

Investigating the Effect of Different Virtual Environments on the Perception of Soundscapes: A Pilot Study on the Influence of Visual Display Technologies

Özlem Gök Tokgöz¹, M. Ercan Altinsoy²

¹ Chair of Acoustics and Haptics, TU Dresden, ozlem.goek_tokgoez@tu-dresden.de

² CeTI, Chair of Acoustics and Haptics, TU Dresden, ercan.altinsoy@tu-dresden.de

Abstract

Virtual reality (VR) is a significant factor in the visualization of places and the replication of real-world environments within controlled laboratory settings. Virtual reality technology allows individuals to pre-experience decisions, creating visually and aurally pleasing environments. For this reason, different virtual environment concepts are frequently used in urban studies. Limited studies have investigated the ecological validity of soundscape studies in various virtual environments, but no comparison has been made between these environments with different audio-visual features. This study is aimed at understanding the difference between different virtual environments in soundscape studies. In this study, Immersive Virtual Reality (IVR) and Multi-Modal Measurement Laboratory (MMM) conditions are compared. In MMM, visual data is limited, and the interaction of the participants is less. However, it is more comfortable and easier to adapt than the IVR headset. In the context, locations within the urban area were determined, and the audio and video recordings were collected. The survey was conducted in accordance with ISO/TS 12913-2:2018, with additional considerations given to themes such as plausibility, and personal comfort. The survey was conducted in both areas. As a result of the study, the differences and similarities of soundscape studies applied in different virtual environments were evaluated.

Introduction

Virtual reality (VR) is an advancing technology that has experienced substantial growth in recent years. It has a wide range of applications in various industries. VR has also been integrated into the fields of urbanization and architecture. This tool is valuable for investigating different urban environments and generating designs that promote user engagement. Moreover, VR enables the assessment of residents' perceptions of urban issues and their degrees of pleased or discomfort. However, the representation of urban areas using virtual reality equipment is complex.

Urban areas are characterized by a multitude of visual and audible stimuli. Because of the dynamic nature of urban environments and VR representations, the presence of numerous actors poses a significant challenge. Therefore, it is crucial to investigate the ecological validity of research that focuses on the acoustics of urban environments. Ecological validity refers to the extent to which test performance accurately predicts behaviors in real-world circumstances. In order to obtain reliable results on user perception, urban environments should be as close to the real environment as possible. In studies on virtual reality, we can talk about two important concepts that will enable us to get reactions close to

the real environment reaction: place illusion and plausibility illusion [1]. Place illusion is the strong illusion of being in a place in spite of the sure knowledge that you are not there. Plausibility illusion refers to the illusion that the scenario being depicted is actually occurring. These two concepts lead to realistic behavior in virtual environments.

The accuracy of the visual and auditory representations in the experimental setting is crucial for creating realistic presentations. Immersive virtual reality (IVR) is particularly notable in this context. The IVR experiments, known for their strong sense of place and plausibility illusion, are extensively utilized across various disciplines. IVR is also employed in the depiction of urban environments. IVR aims to provide consumers with a comprehensive 360-degree experience by transporting them into digital, three-dimensional surroundings. By utilizing technologies like VR headsets and tracking controllers, we can engage in and navigate virtual environments with a high degree of realism. In IVR studies, 360-degree video recordings can enhance soundscape efforts by offering a human-centred and multi-sensory viewpoint, therefore yielding more accurate data on the interplay between the visual and auditory surroundings [2]. However, realistic representations are not exclusively generated in IVR environments. There are other investigations conducted in specialized laboratories to explore ecological validity. The Multimodal Measurement Laboratory (MMM) at Dresden University of Technology is an example of a virtual environment. This lab combines VR equipment for multiple modalities (auditory, tactile, vestibular, and visual) and is capable of presenting high-performance, interactive simulations. Additionally, this laboratory conducted research in the fields of urban studies [3].

The objective of this study is to comprehend the distinctions between two distinct virtual settings (IVR and MMM) that possess varying benefits and drawbacks in the examination of soundscapes. To achieve this objective, audio and video recordings were captured in an urban environment to represent the two environments. Additionally, a questionnaire was administered to evaluate the acoustic, visual, and virtual aspects of the surroundings.

Methodology

The pilot study was conducted in a residential area in Dresden. The video and audio recordings utilized in the experiment were captured within the green region situated amongst the residential complex and the highway.

Preparation of the Field Recordings

As part of the study, video and audio recordings were conducted in the field. Figure 1 shows a photograph captured from a 360-degree video recording of the location.



Figure 1: A photograph captured from a 360-degree video recording

Two separate recordings of around 10 minutes each were taken. Since it is thought that people can experience the environment by sitting, the video recording was taken from a height of 1,30 cm above the ground. Both of these recordings were captured using the Insta360 X3 camera.

The first recording was conducted in a 360-degree field of view and configured at a resolution of 4K, with a frame rate of 30 frames per second. Zylia Pro ambisonic microphone was used for sound recording.

The second recording was conducted in the same location roughly 10 minutes subsequent to the initial recording. The same camera was utilized for the second recording, capturing a single lens image on this occasion. Once again, the video resolution is 4K and the frame rate is 30 frames per second. The sound measurements were made with the HEAD acoustics SQuadriga II equipped with binaural microphone.

Preparing the Scenes

In the study, scenes were prepared separately for two different environments. The video and audio recordings intended for MMM and IVR settings were converted from around 10-minute videos to 3-minute videos.

The first virtual environment was MMM, which integrates VR technology for several sensory modalities and has the ability to deliver advanced, interactive simulations [4]. This laboratory is specifically designed for conducting virtual environment tests using 500 loudspeakers and room acoustics that conform to the specifications outlined in DIN 15996. Full HD visual processing is done on an acoustically transparent screen. In the MMM, binaural recordings were used. The lab uses an ISONO Spatial Audio Workstation (SAW), which makes it possible to manipulate intricate sound scenarios. Nuendo4 software was used for audio and video synchronization and arrangements.

The IVR is the second virtual environment. IVR experiments were performed with a head-mounted display (HMD) and a headset. Meta Quest Pro VR Headset was used as HMD, and Beyerdynamic DT 990 Pro headphones were used for audio. A 360-degree video was first prepared for the experiment. Due to the image's 360-degree nature, the microphone and tripod that were captured in the recording using the Adobe After Effects program were removed from the video. The sound file, which was recorded by the Zylia Pro ambisonic microphone, was exported as first-order ambisonic using the Reaper digital audio workstation (DAW). After that, the recorded material was imported into Adobe Premiere Pro to synchronize the audio and video.

Questionnaire Structure

The questionnaire is divided into two sections. The initial segment addresses the auditory and visual environment, while

the following segment investigates into the virtual environment. The questionnaire form was created using Google Forms and the same inquiries were used in both fields. The questions from both parts are clarified in Table 1.

Table 1: Questionnaire Structure

A. Questions about Visual and Auditory Environment	
1.	How would you describe the present visual environment?
2.	For each of 8 scales below, to what extent do you agree or disagree that the present surrounding visual environment is...?
3.	How would you describe the present surrounding sound environment?
4.	For each of 8 scales below, to what extent do you agree or disagree that the present surrounding sound environment is...?
B. Questions about Virtual Environment	
1.	To what extent did you experience a sense of "being there" inside the virtual environment?
2.	To what extent did you experience the feeling of "immersion" in the virtual environment?
3.	How strong was our "sense of inclusion" in the virtual environment?
4.	In general, how realistic did the virtual environment appear to you?
5.	How plausible is the setting in which you find yourself in the experiment?
6.	What was your overall comfort level in this experience?
7.	What was your overall enjoyment level in this experience?

The initial segment of the research focused on the visual and auditory environment. This section consists of four distinct questions. The questions in this section are derived from the ISO/TS 12913-2 standard. Participants were instructed to assess the visual and audio environment using a five-point scale.

In the second part, participants were asked about their feelings in the virtual setting. The questions in this section were formulated based on research conducted in virtual reality. The first three questions in the questionnaire pertain to sentiments associated with 'being there', 'immersive', and 'sense of inclusion'. The following two questions pertain to the realism and plausibility of the scene. Furthermore, participants were asked questions focusing on the experience itself, such as 'comfort' and 'enjoyment' levels.

Evaluation of Questionnaire Results

The questionnaire results are analyzed and discussed in two separate sections. Initially, the impact of the visual and audio surroundings on the produced scenes was assessed, followed by an evaluation of the virtual environment. A total of twenty-four individuals took part in the questionnaires with MMM and HMD. There were six individuals who participated in

both studies as partners. However, these six people participated in these experiments on different days. In the MMM experiment, there are 12 male and 3 female participants. The age range of the participants is between 25 and 44 (mean age = 35.38, standard deviation = 4.77). In the HMD part, there are 13 male and 2 female participants, and their age range is between 22 and 44 (mean age = 29.4, standard deviation = 5.47).

Visual and Sound Environment

The assessment of the survey results commenced with an inquiry regarding the overall assessment of the visual and auditory environment. In this question, people describe the present surrounding environment using a 5-point scale (1-very bad, 5-very good). Figure 1 presents the average values for two distinct environments.

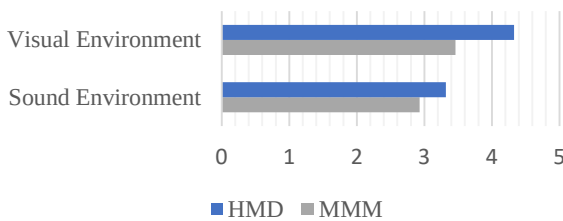


Figure 1: Results of the visual and sound environment descriptions of the participants

In the evaluation of the visual environment, the mean value of HMD is 4.33 and the mean value of MMM is 3.46. The sound environment was evaluated, and the average value of HMD was found to be 3.32, while the average value of MMM was 2.93. Based on the average values, it can be concluded that the participants find the HMD environment to be more pleasant.

In the second part of the study, questions were asked about the perceived affective qualities related to visual and auditory. Regarding the visual aspect, while various adjectives are commonly used to evaluate visual environments [5], certain adjectives have been excluded for the purposes of this study. The answer options range from "5-strongly agree" to "1-strongly disagree".

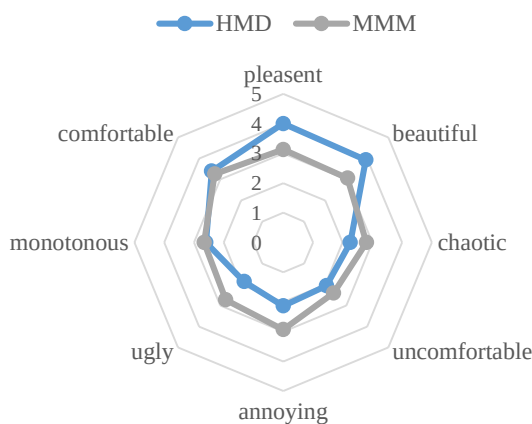


Figure 2: Questionnaire results to visual affective quality

For visual affective quality, the single arithmetic values of the eight features are shown in Figure 2. Visually, the HMD was

determined to be more pleasant, and beautiful. Additionally, it is less annoying, ugly, and uncomfortable. The results suggest that the 360-degree view experienced on the HMD influenced the evaluation of the environment and had a more positive impact for the users.

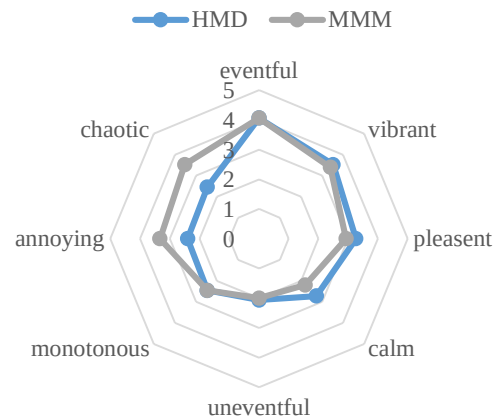


Figure 3: Questionnaire results to auditory affective quality

For the auditory environment, eight adjective pairs and a 5-point scale were given in the question as defined in ISO/TS 129913-2. Figure 3 shows the single arithmetic values of the eight attributes for auditory affective quality. Acoustically, the HMD is perceived as less annoying and chaotic than the MMM. Nevertheless, there was no significant difference in the average values of these two virtual environments. From an acoustic point of view, the environments have the same characteristics.

Virtual Environment

First, participants were asked to assess their sense of "being there," the level of immersion they experienced, and their "sense of inclusion." In this question, people express their opinion by using a 5-point scale (1-not at all, 5-totally). In Figure 4, the average values are given.

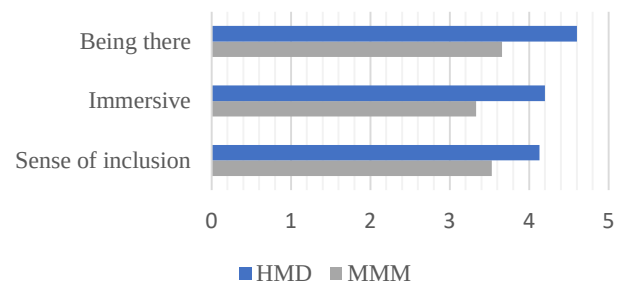


Figure 4: Results of the questions on "being there", "immersive" and "sense of inclusion".

Upon analyzing the mean values, it becomes evident that the HMD is more effective in eliciting sensations of "being there", "immersive" and "sense of inclusion". Given the specific attributes of the VR Headset and the controlled conditions of the experiment, this outcome was expected. Another question related to the "realistic" and "plausibility". Due to the inclusion of video in the recordings, it was anticipated that the experiment would possess a high level of realism and be plausible. Figure 5 shows the mean values.

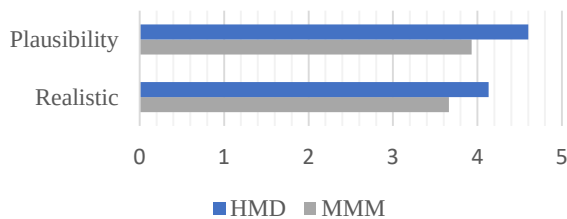


Figure 6: Results of the questions on “how realistic” and “plausibility”.

In the last part, participants were asked to assess the level of “comfort” and “enjoyment” under the HMD and MMM conditions. In this question, people give their opinion using a 5-point scale (1-very bad, 5-very good). Figure 6 presents the average values of the participants.

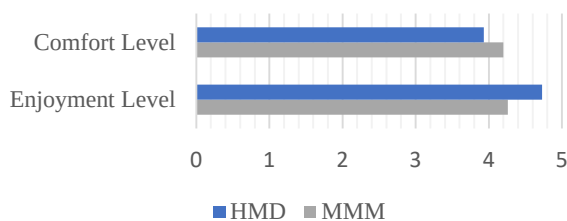


Figure 6: Results of the questions on comfort and enjoyment level.

Based on the average values, participants expressed a higher level of comfort in MMM compared to HMD. Nevertheless, the overall level of enjoyment was more in the HMD.

Discussion

After analyzing this plot study, a brief evaluation of the IVR and MMM tests highlighted the advantages and disadvantages of the various virtual environments.

In preparation, obtaining 360-degree video and ambisonic audio recordings is more challenging compared to other audio and video recordings. To not break plausibility with the experiment participants, it is necessary to remotely capture video and audio recordings or discreetly conceal them in an appropriate location.

Once the recordings have been obtained, the process of preparing for the experiment is still difficult. Given that the image covers a full 360 degrees, any technical equipment, such as tripods and microphones, that appears within the frame must also undergo cleaning through video editing. Following the completion of the video editing process, more software is required to synchronize the audio and video components.

In experiments, unfortunately, the use of VR headsets is not sufficiently user-friendly. It takes time to teach and explain its use. Not everyone accepts it at the same time. For this reason, many IVR studies generally show that the participant group is predominantly composed of young individuals. Therefore, we can claim that it is far from covering the entire society. Furthermore, due to the heaviness and discomfort of VR headsets, as well as potential side effects, it is necessary to restrict the time of the experiment and carefully choose individuals.

In MMM studies, the visual component is limited, in contrast to IVR studies. Consequently, the processes of preparation and implementation are simple. However, due to its specialized nature, the laboratory can only be used for limited research in the field of urban studies.

Conclusion

VR is a rapidly growing technology that has been extensively utilized in the field of soundscape research in recent years. With the proliferation of HMD technologies, it is obvious that we will see a lot of IVR and soundscape studies in the future. The study focused on investigating the variations among different virtual environments and their impact on evaluations of the visual and auditory environments. The study concluded that HMD technologies offer a significant level of immersion. However, it has also been observed to result in numerous complications during the setup of the sets and the execution of the studies. MMM, on the other hand, is a uniquely designed setting that provides limited use. Nevertheless, it has the advantage of accommodating a variety of participants and making the preparation and implementation of experiments easier. It is crucial to conduct a comprehensive examination of the pros and cons of virtual environments and select virtual environments that align with the stated objective in order to promote their wider use in research.

Acknowledgement

This project is co-funded by the European Union and co-financed from tax revenues on the basis of the budget adopted by the Saxon State Parliament.

References

- [1] M. Slater, “Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments,” *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 364, no. 1535, pp. 3549–3557, Dec. 2009, doi: 10.1098/rstb.2009.0138.
- [2] H. A. Dogan Iseri, A. Özcevik Bilen, S. Boucherit, A. Pascale, M. Masullo, and L. Maffei, “Kentsel Ses Planlamasi icin 360 Derece Sesli video kayitlari Kullaniminin Avantaj Ve Dezavantajlari,” *15th Natl. Acoust. Congr. Turkey*, Nov. 2023.
- [3] Ö. Gök Tokgöz and M. E. Altinsoy, “Comparison of Soundscape Assessment Results On-Site and in the Laboratory,” *Proc. 10th Conv. Eur. Acoust. Assoc. Forum Acusticum 2023*, pp. 1055–1061, Jan. 2024, doi: 10.61782/fa.2023.0759.
- [4] M. E. Altinsoy, U. Jekosch, S. Merchel, and J. Landgraf, “Progress in Auditory Perception Research Laboratories - Multimodal Measurement Laboratory of Dresden University of Technology,” *AES 129th Conv.*, 2010.
- [5] J. A. Russell and G. Pratt, “A description of the affective quality attributed to environments.,” *J. Pers. Soc. Psychol.*, vol. 38, no. 2, pp. 311–322, Feb. 1980, doi: 10.1037/0022-3514.38.2.311.