The Effect of Motion-based Video Representation for on Perception of Urban Soundscape in Multimodal Measurement Laboratory Setting

Özlem Gök Tokgöz[1], M. Ercan Altinsoy[2]
[1] *Chair of Acoustics and Haptics, TU Dresden, oezlem_goek.tokgoez@mailbox.tu-dresden.de*
[2] *CeTI, Chair of Acoustics and Haptics, TU Dresden, ercan.altinsoy@tu-dresden.de*

Abstract

Virtual reality technologies are actively used in solving and understanding many urban problems. In re-experiencing urban areas in virtual reality, one of the goals is to produce scenes that are as close to reality as possible. In these virtual reality environments consisting of limited visual and auditory data, the relationship between these data is important. Visual components affect our acoustic perception. This study is aimed at investigating the effects of motion-based video display in virtual reality on urban soundscape evaluations. Audio and video recordings were made in two urban settings, one with heavy traffic and the other in a natural park. In addition, recordings of camera movements in these areas were also taken. Multi Modal Measurement Laboratory (MMM) at the Dresden University of Technology was used for this study. In this laboratory, virtual reality equipment for multiple modalities is combined, and visual data is projected onto a large screen (3.5m). Video and audio recordings were re-prepared for the MMM, and a questionnaire was applied to the participants on their perceptions. As a result of the study, an evaluation has been conducted to determine how the motion in the video affects our evaluations of the auditory and virtual environments.

Introduction

In recent years, virtual reality technologies have been increasingly used to simulate urban environments and develop more efficient solutions for urban problems. It is a useful tool for exploring various urban scenarios and cre ating designs that encourage user involvement. Furthermore, virtual reality allows for the determination of citizens' perceptions of urban problems and their levels of pleasantness or discomfort.

Virtual reality studies are especially important for urban planners and architects, as they can play an important role in the design decisions of indoor and outdoor spaces. Virtual reality tools are used in the design of important areas such as highways, urban green spaces, and historical areas. Virtual reality studies in these areas with different physical characteristics are carried out with different purposes and focuses. There have been numerous studies on the influence of environmental noise because it is a growing problem in our modern society and can be extremely disruptive to everyday life. Virtual reality technology is used to reduce environmental noise and design desirable acoustic environments. For this purpose, various topics such as highway planning[1], the effect of noise barriers[2], the effect of the aircraft [3] are examined. It is important for these studies to prepare dynamic and realistic scenes and to develop a vision for the future. Studies carried out in historical areas aims to better understand the past and make conservation decisions or document the current situation. Acoustic performances should be documented as well as visual animation of areas such as churches, theaters and concert halls, which are seen as cultural heritage. Studies on green areas mostly focus on restorativeness effect in virtual environments[1],[5]. In these studies, carried out in green areas, it has been observed that the virtual environment is not as effective as the real environment, especially for the restorativeness effect. Although restorativeness effect is less than real areas, still it is a promising finding since the number of people unable to visit natural environments increases due to urbanization. All these studies take into account the perception of the user, no matter how different they are carried out with different purposes and focuses.

Virtual reality studies, in which the user's perception can be used actively, help make cities smarter, more environmentally friendly, and more sustainable. In order to obtain reliable results on user perception, these environments should be as close to the real environment as possible. However, urban areas contain different audio and visual data; therefore, it is not easy to create these areas. There are also different requirements for the animation of these areas in the virtual environment. And it is known that our visual and auditory perceptions are affected by each other and change the user's perception. Therefore, it is important to consider both visual and auditory elements when designing the animation of virtual environments. This can greatly enhance the user's overall experience and immersion in the virtual world.

In soundscape studies, in which user perception is investigated extensively, the relations between visual and auditory environments are also investigated. The positive and pleasing effect of a place depends on various physical and individual factors as well as the green and blue elements in the environment. Studies have shown that green areas have a positive effect on our acoustic perception in urban environment [6]. For this reason, there are many studies such as acoustic properties, design, user perceptions of green area. It has been observed that the blue elements (e.g., sea, lakes, and rivers) in the environment also have an important role. Studies have shown that the water element has positive visual and auditory effects on people's perception [7]. People's backgrounds, experiences, or situations specific to their culture can also change their perceptions [8]. However, there are many areas in the city where the effect of the visual effect on our acoustic perception can be investigated. For this reason, there is a need to investigate more systematic design strategies for soundscape and visual elements that can be sustainable and integrated into long-term planning for the future of urban areas.

This study aimed to investigate the effect of visual and auditory data on each other in a virtual environment. It was uncertain whether the increase in visual data would affect our auditory perception and our immersive or being there feeling. For this purpose, audio and video recordings were taken in the two different fields. The camera was panned in on a part of each recording, and the participants were perceiving the environment from a wider angle. In order to evaluate the

effect of this pan movement of the camera on user perception, the recordings were again divided into different scenes, with and without camera movement. A survey study was conducted on these different groups. At the end of the study, the effect of camera movement was evaluated from different aspects, including being there and immersive feeling.

Methodology

In this study, the intersection area, which is near a heavily traffic and a natural park in Dresden, was selected. Audio and video recordings were taken in two different urban areas. The acoustic characteristics of the areas are different from each other. A heavily trafficked area called Straßburg Platz is an intersection area for many vehicles, such as trams, and buses. A natural park area called Carolasee is a pond located in a natural garden. There is a fountain in the middle of the area, surrounded by trees.

Preparing the Scenes

This study conducted in the Multi Modal Measurement Laboratory (MMM) which combines VR equipment for multiple modalities and is capable of presenting high-performance, interactive simulations [9]. In this laboratory, audio and video recordings were prepared for the virtual environment, with 500 speakers and room acoustics created according to DIN 15996. Visual processing is done in full HD on an acoustically transparent screen.

In order to prepare the audio and video recordings recorded for MMM, approximately 10 minutes of recordings taken in both areas were converted into two 2.5-minute HD videos. The videos were also divided into two groups. Four different scenes were prepared, with and without camera movement. In videos with camera movement, the camera panned. The scenes and their group are shown in Figure 1.

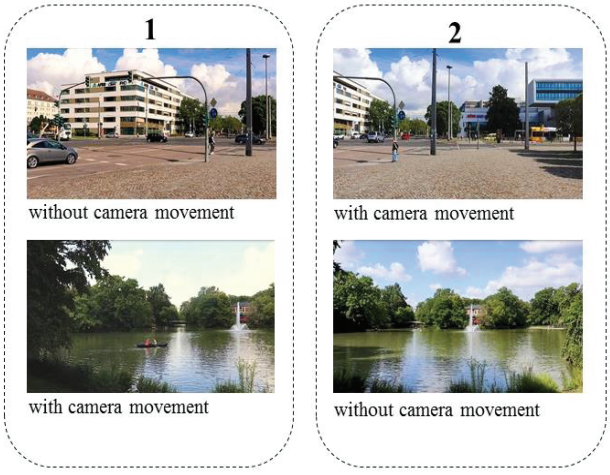

Figure 1: Scenes with and without camera movement

The videos in different areas with and without camera movement were grouped as the first and second and experienced by different participants. Since it is thought that people can experience the environment by sitting in the laboratory, the video recording was taken from a height of 1-30 cm from the ground. The prepared video is shown on the screen in front of the person in the MMM. Omni directional and binaural audio recordings were made with SQuadriga II. Binaural recordings were used in the laboratory. IOSONO Spatial Audio Workstation (SAW) is used in the lab, which enables the editing of complex sound scenes. Video and audio arrangements and synchronization are done in Nuendo4 software.

Survey Structure

The survey consists of two parts. The first part is related to the sound and visual environment, and the second part is related to the virtual environment. The questions from the two parts are explained in Table 1.

Table 1: Survey Structure

	Questions
1. Questions about Visual and Auditory Environment	How would you describe the present visual environment?
	For each of 8 scales below, to what extend do you agree or disagree that the present surrounding visual environment is…?
	What extend do you presently hear the following four types of sound?
	How would you describe the present surrounding sound environment?
	For each of 8 scales below, to what extend do you agree or disagree that the present surrounding sound environment is…?
2. Questions about Virtual Environment	To what extent did you experience a sense of "being there" inside the virtual environment?
	To what extent did you experience the feeling of "immersion" in the virtual environment?
	What was your overall comfort level in this experience?
	What was your overall enjoyment level in this experience?

The first part of the question, which includes 5 different questions, was about the visual and sound environment. Participants were asked to describe the visual and auditory environment using a 5-point scale. In the second part, participants were asked how they felt in the virtual environment, like `being there´ and `immersive´. In addition, participants were asked questions focusing on the experience itself, such as their 'comfort' and 'enjoyment` levels. The questions were mutually evaluated for these four different scenes. Four different scenes created within the scope of the study were investigated, such as how the participants perceived visual and auditory information and whether these perceptions were affected by camera movement.

Evaluation of Survey Results

The evaluation of the survey results is discussed in two parts. First, the effects of the visual and auditory environments of the prepared scenes and then the virtual environment were evaluated. Twenty-four people participated in the surveys. In the first part, there are 9 male and 3 female participants. The age range of the participants is between 20 and 36 (mean age = 29.58, standard deviation = 5.51). In the second part, there are also 9 male and 3 female participants, and the age range of them is between 25 and 37 (mean age = 30.16, standard deviation = 4.34).

Visual and Sound Environment

The evaluation of the survey results started with the question on the general evaluation of the visual and sound environment. In this question people describe the present surrounding environment with using 5point scale (1-very good, 5-very bad). In Figure 2, the average values are given for four different scenes.

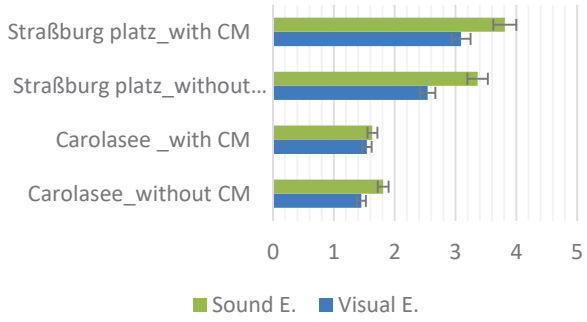

Figure 2: Results of the visual and sound environment descriptions of the participants

Carolasee, the natural park area, was found to be visually and audibly good. Averages of Carolasee are: with camera movement, visual: 1.54, auditory: 1.63; without movement, visual: 1.45, auditory: 1.81. Straßburg Platz, which is the intersection area, was found to be worse than Carolasee. Averages of Straßburg Platz are as follows: with camera movement, visual: 3.09, auditory: 3.81; without movement, visual: 2.54, auditory: 3.36. In light of these results, the visual environment is found to be more pleasant than the sound environment; this difference is more noticeable, especially in the intersection area.

In the other part of the study, questions were asked about the perceived affective quality. In the perceived affective quality part, which consists of two parts as visual and auditory, eight adjective pairs and a 5-point scale were given in the question on sound, as defined in ISO/TS 129913-2. For the visual part, although many different adjectives are used for visual environment evaluations[10], some adjectives have been removed within the scope of the study.

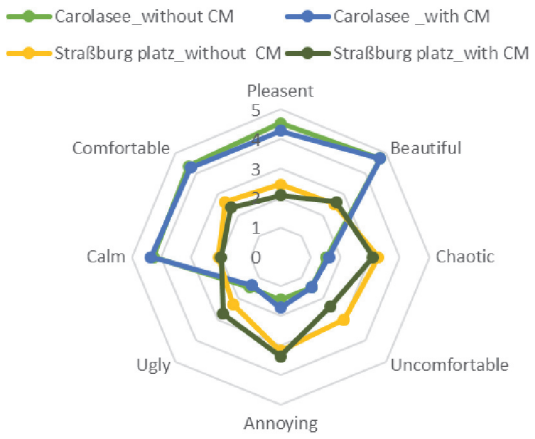

Figure 3: Survey results to visual affective quality

The answer options range from "strongly agree" to "strongly disagree", with a value between five and one. For visual affective quality, the single arithmetic values of the eight features are shown in Figure 3. Visually, Carolasee, the park area is perceived as beautiful, calm, pleasant and comfortable. The assessment of the intersection area is more difficult. Adjectives with an average value above 3 are just annoying:3.36 and chaotic:3.09. The groups with and without camera movement are almost indistinguishable from each other. According to the results, we can say that the camera movement is not very effective on the visual affective quality.

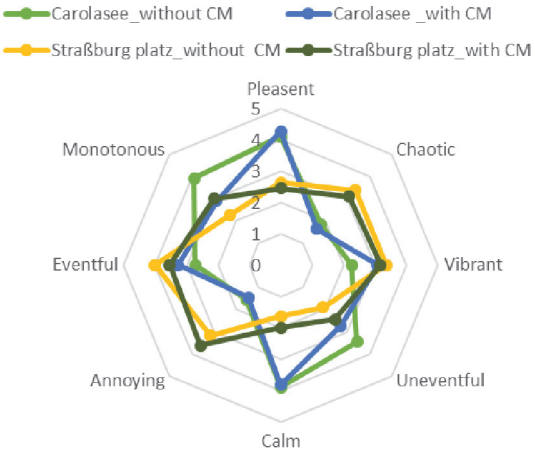

Figure 4: Survey results to auditory affective quality

For auditory affective quality, the single arithmetic values of the eight features are also shown in Figure 4. Acoustically, the park area is perceived as calm, pleasant, and monotonous. The intersection area is perceived as chaotic, vibrant, and annoying. According to the results of the Carolasee, the effect of different scenes with and without movement on the evaluation of the sound environment differs in 'monotonous' and 'uneventful' adjectives.

Virtual Environment

First, participants were asked to assess the sense of "being there" and the level of immersion they experienced while in the virtual environment, under MMM conditions. In this questions people theır opinion with using 5point scale (1-Not at all, 5-totally). In Figure 5, the average values are given for four different scenes.

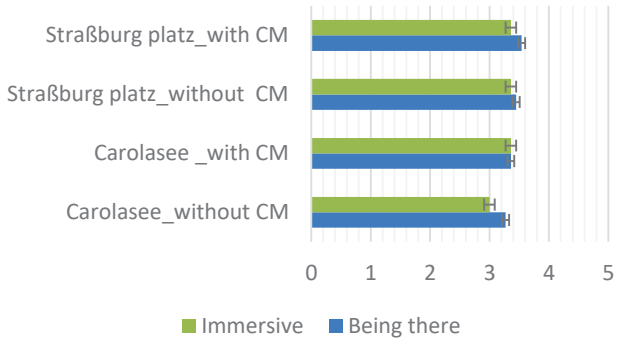

Figure 5: Results of the survey question on "being there" and "immersive".

One of the aims of the study was to investigate the effect of video motion on the 'being there' and 'immersive' feeling. For 'being there' there was a no significant difference in the scores with camera movement (M=3.4, SD=0.9) and without (M=3.18, SD=0.79) conditions; t (42) =0.83, **p = 0.380**. It was seen that the camera movement did not have a significant effect on 'being there' feeling. For 'immersive' feeling scores are with camera movement (M=3.2, SD=0.86) and without (M=3.0, SD=0.75) conditions; t (42) =0.92, **p = 0.360**. It was seen that also the camera movement did not have a significant effect on 'immersive' feeling.

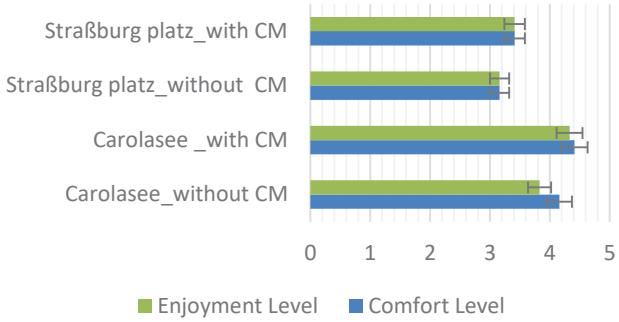

Figure 6: Results of the survey question on comfort and enjoyment level.

Secondly, participants were asked to assess the level of "comfort" and "enjoyment" under MMM conditions. In this questions people their opinion with using 5point scale (1-very bad, 5- very good). In Figure 6, the average values are given for four different scenes. According to these results, it is seen that the participants have a higher level of comfort and enjoyment in the natural park (Carolasee).

Conclusion

Based on the study conducted on camera movements in virtual environments using projection, it can be concluded that the camera movements do not significantly impact the user experience. However, it is worth noting that other factors, such as the quality of the projection, image resolution, and user familiarity with the environment, may play a significant role in determining the experience. In addition, it is thought that the low level of interaction in the pan movement may have an impact on the results. In subsequent studies, it would be beneficial to increase the quantity of similar studies and broaden their reach.

References

[1] F. Ruotolo *et al.*, "Immersive virtual reality and environmental noise assessment: An innovative audio–visual approach," *Environ. Impact Assess. Rev.*, vol. 41, pp. 10–20, 2013, doi: https://doi.org/10.1016/j.eiar.2013.01.007.

[2] G. M. Echevarria Sanchez, T. Van Renterghem, K. Sun, B. De Coensel, and D. Botteldooren, "Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space," *Landsc. Urban Plan.*, vol. 167, pp. 98–107, 2017, doi: https://doi.org/10.1016/j.landurbplan.2017.05.018.

[3] A. K. Sahai, F. Wefers, S. Pick, E. Stumpf, M. Vorländer, and T. W. Kuhlen, "Interactive simulation of aircraft noise in aural and visual virtual environments," *Appl. Acoust.*, vol. 101, pp. 24–38, 2016.

[4] V. P. Senese *et al.*, "The Influence of Personality Traits on the Measure of Restorativeness in an Urban Park: A Multisensory Immersive Virtual Reality Study," 2020.

[5] T. Nukarinen *et al.*, "Physiological and Psychological Restoration in Matched Real and Virtual Natural Environments," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, Apr. 2020, pp. 1–8, doi: 10.1145/3334480.3382956.

[6] H. Nordh, T. Hartig, C. M. Hagerhall, and G. Fry, "Components of small urban parks that predict the possibility for restoration," *Urban For. Urban Green.*, vol. 8, no. 4, pp. 225–235, Jan. 2009, doi: 10.1016/j.ufug.2009.06.003.

[7] J. Y. Jeon, P. J. Lee, J. You, and J. Kang, "Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds," *J. Acoust. Soc. Am.*, vol. 127, no. 3, pp. 1357–1366, Mar. 2010, doi: 10.1121/1.3298437.

[8] B. Kiser and D. Lubman, "The soundscape of church bells - sound community or culture clash," in *Euronoise*, 2008, pp. 9434–9437.

[9] M. E. Altinsoy, U. Jekosch, S. Merchel, and J. Landgraf, "Progress in Auditory Perception Research Laboratories - Multimodal Measurement Laboratory of Dresden University of Technology," 2010.

[10] 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.