Three-dimensional Nonvisual Directional Guidance for People with Visual Impairments

SeungA Chung, Kyungyeon Lee, Sohyeon Park, Uran Oh

Department of Computer Science and Engineering

Ewha Womans University

Seoul, South Korea

{ewhacsa, ruddus716, cookiedough0911}@ewhain.net, uran.oh@ewha.ac.kr

Abstract-Conveying directional feedback is important for individuals who are blind or have limited visual acuity. However, most studies have focused on supporting two-dimensional guidance. In this work, we investigated the effects of different nonvisual feedback conditions for providing directional guidance in a three-dimensional space. We conducted a user study with six people who are blind or have low vision to investigate the effects of stereo sound (on vs. off) and feedback modalities (beeping vs. vibration vs. beeping+vibration). Participants were asked to point a series of virtual targets randomly appeared around them in 3D with a laser pointer as quickly as possible. Findings suggest that the presence of beeping sound have better performance in terms of task completion time and travel distance compared to when vibration feedback was provided without beeping sound, which was the least preferred condition. In addition, we found that the presence of stereo sound has no significant effect on the performance although it is preferred by most participants. This work can contribute to 3D navigation for people who are blind or have limited visual acuity.

Index Terms—directional guidance, audio feedback, haptic feedback, 3D environment, visual impairment

I. INTRODUCTION

Directional guidance can be a great help for people with visual impairments for spatial navigation or explorations. Thus a number of researchers investigated providing directional feedback via a nonvisual channel to individuals who are blind or who have limited visual acuity [5]-[7], [9], [12]-[14]. For instance, audio feedback was investigated for conveying coordinates information or direction on a touchscreen device to people with visual impairments [9], [14]. In addition, Strachan et al. [13] proposed GPSTune, a portable navigation system, that provides audio feedback which informs the remaining distance to the target destination with different levels of volume, and the walking direction with panning sound. Meanwhile, some researches studied haptic feedback with vibration [5], [6]. Hong et al. [6], for example, proposed a wristlet with vibration motors to guide the hand of people with visual impairments to find a specific target location on a 2dimensional surface such as a paper (e.g., printed map). Ertan et al. [5] also designed a vest with four vibration motors where each motor is mapped to one of the four cardinal direction convey walking directions towards the target destination while navigating a route. While a number of studies that aimed to provide directional guidance to people with visual impairments

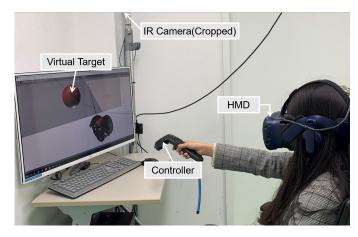


Fig. 1. The experiment setting for the virtual target pointing task in the study. Participants wore a head mounted device and held a controller with their dominant hand. Note that the monitor in the figure was set for the experimenter.

have focused on physical environments, informing directional guidance in a virtual space has been studied recently as well [11], [16].

Yet, the directional guidance supported by most of these studies for spatial navigation are limited to 2-dimensional space such as navigating a physical route with step-by-step instructions and following a path with a finger to reach a specific coordinate on a touchscreen. Thus, little has studied how design 3-dimensional directional guidance with nonvisual feedback for people with visual impairments beyond camera aiming assistance [2], [15].

To identify the effects if different feedback designs for conveying directional guidance with nonvisual feedback in 3D space, we designed and conducted a single-session withinsubject study with 6 participants with visual impairments. As shown in Fig. 1, participants were instructed to point a series of virtual targets that appears in a row at a random direction with a controller. We examined two factors in this study: (1) stereo sound originating from the target object, and (2) feedbackmode varying the type and combination of feedback modalities; beeping vs. vibration vs. beeping+vibration. The second factor is a proximity-based periodic feedback

where the frequency increases (the period shortens) as users' pointing direction get closer to the direction of the target object. We found that presence of beeping sound (beeping & beeping+vibration) have a statistically positive effect on task completion time over vibration feedback, and that travel distance, the length of the trace of the laser pointer during the task, as significantly shorter with beeping only compared to when only vibration feedback was provided. In terms of subjective preference, participants preferred the inclusion of beeping feedback reflecting the performance results. Interestingly, while stereo sound was found to have no significant effect on performance unlike the findings from a prior study with blind-folded sighted participants [3], it is preferred by most participants. The lessons learned from this study can contribute to the design of nonvisual 3-dimensional directional guidance for people with visual impairments such as finding a specific object on a shelf and underwater swimming.

II. RELATED WORK

A. Directional Guidance with Auditory Feedback

Many research studying the guidance in a 2D surface had been conducted. Leplâtre and Brewster [8] conducted research to distinguish whether providing audio feedback can help users easily navigate through a complicated menu structure in a mobile user interface. They discovered that by mapping around 150 different sounds to each individual function, participants assisted with sound feedback were able to complete navigating the menu task more successfully than other participants. Also, Oh et al. [9] examined various sound parameters such as pitch, volume, and stereo sound when teaching people with visual impairments the 2D gestures on touchscreen devices and discovered that it provides the best performance when stereo sound and pitch are each mapped to x and y coordinates respectively. Similarly, Su et al. [14] introduced Timbremap which is a system that navigates users on a touchscreen through auditory feedback to assist people with visual impairments to explore a floor plan or a map. High pitched sound indicated upward direction and the low pitched sound downward. On the other hand, Strachan et al. [13] presented GPSTune, a handheld system that enables users to way-find through audio. It guides users with audio feedback with the purpose of reducing their cognitive overload; volume indicated distance and panning sound to convey users the direction. Furthermore, Zhao et al. [16] presented a wearable VR controller, Canetroller, and introduced the potential of using audio feedback and haptic feedback together to assist people with visual impairments navigate.

Although it seems promising, these studies are limited to providing 2D directional guidance on a 2D surface or navigating with cardinal directions. Studies that rely on providing 3D directional information do exist such as, assisting people with visual impairments to properly aim a camera [2], [15]. However, they did not compare the task performance between haptic and audio feedback which is a distinguishable difference of our study from theirs.

B. Directional Guidance with Haptic Feedback

There are also several research that studied 2D guidance. Stearns et al. [12] developed an optical character recognition system that includes haptic and auditory feedback to assist people with visual impairments read each line in printed texts sequentially using a finger-mounted camera. Similarly, Hong et al. [6] introduced a haptic device worn at wrists to guide hands to trace paths on a 2D surface (e.g., a piece of paper, touchscreen). Some conducted research on studying haptic displays that provide directional information for way-finding [4], [5], [7], [10]. For example, Van Erp et al. [4] introduced a vibrotactile waist belt to assist people with waypoint-oriented navigation when they are in a visually limited situation. The main idea of this system was that distance is transformed into vibrational rhythm and the direction indicated a map that shows the location of the vibration. Likewise, Ertan et al. [5] proposed a wearable haptic navigation guidance system that reduces the amount of auditory attention required for people with visual impairments. This system has micromotors embedded in the back of the vest in a 4-by-4 array form and offers the user five types of instructions: cardinal directions and stop. Unlike most of the studies that use haptic feedback to provide directional guidance in a horizontal space for the purpose of way-finding, Katzschmann et al. [7] designed a smart white cane that has the ability to inform the location of the obstacles in both horizontal (left, right) and in vertical directions (high, low) and its distance from the user using haptic feedback. Furthermore, NaviRadar [10] used vibration feedback to provide directional information to users such as which way to walk and the remaining distance until the next crossing through differences in duration, rhythm, intensity, and roughness of the vibration feedback.

Even though providing distance information is beyond the original goal of this study, we intend to give 3D directional guidance in 3D space through the use of haptic and audio feedback. We had previously conducted a user study with 12 blind-folded participants to explore the effects of various feedback for 3-dimensional directional guidance [3]. We discovered that proximity-based discrete beeping feedback improves performance. It especially improved the task completion time and travel distance, regardless of combination with haptic feedback. Also, stereo feedback generated from where the target is located was effective in delivering the directional information. However, a limitation of our study was that the user study was conducted with sighted people instead of with people with visual impairments. Even though the participants were blind-folded, the difference between people with visual impairments and blind-folded sighted people still exists and this disparity could affect the findings such as task performance or preference. Moreover, we mapped the frequency of the proximity-based feedback (e.g., a beeping sound and vibration feedback) to a discrete range of the distance between the ray cast from the controller and the target. For such reasons, we decided to conduct a further study with idealized feedback designs to collect and examine the data

TABLE I FEEDBACK CONDITIONS VARYING Stereo and Mode were examined in this study.

Condition	Stereo	Beeping	Vibration
Beeping only w/ Stereo	On	On	Off
Vibration only w/ Stereo	On	Off	On
Beeping+Vibration w/ Stereo	On	On	On
Beeping only w/o Stereo	Off	On	Off
Vibration only w/o Stereo	Off	Off	On
Beeping+Vibration w/o Stereo	Off	On	On

TABLE II
PARTICIPANTS' AGE, GENDER, VISUAL ACUITY, ONSET YEARS OF VISUAL
IMPAIRMENTS, AS WELL AS PRIOR EXPERIENCE WITH VR.

PID	Age	Gender	Visual Acuity	Onset	VR
1	24	Male	Blind	Since birth	No
2	28	Male	Blind	Since birth	No
3	25	Male	Low vision	3 years	Yes
4	44	Male	Blind	30 years	No
5	45	Male	Low vision	Since birth	Yes
6	27	Female	Blind	10 years	No

of our previously intended target users, people with visual impairments.

III. USER STUDY

We conducted a target pointing study with 6 participants who are blind or have low vision to investigate how to design nonvisual feedback for informing specific directions in a 3D space focusing on audio and haptic feedback. It was an hourlong single-session within-subjects study.

A. Conditions

We had two factors: (1) presence of stereo sound generated from the target object (Stereo; 2-level: stereo on vs. stereo off), and (2) feedback modality varying the type and the combination of audio and haptic feedback (Mode; 3-level: beeping vs. vibration vs. beeping+vibration) as shown in Table I).

- *Stereo*: A pleasant music [1] is played in a loop constantly in stereo mode where the source of the sound is set to the location of the target.
- **Beeping**: Beeping sound is played in mono mode where the source is set to the top of participants' head. The beeping frequency increases as the pointing direction gets closer to the target.
- *Vibration*: The handheld controller vibrates with different frequencies similarly to the frequency of *Beeping*.

B. Participants

Six participants were recruited for this study. The demographic information is shown in Table II. The average age was 32.2 (SD = 9.7) ranged from 24 to 45. Four participants were completely blind and the remaining two participants had low vision. When asked about their prior experience with virtual reality (VR), two out of six participants reported that they have tried virtual reality before while the other 4 participants have

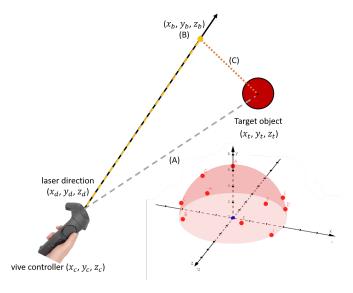


Fig. 2. (A) indicates the distance between the center of the target and that of the controller. (B) indicates a point in the direction of the laser from the controller (D) that is away from the controller with the distance of A. Finally, (C) is the distance between the center of the target and the point B. Red dots on the hemisphere at the bottom right shows locations of the targets that have equal distances from the origin where participants were seated.

not. All of them reported that they are right-handed, and that they do not have hearing difficulties.

C. Apparatus

For the experiment, we used a HTC VIVE Pro Eye, a head-mounted display (HMD), and a handheld controller. The head-mounted device was used (1) to track and log the orientation of the participants' head, and (2) to convey audio feedback (*i.e.*, stereo and beeping sound) via its built-in headset. Similarly, a handheld controller was used not only to monitor the pointing directions but to generate vibration feedback while participants are holding it with their dominant hand (*i.e.*, right hand) to complete the task.

As for the software, we have built a virtual reality application using Unity (version: 2019.2.17f1). While we used Audio Spatializer SDK by Unity to implement stereo sound. We defined the frequency of the beeping sound and vibration based on the distance between the center of the target and a point B, labeled as C in Fig. 2:

$$C = \sqrt{(x_t - x_b)^2 + (y_t - y_b)^2 + (z_t - z_b)^2}$$
 (1)

where,

$$B = (x_b, y_b, z_b) = (x_d, y_d, z_d) \times A$$
 (2)

and,

$$A = \sqrt{(x_t - x_c)^2 + (y_t - y_c)^2 + (z_t - z_c)^2}$$
 (3)

Note that the x-, y-, and z-coordinates of the controller and the target are denoted as (x_t, y_t, z_t) and (x_c, y_c, z_c) , respectively. Also, the direction of the laser emitted from the controller is denoted as a vector $D(x_d, y_d, z_d)$.

The frequency increases as C decreases for both beeping and vibration feedback. In addition, the pitch for beeping sound was set to reciprocal of the half of distance so that as the laser raycast gets closer to the target, it plays higher pitch sound relative to the original sound. While we updated the frequency and pitch of the beeping sound using Unity AudioMixer, we manually set the frequency of the vibration feedback to correspond to that of the beeping sound.

In all conditions, we provided additional sound and vibration feedback to indicate that participants has reached the target. When the controllers' laser raycast has just entered a target, the application played a chime sound. In addition, while the laser raycast is pointing at the target, the controller kept vibrating. A ding-dong sound played when participants push the button on the controller while pointing at the target to confirm that they have successfully found the target. The size of the targets were all the same shape and size (*i.e.*, a sphere with the radius of 0.5m). The distance between each target and the participant was also identical (*i.e.* 6m) where the coordinates were predefined by sampling 10 coordinates from the surface of the upper semi-spheres as shown in Fig. 2.

The application was run on a desktop computer with a AMD Ryzen 7 1700 (CPU) with a 16GB of RAM and a RTX2080 graphic card. The log was saved with timestamps for all trials for the analysis.

D. Procedure

After signing the consent form, participants' personal information such as age and gender were collected. Then, after explaining briefly about our study, we asked the participants to sit on a chair where the location is fixed but enables participants to rotate 360 degrees to changing the direction they are facing as shown in Fig. 1. Next, we asked them to wear a head mounted device, and handed the controller to their dominant hand. They were instructed to listen to the directional guidance and point the a virtual target with their controller and push its button to confirm the direction as fast as possible. Each participant tried all six conditions where the order was counterbalanced using a balanced Latin square design. For each condition, we first provided a practice session with a single target to help participants get familiar with each feedback followed by the actual task with 10 targets in a row that appeared at one of the 10 predefined locations in a random order. After completing all conditions, we collected subjective feedback.

E. Data and Analysis

Overall, we collected 360 data (6 participants × 10 targets × 6 feedback conditions). However, for the analysis, three values beyond 3SD from the mean were excluded for task completion time. We used two-way ANOVA to evaluate the interaction effect between *Stereo* and *Mode* in terms of task completion time and travel distance for finding each target. For post-hoc analyses, pairwise comparisons were used. In addition, participants' subjective responses such as preference were collected.

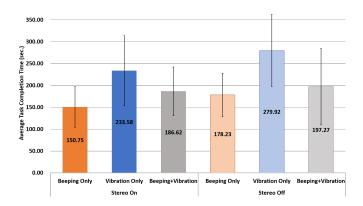


Fig. 3. The average task completion time per condition in seconds. Error bars indicate standard errors.

IV. FINDINGS

A. Task Completion Time

The result of task completion time is shown in Fig. 3. Two-way ANOVA with factors of *Stereo* and *Mode* revealed that the differences between three feedback conditions were statistically significant ($F_{(2)} = 10.62$, p < .001). Pairwise posthoc tests showed that participants were significantly faster with beeping only and beeping+vibration compared to vibration only condition (p = .004 and p = .013, respectively). On the other hand, there was no significant difference between beeping only and beeping+vibration feedback. The main effect of *Mode* and the interaction effect between the two factors were not found to be significant.

B. Travel Distance

In addition, we analyzed the length of the laser trace which is computed based on the sum of all Euclidean distances between every two successive coordinates of B in Fig. 2. Similar to the results of task completion time, there were no significant interaction effect between Stereo and Mode nor the main effect of Stereo, but the main effect of Mode was found to be significant ($F_{(2)} = 7.95$, p = .002. Pairwise posthoc analyses revealed that participants tend to have shorter traces with Stereo than Stereo than Stereo to have significant differences.

C. Trace Analysis

We further analyzed participants' laser traces to check if there is characteristics when trying to point the direction towards a target in 3D space. As shown in Fig. 5, we found that participants have a tendency to scan the environment in a horizontal direction first then adjust the pointing direction with stereo feedback as in prior study [3]. We assumed that this behavior is due to the nature of the stereo sound where left and right is easier to distinguish than top to bottom. Interestingly, however, we could observe the similar behavior even when stereo feedback was not present.

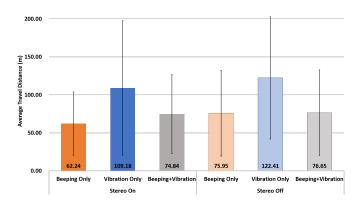


Fig. 4. The average travel distance per target for each condition in meters. Error bars indicate standard errors.

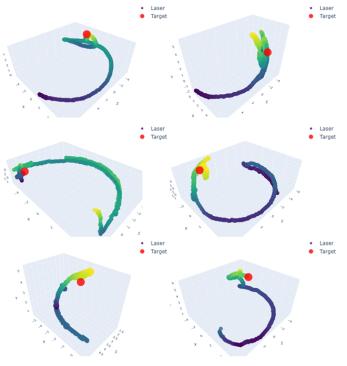


Fig. 5. Example traces during target pointing task from each participant; P1 to P6 from top to bottom, left to right) with stereo and beeping sound feedback. Red circles indicate the target.

D. Preference

After completing the task with all conditions, we collected participants' preferences and the results are shown in Fig. 6.

1) Most and least preferred: All participants chose conditions with beeping sound with or without vibration feedback as their favorite. On the other hand, All of them reported that their least preferred condition was vibrationonly condition especially when stereo feedback is not provided. Participants who preferred beeping sound without vibration feedback believed that vibration feedback is rather redundant because the information that is derived from each feedback is quite similar and the beeping sound is more instinctive than vibration feedback. On the contrary, the other participants preferred to



Fig. 6. The number of responses for the most (dark green) and least (light green) preferred feedback conditions (*N*=6).

have different kinds of information given simultaneously since they could rely on more than one sense.

2) Presence of stereo sound: Their preference towards stereo feedback was equally divided. However, five out of six participants commented that having stereo feedback did help them in estimating the direction of where the target is based on their head orientations, especially with the horizontal directions (e.g., left or right), reflecting the trace analysis results. On the other hand, the remaining P4 mentioned that stereo sound was distracting which made it harder for him to find the correct pointing direction and thus decided to ignore the stereo feedback. In addition, although not tested, all participants noted that stereo sound only without any other proximity-based feedback (i.e., beeping feedback or vibration feedback) will not be sufficient for completing the task.

V. DISCUSSIONS

The findings and lessons learned from the study are summarized below.

A. People with and without visual impairments

We expected that the findings of our study conducted with people with visual impairment share similarities when compared to the study with blind-folded sighted participants in terms of the task performance [3]. However, there were some differences except that we could also confirm that the performance is the lowest when only vibration feedback was given. For instance, the presence of stereo sound feedback did not have positive effect on the task performance (*i.e.*, task completion time and travel distance) with our participants while it had significant effect on the performance for sighted peers. Likewise, while blind-folded sighted participants preferred stereo sound, PVI were neutral about it. Furthermore, it did not show a statistical difference even when the vibration feedback was provided with the beeping sound compared to when only the beeping sound was given unlike the prior study.

Although direct comparison cannot be made as our feedback design is slightly different, the performance and preference between two groups could be due to the varying level of exposure to audio feedback. To be more specific, as people with visual impairments are more familiar with perceiving and recognizing audio feedback than sighted participants in general. In addition, they are more used to audio feedback than haptic feedback. On the other hand, as blind-folded sighted participants could be unfamiliar with both sound and haptic feedback, having multiple feedback modalities could be considered as useful.

B. Two-Step Guidance

Reflecting the prior study with blind-folded sighted participants [3], we could monitor the behavior of horizontal scanning of the surroundings to find a target in 3D space from the participants with visual impairments. This suggest that two-step directional guidance could actually be an intuitive approach for conveying 3D directional guidance. To be more specific, we inform the users the horizontal direction as the first step, then provide the directional guidance in vertical direction as the next step. This would be also preferred by people with visual impairments or anyone who prefer single feedback modality as we can convey one direction at a time instead of delivering both horizontal and vertical directions at once. Still, further investigation is needed as this two-step process may require additional overhead as it could only guide users to follow the path based on Cartesian distance rather than Euclidean distance which would always be longer if not the same.

VI. LIMITATIONS

Our study has several limitations, mainly as a single-session study. First, there were only six participants and thus larger sample size is needed to show that the findings of this study can be generalized. Also, while we were not able to find noticeable performance difference between participants who are blind and who have low vision, visual feedback from the HMD device could have affected the results. In addition, while we ran the experiment in a virtual environment for accurate tracking of participants' head orientation and pointing directions relative to the target object, one should consider how a target guiding feedback system can be implemented in natural settings such as identifying the target, tracking users' head and hand with minimum use of hardware devices.

VII. CONCLUSION AND FUTURE WORK

To better understand the effects of different nonvisual feed-back design, we conducted a user study with six participants with visual impairments where task was to point a virtual target in a 3D space as quickly as possible. Similar to the prior findings, our findings showed that all participants showed better performance in terms of the task completion time and travel distance with beeping sound than vibration feedback. However, unlike the previous study results with blind-folded sighted participants, stereo sound where the source is location of the target was not found to have significant effect on the performance. Yet, it is preferred by most participants. Moreover, we confirmed that participants' tendency to find

the direction of the target with horizontal scanning. Based on the findings learned and confirmed from this study, we plan to extend this work by conducting this study with more number of participants focusing on supporting two-step 3D directional guidance with proximity-based audio feedback with various parameters such as volume and pitch.

REFERENCES

- Ping.wav. https://freesound.org/people/edsward/sounds/341871/. Accessed: 2021-01-25.
- [2] D. Ahmetovic, D. Sato, U. Oh, T. Ishihara, K. Kitani, and C. Asakawa. Recog: Supporting blind people in recognizing personal objects. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–12, 2020.
- [3] S. Chung, K. Lee, and U. Oh. Investigating three-dimensional directional guidance with nonvisual feedback for target pointing task. In 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 206–210(to be published), 2020.
- [4] J. B. V. Erp, H. A. V. Veen, C. Jansen, and T. Dobbins. Waypoint navigation with a vibrotactile waist belt. ACM Transactions on Applied Perception (TAP), 2(2):106–117, 2005.
- [5] S. Ertan, C. Lee, A. Willets, H. Tan, and A. Pentland. A wearable haptic navigation guidance system. In *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No. 98EX215)*, pp. 164–165. IEEE, 1998.
- [6] J. Hong, A. Pradhan, J. E. Froehlich, and L. Findlater. Evaluating wrist-based haptic feedback for non-visual target finding and path tracing on a 2d surface. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 210–219, 2017.
- [7] R. K. Katzschmann, B. Araki, and D. Rus. Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(3):583–593, 2018.
- [8] G. Leplâtre and S. A. Brewster. Designing non-speech sounds to support navigation in mobile phone menus. 2000.
- [9] U. Oh, S. Branham, L. Findlater, and S. K. Kane. Audio-based feedback techniques for teaching touchscreen gestures. ACM Transactions on Accessible Computing (TACCESS), 7(3):1–29, 2015.
- [10] S. Rümelin, E. Rukzio, and R. Hardy. Naviradar: a novel tactile information display for pedestrian navigation. In *Proceedings of the 24th* annual ACM symposium on User interface software and technology, pp. 293–302, 2011.
- [11] A. F. Siu, M. Sinclair, R. Kovacs, E. Ofek, C. Holz, and E. Cutrell. Virtual reality without vision: A haptic and auditory white cane to navigate complex virtual worlds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2020.
- [12] L. Stearns, R. Du, U. Oh, C. Jou, L. Findlater, D. A. Ross, and J. E. Froehlich. Evaluating haptic and auditory directional guidance to assist blind people in reading printed text using finger-mounted cameras. ACM Transactions on Accessible Computing (TACCESS), 9(1):1–38, 2016.
- [13] S. Strachan, P. Eslambolchilar, R. Murray-Smith, S. Hughes, and S. O'Modhrain. Gpstunes: controlling navigation via audio feedback. In Proceedings of the 7th international conference on Human computer interaction with mobile devices & services, pp. 275–278, 2005.
- [14] J. Su, A. Rosenzweig, A. Goel, E. de Lara, and K. N. Truong. Timbremap: enabling the visually-impaired to use maps on touch-enabled devices. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, pp. 17–26, 2010.
- [15] M. Vázquez and A. Steinfeld. An assisted photography framework to help visually impaired users properly aim a camera. ACM Transactions on Computer-Human Interaction (TOCHI), 21(5):1–29, 2014.
- [16] Y. Zhao, C. L. Bennett, H. Benko, E. Cutrell, C. Holz, M. R. Morris, and M. Sinclair. Enabling people with visual impairments to navigate virtual reality with a haptic and auditory cane simulation. In *Proceedings of* the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–14, 2018.