Augmenting the dark: Exploring assistive micro-guidance in sonified mixed reality

Renan Guarese[1,2], **Ron van Schyndel**[1], **Deb Polson**[1], **Fabio Zambetta**[1]

[1]RMIT – Royal Melbourne Institute of Technology
124 La Trobe St, Melbourne VIC 3000, Australia

[2]KTH – Royal Institute of Technology
Brinellvägen 8, Stockholm 114 28, Sweden

{renan.guarese,ron.vanschyndel,debra.polson,fabio.zambetta}@rmit.edu.au

Abstract. *This thesis proposes a series of user evaluations of spatialized sonification methods rendered as AR in simulated and real-life scenarios. It proposes and promotes next-generation micro-guidance methods for low-visibility and vision-impaired (VI) scenarios. In 2D hand-guidance, results (N=47) outlined that sound spatiality methods had the most promising performance in time taken and distance from target. When assessing vertical hand-guidance in a 3D task (N=19), results indicated a significantly higher accuracy for a novel height-to-pitch method. Finally, a significant disparity was found between VI (N=20) and sighted (N=77) people regarding sighted people's empathy with the VI community. After an AR blindness embodiment experience, sighted people's (N=15) empathetic and sympathetic responses towards said community significantly increased. Ultimately, this thesis evaluates how audio AR can help users to have accurate and safe performances in day-to-day manual tasks.*

Resumo. *Esta tese propõe uma série de avaliações de usuários de métodos de sonificação espacializados renderizados como Realidade Aumentada (RA) em cenários simulados e reais. São propostos métodos de micro-orientação para cenários de baixa visibilidade e deficiência visual (DV). Na navegação manual 2D, os resultados (N=47) destacaram que os métodos de espacialidade sonora tiveram o desempenho mais promissor no tempo gasto e na distância ao alvo. Ao avaliar a navegação manual vertical em uma tarefa 3D (N=19), os resultados indicaram uma acurácia significativamente maior para um novo método relacionando altura com tom. Finalmente, encontramos uma disparidade significativa entre pessoas com DV (N=20) e pessoas videntes (N=77) no que diz respeito à empatia das pessoas videntes com a comunidade com DV. Após uma experiência de incorporação da cegueira em RA, as respostas empáticas e simpáticas das pessoas videntes (N=15) em relação à referida comunidade aumentaram significativamente. Em suma, esta tese avalia como o áudio em RA pode ajudar usuários a ter desempenhos precisos e seguros em tarefas manuais do dia-a-dia.*

1. Introduction

Following the modality-centric representation of guidance proposed by Gacem et al. [Gacem et al. 2014] as a taxonomy, this thesis is motivated by the exploration of a gap

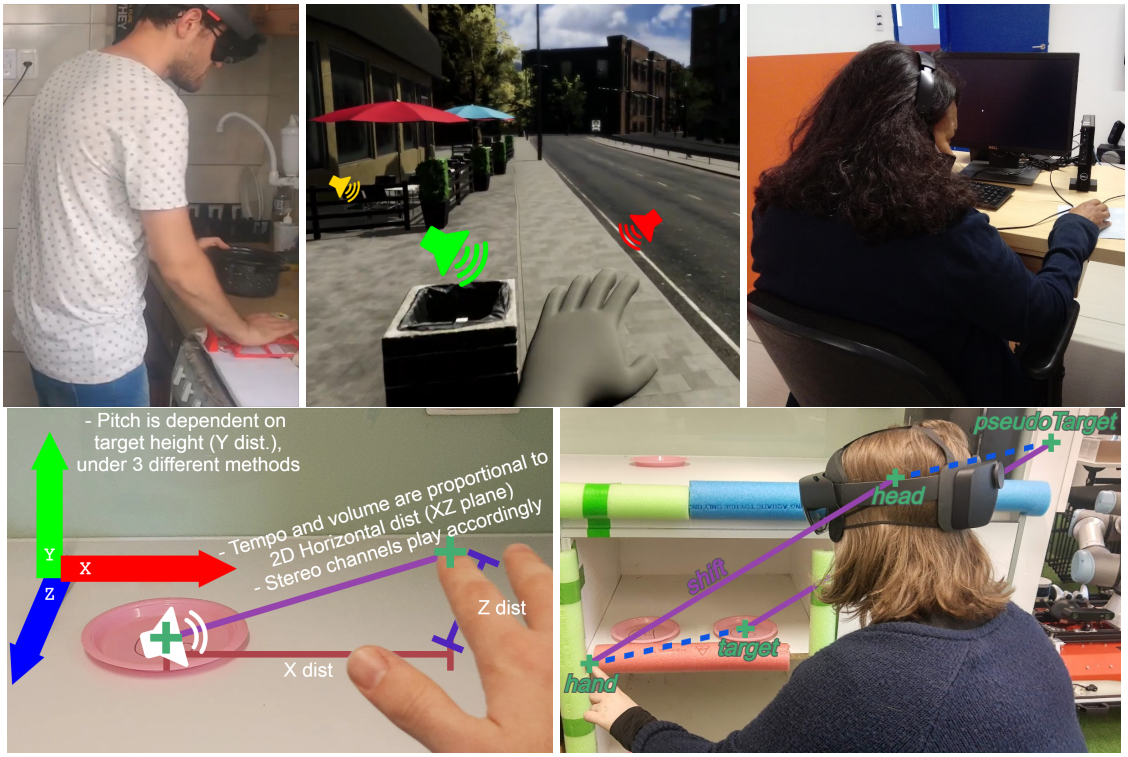

Figure 1. Top-left: a blindfolded user in a 2D guidance task receiving spatialized sound feedback from an AR HMD. Top-center: representation of spatialized sonification. Top-right: a VI user evaluating 2D sonification methods. Bottom-left: representation of a 3D sonification interface. Bottom-right: a blindfolded user in a 3D guidance task receiving spatialized sound feedback from an AR HMD. Vectors represent the 3D transformation performed on target audio beacons in order to attain hand-based spatial sound.

in auditory egocentric micro-guidance. Given the granularity nomenclature of such design space, *micro* guidance - as opposed to *macro* [Jacobson 1996] - involves pointing the user to the precise location of an object that is already within their field of view, "such as finding a jar of jam when staring at an open refrigerator" [Gacem et al. 2014].

Within the concepts of auditory interfaces, sonification is the technique of rendering sound in response to data and interactions, without relying on speech or music [Hermann et al. 2011]. Through the use of sensor fusion [Fritsche. et al. 2016], modern-day AR headsets are able to run spatial localization and mapping algorithms [Durrant-Whyte and Bailey 2006] that help in locating themselves in space. Given that these are wearable devices, egocentric sonification interfaces can be developed to inform users of their required conditions: their current space, their place in space, and how any action alters their position relative to the known locations in the environment [Jacobson 1996].

2. Research questions

RQ1: *Are VI people's performances affected by different sonification patterns in 2D hand-based micro-guidance?*

Several studies have successfully explored macro-guidance with the use of GPS and turn-by-turn speech feedback [Khan et al. 2021] in outdoor settings. In indoor set-

tings, where GPS cannot be relied on, speech has also been investigated to some extent as an audio navigation interface [Liu et al. 2018, Coughlan et al. 2020]. Speech allows direct guidance through semantics, with promising results, granted that only a limited number of cue options are available [Guerreiro et al. 2023]. However, speech is prone to misinterpretation and to the blocking of natural environmental audio [Mascetti et al. 2016], being arguably more limited than sonification [Hermann et al. 2011], which is deemed as the appropriate interface in less restrictive environments [Guerreiro et al. 2023]. Given the scarcity of studies using sonification as an egocentric micro-guidance interface for people who are VI, this is believed to be a valuable research gap, being explored in **RQ1**.

2.1. Overview of Methods and Results

A user evaluation of seven sonification methods is conducted in 2D manual micro-guidance tasks, which can be used as building blocks for spatialized audio in AR to model next-generation guidance aids for people who are VI. The methods are tested in comparable interactive sonifications of 2D positions in a series of hand-navigation assessments with VI and blindfolded sighted users, to validate the different approaches in environments without any visual feedback. Results (N=47) highlight that dynamic tempo and spatiality can be useful resources in sonified micro-guidance, and that users accustomed to faster-than-regular audio speed replay tend to have more precise performances, while musical literacy only affects performance on methods highly dependent on aural skills [Guarese et al. 2022]. Ultimately, this thesis corroborates the notion that sonification may help VI users perform better in day-to-day 2D navigation tasks, e.g. retrieving items from a pantry, handling kitchen appliances, and properly discarding trash.

RQ2: *Are people's performances affected by different vertical sonification patterns in 3D hand-based micro-guidance in low-visibility scenarios?*

Given the historic focus on macro-navigation in audio interfaces, most tasks that have been explored were topologically simple enough to be interpretable in a 2D space, such as exocentric map-based guidance [Gacem et al. 2014]. Although essential in indoor micro-guidance, verticality is an aspect rarely introduced into audio guidance studies [May et al. 2019]. Following the standards in height-to-pitch sonification [Hermann et al. 2011], the few studies that explore the vertical axis map it as a direct relationship of the audio pitch to object elevation [Marquardt et al. 2018, Hu et al. 2022], without exploring any alternatives. With the intention to explore this research gap in low-visibility conditions, this study compares multiple vertical audio feedback methods in **RQ2**.

2.2. Overview of Methods and Results

An evaluation of pitch-based sonification methods via user experiments in real-life scenarios is proposed, particularly for vertical guidance, with the aim of standardizing the use of audio interfaces in AR in micro-guidance tasks. Building on the literature on assistive technology for people who are VI, it is aimed to generalize their applicability to a broader population and in different use cases. A proposal and evaluation of sonification methods for vertical guidance is made in a series of hand-navigation assessments with users without visual feedback. Including feedback from a visually impaired expert in digital accessibility, results (N=19) outline that methods that do not rely on memorizing pitch have the most promising accuracy and self-reported workload performances [Guarese et al. 2024].

Ultimately, it is argued that audio AR can improve user performance in different 3D guidance scenarios, from video games to finding objects in a pantry.

RQ3: *Can sighted people's empathy with VI people be affected by an embodiment of difference AR experience?*

Taking into account the difficulty of recruiting people who are VI, researchers oftentimes include blindfolded sighted users into their experiments [Liu et al. 2018, Coughlan et al. 2020]. This can be helpful in order to analyze the performance and behavior of people with or without lived experience, and to assess the statistics of a greater population in the results. Noticing how sighted participants would try to empathize with the embodied condition of blindness during these trials, it was decided to explore this scenario from an empathic computing perspective [Billinghurst 2021]. VR has been popularly adopted as the "ultimate empathy machine" by evangelists [Myers 2018] since it enables users to be immersed in another person's view. This has led researchers to assume such experiences affect people's cognitive empathy [Schrier and Farber 2021] prior to proper evidence supporting such claims. In an effort to tighten this gap, this thesis explores how embodied AR experiences affect empathy in **RQ3**.

2.3. Overview of Methods and Results

To promote empathy with people who have disabilities, a multi-sensory interactive experience is developed that allows sighted users to embody having a visual impairment while using a novel assistive technology. The experiment involves blindfolded sighted participants interacting with a variety of sonification methods in order to locate targets and place objects in a real kitchen environment. An inquiry is made about the perceived benefits of increasing said empathy from the VI community. To test people's state empathy, an established emotional response scale [Escalas and Stern 2003] was adapted to gather sighted people's self-reported and perceived empathy with the VI community, respectively from sighted (N = 77) and VI people (N = 20), exposing a significant disconnect between these demographics [Guarese et al. 2023]. Results from re-testing sighted people's empathy after the experiment (N = 15) reveal that their empathetic and sympathetic responses tend to significantly increase. Furthermore, survey results suggest that the VI community believes the use of these empathy-evoking embodied experiences may lead to the development of new assistive technologies.

References

Billinghurst, M. (2021). Empathic computing and human robot interaction. In *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '21, page 5, New York, NY, USA. Association for Computing Machinery.

Coughlan, J. M., Biggs, B., Rivière, M.-A., and Shen, H. (2020). An audio-based 3d spatial guidance ar system for blind users. In Miesenberger, K., Manduchi, R., Covarrubias Rodriguez, M., and Peňáz, P., editors, *Computers Helping People with Special Needs*, pages 475–484, Cham. Springer International Publishing.

Durrant-Whyte, H. and Bailey, T. (2006). Simultaneous localization and mapping: part i. *IEEE robotics & automation magazine*, 13(2):99–110.

Escalas, J. E. and Stern, B. B. (2003). Sympathy and Empathy: Emotional Responses to Advertising Dramas. *Journal of Consumer Research*, 29(4):566–578.

Fritsche., P., Kueppers., S., Briese., G., and Wagner., B. (2016). Radar and lidar sensorfusion in low visibility environments. In *Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics - Volume 2: ICINCO*, pages 30–36. INSTICC, SciTePress.

Gacem, H., Bailly, G., Eagan, J., and Lecolinet, E. (2014). A design space of guidance techniques for large and dense physical environments. In *Proceedings of the 26th Conference on l'Interaction Homme-Machine*, IHM '14, page 9–17, New York, NY, USA. Association for Computing Machinery.

Guarese, R., Pretty, E., Fayek, H., Zambetta, F., and van Schyndel, R. (2023). Evoking empathy with visually impaired people through an augmented reality embodiment experience. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pages 184–193.

Guarese, R., Pretty, E., Renata, A., Polson, D., and Zambetta, F. (2024). Exploring audio interfaces for vertical guidance in augmented reality via hand-based feedback. *IEEE Transactions on Visualization and Computer Graphics*.

Guarese, R., van Schyndel, R., and Zambetta, F. (2022). Evaluating micro-guidance sonification methods in manual tasks for blind and visually impaired people. In *OzCHI '22: Proceedings of the 34th Australian Conference on Human-Computer Interaction*.

Guerreiro, J., Kim, Y., Nogueira, R., Chung, S., Rodrigues, A., and Oh, U. (2023). The design space of the auditory representation of objects and their behaviours in virtual reality for blind people. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2763–2773.

Hermann, T., Hunt, A., and Neuhoff, J. (2011). *The Sonification Handbook*. Logos Publishing House.

Hu, X., Song, A., Wei, Z., and Zeng, H. (2022). Stereopilot: A wearable target location system for blind and visually impaired using spatial audio rendering. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:1621–1630.

Jacobson, R. D. (1996). Talking tactile maps and environmental audio beacons: An orientation and mobility development tool for visually impaired people. *Proceedings of the ICA Commission on maps and graphics for blind and visually impaired people*, 96:1–22.

Khan, S., Nazir, S., and Khan, H. U. (2021). Analysis of navigation assistants for blind and visually impaired people: A systematic review. *IEEE Access*, 9:26712–26734.

Liu, Y., Stiles, N. R., and Meister, M. (2018). Augmented reality powers a cognitive assistant for the blind. *eLife*, 7:e37841.

Marquardt, A., Kruijff, E., Trepkowski, C., Maiero, J., Schwandt, A., Hinkenjann, A., Stuerzlinger, W., and Schöning, J. (2018). Audio-tactile proximity feedback for enhancing 3d manipulation. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, VRST '18, New York, NY, USA. Association for Computing Machinery.

Mascetti, S., Picinali, L., Gerino, A., Ahmetovic, D., and Bernareggi, C. (2016). Sonification of guidance data during road crossing for people with visual impairments or blindness. *International Journal of Human-Computer Studies*, 85:16–26. Data Sonification and Sound Design in Interactive Systems.

May, K. R., Sobel, B., Wilson, J., and Walker, B. N. (2019). Auditory displays to facilitate object targeting in 3d space. In *Proceedings of the 2019 International Conference on Auditory display*, pages 155–162. International Community for Auditory Display.

Myers, M. (2018). The empathy machine? virtual reality and the search for a solution to violence. In *Griffith Review*. Accessed: 2024-04-08.

Schrier, K. and Farber, M. (2021). A systematic literature review of 'empathy' and 'games'. *Journal of Gaming & Virtual Worlds*, 13(2):195–214.