The Effect of Spatial Audio on the Virtual Representation of Personal Space

Lauren Buck* Trinity College Dublin
Mauricio Flores Vargas† Trinity College Dublin
Rachel McDonnell‡ Trinity College Dublin

Abstract

Auditory feedback is an important element in the perception of one's surroundings. Vocal intonation betrays emotion, step frequency determines urgency, and volume gives way to proximity. While it is widely understood that sound is an important element in game design and film, research on how auditory feedback affects interactions taking place in virtual reality experiences is lacking. In this work, we propose an experiment in which we will begin to uncover the effect of auditory feedback, particularly spatial audio, on the preservation of personal space in an immersive virtual environment. Users will be exposed to visual and auditory stimuli and asked to report both interpersonal and peripersonal space. Personal space is dynamically responsive to circumstance, and is notably malleable by the presence of auditory stimuli in the real world. Consequently, we expect personal space to dynamically change in the presence of auditory stimuli in an immersive virtual environment. There is no prior study into how auditory feedback affects the maintenance of this space in virtual reality.

Index Terms: Human-centered computing—Interaction design—Epirical studies in interaction design; Human-centered computing—Interaction paradigms—Virtual reality

1 Introduction

It is generally known that sound carries localisation information and emotional expression. Vocal intonation is often thought of as a doorway to understanding an individual's emotional state, and even music and ambient sound have some emotional design about them. For example, listening to a certain song can evoke a feeling of happiness and hearing heavy winds during a storm can inspire fear. Sound changes the way we perceive and interact with our environment on a daily basis, and it is at the heart of the human experience. Sound design has long been recognized as an important component of the video game [6] and film [7] creation process. However, how sound affects the way users perceive interactions taking place in virtual reality (VR) has suffered a lack of attention from the VR research community, and there are many challenges to address [13, 14]. Researchers in our community have primarily directed focus to the visual and tactile feedback users experience in immersive virtual environments. We attempt to begin to fill that gap with work on understanding how spatial audio affects the virtual representation of personal space.

Spatial audio, or the sound signals that are perceived to have a distinct direction and distance from oneself [17], has a rich catalog of prior work. A large part of that work seeks to address the appropriate, personalised head related transfer function (HRTF), or the particular spectral cues that result from the filtering of sound waves with one's body, to create a condition in which optimal spatial audio is perceived. Yang et al. [17] point out that many studies have found spatial audio to possess great potential in many real life applications, including museum guides, spatial music mixing, and navigational aids. In order to create authentic 3D sounds it is required that the position between the objects and the listener is tracked, and localisation measures calculated.

Sound localisation is thought of as the set of perceptual properties of audio signals that enable the auditory system to estimate the particular angular position (azimuth and elevation) and distance of a sound source [9]. This estimation is calculated through the combination of inter-aural time and level differences, amplitude, and mono-aural cues that result from the corporal attributes of the listener, etc. Neurophysiological studies explain that multisensory neurons, which are responsive to visual and tactile stimuli distinctively within the near space of the body, are also responsive to auditory stimuli [12]. The presence of sound, and the location of that sound, aid the brain in calculating the appropriate bound of the space around the body that neuroscientists refer to as peripersonal space (i.e., the region around the body the brain encodes as functional reaching and grasping space). This space is tightly interwoven with interpersonal space (i.e., the space that two individuals maintain between one another) and we refer to them both (peripersonal and interpersonal) when speaking about personal space. Personal space is dynamic, and it is highly sensitive to the contextual elements of a situation, including the emotional valence of that situation. Auditory elements, since they intrinsically convey emotion and provide a cue about the location of an object or person in space, play a role in determining personal space. Several studies in experimental psychology have investigated the role of sound, especially looming or approaching sound, in the perception of interpersonal and peripersonal space in the real world [1, 5, 8], and in these works sound is shown to have an impact on the personal space representation.

What role auditory elements have in the maintenance of personal space in an immersive virtual environment, we do not know. Research has also shown promise in that audio cues aid in enhancing one's spatial understanding of virtual environments, especially at close distances [10]. However, there are several key technical factors about virtual environments that distinguish them from real world environments. One factor is the allowance for the unconstrained creation of environments that do not follow the physical constraints of the real world. Sound mismatch becomes possible, for example, where the texture of a sound does not match what would be experienced in real life (i.e., you may experience a "soft" sound when touching a rigid object), and conflicting visual and audio cues could have some impact on the spatial understanding of the user. Additionally, sound quality can vary from system to system, and depending on the equipment used could elicit a completely different spatial understanding from one user to the next. Consequently, we have designed an experiment to measure the personal space of a user when spatial audio is present.

Next, in Section 2, we discuss the rationale behind our experiment and the expected outcomes. We then discuss in Section 3 the materials and methods we plan on using to execute our experiment, and then include a brief wrap up in Section 4.

2 Rationale and Hypotheses

As mentioned in the previous discussion, sound localisation has promising potential to heighten a VR user's spatial awareness [10]. It is well known that VR users have difficulty determining egocentric distances to visual stimuli when no auditory feedback is present.

*e-mail: lauren.e.buck.12@gmail.com
†e-mail: floresvm@tcd.ie
‡e-mail: ramcdonn@tcd.ie

Additionally, auditory feedback tends to affect the representation of personal space, and provides some emotional valence that this representation is dependent upon [12]. This leads us to ask the question:

R1 Does auditory feedback from spatial audio affect the representation of personal space in VR?

We believe that, in light of the previous literature:

H1 Spatial audio will enable users to articulate the location of the agents in space more accurately, changing the representation of personal space.

Therefore, we aim to conduct a simple experiment in which we measure both the interpersonal and peripersonal space of the user in two conditions: one in which there is no spatial audio present and one in which spatial audio is present. We believe that spatial audio will give important cues to the user about where an approaching agent is located in space, and will, in turn, either cause the personal space to expand or contract.

Additionally, it is also widely understood that personal space is a continuous ellipse that encircles the body in the real world [15]. Thus, we have chosen to have an agent approach a user from both the front and back. With no visual or audio cues available, it would be sensible to imagine that the personal space representation would not extend behind the body. However, with the addition of auditory cues, we believe that personal space will be restored to a continuous ellipse. Therefore, our final hypothesis is:

H2 Without the inclusion of spatial audio, there will be no personal space representation from behind. But with the inclusion of spatial audio, personal space will be present behind the user. Additionally, the shape of the personal space will differ between the approach conditions (front, back) even when auditory feedback is present.

Our experimental design is further detailed in the next section.

3 Materials and Methods

3.1 Apparatus

Figure 1: The environment and avatars to be used in the experiment as seen from a third person perspective.

The virtual reality system used in this experiment will be an Oculus Quest 2, which consists of a head-mounted display (HMD) and two hand controllers. We have built the 3D immersive virtual environment using Unity Game Engine (version 2020.3.17f1), and all scripts are written in C#. The experiment will be conducted in a virtual room (10 m x 10 m x 3.5 m) that is meant to represent a generic public space (i.e., an art gallery). The room contains two doorways (one directly in front of and one directly behind the user), some paintings and decorative columns and objects. The items in the room were either modelled in Unity or taken from CGTrader. The agent and self-avatar used in this experiment will be the same model, which has been taken from Adobe Mixamo. The agent executes a walking animation that has also been taken from Adobe Mixamo. Tracking data will be used from the HMD and controllers to drive the upper body of the avatar, while the lower body will remain fixed since the user is required to remain seated. Data will be recorded and stored in a text file which will be extracted into a csv file for analysis. A picture of the environment can be seen in Figure 1.

The audio spatialization in the virtual environment is implemented using the Oculus Audio Spatializer (version 32.0.0). An acoustic response has been modelled in relation to the virtual room characteristics using early reflections, late reverberation and attenuation features included in the spatializer to provide further sound direction and distance cues and create a more realistic acoustic environment. The sound files used for the experiment are dry monophonic audio samples of common bodily sounds such as footsteps and breathing. These sounds have been attached with an audio source component to the agent in the respective places (footsteps in the feet, breathing in the face) to match the agent's animation, direction and distance. A set of samples has been recorded for each of the agent's sounds. These samples are triggered randomly using an audio manager script to reduce repetition, and, in the case of the breathing, they loop seamlessly to avoid unwanted disruption. Generic HRTFs are used for sound localisation [2]; participants will undergo a pre-test calibration session to familiarize themselves with the response of the HRTFs and improve their auditory spatial localization.

3.2 Experimental Design

The experiment design is modelled after that of Buck et al. [3], who adopted the experimental paradigm from [16]. In their original work, Serino et al. successfully measured peripersonal space boundaries in a mixed reality environment that were similar to those found in the real world, and this has been replicated by several studies in VR [3, 4, 11]. Users will experience two sets of trials: one in which they will report comfort distance (to obtain measure interpersonal space) and one in which they will be asked to respond to a tactile stimulus (to obtain a measure of peripersonal space). In the set of trials where users will be asked to report comfort distance, a virtual agent will emerge from either doorway and approach the user from the front or back. The user will be asked to press a button when they become uncomfortable with the avatar being in their personal space. In the trials where the users will be exclusively asked to respond to the tactile stimulus, which is a vibration delivered via the handheld controllers, users will press a button when the tactile stimulus fires. The tactile stimulus will fire when the agent reaches a distance of either 0.75 m, 1 m, 1.25 m, 1.45 m or 1.85 m from the user. Users will be asked to report comfort distance a total of 10 times (5 times front, back), and respond to the tactile stimulus a total of 50 times (5 times at each distance, front and back), resulting in a total of 60 trials. The study will be within subjects, with one block of trials requiring users to respond with auditory stimuli and another without auditory stimuli. Thus, users will complete a total of 120 trials.

4 Conclusions

Ultimately, we have designed an experiment in which the effect of spatial audio on the virtual representation of personal space will be tested. We believe that, with the inclusion of spatial audio, the representation of personal space will be altered and should expand and contract in a manner that more closely follows how it is represented in the real world. Validating our hypotheses will open up room for further study on the effect of spatial audio on perception and action in immersive VR, and these findings will give further design direction to developers and researchers on the required environmental elements to produce more realistic interaction. Further study on the emotional valence of sounds (such as sounds that evoke sadness, happiness, anger, excitement, etc.), as well as the quality of sound and sound mismatch and their effects on the personal space would all be imperative to understanding the contribution of spatial audio to the quality of the virtual environment.

ACKNOWLEDGMENTS

This research was funded by Science Foundation Ireland under the ADAPT Centre for Digital Content Technology (Grant No. 13/RC/2106_P2) and RADICal (Grant No. 19/FFP/6409).

REFERENCES

[1] M. Bahadori, R. Barumerli, M. Geronazzo, and P. Cesari. Action planning and affective states within the auditory peripersonal space in normal hearing and cochlear-implanted listeners. *Neuropsychologia*, 155:107790, 2021.

[2] C. C. Berger, M. Gonzalez-Franco, A. Tajadura-Jiménez, D. Florencio, and Z. Zhang. Generic hrtfs may be good enough in virtual reality. improving source localization through cross-modal plasticity. *Frontiers in neuroscience*, 12:21, 2018.

[3] L. E. Buck, S. Chakraborty, and B. Bodenheimer. The impact of embodiment and avatar sizing on personal space in immersive virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, In Press, 2022.

[4] L. E. Buck, S. Park, and B. Bodenheimer. Determining peripersonal space boundaries and their plasticity in relation to object and agent characteristics in an immersive virtual environment. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 332–342, 2020. doi: 10.1109/VR46266.2020.00053

[5] E. Canzoneri, E. Magosso, and A. Serino. Dynamic sounds capture the boundaries of peripersonal space representation in humans. 2012.

[6] K. Collins. *Game sound: an introduction to the history, theory, and practice of video game music and sound design*. MIT Press, 2008.

[7] V. Dakic. Sound design for film and television, 2009.

[8] F. Ferri, A. Tajadura-Jiménez, A. Väljamäe, R. Vastano, and M. Costantini. Emotion-inducing approaching sounds shape the boundaries of multisensory peripersonal space. *Neuropsychologia*, 70:468–475, 2015.

[9] M. Filimowicz. *Foundations in Sound Design for Interactive Media: A Multidisciplinary Approach*. Routledge, 2019.

[10] Y.-H. Huang, R. Venkatakrishnan, R. Venkatakrishnan, S. V. Babu, and W.-C. Lin. Using audio reverberation to compensate distance compression in virtual reality. In *ACM Symposium on Applied Perception 2021*, pp. 1–10, 2021.

[11] H.-S. Lee, S.-J. J. Hong, T. Baxter, J. Scott, S. Shenoy, L. Buck, B. Bodenheimer, and S. Park. Altered peripersonal space and the bodily self in schizophrenia: A virtual reality study. *Schizophrenia Bulletin*, 2021.

[12] D. R. Moore and A. J. King. Auditory perception: The near and far of sound localization. *Current Biology*, 9(10):R361–R363, 1999.

[13] S. Serafin, F. Avanzini, A. De Goetzen, C. Erkut, M. Geronazzo, F. Grani, N. C. Nilsson, and R. Nordahl. Reflections from five years of sonic interactions in virtual environments workshops. *Journal of New Music Research*, 49(1):24–34, 2020.

[14] S. Serafin, M. Geronazzo, C. Erkut, N. C. Nilsson, and R. Nordahl. Sonic interactions in virtual reality: State of the art, current challenges, and future directions. *IEEE computer graphics and applications*, 38(2):31–43, 2018.

[15] A. Serino. Peripersonal space (pps) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience & Biobehavioral Reviews*, 99:138–159, 2019.

[16] A. Serino, J.-P. Noel, R. Mange, E. Canzoneri, E. Pellencin, J. B. Ruiz, F. Bernasconi, O. Blanke, and B. Herbelin. Peripersonal space: an index of multisensory body–environment interactions in real, virtual, and mixed realities. *Frontiers in ICT*, 4:31, 2018.

[17] J. Yang, Y. Frank, and G. Sörös. Hearing is believing: synthesizing spatial audio from everyday objects to users. In *Proceedings of the 10th Augmented Human International Conference 2019*, pp. 1–9, 2019.