Reaching Out: Investigating Different Modalities to Help People with Visual Impairments Acquire Items

Sooyeon Lee, Chien Wen Yuan, Benjamin V. Hanrahan, Mary Beth Rosson, John M. Carroll
Center for Human Computer Interaction, College of Information Sciences and Technology
Pennsylvania State University
University Park, PA 16801
{sul131, tuy11, bvh10, mrosson, jcarroll}@ ist.psu.edu

ABSTRACT
We present a lab study of multiple feedback designs for guiding small-scale arm-and-hand movement for people with visual impairments (PVI), so that they can reach out to and grasp an item on a shelf. Little attention has been paid to the guidance of small-scale arm-and-hand movements by PVI, yet this is an essential element of product acquisition in a grocery shopping task and other similar daily activities. We developed a feedback interface that allowed us to explore two types of auditory feedback (speech and tones), haptic vibration feedback, and a combination of both. The result of the study demonstrated that the multi-modal navigational feedback, specifically speech and haptic, was the most effective and preferred mode for small-scale navigation.

CCS Concepts
• Human-centered computing~Empirical studies in accessibility • Human-centered computing~Accessibility technologies

Keywords
Auditory feedback; haptic feedback; multimodal feedback; people with visual impairment

1. INTRODUCTION
Finding a wanted object and obtaining it with small-scale arm-and-hand motion is a frequently occurring activity in our daily lives; most of us would not think twice about grabbing a bottle from the pantry. Sighted people understand their relative locations to the surroundings and easily adjust their positions to accomplish various tasks. However, for people with visual impairments (PVI), reaching out and grabbing an object is quite challenging, and is even more so when the environment is unfamiliar (e.g., a store rather than a home). Activities like locating a glass of water, reaching out, and grabbing it from the dinner table, or retrieving a product from a grocery shelf requires careful and accurate arm-and-hand navigation guidance. Unless PVI are in a familiar space, they must either depend on assistance for moving their body or parts of their body to carry out such tasks, or use other sensory input (e.g., auditory or haptic) to understand what to do.

ACM ISBN 978-1-4503-4926-0/17/10.
https://doi.org/10.1145/3132525.3134817

Most existing research on PVI assistance has studied large-scale whole-body movement; little investigation has focused on small-scale movements of the arm and hand. Because such movements are both small in scale and brief in time, it is critical that PVI receive timely delivery of helpful information. We have been studying the nature and modality of such feedback as part of a larger project aiming to leverage state-of-the-art computer vision technologies in assisting PVI in daily tasks such as shopping. In this brief paper we report a study of five different directional feedback signals using a combination of sensory modes; we assessed the effectiveness of each signal type in guiding small-scale, arm-and-hand movements to reach to and grasp an object.

2. RELATED WORK
Previous research has investigated the presentation of directional information via the auditory channel (e.g., speech or non-speech like sonification), the haptic modality, or a combination of both [4][6][7]. In general, multi-modal feedback was found to be more effective, as predicted by multi-resource theory [8]. These existing studies examined support of large-scale body navigation, such as information needed in way-finding guidance. We contribute by exploring similar issues for small-scale, arm-and-hand movement navigation needed in object acquisition, such as in VizLens [2], GIST [5], ABBI [9], Kim et al., [3], Hong et al. [1], and Zhao et al. [10].

We extend previous research in three ways: 1) we focus on small-scale arm-and-hand movement; 2) we contrast multiple feedback options; and 3) we study grocery shopping as a context, identified by PVI as one of their most challenging daily tasks.

3. EXPERIMENT

3.1 Prototype and Feedback Design
To study the effectiveness of different modalities in small-scale arm-and-hand guidance, we used a prototype composed of a haptic glove and a bone conduction headset. The former was used for delivering haptic stimulation and the latter for conveying auditory feedback. The bone conduction headset enables PVI to receive not only the directional information but also ambient sounds with which they need to understand the surroundings.

We experimented with five different types of feedback: speech, non-speech, haptic vibration, a combination of speech and haptic, and a combination of non-speech and haptic. Synthesized voice was used for the speech feedback; tones (beeping sounds at different frequencies) were generated as non-speech feedback. Haptic feedback consisted of vibration signals delivered to different locations on the hand (left, right, top, bottom). Each type of feedback was mapped to six arm+hand adjustments (left, right, up, down, stop, reach forward) and two larger-scale movements (step forward, step backward). To make the feedback as intuitive and distinguishable as possible, we used spatial mapping (e.g., signal display on the left location for presenting a direction left) and various patterning with using different parameters such as duration and frequency. All feedback types were delivered in a continuous rhythm (e.g., sending continuous beeps to indicate moving left or right until the participant reaches the correct angle/position).

3.2 User Study

We used a Wizard of Oz (WOz) protocol to simulate computer vision and corresponding feedback; the human agent was required due to the instability of the current computer vision algorithm. We created a grocery shelf mockup in the lab and evaluated the five feedback conditions using eleven PVI (M:5; F:6) recruited from the local chapter of National Federation of the Blind (NFB). A research team member performed as the WOz operator using a pre-defined script.

For each trial, we positioned the participants in front of the product within a reachable distance. We provided training for each type of feedback at the beginning of each trial; we allowed participants to repeat the training as many times as needed. All participants performed 15 trials (3 trials X 5 feedback conditions). We used a Latin-square design to counterbalance the order of feedback conditions experienced by each participant.

We recorded the completion time for each trial along with number of errors made before object acquisition (errors included both incorrect product retrievals and an incomplete task). We also recorded the number of training requests. After the 15 trials, the participants rated each feedback condition on a 7-point Likert scale (1 = Not at all helpful and 7 = Extremely helpful); a semi-structured interview was conducted to gather qualitative data about experiences using each type of feedback.

4. RESULTS

4.1 Performance

A one-way ANOVA on completion times revealed significant differences among feedback conditions: $F(4, 160) = 4.84$, $p < .001$. A post-hoc analysis with Bonferroni confirmed significant differences ($p < .05$) between the vibration-only and the speech + vibration feedback; between tones + vibration and speech + vibration feedback. Another one-way ANOVA test with trial orders showed that no evidence of a learning effect across the three repetitions of each feedback condition.

4.2 Perceived Preference and Helpfulness

The analysis of preference ratings led to the following rank order: 1) speech-only ($M = 6.8$, $SD = 0.4$); 2) speech/haptic combination ($M = 6.4$, $SD = 1.5$); 3) haptic-only ($M = 5.7$, $SD = 1.3$); 4) tones-only ($M = 4.9$, $SD = 1.6$); and 5) tones/haptic combination mode ($M = 4.8$, $SD = 1.7$). Consistent with the performance result, the interviews revealed that a speech/haptic combination was most preferred; speech-only was second; and tones/haptic and haptic-only mode were least preferred.

Note that speech-only was rated most positively, but the interview responses placed it in second place. This may have been due to expectations for lowered effectiveness in noisy settings such as a grocery store. Participants told us they liked the combination of speech and haptic because the haptic feedback reinforced the speech and these two channels complement each other. The participants found the tones and haptic-only harder to discriminate and confusing at times.

5. CONCLUSION

We designed five feedback signals using speech, haptic, and non-speech information. We ran a study to examine effectiveness of each feedback mode for guiding arm-and-hand movement for reaching out and picking up items on grocery shelves. Our mixed method research showed that PVI performed best with a combination of speech and haptic feedback and it was also reported as most preferred in participants' comments. Our work extends the body of work on multi-modal feedback for PVI.

6. ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under Grants No. 1317560.

7. REFERENCES

[1] Hong, Jonggi, Lee Stearns, Tony Cheng, Jon E Froehlich, David Ross, and Leah Findlater. "Evaluating Angular Accuracy of Wrist - Based Haptic Directional Guidance for Hand Movement," n.d., 195–200.

[2] Guo, A., Chen, X. A., Qi, H., White, S., Ghosh, S., Asakawa, C., & Bigham, J. P. (2016, October). VizLens: A robust and interactive screen reader for interfaces in the real world. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (pp. 651-664). ACM.

[3] Kim, K., Ren, X., Choi, S., & Tan, H. Z. (2016). Assisting people with visual impairments in aiming at a target on a large wall-mounted display. *International Journal of Human-Computer Studies*, *86*, 109-120.

[4] Fiannaca, A., Apostolopoulous, I., & Folmer, E. (2014, October). Headlock: a wearable navigation aid that helps blind cane users traverse large open spaces. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility* (pp. 19-26). ACM

[5] Khambadkar, Vinitha. "GIST : A Gestural Interface for Remote Nonvisual Spatial Perception," 2013, 301–10.

[6] Merabet, L., & Sánchez, J. (2009). Audio-based navigation using virtual environments: combining technology and neuroscience. *AER Journal: Research and Practice in Visual Impairment and Blindness*, *2*(3), 128-137.

[7] Paneels, S., Anastassova, M., Strachan, S., Van, S. P., Sivacoumarane, S., & Bolzmacher, C. (2013, April). What's around me? Multi-actuator haptic feedback on the wrist. In World Haptics Conference (WHC), 2013 (pp. 407-412). IEEE.

[8] Wickens, C. D. (1991). Processing resources and attention. *Multiple-task performance*, *1991*, 3-34.

[9] Wilson, G., & Brewster, S. A. (2016, October). Using Dynamic Audio Feedback to Support Peripersonal Reaching in Young Visually Impaired People. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 209-218). ACM.

[10] Zhao, Y., Szpiro, S., Knighten, J., & Azenkot, S. (2016, September). CueSee: exploring visual cues for people with low vision to facilitate a visual search task. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 73-84). ACM.