SoundModVR: Sound Modifications in Virtual Reality for Sound Accessibility

Xinyun Cao
University of Michigan, Ann Arbor
xinyunc@umich.edu

Dhruv Jain
University of Michigan, Ann Arbor
profdj@umich.edu

ABSTRACT

Previous VR sound accessibility work have substituted sounds with visual or haptic output to increase VR accessibility for deaf and hard of hearing (DHH) people. However, deafness occurs on a spectrum, and many DHH people (e.g., those with partial hearing) can also benefit from manipulating audio (e.g., increasing volume at specific frequencies) instead of substituting it with another modality. In this demo paper, we present a toolkit that allows modifying sounds in VR to support DHH people. We designed and implemented 18 VR sound modification tools spanning four categories, including prioritizing sounds, modifying sound parameters, providing spatial assistance, and adding additional sounds. Evaluation of our tools with 10 DHH users across five diverse VR scenarios reveal that our toolkit can improve DHH users' VR experience but could be further improved by providing more customization options and decreasing cognitive load. We then compiled a Unity toolkit and conducted a preliminary evaluation with six Unity VR developers. Preliminary insights show that our toolkit is easy to use but could be enhanced through modularization.

CCS CONCEPTS

• **Human-centered computing** → Accessibility; Accessibility technologies.

KEYWORDS

Accessibility, virtual reality, deaf and hard of hearing, sound, customization

ACM Reference Format:
Xinyun Cao and Dhruv Jain. 2024. SoundModVR: Sound Modifications in Virtual Reality for Sound Accessibility. In *The 37th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '24), October 13–16, 2024, Pittsburgh, PA, USA*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3672539.3686754

1 INTRODUCTION

Prior work in VR sound accessibility for deaf and hard of hearing (DHH) users have focused on substituting sounds with visual or haptic outputs [6, 8, 10], such as closed captions for in-game dialogs or vibrations to represent environmental explosions. While promising for some specific sounds, visual and haptic feedback could lead to information overload [5] and hinder accurate information delivery due to bandwidth differences with auditory stimuli [4, 13].

Moreover, not all DHH users require complete substitution of sounds. Deafness occurs on a spectrum, and DHH individuals could have different hearing levels [1, 18]. Some users can hear sounds to some extent. For these users, the application can use sound modification techniques like increasing volume or shifting frequencies to audible ranges to deliver sound information seamlessly. Indeed, such customization may offer a more intuitive experience than a complete sensory substitution, as indicated by DHH participants in prior evaluations of VR sound substitution systems [6, 9].

In this demo paper, we explore modifying and customizing sounds in VR to better support the needs of users with partial hearing.

2 THE SOUNDMODVR TOOLKIT

Inspired by features explored in previous work, such as sound prioritization [15], directional sound enhancement [14], frequency-specific gain adjustment [2, 12], and spatial sound localization [8, 11], we designed 18 sound modification tools that allow developers to incorporate sound accessibility into their apps. Developers can use any subset of tools for their apps, assigning some to activate automatically during gameplay and some to be manually toggled by the users. Our toolkit is open-sourced on GitHub and is described below. Please refer to our supplementary video for a demo of the tools.

2.1 Prioritization Tools

Speech Prioritization (PT1) lowers the volume of co-occurring environmental sounds during important speech.

Group Prioritization (PT2) allows users to focus on sounds from one group and reduce the volume of all other surrounding sound groups.

Keyword Prioritization (PT3) allows users or developers to assign keywords to monitor, which, when detected, plays a notification sound. It also restores the volume of the spoken content to its original level if other tools have lowered it.

Direction-Based Prioritization (PT4) amplifies the sounds within the 10-degree arc on each side in the direction the user faces while simultaneously reducing the volume of sounds coming from other directions.

2.2 Parameter Modification Tools

System Frequency/Volume Adjustment (PM1) enables users to shift the frequency range of sounds, as well as adjust the volume system-wide.

UIST Adjunct '24, October 13–16, 2024, Pittsburgh, PA, USA

ACM ISBN 979-8-4007-0718-6/24/10
https://doi.org/10.1145/3672539.3686754

Sound Frequency/Volume Adjustment (PM2) enables users to adjust the frequency and volume of individual sound sources.

Frequency Contrast Enhancement (PM3) adjusts the frequencies of adjacent sound sources, elevating one while lowering the other to enhance their distinction.

Speech Speed Adjustment (PM4) allows users to adjust the speed of individual speech sources.

Beat Enhancement (PM5) boosts the rhythm of music sounds by dynamically increasing and decreasing the volume along with the beats.

2.3 Spatial Assistance Tools

Left-Right Balance (SA1) enables users to adjust the system sound balance to either the left or right, thereby equalizing stereo sounds and amplifying the sound on their preferred side.

Shoulder Localization Helper (SA2) provides auditory cues ("To your left" or "To your right") and captions to indicate the direction of important in-game sounds.

Hearing Range Adjustment (SA3) allows users to adjust the range for sound activation, enabling them to choose desired audio from various spatial sources based on distance.

Sound Distance Assistance (SA4) aids in perceiving distance by modulating sound pitch based on the user's proximity to the source: pitch decreases as distance increases and vice versa.

Live Listen Helper (SA5) isolates the sound from a source when nearby, muting all others. Users can move it around the scene to isolate the sounds they desire.

Silence Zone (SA6) aims to increase contrast between spatial sounds by including a silence zone between them, facilitating better auditory transitions for DHH users while traversing a VR scene.

2.4 Additional Sounds Tools

Smart Notification (AS1) enables users to receive a notification sound at the sound source location when important sounds are played.

Custom Feedback Sounds (AS2) offers users a wider range of sound options for specific actions, including variations in pitch, volume, and style.

Calming Noise (AS3) enables users to select between white noise, pink noise, and rain sounds to add to the VR environment.

3 STUDY 1: USABILITY STUDY WITH DHH USERS

To assess our tools, we conducted a scenario-based evaluation. Specifically, we designed five scenarios to cover various common VR use cases (*e.g.,* gaming, nature, educational, music apps) and conducted a usability study with 10 DHH people.

3.1 Scenarios

Scenario 1: Forest Tour. Our first scenario, based in a forest, contains background sounds of wind, leaves, and birds, three groups of animals making localized spatial sounds, and a tour guide providing important information. We implemented six tools in this scenario: (1) PT1 to prioritize tour guide speech over environmental sounds, (2) PM2 to shift the frequency and volume of each environmental sound, (3) PM4 to adjust the speed and volume of the tour guide's speech, (4) PT4 to prioritize the animal sounds that the user is facing, (5) SA3 to allow users to adjust the range of animals that they hear, and (6) AS1 to play a notification sound when the tour guide speaks.

Scenario 2: Office Convo. This scenario is set in an office with six characters forming three groups of concurrent conversations. We implemented five tools: (1) PT2 to prioritize a certain group of conversations, (2) PM3 to separate two voices close in distance and frequency, (3) PT3 to notify the user when a character mentions a keyword and temporarily increases its volume if lowered, (4) AS3 to add white noise, pink noise, or rain sounds, and (5) SA2 to cue "to your left" for an important sound event originating from the left.

Scenario 3: Shooting Game. This scenario includes enemies shooting the user from various locations, accompanied by movement sounds. We implemented four tools: (1) SA1 to shift all sounds in the game towards the left or right side, (2) PM1 to adjust the volume and frequency of all sounds in the game, (3) PT4 to prioritize the enemy sounds that the user is facing, and (4) SA2 to notify when an enemy starts shooting and whether the enemy is on their left or right.

Scenario 4: Escape Room. This scenario consists of three rooms: a tutorial room, a room where the user finds a speaker playing a sound clue, and a room with a maze leading to an active target sound source. We implemented: (1) SA6 to insert a silence zone between ambient sounds in different rooms, (2) SA5 to let users isolate the clue sound source from other noises, (3) SA4 to change the pitch of the target of the navigation task as the player increases or decreases their distance to the target, and (4) SA2 to inform the user whether the target is on their left or right when they press a button.

Scenario 5: Rhythm Movement. This scenario was inspired from the popular VR game, Beat Saber [16]. Colored cubes move towards the user in sync with music, prompting the user to cut them with the controller. We implemented: (1) PM1 to change the volume and frequency of the music and feedback sounds, (2) PM5 to increase and decrease the music volume in sync with the music beats, (3) AS2 to allow the user to choose the sound notification for correct and incorrect feedback to fit their hearing range, and (4) AS3 to add white noise, pink noise, or rain sound to the ambient scene.

3.2 Method

We recruited 10 DHH participants (P1-P10, six men and four women) through email lists, social media, and snowball sampling. The participants were, on average, 40.10 years of age (*SD*=18.13 years). Eight participants identified as hard of hearing, one as deaf, and one as Deaf. Six participants had profound to severe hearing loss, two had moderate to moderately severe hearing loss, and two had unilateral hearing loss.

The user study was conducted in our research lab and lasted for about two hours. For each scenario, the researcher described the scenario, detailed the tools used in the scenarios, and instructed participants on using each tool. The participants then participated in the scenario and were asked to turn each tool on and off at least once using the tool configuration UI. After experiencing each scenario, the participants removed the VR device and completed a feedback questionnaire rating the scenario's immersiveness, sound information gained, and their experience with the tools on a scale of 1-5. At the end, we conducted a semi-structured interview on the participant's experience. Participants were compensated $50.

For analysis, we used descriptive statistics to summarize the questionnaire data and conducted an applied thematic analysis [3] on the interview transcripts with two coders. The IRR, measured using Krippendorff's alpha [7], was 0.88.

3.3 Findings

We found that in each scenario, the participants' experience with our tools improved greatly compared to without the tools, with average across-scenario reported improvement of 1.96 (*SD*=1.32) for immersion, 1.98 (*SD*=1.04) for sound information gained, and 2.16 (*SD*=1.13) for the overall VR experience on a scale of -3 to 3 (where -3 indicates much worse and 3 indicates much better experience). When comparing individual tools, we found that most tools (14/18) received high ratings (>=3.5 on a scale of 1-5, 5 being best). The four low rated tools included keyword prioritization (PT3) (*mean*=3.20, *SD*=1.32), frequency contrast enhancement (PM3) (*mean*=3.00, *SD*=1.25), calming noise (AS3) (*mean*=2.85, *SD*=1.39) and beat enhancement (PM5) (*mean*=2.50, *SD*=1.08) since they induced distraction (PT3), increased sound processing workload (AS3), was not very useful (PM3), or felt unnatural (PM5).

Subjective comments from the participants support these ratings. For example, P7 said: *"when you turn [the speech prioritization (PT1) feature] on, it is much easier to hear what [the tour guide] is saying"*. Similarly, P6, who has high-frequency hearing loss, explained: *"For the bird [sound], [...] I had a hard time hearing because it was high pitched, so I moved it over [to a lower range using the sound frequency/volume adjustment SP2 tool] and it's really easy to hear"*.

Despite the overall positive experience, almost all participants (*N*=9) expressed feeling distracted or overwhelmed with our tools at times. For example, the custom feedback sounds (AS2) tool provided seven custom notification options for users to choose from, and P5 *"was a little bit overwhelmed by how many choices there were."* For future improvements, participants (*N*=2) recommended using less intrusive alerts and dynamically adjusting tools based on users' focus. For example, P4 explained that the keyword prioritization tool should not release a notification if the user is already focused on the conversation containing the keyword.

Some participants also indicated that they desired further customization of the tools to accommodate their hearing levels. For example, two reported that volume changes performed by PM1 and PM2 tools were too drastic for them and should be configurable. Four participants requested customizable frequency ranges for SA4 and PM3 tools to fine-tune the tool to their specific frequency hearing range. Finally, six participants desired the ability to customize the spatial locations of some tools. For example, P7, who has unilateral hearing loss, suggested allowing repositioning of the shoulder localization helper (SA2) tool to the *"better hearing"* side.

4 STUDY 2: PRELIMINARY EVALUATION WITH VR DEVELOPERS

To understand how VR developers may use our tools in their apps, we compiled a Unity toolkit using our tools and conducted a usability study with six Unity VR developers (three men, two women, one non-binary). The developers were 24.3 years old on average (*SD*=1.8) and were experienced Unity VR developers (average experience of 3.6 years, *SD*=2.2).

For the study, we shared SoundModVR's GitHub repository with developers who asynchronously chose two of their personal Unity VR apps to incorporate the tools into. While incorporating the tools, developers completed an online study questionnaire to rate their experience and respond to open questions. Developers took about 60 to 90 minutes to complete the study and were compensated $30. We analyzed the data using descriptive statistics and thematic analysis with two coders [3]; the IRR, measured using Krippendorff's alpha [7], was 0.92.

Our preliminary findings suggest that all nine developers found the toolkit easy to use. The technical difficulty of implementing toolkits was low (average 1.83 on a scale of 1-5, 5 being the hardest), as also echoed by subjective comments. For example, D2 wrote that the toolkit was *"pretty quick to learn how to use, and the documentation was very thorough"*. Apart from technical implementation, developers also found it easy to incorporate the tools into the conceptual design of existing apps (*i.e.*, what tools to use with what sounds) (average 2.00 on a scale of 1-5, 5 being the hardest) owing to our clear documentation in the GitHub repository.

Despite the positive reviews, developers expressed concerns about the toolkit's implementation workload. For instance, two developers mentioned that the tools increased the cognitive load during development. To enhance this aspect, most developers (*N*=5) said the toolkit could benefit from further modularization, including packaging tools into Unity prefabs [17] and providing visual elements like default UI inside the prefab. Developers also indicated that the tools' audio feedback should allow further configurability so they can tune it to better fit into the app's design aesthetics. As D1 argued, *"it's important that the accessibility features feel like part of the app"*.

5 LIMITATIONS AND FUTURE WORK

Our evaluations show that our toolkit has the potential to improve VR accessibility for DHH users with partial hearing. However, we acknowledge not all DHH people will benefit from enhanced sounds, including people with profound hearing loss and people who are reluctant to use sound information. Nevertheless, based on the diversity of the community [1] and the experience of our hard-of-hearing coauthor, we argue that many DHH users may prefer our approach. Still, future work should continue to study our toolkit with a larger DHH population and through longitudinal, more naturalistic evaluations to incorporate diverse perspectives and use cases.

Moreover, although our VR scenarios covered a wide range of applications, we do not claim that they are exhaustive. Indeed, some VR app genres (*e.g.*, educational and meditation) were not included in our scenario evaluation. We welcome future work to extend the idea of VR sound modification into more diverse scenarios.

As outlined by participants in both studies, our toolkit could support further customization, like configuring the volume change in the prioritization tools or the location of notification tools. We invite future work to consider: (1) what the different dimensions of customization for these tools are, (2) how to achieve the right balance between delivering customization options and maintaining interface simplicity, and (3) whether end-user customization could interfere with the design of the original application.

We also acknowledge that our tools could cause cognitive overload. We believe that developers are best positioned to make these decisions for their apps, and we have included guidelines and best practices for each tool in our repository. Nevertheless, we agree that our toolkit can be further modularized, and these guidelines can be strengthened to include, for example, instructions for beginner developers and tool recommendations for different scenarios.

6 CONCLUSION

Previous work in VR sound accessibility has covered visual and haptic substitutions of sounds. Our work contributes to the first exploration of sound modification technique to help sound accessibility in VR environments. Some VR games and applications include sound accessibility features, but these are one-off efforts. We offered a more extensible and scalable approach by developing a toolkit that can be integrated into any VR app. Our evaluations with 10 DHH users and with 6 VR developers revealed that our toolkit improves the VR experience for DHH users and is easy to use and integrate into VR apps, but can benefit from further modularization, supporting more customization options, and handling cognitive overload.

REFERENCES

[1] Anna Cavender and Richard E. Ladner. 2008. Hearing Impairments. In *Web Accessibility: A Foundation for Research*, Simon Harper and Yeliz Yesilada (eds.). Springer, London, 25–35. https://doi.org/10.1007/978-1-84800-050-6_3

[2] Wouter A. Dreschler, Gitte Keidser, Elizabeth Convery, and Harvey Dillon. 2008. Client-Based Adjustments of Hearing Aid Gain: The Effect of Different Control Configurations. *Ear and Hearing* 29, 2 (April 2008), 214. https://doi.org/10.1097/AUD.0b013e31816453a6

[3] Greg Guest, Kathleen M.MacQueen, and Emily E.Namey. *Sage Research Methods - Applied Thematic Analysis*. Retrieved July 1, 2024 from https://methods.sagepub.com/book/applied-thematic-analysis

[4] Trevor Hogan, Uta Hinrichs, and Eva Hornecker. 2017. The Visual and Beyond: Characterizing Experiences with Auditory, Haptic and Visual Data Representations. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (*DIS '17*), June 10, 2017. Association for Computing Machinery, New York, NY, USA, 797–809. https://doi.org/10.1145/3064663.3064702

[5] Dhruv Jain, Sasa Junuzovic, Eyal Ofek, Mike Sinclair, John Porter, Chris Yoon, Swetha Machanavajhala, and Meredith Ringel Morris. 2021. A Taxonomy of Sounds in Virtual Reality. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (*DIS '21*), June 28, 2021. Association for Computing Machinery, New York, NY, USA, 160–170. https://doi.org/10.1145/3461778.3462106

[6] Dhruv Jain, Sasa Junuzovic, Eyal Ofek, Mike Sinclair, John R. Porter, Chris Yoon, Swetha Machanavajhala, and Meredith Ringel Morris. 2021. Towards Sound Accessibility in Virtual Reality. In *Proceedings of the 2021 International Conference on Multimodal Interaction* (*ICMI '21*), October 18, 2021. Association for Computing Machinery, New York, NY, USA, 80–91. https://doi.org/10.1145/3462244.3479946

[7] Klaus Krippendorff. 2004. Measuring the Reliability of Qualitative Text Analysis Data. *Qual Quant* 38, 6 (December 2004), 787–800. https://doi.org/10.1007/s11135-004-8107-7

[8] Ziming Li, Shannon Connell, Wendy Dannels, and Roshan Peiris. 2022. SoundVizVR: Sound Indicators for Accessible Sounds in Virtual Reality for Deaf or Hard-of-Hearing Users. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (*ASSETS '22*), October 22, 2022. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3517428.3544817

[9] Ziming Li, Kristen Shinohara, and Roshan L Peiris. 2023. Exploring the Use of the SoundVizVR Plugin with Game Developers in the Development of Sound-Accessible Virtual Reality Games. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (*CHI EA '23*), April 19, 2023. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3544549.3585750

[10] Mohammadreza Mirzaei, Peter Kán, and Hannes Kaufmann. 2020. EarVR: Using Ear Haptics in Virtual Reality for Deaf and Hard-of-Hearing People. *IEEE Transactions on Visualization and Computer Graphics* 26, 5 (May 2020), 2084–2093. https://doi.org/10.1109/TVCG.2020.2973441

[11] Mohammadreza Mirzaei, Peter Kán, and Hannes Kaufmann. 2021. Multi-modal Spatial Object Localization in Virtual Reality for Deaf and Hard-of-Hearing People. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, March 2021. 588–596. https://doi.org/10.1109/VR50410.2021.00084

[12] Jan Rennies, Dirk Oetting, Hannah Baumgartner, and Jens-E. Appell. 2016. User-Interface Concepts for Sound Personalization in Headphones. August 19, 2016. Audio Engineering Society. Retrieved January 28, 2024 from https://www.aes.org/e-lib/browse.cfm?elib$=$18370

[13] Michael Richardson, Jan Thar, James Alvarez, Jan Borchers, Jamie Ward, and Giles Hamilton-Fletcher. 2019. How Much Spatial Information Is Lost in the Sensory Substitution Process? Comparing Visual, Tactile, and Auditory Approaches. *Perception* 48, 11 (November 2019), 1079–1103. https://doi.org/10.1177/0301006619873194

[14] Todd Andrew Ricketts. 2001. Directional Hearing Aids. *Trends in Amplification* 5, 4 (December 2001), 139–176. https://doi.org/10.1177/108471380100500401

[15] B. Shirley, L. Ward, and E. T. Chourdakis. 2019. Personalization of object-based audio for accessibility using narrative importance. (January 2019). Retrieved January 28, 2024 from https://qmro.qmul.ac.uk/xmlui/handle/123456789/59636

[16] Ler digital studio. Beat Saber - VR rhythm game. Retrieved March 7, 2024 from https://beatsaber.com/

[17] Unity Technologies. Unity - Manual: Prefabs. Retrieved April 23, 2024 from https://docs.unity3d.com/Manual/Prefabs.html

[18] Alina Zajadacz. 2015. Evolution of models of disability as a basis for further policy changes in accessible tourism. *Journal of Tourism Futures* 1, 3 (January 2015), 189–202. https://doi.org/10.1108/JTF-04-2015-0015