ECHOEXPLORER™: A GAME APP FOR UNDERSTANDING ECHOLOCATION AND LEARNING TO NAVIGATE USING ECHO CUES

Wenyu Wu, Rudina Morina, Arley Schenker, Alexander Gotsis, Harsha Chivukula, Madeline Gardner, Felix Liu, Spencer Barton, Steven Woyach, Bruno Sinopoli, Pulkit Grover and Laurie M. Heller

Carnegie Mellon University
Pittsburgh, Pennsylvania

ABSTRACT

Echolocation -- the ability to detect objects in space through the perception of echoes from these objects -- has been identified as a promising venue to help visually impaired individuals navigate within their environments. The interest is in part because a proof-of-concept exists: certain visually impaired individuals are able to navigate using active echolocation. Why, then is echolocation is not in more widespread use among visually impaired individuals? It is possible that a lack of systematic echolocation training platforms has impeded individuals in picking up this skill. We designed a game-application that serves as a training platform for individuals, sighted or not, to train themselves to echolocate. Preliminary testing from both sighted and visually impaired individuals showed that users uniformly understood the game, although their enjoyment of the game was mixed. Although a number of game features could be improved, it is a promising training tool prototype for individuals learning to use echo cues for navigational purposes.

1. INTRODUCTION

Visual impairment is a significant and growing problem. The need to address this medical problem has led to the development of many systems that attempt to at least partially restore the visual sense. Many sensory substitution systems have been implemented as navigational aids for visually impaired individuals, such as those that use tactile inputs [1] or sounds [2] to substitute for or supplement vision. Our approach in particular has been inspired from studies that have shown the ability of some blind individuals to exhibit auditory navigation skills using echolocation. There is some evidence that this may be, in part, due to cortical reorganization and recruitment of visual areas for auditory processing [3].

https://doi.org/10.21785/icad2017.040

Echolocation -- the ability to locate objects in one's surroundings by sensing the reflections of an emitted sound -- is a promising venue for conveying visual information and aiding navigation of visually impaired individuals. Echolocation includes both passive and active processes, via sensing existing sound reflections and emitting one's own referent sound, respectively.

This raises a profound scientific question: *Why, then is active echolocation not in more widespread use among visually impaired individuals?* It is possible that this is due to lack of training. Currently, echolocation training is typically limited to self-training by trial and error or laborious training with another echolocator. Some research has used auditory virtual reality to simulate acoustic environments for studying echolocation [4], [5]. It is possible that many more individuals could train themselves to echolocate if a systematic platform were available to enable them to do so.

This work focuses on developing a mobile phone game-application that attempts to train individuals, sighted or not, to echolocate. The goals of the application are manyfold. First, it will help us assess whether echolocation is a skill that is amenable to being learned by sighted and visually impaired adults. Second, it will help obtain the science for developing systematic techniques to train people to echolocate. For instance, the artificial environment of a game-app can be made free of the audio and echo clutter of the real world, as well as make use of *physically unrealistic* echoes that one would not encounter in the real world but are nevertheless useful for training purposes.

2. BACKGROUND ON ECHOLOCATION AND AUDITORY PERCEPTION

Conveying spatial information through echoes relies on a few important aspects: (1) the time between the reference sound (such as a mouth click) and the echo, known as *echo delay*, (2) the difference between the echo timing for the right and left ear, known as the *interaural time difference* (ITD), (3) the difference between the sound levels entering the ears, known also as the *interaural level difference* (ILD), and (4) the angle-dependent alterations in the frequency spectrum produced by the head and external ears (pinnae) known as the head-related transfer function (HRTF). HRTFs help to externalize sounds presented over headphones and also provide some vertical location cues.

Echo delay is important in conveying object distance. If an object is further from the observer, the echo takes longer to reach the observer. Likewise, ITDs and ILDs are important in conveying direction information. If, for example, an object is located to the right of the observer, the echo will reach the right ear first [6]. Echo level is not as robust a cue to distance as echo delay because it varies not only with distance, but also with frequency and the absorption characteristics of the surrounding reflective surfaces.

3. GAME DESIGN AND IMPLEMENTATION

The goal of this application is to provide a training platform for learning echolocation. We designed a game that requires the user to navigate through various mazes using simulated echoes. An avatar is used to represent the current location of the user in the maze. At any time, the user can tap to instruct the application to play echoes based on the current location of the avatar within the maze. The application plays a pre-generated click followed by a realistic echo that conveys spatial information about the maze through echo delays, ITD, and/or ILDs. For instance, if the user is facing a close wall straight ahead, the referent click and the resulting echo will be heard in quick succession.

Mazes are carefully chosen such that specific navigation skills, such as learning when obstacles are straight ahead, to the left, or to the right of the avatar, can be learned in sequence. As the game progresses, the mazes and auditory cues become more sophisticated and thus more difficult to navigate. Throughout gameplay, the app collects data about the performance of the user, and these data are eventually transmitted in a secure, confidential manner to a server. This includes metrics like the number of times the user hits an obstacle and the time it takes the user to complete a particular maze. These data provide the foundation that will allow us to improve on the scientific understanding of echolocation.

To ensure user friendliness for visually impaired individuals, feedback from the Disability Services Center at Carnegie Mellon University as well as the Blind and Vision Rehabilitation Services of Pittsburgh was incorporated in design of the project, and will continue to inform our improvement of the app.

3.1 Programming environment

We used Unity, a cross-platform game engine which provided us with a software framework for the creation and development of the game [7]. Unity takes care of the majority of low-level infrastructure, so the developer can focus primarily on game design. Our current development prioritizes the Android version, but we expect the game to be available for iOS in the future.

Although Unity is well-suited to this project, it has some drawbacks. For example, Unity is incompatible with many smartphone accessibility applications, and its user interface has been designed for sighted users and developers. In addition, several of the control gestures used in the game are also meaningful in accessibility applications, which may be confusing. However, it was our conclusion that Unity's advantages far outweigh its disadvantages. Future testing will inform the best way to overcome the aforementioned drawbacks.

3.2 Maze design

The player's avatar (visually represented as an arrow) is inside a maze with walls and walkable areas (e.g., Figure 1). A green exit sign indicates the goal location that the avatar needs to reach to complete each level. This visual display is turned off in gameplay but can be helpful in debugging. The unit for player movement is one tile, and the entire maze is a 9 x 9 corridor-based grid system. The first few mazes that a user completes are simple to navigate, providing one novel echo type and/or requiring one novel navigational input (such as turning).

During this tutorial phase, users are given voice instructions drawing their attention to certain echo cues, providing hints after idling time or a crash into a wall, and familiarizing users with the gestures necessary for

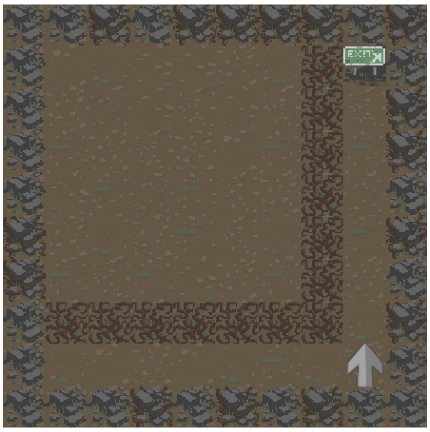

Figure 1. An example maze in the game, with two hallways. Player's avatar is indicated by the silver arrow. The level's exit is indicated by the green rectangle.

gameplay. This ensures that the user learns some basic navigational skills and begins to understand how echoes sound in simple scenarios. As the game progresses past the tutorial, the mazes become more sophisticated by using more complicated junctions and additional hallways. The game incorporates "structured randomness" to enable infinite gameplay levels.

3.3 Echo design and generation

Rather than generating echoes in real time, the game uses previously generated echoes that are loaded into the game for each physical situation that the avatar finds itself in. This avoids the computational overhead of generating sophisticated echoes in real time. Echo delay between the reference click onset and the first arriving echo was a reliable cue to the exact distance between the avatar and the wall ahead. In addition, we enhanced the sense of immersion in the game by using acoustic simulation and HRTFs. We used an individual HRTF [8]; however, combinations of HRTFs may have different acoustic properties than individual HRTFs and may be worth using in future versions [9]. Echoes varied according to distance from each of the surrounding walls and interactions at different types of hallway junctions so that users could distinguish left from right turns. These ambient cues to hallway location were updated every 2-3 blocks because of the many stored sounds required for all possible maze configurations. The exit location is represented as stairs, indicating that users must climb the stairs to get to the next level. The echoes generated adjacent to a set of stairs were judged to be sufficiently different from those adjacent only to walls, and were thus chosen to uniquely cue the exit location.

We used Odeon, a room acoustic simulation and measurement software, to artificially create echoes (Odeon A/S, Kongens Lyngby, Denmark). Users can input specific 3D room designs from which Odeon generates sounds using a ray tracing algorithm. We created several different hallways in Odeon to use in the game. In Odeon, we set the materials of the walls to 13mm plasterboard on frame, with a 100mm empty cavity. The floor and stairs materials were set to a wooden floor on joists. Within each of these rooms, a virtual cardioid sound source emitted a mouth click that had been previously recorded using a binaural microphone in our laboratory. The echoes produced by that click, but not the original click, were recorded from a location directly behind the sound source. We produced a set of echoes in response to this click for each tile location in the room. Using our own software written in Matlab™, these echoes were then placed at the appropriate time delay after the source click based on the distance from the front wall. We also used Matlab to further increase the loudness of each set of echoes to make the various echo cues obvious. Every 15 levels in the game, echo volume relative to initial click volume decreased by 2 dB to keep the game challenging. Further details regarding the creation and effectiveness of these sounds can be found in another paper in these proceedings [10].

3.4 Interface design and accessibility

Given the sophistication of the echoes in the game, users must wear headphones to play. Users navigate the game using a series of control gestures, which are introduced via a set of voice instructions in the tutorial phase of the game. Users have access to a menu, read vocally, through which they can play the tutorial, play the main game, or hear gesture instructions, and they can access this menu at any time. Within levels, the perspective of the game is first-person, so users swipe up to move one block forward no matter what compass direction they are facing. At each location, users tap the screen with one finger to hear the echo corresponding to that location. To turn right or left, users swipe in that direction. After swiping, to confirm their move, users hear a swoosh sound in the same ear as the direction they swiped. If they swiped up, users hear the sound in both ears. This sound is sufficiently different from the echoes. If a user crashes into a wall, they hear a crunch sound. When a user believes they have reached the exit location, denoted by a stairs echo, they use two fingers to tap once. If correct, users hear a congratulatory sound before starting the next level.

Though this game is intended for visually impaired individuals, a minimal amount of visual information is left for development purposes and to orient users with normal or partial vision. However, the game is designed to be played with all visuals turned off. This is expected to provide improved echolocation training by nearly equating the cues available to sighted and visually impaired individuals.

4. EVALUATION AND CONCLUSION

Survey data from 6 sighted participants using a beta version of the game indicated that it was initially well liked. After 15 hours of game play, participants were, on average, neutral about how fun it was and uniformly said that they would prefer not to continue playing the game. Data regarding the improvement of these individuals on the game and on a natural echolocation task are reported elsewhere in these proceedings [10]. All participants responded that they understood how to play the game and did not find the game frustrating. The control gestures were rated moderately easy to use.

Additionally, 8 visually impaired participants used an early version of the app for 5-10 minutes and provided survey data on their experience. On average, these users would play the game on their own (5.3 out of 7) and said it was fun (5.1 out of 7). They also provided qualitative feedback about that version of the game, which we took into consideration while making the current version. In combination with the ratings from the sighted users, these responses suggest that although the game could be made more engaging, it has promise as a convenient resource for supplemental echolocation training.

In conclusion, EchoExplorer™ is an app in development that has the potential to be an easy, accessible way for people to increase their awareness of echo cues. Future versions of the app may incorporate scoring or more complex obstacles to increase users' enjoyment of the game, as well as improve its compatibility with accessibility applications. If users find the game fun, they may use it for longer than they would a more boring training task. We hope to make future versions of this game useful to individuals who want to use echoes to be more aware of their environment for the purposes of navigation.

5. ACKNOWLEDGMENTS

We thank Sarah Kwan, Tejal Kudav, Kiran Matharu, Daya Lee, Jessica Kwon, Jacqueline Hon, Alan Lu, Aaron Steinfeld, Jeff Bigham, Chieko Asakawa, Catherine Getchell, Art Rizzino, and participants at the Pittsburgh Blind and Vision Rehabilitation Services, for helpful discussions, feedback, and pointers. We also thank Google Inc., the NSF Center for Science of Information (CSoI), CMU's SURG program, and the NSF REU program for their generous support.

6. REFERENCES

1. Y. Danilov, M. Tyler. Brainport: an alternative input to the brain. J Integr Neurosci. 2005;4: 537–550.
2. L. Kay. A sonar aid to enhance spatial perception of the blind: engineering design and evaluation. Radio and Electronic Engineer. 1974;44: 605.
3. O. Collignon, P. Voss, M. Lassonde, F. Lepore. Cross-modal plasticity for the spatial processing of sounds in visually deprived subjects. Exp Brain Res. 2009;192: 343–358.
4. B. F. G. Katz and L. Picinali, "Spatial audio applied to research with the Blind.," in *Advances in Sound Localization*, 2011, pp. 225–250.
5. D. Pelegrin-Garcia, M. Rychtáriková, C. Glorieux, and B. F. G. Katz, "Interactive auralization of self-generated oral sounds in virtual acoustic environments for research in human echolocation," in *Proceedings of Forum Acusticum 2014*, 2014.
6. J. Schnupp, I. Nelken, A. King. Auditory Neuroscience: Making Sense of Sound. MIT Press; 2011.
7. R. H. Creighton. Unity 3D Game Development by Example: A Seat-of-Your-Pants Manual for Building Fun, Groovy Little Games Quickly. Packt Publishing Ltd; 2010.
8. E. De Sena, N. Kaplanis, P. A. Naylor, T. van Waterschoot. Large-scale auralised sound localisation experiment. AES 60th International Conference 2016.
9. L. S. R. Simon, N. Zacharov, and B. F. G. Katz, "Perceptual attributes for the comparison of head-related transfer functions," *J. Acoust. Soc. Am.*, vol. 140, no. 5, pp. 3623–3632, 2016.
10. L. M. Heller, A. Schenker, M. Gardner, F. Liu, "Evaluating two ways to train sensitivity to echoes," in *ICAD Proceedings 2017*, 2017.