Towards Inclusive Conversations in Virtual Reality for People with Visual Impairments

Markus Wieland
markus.wieland@ifi.lmu.de
LMU Munich

Tonja Machulla
tonja.machulla@phil.tu-chemnitz.de
TU Chemnitz

Figure 1: Providing gaze via different sensory modalities in an VR space. Left: picture from a VR space where two avatars are talking to each other. Right: transformation of gaze via a high-contrast visual ray, auditive, and tactile cues

ABSTRACT

Current mainstream social Virtual Reality (VR) spaces pose barriers to the equal participation of people with visual impairments (PVI) in social interactions. At present, VR is first and primarily a visual medium with a strong emphasis on the visual design of the VR scene and the available avatars. If social communication cues, such as non-verbal communication, are available at all, they are often not provided in a form accessible to PVI. Such cues are essential in social interactions to successfully participate in social interactions and experience a conversation in VR as realistic. Here, we summarize previous research regarding specific requirements for social VR spaces to be accessible to PVIs. We describe how people with disabilities recognize and identify potential conversational partners and how non-verbal communication works between PVI and sighted people. Our goal was to provide an overview of valuable features that can be implemented for inclusive conversations in a social VR space.

Veröffentlicht durch die Gesellschaft für Informatik e.V.
in K. Marky, U. Grünefeld & T. Kosch (Hrsg.):
Mensch und Computer 2022 – Workshopband, 04.-07. September 2022, Darmstadt

https://doi.org/10.18420/muc2022-mci-ws11-467

KEYWORDS

visual impairments, non-verbal communication, virtual reality

1 INTRODUCTION

Virtual Reality (VR) is used nowadays not only for gaming purposes but also for social encounters. Social VR spaces like Meta's Horizon Worlds[1] contribute to meeting and talking in a virtual world and playing games together. The users create three-dimensional avatars with which they interact with the virtual world. For immersive social interactions in VR, the design of embodied avatars is increasingly important [1]. This also includes various forms of social communication, such as addressing another person and non-verbal communication. Indeed, these forms of social communication are mostly not integrated into social VR spaces, even though today's technology is there to integrate them. Instead, social VR spaces focus on improving avatars' visual representation and the virtual world. However, people with visual impairments (PVI) do not benefit from visual fidelity enhancement and require the integration of social communication cues into a conversation in a virtual world [5, 6]. There are already tools that support PVI in different VR scenarios. For example, there is a blind cane that can be used to explore the virtual world and provide three different types of feedback: physical resistance when the cane hits a virtual object; vibrotactile feedback when the cane hits different objects or surfaces; and 3D spatial audio [7]. In addition, a tool can be integrated into VR games

[1] https://www.oculus.com/horizon-worlds

and offers various functions such as magnification lens, edge enhancement, and text-to-speech output [8]. Nevertheless, these tools focus on navigation and interaction with objects rather than social interactions in VR. Moreover, since the behavior of PVI in conversations differs from that of sighted people, special approaches are required [3]. This paper provides an overview of potential features for conversations in VR for PVI. We summarized findings from previous research that have examined how PVIs detect and identify others and how non-verbal communication is perceived and provide possibilities for inclusive social VR spaces.

2 DETECTING AND IDENTIFYING OTHERS AND ACTIVITIES

To find out how PVI recognize and detect others, the authors conducted an online survey with 15 PVI [5]. This revealed that auditory methods were used most frequently followed by touch, smell, and vision. In order to detect the presence of others fuzzy sounds were made approaching and whether voices or breathing could be heard. Instead of listening, PVI also greeted into a room and waited for a response to detect the presence of others. Regarding the sense of touch, PVI used airflow perception and the wind that the clothes make under certain conditions. PVI used the smell to estimate the distance to the other person. The residual vision of the PVI was used to detect silhouettes. The auditory channel is used for various methods to detect other people.

In order to identify familiar people, the auditory sense is the dominant perceptual channel. PVI listen especially to the voice to recognize the person. The other person must speak a complete sentence because a simple greeting is not enough to recognize the person. Further, the way of breathing and particular sounds made by another person (e.g., laughter) are used. A few PVI only uses Silhouette identification and motion. Additionally, PVI define precise meeting points and use context (e.g., work) to identify familiar people. Thus, known persons are recognized predominantly via the auditory.

Other persons' activities are identified auditorily, visually, and via the general atmosphere of the environment. For example, attention is paid to sounds that people or materials emit before an activity happens. In addition, PVI use the content of the conversation to infer activities. Activities already in progress are detected visually by the movement of the human silhouette. The atmosphere of the environment is used to find out which activities are mainly taking place here (i.e., tavern).

Taken together, the detection and identification of other people and their activities are mainly made through the sense of hearing. This involves paying attention to various auditory strategies, such as listening for sounds, breathing, or distinctive sounds from a familiar person. PVI with residual vision use motion and human silhouettes to detect other people or activities. Other senses, such as smell or touch, are also used by PVI to gather information about other people, but these are not reliably available in every situation.

3 NON-VERBAL COMMUNICATION

A better understanding of how PVI use and perceive nonverbal communication in a conversation can help make these signals perceivable in a VR conversation for PVI. In the following section, we examine the work of [4], who conducted interviews with 20 PVI, and [6], who interviewed seven visually impaired people and their four sighted partners about the use and perception of non-verbal signals during a conversation.

Gaze Both of these works showed that the gaze for PVI is important in a conversation, although it is usually not perceivable. [6] said that all PVIs try to look their interlocutor in the eye to signal a willingness to receive something. In the study of [4], the PVI mentioned that the missing of the gaze leads to disadvantages in conversations with sighted persons because the gaze is sometimes used as a signal for turn-taking. This can accelerate communication in a conversation. Furthermore, [6] asked if PVI knew of situations in which people communicated only by gaze, and they all agreed but would use other cues. [4] reported that the PVI also had an understanding of gaze for sighted people. However, it is simply not perceivable. This shows that PVI have an idea of the meaning of gaze for sighted people and have learned this (i.e., in school). The sighted partners in the study of [6] mentioned that they and their partners always try to look each other in the eye. One sighted partner even stated that they missed eye contact with their partner (the PVI used to have better vision and eye contact was possible).

Facial Expressions Facial expressions are not perceivable; sometimes it depends on the individual residual vision of the PVI. Under certain circumstances, such as good lighting, the interlocutor is close or known, and facial expressions can be recognized. Since facial expressions are used to detect emotions, PVI have difficulty detecting the latter directly. [6] mentioned that the PVI inquire emotions directly or they perceive it through other signals such as voice, breathing, sighing, and posture.

Gestures Both publications reported that PVI in most cases do not perceive gestures during a conversation. However, [6] described that if the gestures are presented very expansively and expressively, the PVI is in a close range, and there are good lighting conditions, it can be possible to perceive a particular gesture. Regarding the use of gestures, [6] said that PVI use pointing gestures only when the place of an object is known, or they use gesticulations to express themselves in a conversation. Two of the sighted partners mentioned that they use fewer and no gestures during a conversation with PVI.

Hearing The sense of hearing is used to obtain various information about the interlocutor. According to [6] PVI uses the sense of hearing to recognize behavior and mood and draw conclusions about the interlocutor. Non-communicative cues such as noises are perceived as distracting because they block the auditory channel, but other PVI also use these cues to infer the behavior of the interlocutor (i.e., nervous behavior) [6]. In the interview of [4], the auditory channel was identified as the channel most used by PVI during a conversation.

Touch The sense of touch is only used by friends, socially appropriate, and as a positive signal during a conversation [4, 6]. One PVI even used handshakes and hugs to find out how serious the communication is meant [6].

In the categories of non-verbal communication presented, it is clear that all sensory modalities convey various information for PVI about the interlocutor. On the other hand, not all channels are perceivable (e.g., gaze) but are still considered important in a conversation by PVI. The sense of hearing seems to be the most

crucial channel. Here, information such as voices, sounds, and noises are processed to infer the interlocutor's behavior.

4 POSSIBILITIES TOWARDS AN INCLUSIVE SOCIAL VR SPACE

Due to the possibilities of VR, it is possible to make various detection and identifying mechanisms and non-verbal cues perceivable for PVI.

Detect, Identify others and activities There could be automatic verbal output from other avatars near the PVI in the VR space. Similar to the work of [2], they divided the space around an avatar in VR and made different sounds when another avatar approached. In addition, the incorporation of supposedly small sounds such as footsteps or the rustling of clothing can help PVI to detect other persons in a VR space. Further, implementing a friends function shows on request whether friends are currently in the same VR room.

Non-verbal communication Gaze can be represented via other sensory modalities because it is not perceptible, but it is perceived as important (Figure 1). Facial expressions can be presented auditorily (e.g., happy background music during a smile) or with automatic magnification for PVI who use their remaining vision. Gestures can be displayed more expressively so that the movement can be perceived accompanied by airflow sounds. For hearing, spatial listening can be implemented along with sounds made by the interlocutor. Since there is currently no technology for mainstream VR applications that can integrate different types of touches, it remains with the controllers' vibration. Nevertheless, these can be used to display handshakes or hugs in the VR world.

Stated by [4], it is not helpful to implement all features simultaneously, as this could lead to stimulus overload for PVI. The functions could behave contextually. For example, during a simple conversation in VR, the gaze function could be activated automatically. The gesture function could also be activated when playing smaller multiplayer games in social VR. When multiple avatars are talking in VR, spatial hearing could be enabled.

In this work, we summarized results from the field of social communication between PVI and sighted people. In particular, detecting, identifying others and their activities, and non-verbal communication. Our goal is to provide an overview of how social VR spaces can be designed to allow equal participation for PVI.

REFERENCES

[1] Guo Freeman and Divine Maloney. 2021. Body, Avatar, and Me: The Presentation and Perception of Self in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (Jan. 2021), 1–27.

[2] Tiger F. Ji, Brianna R Cochran, and Yuhang Zhao. 2022. Demonstration of VRBubble: Enhancing Peripheral Avatar Awareness for People with Visual Impairments in Social Virtual Reality. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. ACM, New Orleans LA USA, 1–6.

[3] Nigel J Kemp and Derek R Rutter. 1986. Social interaction in blind people: An experimental analysis. *Human Relations* 39, 3 (1986), 195–210. Publisher: Sage Publications Sage CA: Thousand Oaks, CA.

[4] Shi Qiu, Pengcheng An, Jun Hu, Ting Han, and Matthias Rauterberg. 2020. Understanding visually impaired people's experiences of social signal perception in face-to-face communication. *Universal Access in the Information Society* 19, 4 (Nov. 2020), 873–890.

[5] Lauren Thevin and Tonja Machulla. 2020. Guidelines for Inclusive Avatars and Agents: How Persons with Visual Impairments Detect and Recognize Others and Their Activities. , 164–175 pages. https://doi.org/10.1007/978-3-030-58796-3_21 Series Title: Lecture Notes in Computer Science.

[6] Markus Wieland, Lauren Thevin, Albrecht Schmidt, and Tonja Machulla. 2022. Non-verbal Communication and Joint Attention Between People with and Without Visual Impairments: Deriving Guidelines for Inclusive Conversations in Virtual Realities. In *Computers Helping People with Special Needs*, Klaus Miesenberger, Georgios Kouroupetroglou, Katerina Mavrou, Roberto Manduchi, Mario Covarrubias Rodriguez, and Petr Penáz (Eds.). Vol. 13341. Springer International Publishing, Cham, 295–304. https://doi.org/10.1007/978-3-031-08648-9_34 Series Title: Lecture Notes in Computer Science.

[7] Yuhang Zhao, Cynthia L. Bennett, Hrvoje Benko, Edward Cutrell, Christian Holz, Meredith Ringel Morris, and Mike Sinclair. 2018. 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*. ACM, Montreal QC Canada, 1–14.

[8] Yuhang Zhao, Edward Cutrell, Christian Holz, Meredith Ringel Morris, Eyal Ofek, and Andrew D. Wilson. 2019. SeeingVR: A Set of Tools to Make Virtual Reality More Accessible to People with Low Vision. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14.