar environments and plan trips. While tactile (2D) and 3D maps offer non-visual map access to people who are blind or visually impaired (BVI), this access is greatly enhanced by adding interactivity to the maps: when the user points at a feature of interest on the map, the name and other information about the feature is read aloud in audio. We explore how the use of an interactive 3D map of a playground, containing over seventy play structures and other features, affects spatial learning and cognition. Specifically, we perform experiments in which four blind participants answer questions about the map to evaluate their grasp of three types of spatial knowledge: landmark, route and survey. The results of these experiments demonstrate that participants are able to acquire this knowledge, most of which would be inaccessible without the interactivity of the map.

Keywords: Assistive Devices, Accessibility, Augmented Reality, Audio Labeling, Visual Impairment, Blindness, Low Vision.

1 State of the Art and Related Technology

Tactile (2D) and 3D maps offer non-visual map access to blind or visually impaired (BVI) people, filling an important gap left by available navigation tools. GPS-enabled smartphone apps like Google Maps offer fully accessible turn-by-turn directions, but they don’t convey the spatial layout of the environment to users who can’t see the mobile device screen, which has been shown to hinder navigational understanding [1]. Other GPS-enabled apps such as BlindSquare[1] and Microsoft SoundScape[2] give real-time information about the user’s surroundings but don’t replace the information contained in a map – which can be explored freely offline, in a comfortable setting (such as one’s home or office), and without having to be physically present in the region represented by the map. Similarly, while verbal descriptions of routes can be read aloud using a screen reader, they don’t convey spatial layout information except in very simple cases.

---

[1] https://www.blindsquare.com/about/
[2] https://www.microsoft.com/en-us/research/product/soundscape/

An alternative approach to providing non-visual map access is the vibro-audio map (VAM) [2], which displays geographic information on a tablet device by issuing vibrations and sounds when the user's finger contacts a geographic feature such as a path or junction on the tablet screen. Digital auditory maps [3], which allow a user to explore a map on a computer (e.g., moving around the map by pressing the left, right, up and down arrow keys) and hear a combination of speech and sounds to indicate geographic information, are another promising approach to accessing maps non-visually. However, both of these approaches have the disadvantage that the user is unable to feel the map with their hands and explore the shapes of objects and routes on the map tactilely, which is what BVI users desire in an ideal map [4].

Raised-line tactile maps (such as the commercially available and customizable TMAP [5]) offer direct access to the spatial information encoded in a map, but important semantic information such as the names of streets and buildings is typically represented using braille. Space limitations usually necessitate the use of braille abbreviations coupled with an extensive key; users often have difficulty interpreting these abbreviations, which force them to alternate between the map itself and the key, thereby disrupting the process of exploring the map. 3D tactile maps are an alternative that facilitate easier learning of the environment represented by the map [6], but placing braille on 3D maps is time-consuming and expensive. Moreover, braille is inaccessible to the large majority of people with visual impairments who don't read it[3].

An effective way to circumvent the limitations of braille labels is to substitute them with audio labels, i.e., by making the map interactive, which may be accomplished in several ways. For instance, tactile maps may be overlaid on a touch-sensitive tablet (e.g., Touch Graphics T3[4]). Alternatively, special 3D models can be built with touch-sensing capabilities (such as those made by Touch Graphics), but these require a significant amount of custom hardware. By contrast, computer vision may be used to add audio labels to virtually any tactile or 3D map [7,8], with the benefit that an existing map can be used with little or no modifications to the map itself.

While studies such as [9] have explored the benefits of adding interactivity to tactile maps, little or no work has been done to our knowledge focusing on the impact of adding interactivity to 3D maps. Our study focuses on a 3D map of a playground, using audio labels provided by the CamIO system for audio labeling [8], extending the co-design in [4] and the work performed in [10]. This map was originally a cardboard prototype that allowed BVI participants to give feedback on what else they wanted in a 3D model. They requested more information around pathways, and details on the small walls that were located around the playground. An early version of CamIO was also used in this version to digitally label objects. The new version of the model (based on the 3D-printed prototype shown in Fig. 1) includes all the requested tactile features, and a significantly improved CamIO.

We explore the use of an interactive 3D map, in which pointing a handheld stylus at a location on the map triggers the announcement of a specific audio label about that location (for instance, pointing at a small rocking horse structure triggers the an-

---

[3] https://nfb.org/images/nfb/documents/pdf/braille_literacy_report_web.pdf

[4] https://www.touchgraphics.com/education/t3

nouncement “rocking horse”). The map is a 1/100 scale 3D-printed nylon model of an actual playground, and the interactive system contains over seventy audio labeled features, including playground structures and features such as paths. We conducted a pilot experiment with one blind participant to assess the functionality of the system and the design of the formal experiment, and made improvements based on this pilot experiment to both the functionality and the experimental design. Formal experiments were then conducted with four additional blind participants, who were tasked with answering questions about the 3D map, demonstrating their grasp of three types of spatial knowledge: landmark, route and survey.

2 Overview of Interactive 3D Map System

This section summarizes the functioning of the interactive 3D map system. Specifics about the user studies are presented in subsequent sections.

Fig. 1. 3D playground map shown with stylus held by user to point at a feature on the map.

Our study is based on a 3D-printed map (see Fig. 1) representing an actual playground in Palo Alto, California. The map measures roughly 59 cm x 76 cm horizontally, with the highest point about 14 cm. It is a 1/100 scale model of the playground containing over seventy features, such as “disk swings”, “climbing loops”, “playhouse”, “Ava’s bridge” and “climbing giraffe”. The 3D map was designed so that the shapes of the features are as familiar as possible and could either be recognized by touch, or are intuitive enough to be easily remembered once they are introduced.

To make the 3D map interactive, we used the CamIO system [8] to create audio labels for the features on the map. Using this approach, an iPhone is mounted rigidly above the 3D map so that the iPhone camera views it in its entirety. Computer vision algorithms run on the iPhone to analyze the scene, specifically determining where the tip of a handheld stylus (Fig. 1) lies in relation to the 3D map. Whenever the stylus tip touches a feature of interest (referred to as a “hotspot” in [8]), the corresponding audio label is read aloud using text-to-speech.

Next we describe the stylus in more detail. The stylus is made of a foam cube 3 inches wide, covered with barcode marker patterns printed on paper, and attached to a stick for the user to hold. Computer vision algorithms estimate the *pose* (3D translation and 3D rotation) of the stylus and of the 3D map (which is also framed with printed barcode marker patterns for this purpose) in each camera frame, allowing the system to track the 3D location of the stylus tip relative to the map. The stylus was designed with barcode marker patterns large enough to be clearly resolved by the camera, but small enough to be mounted on a pointing stick that is easily grasped. (We are currently experimenting with a more durable and compact 3D-printed stylus with a 2-inch-wide cube on top.)

Since the actual playground is organized into seven zones, which are groupings of related playground structures (such as the “Tot Zone” and “Swinging and Swaying Zone”), we created special audio labels for these zones. The audio label for each zone is triggered whenever the stylus tip falls inside a volume of space roughly 12 – 20 cm above the structures contained in the zone.

3 Pilot Experiment

Before embarking on formal experiments with BVI participants, we conducted a pilot experiment with one BVI participant (female, age 29, blind with very limited form perception) to determine how well the 3D interactive map works and to assess if any changes should be made to the system or our prototype experimental design.

3.1 Procedure

At the beginning of the pilot experiment, we briefly described the overall layout of the map and explained how to use the stylus to trigger the audio labels. Next we asked the participant to spend 5 minutes familiarizing herself with the map, and then asked her to perform three groups of tasks to assess three specific types of knowledge that can be obtained from maps [9,11]: (a) *landmark knowledge*, which is the awareness of specific locations in an environment; (b) *route knowledge*, which is knowledge about how to traverse one or more specific routes; and (c) *survey knowledge*, which is knowledge of how landmarks and routes relate to one another spatially. We note that survey knowledge is useful for creating robust mental maps of the environment, which facilitate independent navigation of the environment [9]. Spatial knowledge was assessed in multiple trials asking the participant to locate specific locations on the map, trace walking routes from one location to another, and specify cardinal directions (north/south/east/west) from one location to another.

Here we describe the spatial knowledge questions in detail. The landmark knowledge tasks asked the participant to find all seven zones, and each task required them to identify two structures in each zone by name (except for one specific zone that contained few structures). For each route task, the participant was given a starting feature and destination feature, and was told to find each feature and to trace the shortest walking path from the start to destination using their finger or the stylus (without walking over walls or through structures or other barriers). Finally, each survey task specified a starting feature and destination feature, and the participant was asked to find these features and then to indicate the direction (as the crow flies, i.e., along a straight path irrespective of barriers such as walls) that the destination was located in relative to the start. The starting and destination features were chosen so that the directions connecting them were always aligned to the four cardinal directions defined by the rectangular map; the participant was allowed to specify the direction by saying "north", "south", "east" or "west" or else "12 o'clock", "3 o'clock", "6 o'clock" or "9 o'clock".

Next the participant completed an SUS questionnaire [12] (with its language adapted to BVI participants [9]). Finally, the experimenter conducted a semi-structured interview in which the participant was asked what works well/needs improvement/works poorly in the system, how they use the stylus, how much information the shapes on the 3D map provide by themselves (without the audio labels) and how they might want descriptive information about each object on the map presented.

3.2 Results

The participant was indeed capable of finding some map features – and of reasoning about them spatially when the features were correctly located. However, in this pilot study the experimenter gave hints to her because it was very difficult for her to find many features (especially given that there are over seventy of them on the map and the system offered no way of guiding her to each feature). Indeed, the SUS score was 60, which indicates that the system was only partially usable. In the semi-structured interview, the participant reported that the overall concept of the interactive 3D map was sound, and that the features were well labeled, but that it was easy to get lost in the map, and noted confusion relating to certain features (such as slides) being present in multiple zones. She noted that she tended to alternate between exploring with both hands (without holding the stylus), and then picking up the stylus to obtain information about a specific structure. In addition, she suggested that detailed descriptions of a structure could be triggered by having the stylus tip dwell on the structure after the name was announced.

To address the challenge of having to search among over seventy structures, in the formal experiments we modified the zone announcements to include not only the zone names but also a listing of important features in each zone. We noticed that the zone announcements were not always triggered reliably, so we improved the CamIO software in such a way that not only increased how reliably zone announcements were triggered but also decreased the incidence of announcements being halted prematurely (due to noise in the estimation of the stylus tip location). Based on the performance of the participant in the pilot experiment, we set time-out periods for each task in the formal experiments (described in the next section).

Finally, the SUS was omitted in the formal experiments due to the difficulty of evaluating only the interactivity of the system (as opposed to factors beyond our control such as the design of the actual playground itself), and also to save time.

4 Formal Experiments

After the pilot experiment, we conducted formal experiments to confirm the usability of the improved interactive map system and to assess how using the system affects spatial learning and cognition.

4.1 Procedure

The formal experiments were conducted with four additional BVI participants ranging in age from 31-75 years old (2 male, 2 female, all blind: three with no form perception and one with some form perception in one eye). These experiments were similar to the pilot experiment, except that the audio labeling of zones was improved (as described in the previous section) and no hints were given to the participant during the formal trials. In the brief training phase of the experiment, we explained to the participants that the zone announcements functioned as directory listings, and could be used to locate specific structures more efficiently than exhaustively searching the entire map. We set the following time-out periods for the formal tasks: 2 minutes for each landmark task, 3 minutes for each landmark/route task and 3 minutes for each landmark/survey task.

4.2 Results

The results of the formal trials demonstrate that participants were able to acquire spatial knowledge (which would likely be inaccessible without the audio labels due to the absence of braille on the 3D map) *without hints* from the experimenter. (We speculate that this success was partly due to the listing of important zone features in the zone announcements, which enabled participants to search for features more efficiently than exhaustively searching the entire map.) Specifically, out of 23 total questions about spatial knowledge for each participant, the numbers of correct responses for the four participants were 22, 19, 14 and 21; we note that nearly all incorrect responses occurred when trials timed out (i.e., when the participant was unable to find a given feature in the allowed time).

We analyzed one component of this spatial knowledge statistically, the cardinal direction estimates, which relate specifically to survey knowledge. These estimates were reported in all trials that didn't time out, and they were *correct every time they were reported*. For each participant, the resulting 100% correct cardinal direction estimate rate is well above the chance success rate of 25% (implied by four possible cardinal directions). This is confirmed by a two-tailed binomial test [13] (with null hypothesis probabilities of 25% for correct estimates and 75% for incorrect estimates) for each participant, with all p-values equaling 2.5×10^{-4} or lower, ruling out the null hypothesis that the cardinal directions were estimated by chance.

In the semi-structured interviews, participants expressed positive feedback about the overall approach of the interactive 3D map, including the usefulness of the zone announcements, and noted that many of the structures had shapes that made them recognizable, or at least easy to remember once learned. Three participants said they alternated between exploring by hand without the stylus and using the stylus to get information about a structure of interest, while another said he preferred to hold the stylus while exploring with both hands. They described some issues that need improvement, including the difficulty fitting the stylus tip in some crowded locations of the map; important items (such as the drinking fountain) that have audio labels but are difficult to find since they are not listed in any zone announcement; audio label announcements that sometimes stutter or halt prematurely; how the large cube on top of the stylus sometimes makes it difficult to hold the stylus; and the need for explicit guidance to specific destinations. When asked for an appropriate way to trigger the announcement of detailed information about a feature, some suggested having the stylus tip dwell at a feature, while others suggested a stylus double-tap gesture.

5 Conclusions and Future Work

We have described experiments with BVI participants using an interactive 3D map that is a scale model of an actual playground. These experiments not only confirm the usability of the interactive map but also demonstrate the ability of the participants to non-visually learn spatial information represented by the map, assessed in terms of landmark, route and survey knowledge. The approach highlights the advantages of creating interactivity with computer vision, which allows audio labels to be created for an existing 3D model without having to modify the model physically (aside from attaching barcode markers to the corners of the model). We note that the audio labels can include as much information as desired, and they are fully accessible even to BVI people who don’t read braille.

Future work will include exploring the implementation of hand/finger tracking algorithms to eliminate the need for a stylus; offering audio labels in multiple languages; experimenting with an interface accessible to deaf-blind users; and providing audio instructions to guide the user to a desired destination on the map (perhaps using an approach similar to [14]). A bronze version of the interactive map is planned for permanent installation (including a stylus and rigidly mounted iOS device) at the actual playground. After the map installation is complete, we will perform user experiments to assess the impact that the interactive 3D map has on physical navigation of the playground.

6 Acknowledgments

The authors gratefully acknowledge support from NIH grant 5R01EY025332 and NIDILRR grant 90RE5024-01-00. We would like to thank the Magical Bridge Foundation for funding the development of the 3D-printed map.

References

1. Giudice, N.A. (2018). “Navigating without vision: Principles of blind spatial cognition.” In Handbook of Behavioral and Cognitive Geography. Edward Elgar Publishing.
2. Giudice, N.A., Guenther, B.A, Jensen, N.A., Haase, K.N. “Cognitive Mapping Without Vision: Comparing Wayfinding Performance After Learning from Digital Touchscreen-Based Multimodal Maps vs. Embossed Tactile Overlays,” Front. Hum. Neurosci., vol. 14, p. 87, 2020, doi: 10.3389/fnhum.2020.00087.
3. Biggs, B., Coughlan, J.M., Coppin, P. “Design and Evaluation of an Audio Game-Inspired Auditory Map Interface.” 25th International Conference on Auditory Display (ICAD 2019). Northumbria University, Newcastle-upon-Tyne, UK. June 2019.
4. Biggs, B. 2019. “Designing accessible nonvisual maps.” OCAD University. Retrieved from http://openresearch.ocadu.ca/id/eprint/2606
5. Biggs, B., Pitcher-Cooper, C., Coughlan, J.M. “Getting in Touch with Tactile Map Automated Production: Evaluating Impact and Areas for Improvement.” Journal on Technology and Persons with Disabilities. Vol. 10. 2022.
6. Holloway, L., Marriott, K., & Butler, M. (2018, April). “Accessible maps for the blind: Comparing 3D printed models with tactile graphics.” Proceedings of the 2018 CHI conference on human factors in computing systems (pp. 1-13).
7. Shi, L., McLachlan, R., Zhao, Y., Azenkot, S. (2016, October). “Magic touch: interacting with 3D printed graphics.” In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (pp. 329-330).
8. Coughlan, J.M., Shen, H., Biggs, B. “Towards Accessible Audio Labeling of 3D Objects.” Journal on Technology and Persons with Disabilities. Vol. 8. 2020.
9. Brock, A.M., Truillet, P., Oriola, B., Picard, D., Jouffrais, C. “Interactivity improves usability of geographic maps for visually impaired people,” Human–Computer Interact., vol. 30, no. 2, pp. 156–194, 2015.
10. Biggs, B., Coughlan, J. M., & Coppin, P. (2021). “Design and evaluation of an interactive 3D map.” Rehabilitation Engineering and Assistive Technology Society of North America, 2021.
11. Siegel, A.W., White, S.H. “The development of spatial representations of large-scale environments,” Adv. Child Dev. Behav., vol. 10, pp. 9–55, 1975.
12. Brooke, J. (2013). “SUS: a retrospective.” Journal of Usability Studies, 8(2), 29-40.
13. Howell, D.C. Statistical Methods for Psychology. (2012). Cengage Learning.
14. Coughlan, J.M., Biggs, B., Rivière, M.-A., Shen, H. “An Audio-Based 3D Spatial Guidance AR System for Blind Users.” 17th International Conference on Computers Helping People with Special Needs (ICCHP '20). Sept 2020.