Creating a Virtual Reality Orchestral Concert Experience With 3D Audio

Per Magne Røsvik

Master's thesis in Software Engineering at

Department of Computing, Mathematics and Physics,

Western Norway University of Applied Sciences

Department of Informatics,

University of Bergen

June 2020

Western Norway University of Applied Sciences

Abstract

With the emergence of Virtual Reality (VR) and recent advancements on both the hardware and software side, VR is now available to the masses and a plethora of new experiences can be made and explored by people and researchers. In this thesis, we have created a VR orchestral concert experience in collaboration with Bergen Philharmonic Orchestra. Additionally, we have explored how 3D audio works in the context of VR, and we have provided a novel method for creating 3D audio for a VR orchestral concert experience, by using a recording of the orchestra performing Tchaikovsky's Symphony No. 5, where we had individual audio tracks from each of the orchestra's different instrument groups. With the game engine Unity and a spatialization Software Development Kit (SDK) from Oculus, we created the 3D audio from the recording of the orchestral concert. The concert was also recorded with a spherical 360° camera, placed directly in front of the conductor. Allowing the user to experience the concert from a unique perspective, which is not normally possible. The results from our user testing show us that the VR experience we have created, proved to be quite immersive, and the users noted that they could perceive the 3D audio. This, in turn, adds validity to our method for creating the 3D audio in our VR experience.

That was actually really close
to what it's like to stand
on the conductor's podium

BJØRN SAGSTAD

Norwegian Conductor

Preface

This project was a collaborate effort between Western Norway University of Applied Sciences (HVL) and Bergen Philharmonic Orchestra.

Enthusiasm surrounding the project is large, and everyone who hears about it gets excited, and wonders when the Virtual Reality (VR) experience we have created will be released to the public.

I was given the opportunity to present the project at *Forsking i fyr og flamme* at HVL, alongside two fellow master's thesis projects related to VR, on April 30, 2020.

The COVID-19 pandemic outbreak undoubtedly affected all of our lives the spring of 2020, and it also affected this project. Certain tests that were planned in order to evaluate our results, had to be cancelled due to health and safety concerns.

Acknowledgements

Contents

List of Figures

viii

List of Tables

List of source codes

Glossary

3D audio Refers to audio that surrounds the user in a 360° sphere.. i, vii, viii, 2–6, 8–10, 12, 20, 21, 29, 31, 41–43, 46–48, 51, 52, 55, 58, 59, 61, 62

6DOF Six degrees of freedom; refers to the freedom of movement of a rigid body in three-dimensional space.. 16

ambisonics Refers to either, the full-sphere surround sound format, or the microphones used to record audio in said file format. vii, 4, 5, 9, 10, 31, 40, 62

Android A mobile operating system by Google, based on a modified version of the Linux kernel.. vii, 16, 17, 39, 49, 56

Asset In video game development, an Asset is a resource or binary file containing data, such as textures, 3D models, audio files and more.. 27, 29, 39

spatial audio Synonym to 3D audio. 58

spatialization Spatialization refers to the creation of spatial audio, or 3D audio, from monophonic audio sources.. i, viii, 2, 3, 5, 9, 12, 31, 33

XR Umbrella-term for Augmented Reality (AR) and Virtual Reality (VR). vii, 13, 19, 22, 23, 28, 31

Acronyms

2D Two-dimensional space. 4, 17, 22, 50

3D Three-dimensional space. vii, viii, 6, 8, 9, 11–13, 15, 17, 18, 22, 24–29, 31, 34–36, 40, 42, 50, 56, 57, 59, 61

AAX Avid audio eXtension. 12

API Application programming interface. 17

apk Android Application Package. 39, 49

BIM Building Information Model. 24

CPU Central processing unit. 33, 34, 56, 57

DAW Digital audio workstation. 12, 21, 62

FPS Frames Per Second. 22, 56, 57

GPU Graphical processing unit. 16, 57

HMD Head-mounted Display. vii, 13–16, 18, 28, 36, 52

HRTF Head-Related transfer function. 11, 12, 48

MVP Minimum viable product. 3, 21, 31

ONSP Oculus Native Spatializer. 9, 12, 24, 25, 31, 33, 34, 42, 48, 56, 57, 61, 62

OS Operatingsystem. 16

SDK Software Development Kit. i, 2, 3, 9, 12, 17–19, 29, 58

UI User Interface. 36

VR Virtual Reality. i–iii, vii, viii, 1–6, 8, 10, 12–23, 27, 28, 31, 34, 36, 37, 39, 40, 42, 43, 45, 48, 50–52, 55–62

VST Virtual Studio Technology. 12

Chapter 1

Introduction

This chapter acts as the overall introduction to the thesis. It outlines the projects origins, motivation, goals and research questions. Additionally, related work to the the thesis is presented.

1.1 Thesis Outline

Introduction: The first chapter introduces the motivation and context for this thesis and project. It contains an overview our goal and research questions, as well as related work.

Background: The background chapter introduces the theoretical background of topics and techniques used in this thesis for the development of our VR experience.

Design and Solution: This chapter describes how our VR experience was built and developed.

Analysis and Assesment: Here we describe how we tested our experience with the test groups, and provide the results from the testing. A numerical performance test is also described and presented, in addition to the user testing.

Conclusion: Here we summarize our work and goals in relation to the research questions.

Further Work: Finally, we discuss and describe potential future additions or improvements to our VR experience.

1.2   Motivation

With the emergence of Virtual Reality (VR) comes new opportunities for people to experience a vast range of activities, that normally might not be accessible to them, from the comfort of one's own home. For this master's thesis the goal is to look into the 3D audio aspect of VR, and ultimately create an orchestral concert experience.

Bergen Philharmonic Orchestra [24] wishes to extend their offering of music and culture with the usage of VR. A very important aspect of this is the audio, and how we can translate the experience of going to an orchestral concert to VR. Additionally, in VR it will be possible to let the user to sit in the orchestra itself, alongside all the other musicians, giving the user an unique experience. Furthermore, it will be possible to expand the VR experience to include multiple viewing positions of the orchestral concert. Bergen Philharmonic Orchestra wishes to bring this experience to people who normally cannot attend their concerts, and to show people what an orchestral concert is like to experience as a musician.

In order to recreate the experience of going to an orchestral concert we want to look into the usage of 3D audio in order create a believable audio experience from the concert hall. To create 3D audio we will be using a recording of the orchestra, where microphones are placed around in the different instrument groups of the orchestra. Then, we will use a game engine and spatialization SDKs in order to spatialize the audio and to calculate how the user's ears should perceive the audio from the orchestra depending on the position of the user. Additionally, we will look into the visual part of the experience using spherical 360° video recordings of the concert.

The reason why we are doing this, as opposed to recording 3D audio using specialized microphones used to record 3D audio [25, p. 32-33], is that Bergen Philharmonic Orchestra together with Gunnar Herleif Nilsen from NRK [26] already has the equipment and extensive knowledge of recording the orchestra. By using the proposed method there is no need for new recording equipment nor the knowledge to use the new equipment. Additionally, the proposed method will allow us to use one audio recording, for any number of 360° or 180° video recordings of the same concert, from cameras placed wherever we want the user to experience the concert from. We can also expand the experience with the functionality to move from position to position during the concert, to experience different viewpoints. Lastly, we can also add the possibility to move around in the 3D model of Griegsalen, and experience how the audio differs from position to position in real time.

1.3 Goal

The goal of the thesis is to determine the feasibility of using the audio from multiple microphones placed around the different instrument groups in the orchestra, in order to create 3D audio using a game engine and spatialization SDK. The feasibility is determined by verifying and validating the resulting VR experience using quantitative and qualitative methods and allowing a group of people to view the experience, and afterwards answer a questionnaire about the experience.

1.4 Research Questions

- What is the feasibility of using the audio recording from multiple microphones, along with a game engine and spatial audio SDK in order to create believable 3D audio of an orchestral concert?

- How believable and well perceived by user of the experience is the resulting 3D audio and VR orchestral concert?

1.5 Methodology

In order to answer the research questions, a Minimum viable product (MVP) of an VR orchestral concert experience is designed and implemented. Deciding what sort of an experience this should be, and what Bergen Philharmonic Orchestra [24] wants the public to experience with a VR experience like ours, was discussed in meetings with key people from Bergen Philharmonic Orchestra during the early stages of the project. We then started preparing for the day of recording an orchestral concert performed by Bergen Philharmonic Orchestra. The design and implementation of our VR experience is described in Chapter 3. After the VR experience was developed and implemented we started user testing, as described in Chapter 4, in order to evaluate the resulting VR experience. The VR experience was developed iteratively. I held regular meetings with both my supervisors and our external contact from Bergen Philharmonic Orchestra, Ingeborg Ekeland, were we discussed the progress of the development of the VR experience and what the next steps in development should be. The main platform used for the development and testing of the VR experience was the Oculus Quest [16], a VR system made by Oculus. Throughout the project, the VR experience was continuously tested and used by the author as new features were added or improved.

1.6 Scope and Limitations

This thesis is primarily rooted in the software engineering field, however in many areas of this thesis we are crossing over into the audio engineering field due to our focus on 3D audio in VR. Terms and techniques used in the audio engineering field and music production will only be briefly explained and not explored in-depth.

1.7 Related Work

There are many examples of using 3D audio recording and 360° video from orchestras and other music venues. However, most of them are 2D 360° videos uploaded to YouTube [27], which gives an idea of the concept. A more notable example is Philharmonia Orchestra's *Virtual Orchestra* [28] which features higher resolution video and ambisonics audio. As seen in the video posted by Philharmonia Orchestra [28], they used special microphones in order to record the audio, as seen in figure 1.1. Philharmonia Orchestra also has an experience where users view a VR movie, while an speaker array with 18 speakers, placed around the users, are playing a recording from the orchestra [29]. Our VR experience will be quite similar to Philharmonia Orchestra's *Virtual Orchestra* [28], with one of the key differences being the way we create the 3D audio.

In the paper *Immersive orchestras: Audio processing for orchestral music VR content* [30], they create a proof of concept of using multiple audio tracks from a recording of an orchestra, in order to create what they call *acoustic zoom effect*. Their result is a prototype that allows the user to emphasize the audio from an instrument by looking at it, and then hearing only that instrument. What we are doing is similar, in that we are using multiple audio tracks from several instrument groups. Despite this, we are using them to create immersive 3D audio depending on the position the user is experiencing the concert from.

Figure 1.1: Screenshot from Philharmonia Orchestra's video [1] about their VR experience, showing the Omni Binaural Microphone [2], made by 3DIO [3], that they used to record the audio.

In the paper *The plausibility of a string quartet performance in virtual reality* [31], they discuss creating a string quartet performance in VR and how stereo vs 3D audio affect player immersion. The paper mainly focuses on player immersion in a virtual environment using the concepts of *being there*, or *Place Illusion (PI)*, and *Plausibility (Psi)*, the illusion that the events in the virtual environment are really happening [31] in order to gauge various levels of immersion in respect to the overall virtual environment. Some of the aspects include the spatialization of the audio from a string quartet performance, and the aurealization of the audio in the room. I.e, how the audio travels around the room. We are using some different tools and techniques in order to achieve this on a bigger scale.

Other similar VR examples include MelodyVR [32] which provides live music performances from popular artists. In which the user views a live recorded 360° video along with live audio recorded from an ambisonics microphone. Oculus Venues [33] is another example, where the users are placed in a virtual cinema where they can watch a big virtual screen with videos of concerts and other events.

Outside of VR, Dolby Atmos [34] is a great example of 3D surround sound in which there are several speakers around the user and the audio can move around the user, similar to how 3D audio works in VR. Sony has created a new spatial audio technique for headphones, called 360 Reality Audio [35], using a method they call object-based spatial audio technology. The audio from different instruments and vocals are placed in a 360° spherical sound field by a producer. However, there is no head-tracking involved, meaning the audio direction does not change when the user turns their head, which however is the case for our project. Users can also scan their ears with a phone app in order to get a personalized head-related transfer function [36, p. 41 - 67], a mathematical function that calculates how the audio from a point source in space should hit the ears of the user, in order to improve the audio. Sony's 360 Reality Audio is currently only available through selected music streaming services.

1.8   Pyotr Ilyich Tchaikovsky

Pyotr Ilyich Tchaikovsky, born May 7, 1840 and died November 6, 1893, was a Russian composer who also played the piano [37]. Tchaikovsky is cited as being the most popular and famous Russian composer [38]. His works include several symphonies, operas, ballets, concertos and much more. Including famous works like *Swan Lake* [39], *The Nutcracker* [40] and *Violin Concerto in D Major* [41].

1.8.1   Symphony No. 5 in E minor

Tchaikovsky's Symphony No. 5 in E minor, was composed and orchestrated between May and August in 1888 [42]. The symphony consists of four movements:
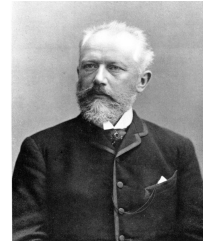
Figure 1.2: Tchaikovsky, aged 46. Photographed by Charles Reutlinger in Paris, 9 June, 1886 [4].

- I. Andante—Allegro con anima (E minor)

- II. Andante cantabile con alcuna licenza (D major)

- III. Valse. Allegro moderato (A major)

- IV. Finale. Andante maestoso–Allegro vivace (E major).

While Tchaikovsky initially considered the symphony a failure, it was enthusiastically received by his friends in Moscow [42], and has become one of his most popular symphonies. The symphony was first performed November 17, 1888 in Saint Petersburg by the Saint Petersburg Philharmonic Society, conducted by Tchaikovsky himself. Bergen Philharmonic Orchestra performed the symphony on November 14, 2019, conducted by chief conductor Edward Gardner [43].

Chapter 2

Background

This chapter introduces the theoretical background to the topics, tools and techniques used in this thesis.

2.1 3D Audio

3D audio, also called spatial audio, is about creating a convincing sound-field surrounding the user, as visualized in figure 2.1. Humans are able to pinpoint where a sound is coming from, and we need to mimic this ability in virtual 3D environments in order to create convincing 3D audio for VR. In the last couple of years, interest in 3D audio has increased due to the more common availability of VR and Augmented Reality (AR) systems and applications [25, p. 1]. When creating VR content and experiences, the audio can be just as important as the visuals, especially in our case [31, 44, 45].

There are multiple ways of creating 3D audio. One method is to use special microphones which records the audio from multiple directions. One such microphone type are binaural microphones. Binaural microphones aim to simulate the head of a person, recording the audio in two channels. [46] [25, p. 31] One for the left ear, the other for the right ear. However, normal binaural recordings are not suitable for VR because they lack positional awareness of the user. This means that the binaural recording, and subsequent binaural audio is only believable if the user sits still in one place. In VR, however, the user moves around and turn their head; thus the perceived audio becomes wrong from the user's perspective if it is paired with visuals, like a video. This is due to the fact if a sound source is supposed to be to the left of the

user, and that is reflected in the binaural audio by primarily playing the sound in the left channel, if the user now turns around or the sound source is positioned differently, the binaural recording is no longer correct from the user's perspective.

Figure 2.1: Visualizing 3D audio [5].

Therefore, you can use ambisonics microphones [47] instead, such as the RØDE NT-SF1 [6] or Sennheiser Ambeo [7], seen in figure 2.2. A first-order ambisonics microphone, sometimes referred to as a soundfield microphone, is actually four microphones positioned together in the shape of a tetrahedron in order to record the audio from different directions [25, p. 32-33] [47]. This allows the audio to stay believable from the user's position even if the user turns their head, and the user can tell which direction the audio is coming from. This technique was mostly developed by Michael Gerzon, during the seventies [48]. Higher-order ambisonics microphones, using more channels and microphones, exist as well. The higher order of the ambisonics, the more accurately the user can tell the direction of the audio [49]. In order to playback audio recordings done by ambisonics microphones, one needs software capable of decoding the ambisonics audio format.

Another method is simulating the 3D audio using game engines and spatialization SDKs. Multiple audio sources are placed in a 3D environment in a game engine, and with the usage of 3D audio spatialization SDKs like Oculus Native Spatializer (ONSP) [50] or Steam Audio [51] one can spatialize each audio source such that the user can correctly perceive its

Figure 2.2: The RØDE NT-SF1 microphone [6] (left), and the Sennheiser Ambeo [7] (right), two examples of first-order ambisonics microhpones.

position in the virtual environment. This is the method we will be using for our VR experience, as it will allow us to use a single audio recording from the microphones, that are already used when a audio recording is done of a orchestral concert performed by Bergen Philharmonic Orchestra [24]. Had we used ambisonics microhpones for the audio recording, the user would be locked to where the microphone was positioned during the concert. This would also make the audio recording setup unnecessarily complex, as they are already recording the audio for live-streaming and radio broadcasts. With our method, we will use the audio from the microphones they are already using for the live-stream and radio broadcast. Additionally, we can potentially add the feature allowing the user to *walk* or *fly* around in a virtual version of Griegsalen, and experience how the audio changes depending on the position within the concert hall. Furthermore, we can add any number of 360° or 180° cameras to the recording setup, and add a feature allowing the user to change positions during the experience. Had we used ambisonics microphones, then one such microphone would need to positioned together with each camera, and the recorded audio would subsequently needed to be produced and mixed. This would then add a lot of extra work for any additional camera position one wishes to have in the VR experience, but with our method, the same audio recording and setup in Unity [52] is used for all potential camera positions and positions the user can experience. As the creation of 3D audio is based of the user's position in the virtual version of Griegsalen.
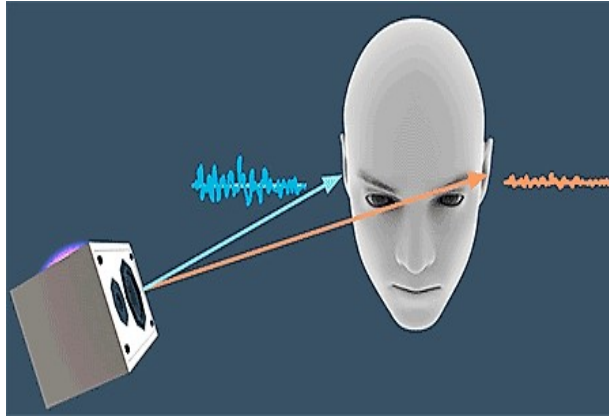
Figure 2.3: How sound waves from a source hits a person's ears [5].

2.1.1   Head-Related Transfer Function

A Head-Related transfer function (HRTF) is a function that describes how a sound from a specific point in a 3D space, arrives at the ears of a person [36,44]. This function is used to trick the user into believing the audio comes from a specific point in space. The audio used in these cases are typically recorded very close to the source. When the audio from the audio sources hits the outer ear, the outer ear causes spectral filtering and minute audio delays before it hits the ear drums [36, p. 41 - 67]. The audio from a source will also hit each ear at different times, as seen in figure 2.3. Every single person has a different set of unique ears, and thus how the audio from the audio source changes when it hits the user's ears is going to be different for everyone. Additionally, a person's head and torso will also have some effect on how the audio will hit both ears differently [44].

Therefore, adding individualized HRTF for each person would be the ideal scenario. However, this is not currently feasible, as measuring the HRTF for a person requires a lab with the appropriate equipment [53]. Any media said to use HRTF, will usually be using a generalized HRTF based on a single person, or a mannequin model [54], or using data from a HRTF database like the CIPIC HRTF database [53]. Using a generalized HRTF is not always ideal, as if the persons measured HRTF differs a lot from the general HRTF used, the person will not perceive the audio correctly as intended. Localization errors often occur, as the person will not be able to accurately pinpoint the location the audio is coming from. The most typical

confusion is whether the sound originates in front of the user, or behind them. Wenzel et al. measured more than 30% (mean value) of front to back confusion when using non-individual HRTF [55]. Michael Abrash [56], Chief Scientist at Oculus, made at prediction at Oculus Connect 5, Oculus' annual developers conference, that personalized HRTF will be able to be generated at home by users in the near future [57]. Sony currently offers this ability with their 360 Reality Audio [35], where users can generate a personalized HRTF by taking a picture of their ears. Despite this, it is only available for use with Sony's system and cannot be incorporated into other applications.

2.1.2   Oculus Native Spatializer

Oculus Native Spatializer (ONSP) [50] is a 3D audio spatializer available for game engines like Unity [52] and Unreal Engine [58], as well as audio solutions like FMOD [59] and WWise [60], and Digital audio workstations (DAWs) like Avid Pro Tools [61] and other VST and AAX capable DAWs. ONSP provides features like spatialization and head-tracking, which transforms monophonic sound sources to make it seem like they originate from a specific point in a 3D environment. Furthermore, it provides real-time reverberation and occlusion simulation based on the geometry of the 3D environment. Reverberation, reverb for short, refers to how sound waves reflect off of different surfaces in a room; before reaching the listener [62, 63]. Methods for simulating the acoustics of a room and calculate reverb, typically involve computationally heavy *ray-tracing* and *path-tracing* [64].

2.1.3   Alternatives to ONSP

Similarly to the aforementioned ONSP, Steam Audio is a 3D audio spatializer available for use in development of VR applications and video games [51]. Like the ONSP, Steam Audio is available for use with Unity [52] and Unreal Engine [65], as well as a standalone SDK for implementing in other game engines or standalone applications. Resonance Audio [5] is yet another spatial audio SDK with a similar set of features as the previous two mentioned, also available for use in Unity, Unreal Engine and on mobile. All three share some common features, but they differ in their implementation and approach to implementing 3D audio [44, 66]. Nevertheless, these are only some of many spatialization SDKs available to developers and music producers, in order to create 3D audio. ONSP was mainly chosen for its promised list of features, and for the fact that it is made by Oculus, promising better integration and performance on the Oculus Quest [16]. As Oculus Quest is one of our main

platforms we are targeting. However, both Steam Audio and Resonance Audio are available to be used on the Oculus Quest as well.

2.2   360° Cameras and Spherical Video

360° cameras, also referred to as omnidirectional cameras, are cameras that, through multiple lenses, captures a full 360° view. In order to capture 360° spherical video, one needs at least two lenses placed in the opposite direction of one another. Nevertheless, any number of lenses in either a circular or spherical configuration can be used to capture spherical video. The basic configuration with two lenses, used in cameras such as the Insta360 One X [67], VUZE XR [8], or GoPro Max [9], use two fish-eye lenses with a wide field of view of over 180°, as seen in figure 2.4. At the overlaps between the lenses, the captured video needs to be stitched together using software, such that the image of each lens blend in with each other. However, objects that are exactly in the overlaps of the lenses may appear distorted, depending of how well the aforementioned stitching was performed.

When using a configuration of several lenses, one can also capture spherical video in stereoscopic 3D, which allow for depth perception in the resulting video, giving a better viewing experience when using a VR Head-mounted Display (HMD). Cameras like the Insta360 Pro 2 [10] uses six lenses instead of two in order to achieve stereoscopic 3D spherical video recording.

Figure 2.4: The Vuze XR [8] (left) and the GoPro MAX [9] (middle), two examples of 360° cameras using two lenses. The Insta360 Pro2 Camera [10] (right), featuring six lenses for high resolution 3D 360° video.

Figure 2.5: The Sword of Damocles [11], created by Ivan Sutherland [12]

2.3   Virtual Reality

Virtual Reality (VR) allows the user to experience highly immersive computer generated simulations, in which the user may also interact with the virtual environment in various ways. VR is experienced through wearing a HMD. Most modern VR systems usually have a pair of tracked controllers as well, which are used for interacting with the virtual environment. VR has come a long way since its early days in the sixties with the *Sword of Damocles*, created by Ivan Sutherland [12], as seen in figure 2.5. This VR HMD was large, bulky and so heavy it needed to be suspended from the ceiling in order to be used. In 2016, the first two modern VR systems ready for consumers, the Oculus Rift [14] and the HTC Vive [15], were released, as seen in figure 2.7. As the popularity and sales of VR increases for each year that passes [68], we are seeing continuous development and improvement of VR systems.
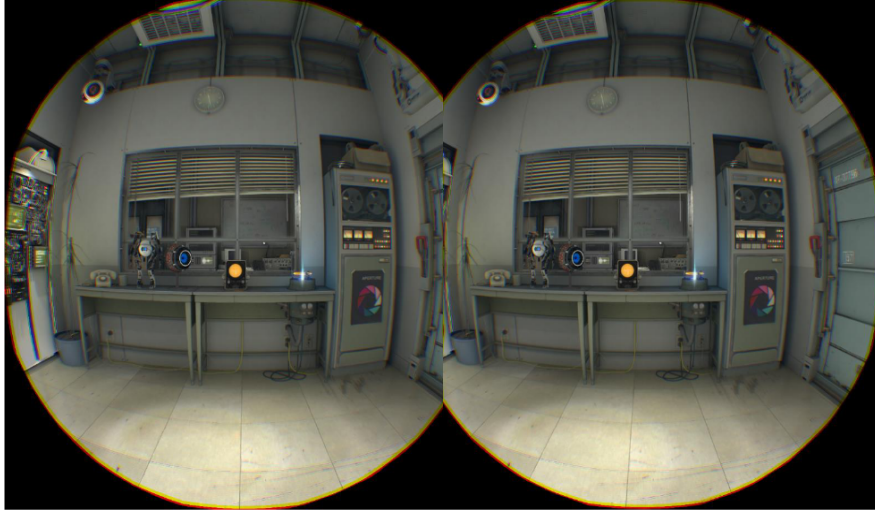
Figure 2.6: Example of how the left and right images are offset and displayed on the HMD [13].

2.3.1 Technology

When wearing a VR HMD, the user's entire field of view is covered by the HMD. Inside the HMD there is one or several screens. In front of the screen(s) there is a pair of lenses, one for each eye. In order to create 3D imaging and the perception of depth, two images are rendered, one for each eye and screen, which are slightly offset from one another, as seen in figure 2.6. The images are slightly offset from another due to the fact that humans have two eyes, which are slightly offset from one another. This means that the left and right eye sees the same thing from a slightly different angle, and together the two images simulate stereopsis, allowing us to see depth [69, p. 921 - 924]. In the case of HMDs with a single screen, the screen is split in two and a different image is displayed on each half of the screen. The need for rendering two different images for each frame increases the computation costs and performance requirements compared to non-VR applications [69]. Another fundamental part of VR systems is head-tracking, such that when the user moves their head, the view in the virtual environment changes in the way the user expects it to when moving their head. Changing the user's view, or moving the user around in the virtual environment, without input from the user is one of the leading cause for VR-sickness [70], including symptoms like nausea and headaches. This is due to the mismatch caused

Figure 2.7: The Oculus Rift [14] (left) and the HTC Vive [15] (right), two PC VR systems.

by the information the user receive through their vestibular system and the visual data displayed in the HMD.

Modern VR Systems

VR system like the aforementioned Oculus Rift and HTC Vive requires a PC with a higher-end GPU in order to be used, and the VR HMDs are tethered to the PC. Although wireless add-ons do exist for both systems. The HTC Vive is a SteamVR [71] VR system. SteamVR was created by Valve [72], and Steam [73] is the largest platform on PC for video games and VR applications.

The Oculus Quest [16] is an entirely standalone VR HMD that does not require a PC to run. It features 6DOF tracking using four cameras attached to the HMD. It has two screens, each with a resolution of 1440x1600 pixels and a refresh rate of 72Hz. The Oculus Quest's OS is Android based. Due to the mobile hardware used, the Oculus Quest is a fair bit weaker than desktop PCs used for PC VR, we are therefore limited in the amount of details and complexity the VR experience can have. We chose the Oculus Quest as our main target platform, as due to its mobile nature, it is more accessible to the general public and easier to use. Despite this, we are still targeting other VR platforms as well, namely PC VR, for our VR experience.

Figure 2.8: The Oculus Quest [16], a standalone VR system running on mobile hardware and Android.

2.4   Game Engines

Creating a game engine requires a lot of resources and time, as there are a lot of complicated systems that needs to be implemented and fitted together. Using a game engine in order to develop video games and other similar applications speeds up development time, as game engines offer the implementation and tools of common and ubiquitous techniques used in the development of said applications. The game engine typically features an editor where the developer can build their 3D or 2D scenes that make up the application. Additionally, it provides an API to standard features, such as rendering, audio, physics and animation. Using a game engine in our project is a no-brainer, as modern, free and commercially available game engines, like Unity [52] and [65], offer all the features we need in order to create our VR experience. Otherwise, features not available directly in the game engine, are accessible through plugins or external SDKs. Game engines also make targeting multiple platform easier, by providing build options for the most common platforms.

2.4.1   Choosing a Game Engine

At the start of the project, we primarily considered two different game engines; Unity [52] and Unreal Engine [65]. Unity is a popular game engine used by both amateurs and professionals for making 2D and 3D video games and experiences. It also has extensive VR support and runs on any modern platform, allowing us to target multiple platforms with our VR experience. Unreal Engine [65] is another powerful and popular game engine, and can also

target multiple platforms. Although Unity and Unreal Engine have a quite similar set of features, Unreal Engine is typically regarded as the better choice in terms of graphical fidelity and post-processing capabilities [74]. Wherein Unity is often cited as the better choice for mobile development [74]. Unreal Engine is also open source, while Unity is not. Unity was ultimately chosen for this project, mainly due to the author's prior experience with developing 3D applications in Unity, but also due to Unity's extensive VR support, and as it is often recommended for development on mobile platforms.

2.4.2   VR Support

VR support is naturally an important factor for our project and the VR experience we have created. The aforementioned game engines both feature extensive VR support. At the start of the project, Unity version 2019.2 was the latest released by Unity. In this version, one had to use different VR SDKs, depending on which VR system one wanted to target. For example, if one wanted to target the Oculus Quest [16], we had to use the Oculus Integration for Unity [75]. Furthermore, if we wanted to target SteamVR [71] HMDs we would have had to use the SteamVR Unity Plugin [76]. Although the various plugins for developing to all the different VR systems do offer some cross-platform capabilities, but as seen in figure 2.9, it used to be a complicated mess targeting multiple VR systems.
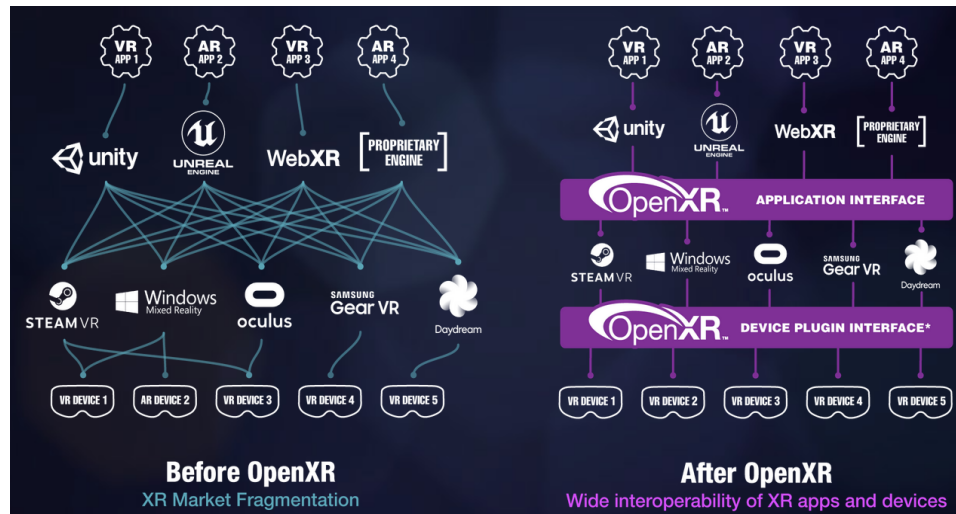
Figure 2.9: The OpenXR standardization effort made by the Khronos Group [17].
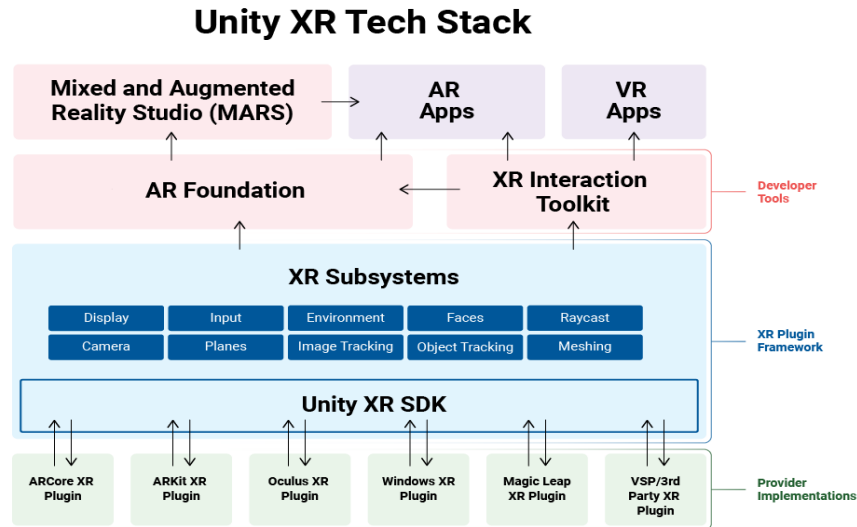
Figure 2.10: The Unity XR Tech Stack in Unity version 2019.3 [18].

The Khronos Group has put a lot of efforts into the standardization of developing for multiple VR systems at once, through the OpenXR [17] initiative, first released July 29, 2019. With the release of OpenXR, and as OpenXR is gradually being supported by the various game engines and VR system manufacturers, the development of cross-platform VR applications has been streamlined, as seen in figure 2.9. As the Oculus Quest was our main target platform for our VR experience, we started development with Unity 2019.2 and the Oculus Integration for Unity. During the project, Unity version 2019.3 was released, and with it, an upgraded plugin-based architecture approach to XR development [18], seen in figure 2.10. The Unity XR Interaction toolkit [77] was also released alongside Unity 2019.3, this toolkit acts as a replacement for the features previously provided by the various VR SDKs like Oculus Integration and SteamVR Plugin. When we upgraded our Unity project to Unity 2019.3, we also replaced the Oculus Integration for Unity with the Unity XR Interaction toolkit. This made targeting multiple different VR systems and platforms easier.

19

Chapter 3

Design and Solution

In this chapter, we will describe the entire process of creating our VR experience of a orchestral concert with Bergen Philharmonic Orchestra [24] performing Tchaikovsky's fifth symphony [78]. From recording the audio and video, to creating the 3D audio, and creating a 360° video player in Unity [52].

3.1   Audio Recording

The audio recording was handled by Gunnar Herleif Nilsen, a sound engineer from NRK [26]. The audio was recorded during a live concert in Grieghallen [19] where Bergen Philharmonic Orchestra [24] performed Tchaikovsky's fifth symphony [78]. At the same time, this concert was produced for a live-stream on BergenPhiLive [79, 80], Bergen Philharmonic Orchestra's live streaming service. A total of 21 microphones were used for the recording of the symphony. Gunnar Herleif Nilsen has extensive domain knowledge of recording orchestras and music production. The microphones are placed in such a way that you have microphones focused on each instrument group, in addition to two main microphones hanging above the conductor. There are also a couple of microphones hanging above the audience. As seen in figure 3.1 and figure 3.2, some of the microphones are hanging from the roof, while some are on microphone stands.
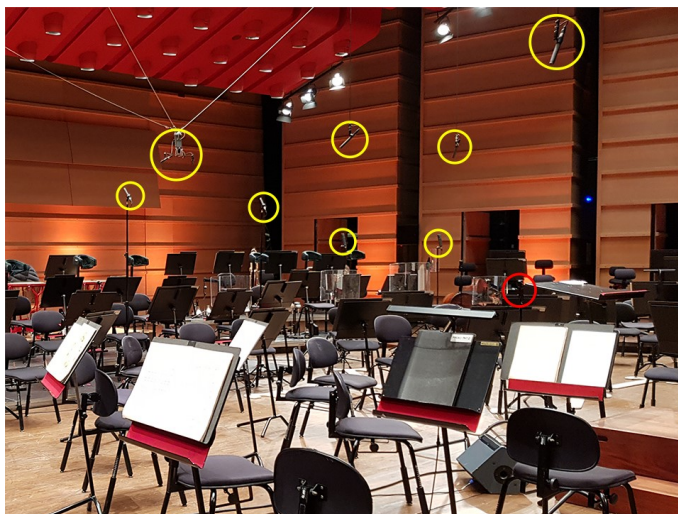
Figure 3.1: The scene in Griegsalen before the concert. Yellow circles are some of the microphones. Red circle is the 360° camera we used for recording the video.

For the purpose of our goal and VR experience, we are only using the audio recording from the microphones focused on each instrument group, in order to create 3D audio. The author of this thesis did not have any influence regarding neither position of the microphones positions nor the recording of the audio.

3.1.1 Preparing the Audio for Use in Unity

The resulting audio file from the recording consisted of 25 tracks with a total size of 10.3 GB. Fifteen of these tracks correspond to one instrument group in the orchestra, and it was these fifteen we used to create the 3D audio, the process of which is explained later in chapter 3.6. The other tracks include the produced stereo mix, the live stream of the concert, and the audio from ambiance microphones. These were not used in the creation of the 3D audio. Tchaikovsky's fifth Symphony [78] consists of four movements, and for the purpose of creating the MVP of our experience, we are limiting the experience to one movement in the symphony. We therefore needed to split the audio between these movements. In order to accomplish this, we used the Digital audio workstation (DAW) REAPER [81]. REAPER allows for easy splitting of the audio files, and rendering the audio at the desired sample rate. Any audio production and mixing can also be performed here,
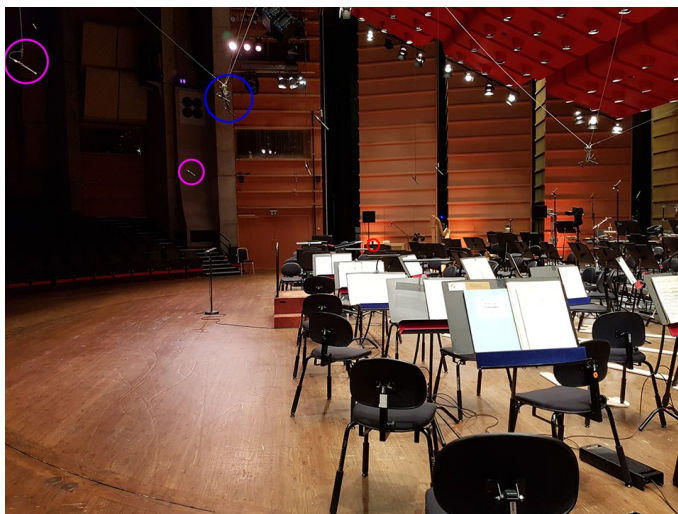
Figure 3.2: Another view of the scene and the microphones. Blue circle are the main microphones, pink circle are ambience microphones. Red circle is the 360° camera, same as before.

before importing the audio into Unity [52], should it be necessary or should one desire to do so. However, the author of this thesis is not a sound engineer, nor a producer, and does not have any knowledge in mixing and producing audio. The audio files have therefore been used as we received them from NRK [26].

3.2 Spherical Video Recording

For the video recording we used the camera Vuze XR [8], as seen in figure 2.4. This camera has a max resolution of 5760*2880 pixels at 30 FPS. It uses two 210° fish-eye lenses in order to record 2D 360° video. The Vuze XR [8] is also capable of recording footage in 180° stereoscopic 3D. The camera was provided by *Senter for nye medier* [82] at Western Norway University of Applied Sciences [83]. The camera was placed between the conductor and the first row of musicians as seen in figure 3.1 and figure 3.2, as this was one of the positions Bergen Philharmonic Orchestra [24] wishes the users to experience in our VR experience. The original plan was to have a second camera as well, placed in a different position within the orchestra, such that the user can choose whichever position the user wants to view the orchestra

from, either before or during the experience and switch between all positions at the user's choice. However, the second camera malfunctioned on the day of recording, therefore the ability to chose multiple positions is not a feature of the current version of our VR experience. Despite this, it can be added in the future, should one desire to expand the VR application with more 360° recordings of Bergen Philharmonic Orchestra [24]. This topic is explored further in chapter 6.

3.2.1 Preparing the Video for Use in Unity

In order to process the footage recorded by the Vuze XR [8], one needs to use Vuze VR Studio [84]. Like the audio, the video footage was divided between the four movements of the symphony, as we are only using the fourth movement in our prototype. As mentioned and explained further later in chapter 3.9, the video bit rate had to be reduced from the original 120 Mb/sec in order to reduced the filesize. The total size of the recorded footage was approximately 55 GB before processing and rendering the video. We first rendered the 360° video using Vuze VR Studio [84] at the original resolution and video bit rate. We then encoded the video following Oculus's recommendations [85] for encoding 360° video for the Oculus Quest [16]. The reduction in resolution and video bit rate does have a slightly noticeable effect on the quality of the video.
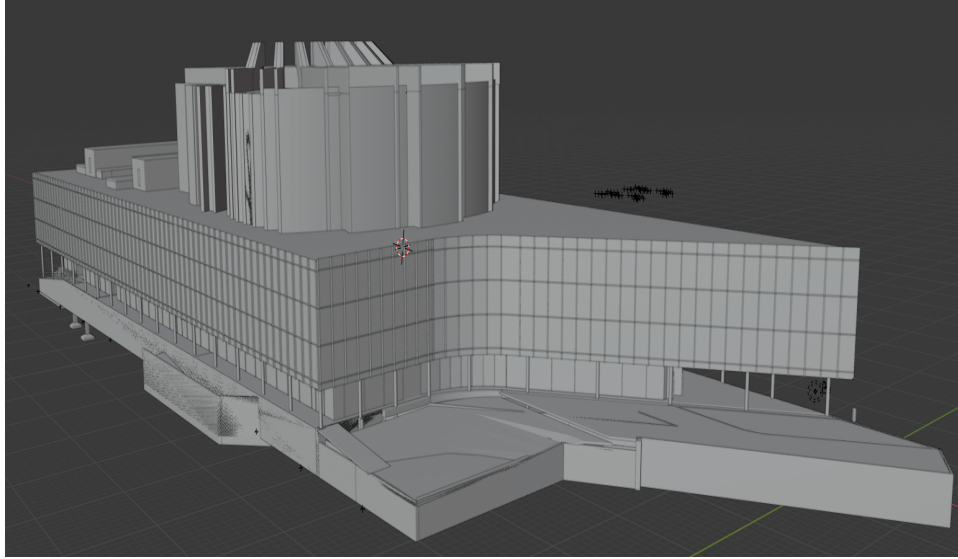
Figure 3.3: The 3D model of Grieghallen [19], as provided by COWI [20].

3.3 3D Model

The 3D model of Grieghallen [19] we are using was originally a Building Information Model (BIM), provided by COWI [20], that contained a lot of details and information about the entire building, as seen in figure 3.3. We only needed the Griegsalen, the main concert hall inside Grieghallen, in our project and not the entire Grieghallen building. The 3D model was used for two different reasons in our project, the first one being a visual representation of the Griegsalen in our Unity [52] scene, allowing us to accurately place the audio sources and the user accordingly, explained in chapter 3.6. The second use is Oculus Native Spatializer (ONSP) using the 3D model for calculating reverb and reflections in real-time, explained later in chapter 3.6.2. The model was then imported into Blender [86] for rework.

Blender [86] is a free and open source 3D creation suite. It has features like 3D modelling, animation, and rendering amongst others. Blender is a popular free alternative to the Autodesk [87] suite of programs, often used in the architecture, construction, media and entertainment industries. The animated movie *Next Gen*, made by Tangent Animation, was made entirely in Blender [88]. Which shows how powerful Blender has become as a 3D modelling, animation and rendering suite.
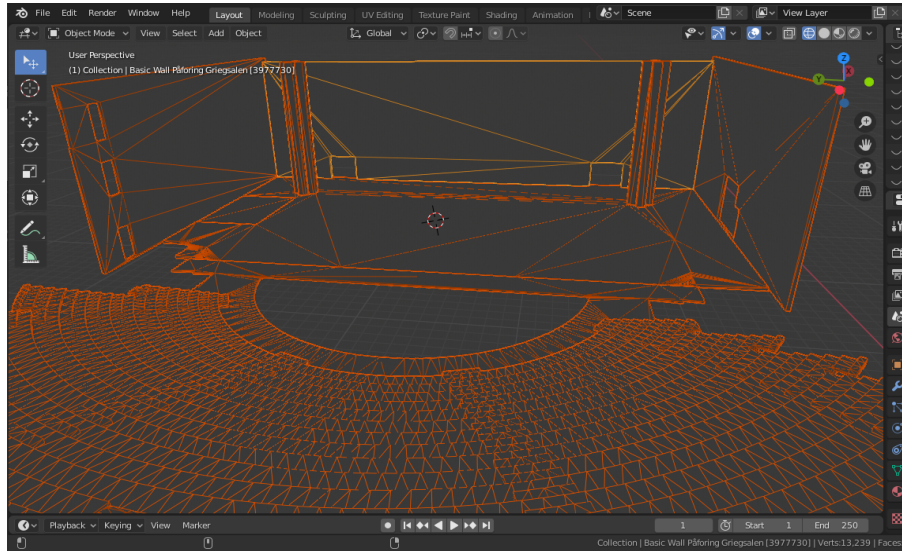
Figure 3.4: A slice of the 3D model before any mesh rework has been done, as shown in Blender. All the orange lines show an edge between vertices.

3.3.1 Reworking the 3D model

In order to reduce the performance impact of the 3D model, it was imported into Blender for rework, before being imported into Unity. Oculus has set some guidelines for performance on the Oculus Quest, these include a maximum of 100 000 vertices and at most 100 draw calls per frame [89]. Although we are not rendering the model in our experience, its mesh and geometry can be used by ONSP in order to calculate reverb, reflections, and sound propagation through the concert hall, which is explained later in chapter 3.6.2. We can also add functionality where the user is allowed to freely move inside the 3D model, and listen to how the audio changes depending on where you are in the concert hall. In that case, using the reworked model is required in order to meet the performance targets set by Oculus [89].

In Blender, any part not connected to the concert hall itself, Griegsalen, was removed. Although the same result could be achieved by using various culling methods in Unity [52], like occlusion culling [90], we decided to remove these areas of the model in order to reduce the file size of the 3D model. Furthermore, we used Blenders tools and features for mesh rework, in order to reduce the amount of triangles and detail in the end model. As seen in

25
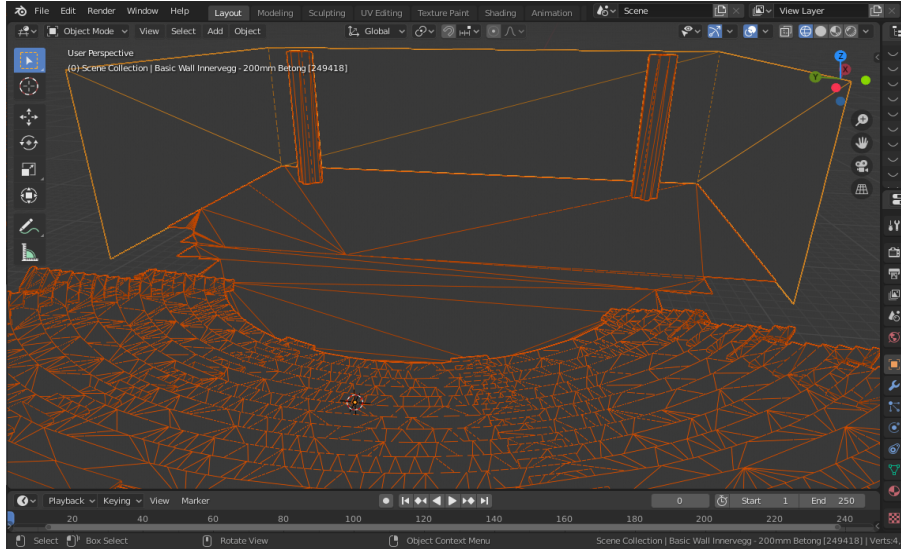
Figure 3.5: The same slice shown before, after mesh rework has been done, as shown in Blender. All the orange lines show an edge between vertices.

figure 3.4 the mesh was complex and not optimized. Using Blender's tools for mesh manipulation, such as *Decimate geometry*, and *Limited dissolve*, the mesh for each object in the model had its amount of triangles and vertices reduced, as seen in figure 3.5. In table 3.1, we observe the total amount of triangles and vertices, as well as the file size, of the 3D before and after rework had been done. Additionally, some objects were merged together in order to reduce the amount of draw calls. While all the chairs in the concert hall do exist in the 3D model, they are disabled in Unity, and not counted towards the total amount of vertices and triangles, as including them makes the total amount of vertices go beyond the 100 000 vertices limit.

Model	Vertices	Triangles	File size
Original	2 059 488	3 765 802	446 MB
Griegsalen only	57 706	103 037	64 MB
After rework	31 445	32 155	26.2 MB

Table 3.1: The amount of triangles and vertices in the 3D model before and after rework

3.4 Unity Scene

In Unity, we have created a scene using a 3D model of Griegsalen as the base. The 3D model was imported into Unity from Blender [86], and is placed in the scene using Unity's 3D coordinate system. The 3D model acts as our base, as we are replicating a real-world experience that took place in Griegsalen, every other Asset is placed inside the 3D model using Unity's 3D coordinate system. The 3D model is scaled according to the real concert hall, giving an accurate representation of the space. However, the 3D model is not a complete and 100 percent accurate representation compared to the real thing, but for our work it is good enough. The overall geometry is correct, but the model is lacking in finer details. In our VR experience, the model is not actually rendered or shown to the user. It was used during development in order to accurately place audio sources and the user's position. In addition to the 3D model, an orchestra layout was provided by Bergen Philharmonic Orchestra showing the seating placements of each musician in the orchestra. This layout was placed in the scene as a sprite and the position was matched up with the position of the 3D model, as seen in figure 3.7. The orchestra layout further aided in the accurate placement of audio sources and the user.
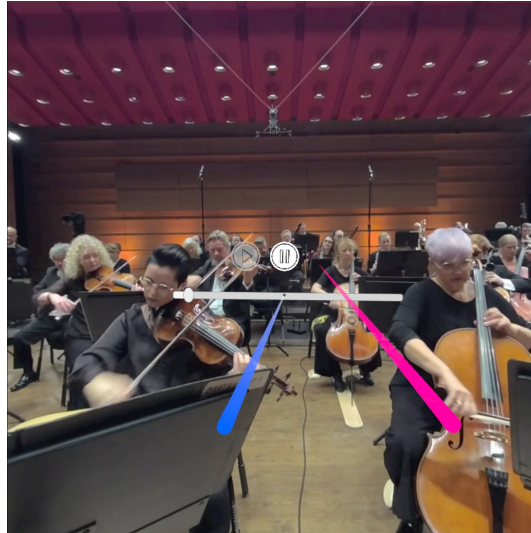
Figure 3.6: Screenshot from the VR experience running on the Oculus Quest.

3.5 Developing for VR in Unity

In order to make our VR experience work on a VR HMD we are using Unity's
XR Interaction Toolkit [77], a cross platform toolkit used for displaying the
Unity [52] application on a VR HMD and for handling the input of the tracked
controllers associated with the particular VR HMD. The XR Interaction
Toolkit allows us to target multiple VR systems at once with the same code,
and application built in Unity. An XR Rig, a Unity game object, handles
displaying our Unity application in VR on the VR HMD. The placement of
this game object in the Unity scene, represents the position of the user in our
application. In addition to the XR Rig, the toolkit requires another game
object, called XR Interaction Manager, to be added. This game object will
handle input from the tracked controllers associated with the VR system.
In our VR experience, the tracked controllers are not visible in VR, and
is instead represented by a pair of pink lasers, that shows the user where
they are currently pointing, as seen in figure 3.6. The laser pointer turns
blue when hovering over interactable objects. The user can interact with
the video player controls using the *trigger* button, found on the back-side of
the tracked controllers. It is the same button for the Oculus Quest version
and PC VR version. Adding 3D models of the tracked controllers associated
with the VR system can be added later.
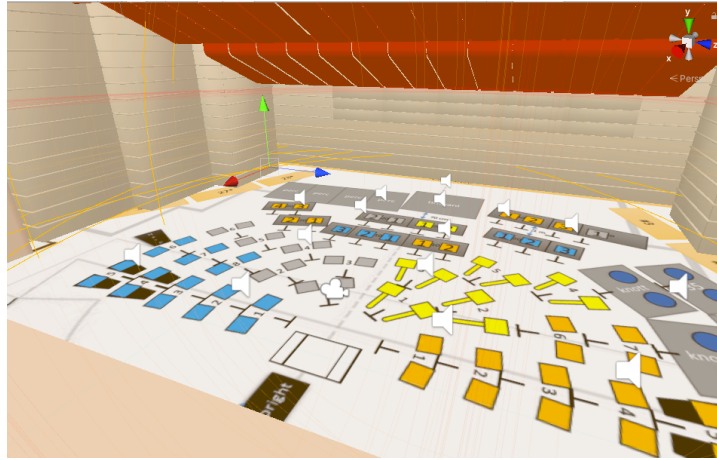
Figure 3.7: Scene in Unity showing the placement of the audio sources, indicated by white speaker symbols. The camera icon in the middle represents the position of the user.

3.6 Creating the 3D Audio

This section breaks down how we create the 3D audio for our experience.

3.6.1 Audio Sources

All the individual sound files are loaded into the Unity project as Assets. As seen in figure 3.7, in order to ensure correct placement of the audio sources, we are using the layout of the orchestra as a reference for where the audio files should be placed in reference to the model. Each audio source is represented by a white speaker symbol, and each audio source correspond to one of the microphones that was used for recording the concert. For example, the audio file for the concertmaster is placed in the scene where the concertmaster is seated according to the orchestra layout. The same process is then repeated for all the audio files and the instrument groups in the orchestra. In figure 3.8, we observe the settings applied to all audio sources. Spatialize is enabled, and spatial blend is set to 3D. This tells Unity to use the selected spatializer SDK in order to spatialize the audio, as opposed to simply playing the audio in stereo. Output is set to the audio mixer for the instrument group the audio track belongs to. The rest of the settings are set to the default settings in Unity. The audio mixer is explained further in chapter 3.6.3.
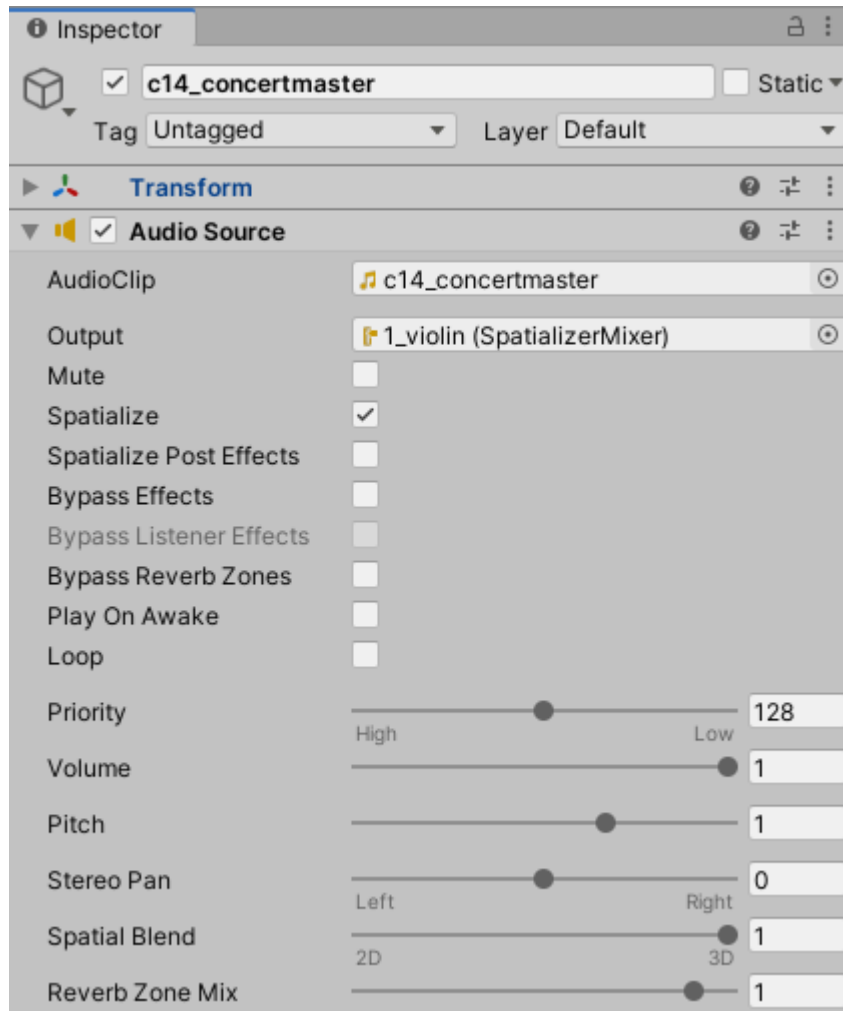
Figure 3.8: The settings applied to all audio sources in the scene. Here the concertmaster is used as an example.

3.6.2 Oculus Native Spatializer for Unity

In order to spatialize all the audio
sources in order to create 3D audio,
we are using Oculus Native Spatializer
(ONSP) [50]. In the audio settings for
the Unity project, ONSP is selected as
the spatializer and ambisonics decoder.
A script called *ONSP Audio Source* is
added to each audio source. This script
is part of ONSP and enables several of
ONSP features to be used on the audio
source, like reverberation and attenua-
tion, as seen in figure 3.9. In the settings
for the script, we can apply gain to the
audio source, or apply gain to the reverb
in order to make the audio sound more
reverbarant. We can also set how far the
audio should travel before being muted,
although in our case all audio sources
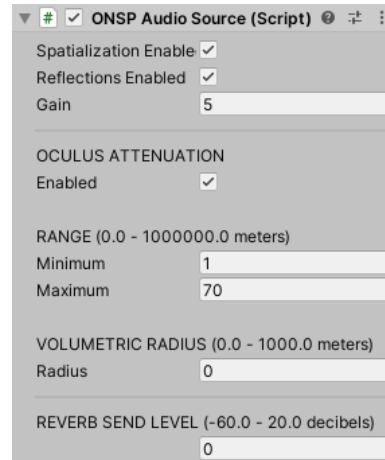
Figure 3.9: The *ONSP Audio Source
script*, and its settings.

should always be heard, their range is thus simply set to be larger than
the 3D model. ONSP handles the spatialization of each audio source in the
scene, and creates 3D audio in real-time depending on the user's position.
In our case, the user's position refers to the placement of the XR Rig in
our scene, which allows our application to be displayed in VR. The user's
position is not static, and can be moved during run-time either by user input
or programmatically. Despite this, for the MVP of our experience the user's
position is static, and only accounts for the user rotating its head. This is due
to the fact we only have one viewing position. Nevertheless, more viewing
positions can be added in future recordings of the orchestra, and more about
this potential feature will be discussed later in chapter 6. ONSP can either
use a box model, or the geometry of your scene, in order to calculate reverb,
reflections, attenuation and occlusion for the audio. For the Oculus Quest
version, we had to use the box model due to performance issues when using
the 3D model, explained later in chapter 4.6. The box model, is a rectangular
box set to the approximate size of Griegsalen, and surrounds the scene. On
the PC VR version, the 3D model is used instead.

3.6.3   Unity's Audio Mixer

In Unity [52], all audio is routed through an audio mixer before being output as audio to the user [21]. Each audio source in our scene is assigned to an audio mixer group, based on which instrument group the audio belongs to. For example, as seen in figure 3.8, the audio from the concertmaster, who leads the first violin section, is assigned to the audio mixer group for the first violin section. Each audio mixer group then routes the audio to the master audio mixer. By separating each instrument group into different audio mixer groups, we can individually control any settings or effects applied to the audio from each instrument group.

As seen in figure 3.10, we are applying a 10 dB increase to the volume level, this was done due to the default volume level sounding quieter than one would except a concert to sound like, requiring the user to put the volume level on the Oculus Quest [16] at the highest setting in order to hear the concert at a normal level. This behaviour is undesirable, however, as we want the user to have more control over the volume levels, such that they can select a volume level suitable to them. The Oculus Quest [16] also warns the user when the volume is raised above a certain threshold. Additionally, to the master mixer we are adding the normalize effect [91] which also applies a constant amount of gain to the audio output, and together with the gain on all individual mixer groups the overall audio level is raised to the desired volume level when the Oculus Quest's volume setting is at its default, or middle, setting. This gives the user more control over the audio volume.
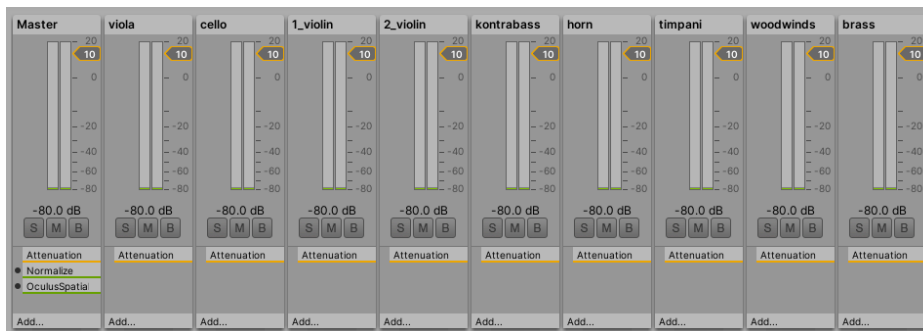
Figure 3.10: All the audio mixer groups shown in Unity [21]. The different mixer group represents one instrument group in the orchestra. Audio from each group is routed through the Master group before the audio is output to the user.

To the master audio mixer, we are adding the effect *Oculus Spatializer Reflection*, this is the last part of ONSP, that spatializes the audio output from all the different mixer groups before outputting the audio to the user, as seen in figure 3.11. Here we can enable or disable reverb and reflections, set the room dimensions when using the box model, and change the reflection coefficients for the walls in the box model. The reflection coefficients were left at their default settings. *Shared reverb attenuation range* refers to the max distance an audio source should travel, before it hits a wall, in order to be included in the reverb. *Shared reverb wet mix*, applies an additional gain to all sound sources, and the higher the value, the more reverberant the audio will sound. *Propagation quality level* can be used to change the quality level of the simulation done by ONSP, higher level means higher quality, but in turn it also means higher CPU usage. Lower quality uses less CPU. We have left it at its default level, but we could play around more with this settings in the future, due to the performance issues we ran into, explained later in chapter 4.6.
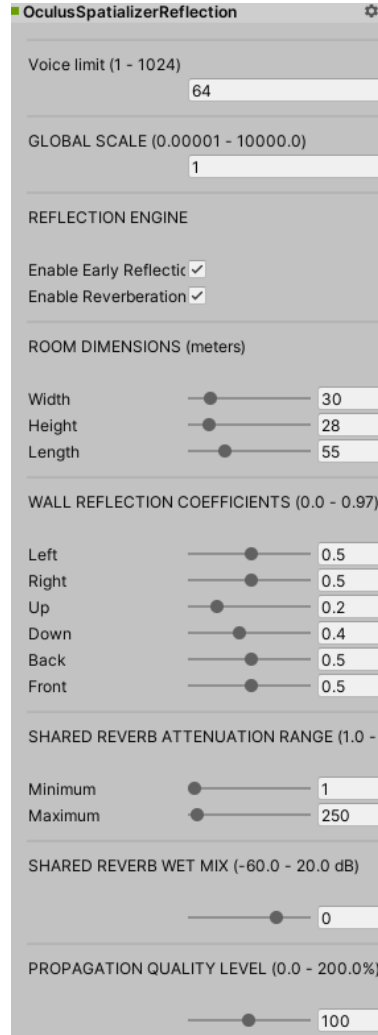
Figure 3.11: Settings used by the *Oculus Spatializer Reflection*, the effect added to the Master output that adds the spatialization to all audio sources.

3.6.4 Materials and ONSP

When ONSP uses the 3D model or scene geometry in order to calculate reverb and reflections, it also has the ability to take the materials of the objects into account [92]. Different materials, like concrete and wood, will have different properties on how they reflect and absorb sound [63]. We have added a preset material to all objects in the 3D model, as seen in figure 3.12. However, we do not know the exact properties of all the different materials that exist in the real concert hall. We are also missing certain parts and details in our 3D model. It is therefore not an entirely accurate representation of the real concert hall, and the calculated reverb and reflections are therefore also not accurate. The amount of people inside the concert hall also affects reverb [62], something our 3D model does not take into account, as there hare no human models inside of it. Calculating reverb and reflections in real-time is a costly computation [92]. The extra CPU usage this adds on the Oculus Quest version of our VR experience, can be seen later in chapter 4.6 and figure 4.8.
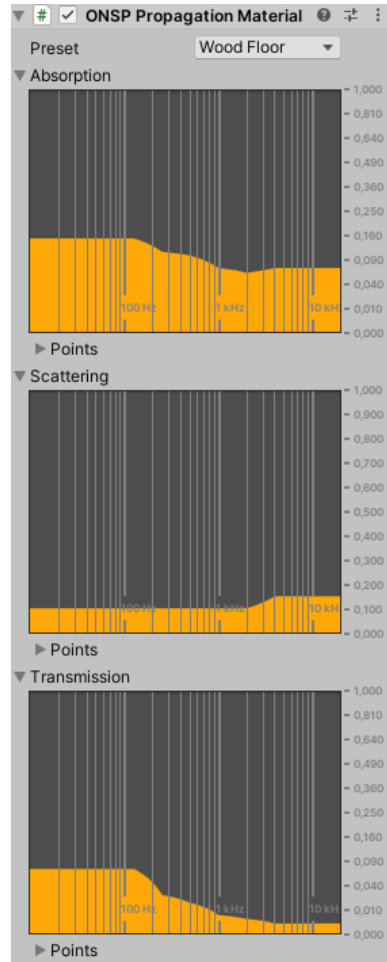
Figure 3.12: The *ONSP Propagation Material* is added to all objects in the 3D model. Here the wooden scene floor is used as an example. It shows how the audio is absorbed and reflected at different frequencies.

Figure 3.13: What the 360° video looks like when playing in a normal video player, as it is formatted using equirectangular projection.

3.7   Playing 360° Video in Unity

In order to play 360° videos in Unity we will be using their built in Panoramic Skybox shader [93]. The 360° video is formatted in an equirectangular projection, seen in figure 3.13. Equirectanglurar projection is an old map technique created by Marinus of Tyre in 100 AD [94, p. 5 - 8], for displaying a sphere on a flat surface. The video itself is rendered to a Render Texture using reverse equirectangular projection. The Panoramic Shader then renders this texture to Unity's built in Skybox as a sphere around the user. A Render Texture is a special texture in Unity that is created and updated at run time, allowing us to render videos or camera views to it [95]. As Unity's Skybox is rendered around the whole scene [96], behind any geometry in the scene, we had to hide the 3D model of Griegsalen. In order to hide our geometry, we have created a separate layer for the 3D model of Griegsalen. In Unity, layers are most commonly used by cameras to render only a part of the scene, or by the lighting system to only illuminate selected parts [97]. The player camera, or XR Rig, is then set to ignore the layer containing the model of Griegsalen.

As seen previously in figure 3.6, we have created an UI for *play* and *pause*. There is also a progress bar that shows progress through the VR experience, and the user can interact with this and move back and forth in the VR experience. The UI was created with Unity's UI package [98]. The buttons and progress bar are placed on a canvas, and are all from Unity's UI package. The canvas is set to be a *world-space* canvas, this means that the canvas exists as a 3D object in our scene. Typically in video games and other applications, the UI is instead drawn on top of the scene as an overlay. In VR applications, we do not want the UI to be rendered as an overlay on top of everything, as that would mean the UI would constantly be right in front of the user's view in the VR HMD. The UI would also move with the user, and it would block out the view of the VR application. Instead, we want the UI to be at a fixed distance from the user, and this is accomplished by using a *world-space* canvas, as previously mentioned. In the build used for testing, our video players does not feature any standard video player functions like *pause* and *seek*. As these were disabled for testing purposes. Nevertheless, we can access all the video player functions via scripts and we can create a UI for these functions and more.

3.8 Synchronizing the Audio and Video

Typically, the audio is a part of the video file, and is synchronized with the video before being rendering into the final version of the video. In our case, however, the audio and video are separated and played back using two separate system. The Unity built in video player for the 360° video, and Unity's audio system for playing the audio. This causes some challenges with synchronizing the audio to the video. As the video player in Unity may have to prepare, by reserving resources or pre-load a few frames [99] before starting playback. All the audio clips are attached to a their own game object in our scene, and is loaded into the memory when the scene loads, and will start to play immediately if we call on the appropriate function within a script. As the audio and video are loaded in at separate times, and when starting the playback they will not start at exactly the same time.

To combat this, the start of the audio and video has been synchronized as best as we can outside of Unity before importing the video and audio files into Unity. Then, we have created a script in C#, the programming language used by Unity, that handles playing the video and audio as synchronized as possible when the user hits *play* in our VR experience, as seen in listing 1. The code for *seek* and *pause / unpause* work in similar matter. When the user hits play, the scripts starts a coroutine [100] that begins by preparing the video player. This allows us to wait until the video is fully prepared, before starting playback of the audio and video. After the video has been prepared, we then start the audio with a delay of x seconds. The audio delay is different on the PC and Oculus Quest version due to different hardware. The delay has to be figured out by a human looking at the video and listening to the audio, to see if the audio and video are synchronized. After Bjørn Sagstad [101] had tested our VR experience, talked about later in chapter 4, he commented that the audio was slightly behind the video. He thus helped us with making the audio more in sync with the video

```
1   public void play()
2   {
3       StartCoroutine(PlayCoroutine());
4   }
5   IEnumerator PlayCoroutine()
6   {
7       mVideoPlayer =
        ↪  GameObject.FindGameObjectWithTag("360videoplayer").GetComponent<VideoPlayer>();
8       mAudioParent = GameObject.FindGameObjectWithTag("audio");
9       mVideoPlayer.playOnAwake = true;
10      mVideoPlayer.Prepare();
11      //Prepare/Wait for 1 sceonds only, have to wait at least 1
        ↪  second due to a bug in Unity
12      WaitForSeconds waitTime = new WaitForSeconds(1.0f);
13      while (!mVideoPlayer.isPrepared)
14      {
15          yield return waitTime;
16          break;
17      }
18      foreach (Transform child in mAudioParent.transform)
19      {
20          if (child.GetComponent<AudioSource>().time == 0)
21
               ↪  child.GetComponent<AudioSource>().PlayDelayed(AudioStartDelay);
22          else
23              child.GetComponent<AudioSource>().UnPause();
24      }
25      mVideoPlayer.Play();
26  }
```

Listing 1: Part of the code in the script for handling the playback of the video and audio.

3.9 Issues with Android and Large Files

On the Oculus Quest, the max size allowed for the Android Application Package (apk) is 1 GB [102]. In order to allow for larger programs than 1 GB, you can have multiple expansion files [103]. Each expansion file can be up to 4 GB in size. These expansion files are stored separate from the the main apk. In Unity, expansion files are created by creating AssetBundles [104]. However, loading a video from an AssetBundle on Android does not work. Thus we cannot include the entire hour long symphony, as the audio and video files end up being a lot larger than 1 GB. This also prevents us from adding new symphonies in the future. In order to be able to include the entire symphony and have the possibility to add new concerts, the video will have to be streamed from an external source. This has a major drawback in that one will need an internet connection in order to use the VR application. How we can solve this with video streaming, and how it will change the implementation of the VR application will be discussed later in chapter 6.

Chapter 4

Analysis and Assessment

For the research par of the project we used a mix of quantitative and qualitative research methods. The VR experience we created was validated using questionnaires and feedback from test users. Defining what a good experience would be, is individual and subjective.

Resulting data from this thesis should provide knowledge on the feasibility and quality of using the audio recorded with multiple microphones in a 3D environment and a game engine to create a VR concert experience, as opposed to using speciality binaural and ambisonics microphones, like Philharmonia Orchestra [28] used, designed to record audio in a format that can be used for 360° videos and VR.

Additionally, it will provide insight into what the test users want from an experience like this. As both Jostein Stalheim [105] and Bjørn Sagstad [101] are educators, they were interested in knowing how experiences like these could potentially be used in education.

4.1 Questionnaire

1. Have you been to a classical concert in Grieghallen before? [Yes, No]

2. What did you think about the audio quality? [1 = Bad, 7 = Excellent]

3. How well did you perceive the 3D audio? [1 = Did not hear any 3D audio, could not tell the direction of where the music from the instruments came from, 7 = 3D audio felt natural, could tell where the music from the instruments came from]

4. What did you think about the video quality? [1 = Bad, 7 = Excellent]

5. How immersed did you feel during the experience? [1 = No immersion, 7 = Very immersed]

6. Did you like the viewing position? Would you have preferred a different position?

7. Do you think experiences like this could be used in education? If yes, in what ways?

8. Do you have any other comments to the experience?

At first, we created a draft of the questionnaire. Then we had a meeting with Jostein Stalheim [105], where we discussed the questionnaire and he provided input and feedback on the questions in the questionnaire. For the questions answered on a linear scale, we are a seven point scale, similar to a Likert-type scale [106] were each question is answered. One end of the scale is negative, and the other end is positive, with a neutral or average point in the middle. Then we have three open ended questions in which the testers can freely answer with their thoughts on the question and the experience.

Figure 4.1: Two people testing our VR experience.

4.2 Testing

In order to gather feedback for our experience, we invited two different groups to test a version of the VR orchestral concert experience. In the version of our experience used for testing, ONSP did not use the 3D model for reverb and reflection calculation in real-time, due to the performance issues explained in chapter 4.6. Video player functionality like *pause* and *seek* were also disabled. In the VR experience, the test users listened to the first six and a half minutes of the fourth movement in Tchaikovsky's fifth symphony, while viewing a 360° video. The audio and video were recorded during Bergen Philharmonic Orchestra's live concert of Tchaikovsky's fifth symphony, November 14, 2019 [78]. The fourth movement was chosen for testing purposes, as during the movement all instrument groups are playing and it contains slower sections as well as more faster sections, giving a lot of variation. The different instrument groups are also not always playing at the same time, which should give the user a more clearer perception of the 3D audio. We cut the movement short after six and half minutes due to time constraints, as it was preferable for us to be able to have several people test the experience over a short period of time. We also considered using the first movement for testing purposes considering it has similar characteristics to the fourth movement, which would have provided us with similar results.

Before starting, the test users were encouraged to look around while viewing the experience, such that they could experience the full effect of the 3D audio, and hear how the audio changes when they look and turn around. In order to fully experience the 3D audio, the test users are wearing a pair of external headphones, as seen in figure 4.1. The built in speakers on the Oculus Quest [16] do not provide the same level of 3D audio as when using a pair of headphones. Additionally, the built in speakers leak audio out into the room, disturbing any other person in the room and also anyone else testing the experience. We tested multiple people at the same time for practical purposes and due to limited time available from each test group.

4.2.1   The Test Groups

We planned for several groups of people to test our experience. All test groups had some sort of relation to music, classical music and education. One test group, henceforth called Group JS, consisted of Jostein Stalheim [105], a composer, accordionist, and professor at Western Norway University of Applied Sciences [83], and eight of his students in the course *FM1-AVM2-10 Music in use 2* [107], which is a part of *Teacher Education - Bachelor of Music* [108]. The students in this group are studying to become music teachers, and have knowledge of playing instruments, composing and arranging music, music history and teaching music amongst others. A colleague of Jostein Stalheim also tested our VR experience as part of the group. The second group, hencefort called Group BS, was Bjørn Sagstad [101], a conductor and associate professor at University of Bergen [109], and four of his students in *Ensemble Conducting* [110]. These students are studying ensemble conducting, and have experience with music instruments, music theory and conducting musicians. We intended to test the experience with a third group, composed of musicians from Bergen Philharmonic Orchestra. However, due to the COVID-19 [111] pandemic outbreak it was not possible to get the tests done with the musicians from the orchestra at this time. An official statement from *Musikkselskapet Harmonien*, the company behind Bergen Philharmonic Orchestra, can be read in Appendix B. This test would have provided valuable data, as the musicians are trained listeners and know exactly what the music should sound like when you are in Griegsalen. We intended to ask the musicians some additional question, including asking about which instrument they play in the orchestra. In addition to the test with the musicians, we also planned for a second more in-depth test with Stalheim and Sagstad, where the entire VR experience would be discussed and its potential usage in education.

4.3 A Note on Headphones

A pair of headphones can sound drastically different to another pair of
headphones, and there is a huge range of headphones available from cheap
to expensive, professional grade equipment. The choice of headphones will
therefore have an effect on the audio quality. Due to limited resources, we
had to use the headphones the author already had on hand. These two
headphones were the Beyerdynamic DT-770 Pro 80 ohms [22] and the Sony
WH-1000XM2 [23], as seen in figure 4.2. The Beyerdynamic DT-770 Pro is
an industry standard headphone in the music industry, and is widely used
in music production and monitoring. While the Sony WH-1000XM2 is a
consumer noise cancelling headphones. The audio quality of both headphones
are generally considered to be *good* [112, 113], but the audio quality of a pair
of headphones is a highly subjective and debated topic. The headphones
used for testing, may have an impact on the results from the user testing.

Figure 4.2: The headphones used. The Beyerdynamic DT-770 Pro [22] to the left,
and the Sony WH-1000XM2 [23] to the right.

4.4 Answers to the Questionnaire

In this chapter we will be summarizing the results to the questions which were answered using a seven point scale, a scale similar to other Likert-type scales [106]. We will also explain the meaning of the results, and factors that may have impacted the results.

Figure 4.3: Question 1: Have you been to a classical concert in Grieghallen before?
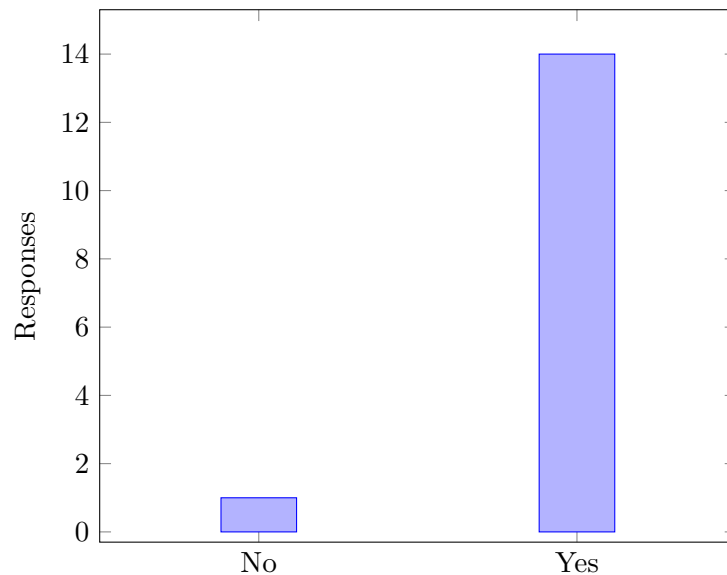
4.4.1 Question 1: Have you been to a classical concert in Grieghallen before?

We asked this question in order to gauge whether having been to a classical concert in Grieghallen before would in any way influence the answers given. However, as we see only one the person had not been to a classical concert in Grieghallen before. Interestingly enough, the only person that had not been to a classical concert in Grieghallen before, answered five, or just above average, on all questions and gave overall positive feedback to the entire VR experience.

Figure 4.4: Question 2: What did you think of the audio quality?
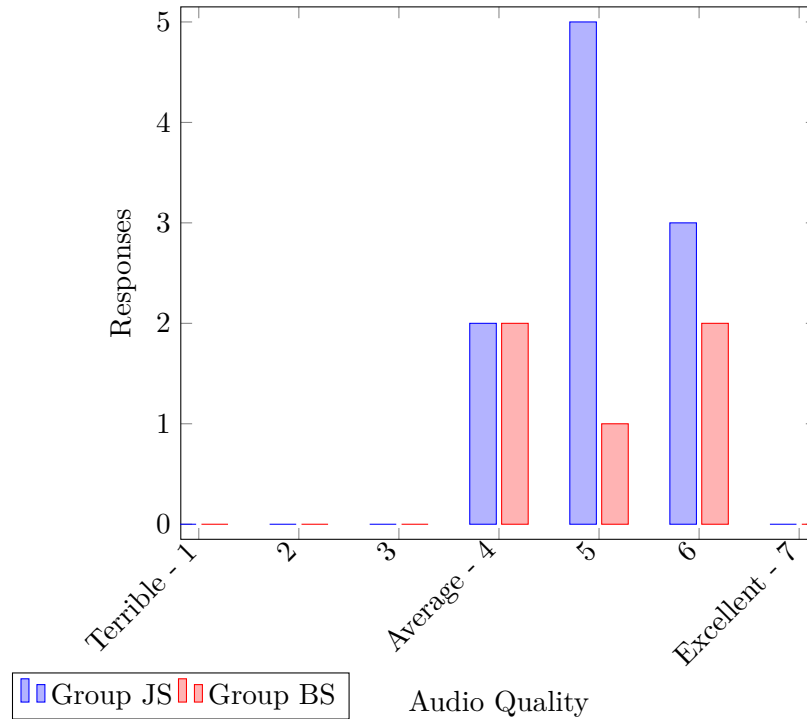
4.4.2 Question 2: What did you think of the audio quality?

This question was asked in order to gauge what the test users thought about the quality of the audio. Here we mean the audio quality itself, I.e compared to the audio from a professionally produced music track, and not the quality of the 3D audio itself. In order to determine whether we should spend more time working on the audio itself, or leave it be and focus on other areas first.

The responses the the audio quality shows us that it is *good* or *great*, but there is room for improvements in the audio quality. A couple of test users noted that the audio sounded *flat*. How we can improve the audio quality will be discussed later in chapter 6. As mentioned previously in chapter 4.3, the headphones we used may have impacted the results of the audio quality. Other factors that are outside of our hands include the sound chip used in the Oculus Quest, and the way Unity's audio system handles audio, and compression of the audio files done by Unity. However, we do not consider these to be major contributors to the loss in audio quality. Nevertheless, we received an overall positive response to the audio quality.

Figure 4.5: Question 3: How well did you perceive the 3D audio?

Percieved 3D audio
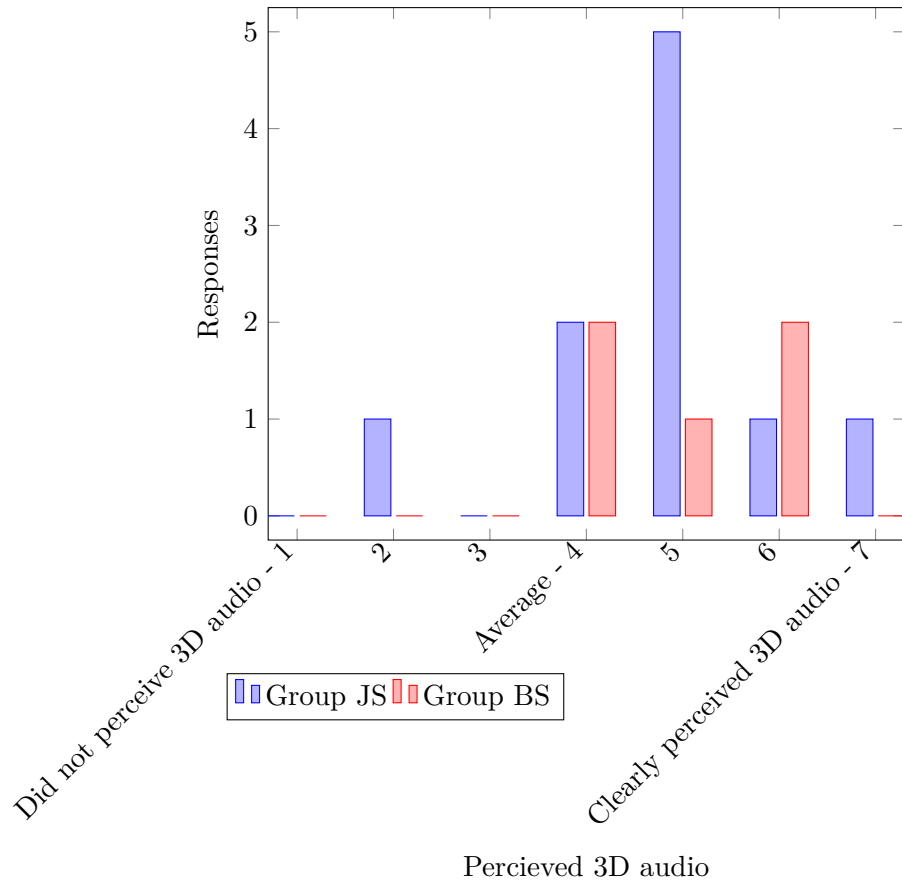
4.4.3   Question 3: How well did you perceive the 3D audio?

We asked this question in order to gauge how well the test users actually perceived the 3D audio we have created using our method. And also to understand if the test users were able to tell the direction of the music, from the different instrument groups in the orchestra, came from.

The answers to how well the testers perceived the 3D audio from our experience is more varied. Some test users were clearly able to tell where the audio from each instrument group came from, whereas other test users struggled distinguishing the direction for each instrument group. Overall, the results show that test users are for the most part able to perceive 3D audio from our experience. There are different factors explaining the varied

results. One factor, is that due to the nature of the generic HRTF, it will not be a good fit for everyone. For some test users their own measured HRTF will differ too much from the generic one used by ONSP, causing the 3D audio effect to not be perceived properly [55]. To improve on this, adding a personalized HRTF for each user will greatly improve the perception of 3D audio and overall audio experience for the user. Despite this, as mentioned in chapter 2.1.1, this is not currently feasible. ONSP does not currently support this feature either, but it may be added in the future [57]. Other factors include the test users' expectations of what exactly 3D audio is, and what it was supposed to sound like. Although the test users were given a brief explanation of what 3D audio is, and what they should listen for during the experience, they were not given an in depth explanation, nor did most of the test users have any prior experience with VR or 3D audio.

Figure 4.6: Question 4: What did you think about the video quality?
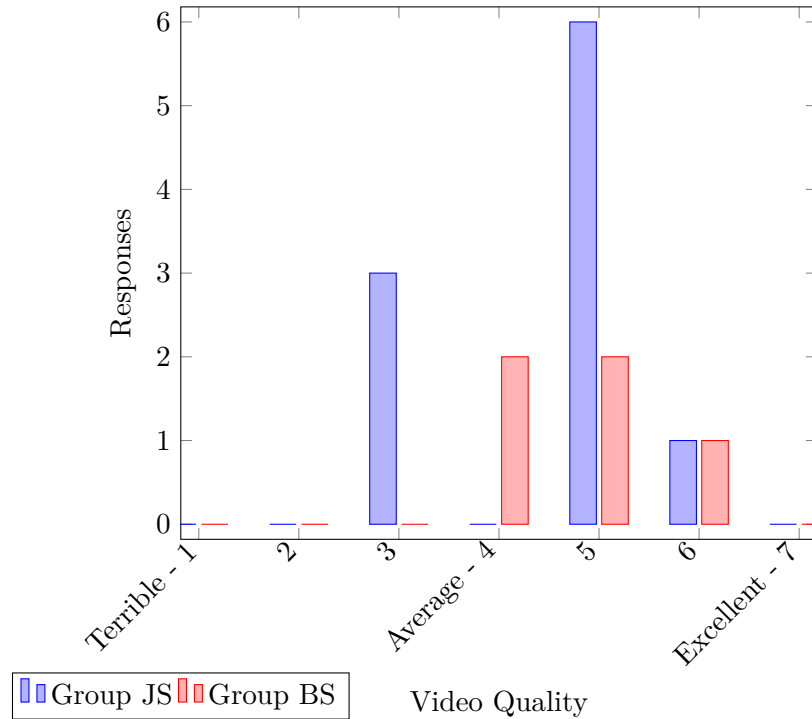
4.4.4 Question 4: What did you think about the video quality?

A simple question in order to gauge what the test users think of the video quality. The video quality is for the most part dependent on the camera itself, but the resolution change and encoding done by us does have an effect on the video quality.

From the answer we can tell that the test users think the overall video quality is *good*, with some test users thinking the video quality is merely *average*. This result is somewhat interesting to the author, as the video resolution and quality had to be reduced from the original size in order to fit inside the Android apk limits, as described in chapter 3.9, resulting in a poorer video quality. The authors personal opinion of the video quality is that the video quality is not great. The original resolution and quality of the 360° video is also worse compared to higher end cameras. At the very least, it shows us that it may not be necessary to use high end cameras for these

kinds of experiences. Despite this, using a 3D 360° camera instead of the 2D 360° we used will provide a better experience, due to the added depth 3D video provides. It is also important to note, that almost all test users had no or little prior experience with VR, they therefore do not know about the quality potential that exists with higher end cameras, higher resolution 360° video and 3D 360° video.

Figure 4.7: Question 5: How immersed did you feel during the experience?
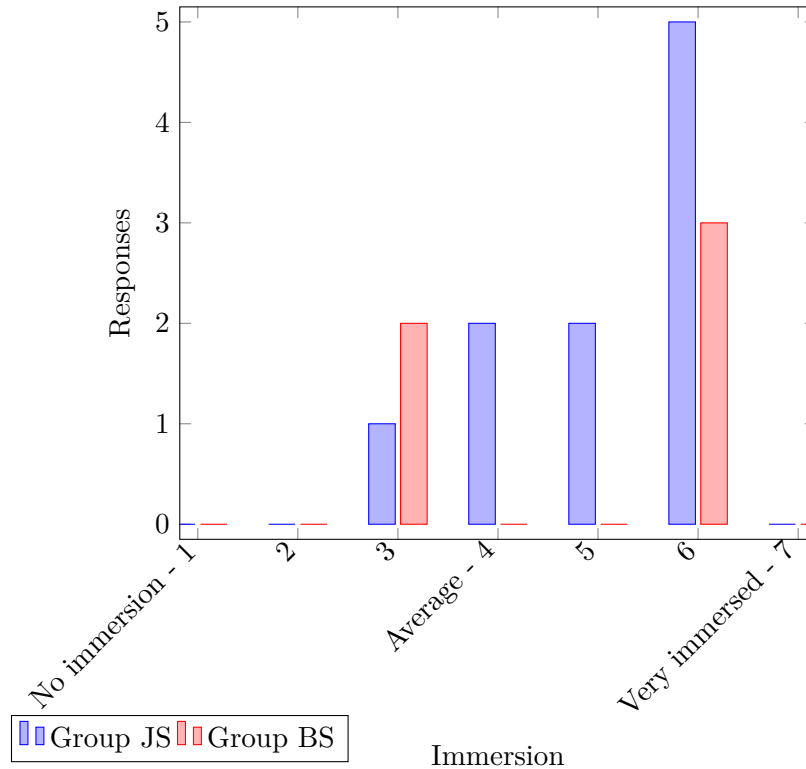
4.4.5 Question 5: How immersed did you feel during the experience?

This question was asked to determine the level of immersion the test users feel during our VR experience. An important aspect of VR is immersion and how much it feels like you are "really there" [31, 44, 45].

A majority of the testers found the experience to be quite immersive, with some test users feeling less immersed. From the results and conclusions of Bergström et al. [31] we know that audio spatialization, I.e 3D audio, plays a role in how immersed someone feels in a VR experience. However, it is not the only factor that plays into how immersed a person feels during a VR experience. Rogers et al. [45] concludes that the visuals are more important than the audio when it comes to immersion. While further suggesting that the impact audio has on immersion is more subjective, as their results were quite varied. Overall, from the results we see that we have achieved a good

level of immersion. This, in turn, shows that we have achieved a good level of audio and video quality, as these are related to immersion in video games and VR experiences. Lastly, after Sagstad had tested our VR experience and taken off the VR HMD, he said, translated from Norwegian: "[...] that was actually really close to what it's like to stand on the conductor's podium," further showing the level of immersion and perceived 3D audio we have achieved.

4.5   Written Feedback

For some questions in the questionnaire, the test users could write freely in either English or Norwegian. Quotes from these answers will be abridged and translated. Un-abridged and un-altered answers to the questionnaire are available in Appendix A

Overall, the responses to the experience were positive. Several people commented "Very cool!" after taking off the Oculus Quest after viewing and listening to the experience.

4.5.1   Question 6: Did you like the viewing position? Would you have preferred a different position?

We asked this question to determine whether the viewing position, right in front of the conductor, provided a good experience or if the testers would prefer another position.

Most people agreed that this was a good position to experience the orchestral concert from. At the same time, most testers said they would like other positions in the orchestra to be available for choosing, such that you can change position during the experience. This was originally an intended feature for the experience. However, it requires multiple 360° or 180° cameras to be setup in different positions in the orchestra on the day of recording. We did initially have two such cameras, but on the day of recording one of them malfunctioned. How this feature could work in practice, will be discussed in chapter 6.

Some testers noted that the perception of distance to the floor and the closest musicians felt "off". This distortion is caused by the fish-eye lenses prominent in 360° cameras, and objects close to the camera will appear distorted.

Examples

- "Cool to have the opportunity to see both the conductor and the orchestra"

- "A bit scary, felt like a long way down to the floor"

- "I liked it, would be cool to also be able to change position"

- "Nice place to sit, would have been interesting if one could move around"

- "Better to sit behind the conductor, in audience"

4.5.2 Question 7: Do you think experiences like this could be used in education? If yes, in what ways?

Stalheim and Sagstad are both educators in addition to their careers in music, and we asked this question as we wanted to explore if this type of experience could have any benefit or uses in education.

Every single tester agreed that this type of experience could have uses in education. Examples include the ability to study a conductor or musician up close, to see how they play their instrument and how they conduct an orchestra. Otherwise a common suggestion was that these types of experiences could used as an overall introduction to classical music.

Examples

- "As an introduction to classical concerts"

- "Musical directing, learning about instrument groups, learning about specific pieces of music"

- "This could be used in observing the conductors technique, and the communication between conductor and orchestra players"

- "The placements of instrumental groups in a orchestral context"

- "Great for learning and experience the repertoire"

4.5.3 Question 8: Do you have any other comments to the experience?

As a final question in the questionnaire, we asked if the testers had any general comments to or about the experience. Here the test users could answer what they thought about experience overall, give suggestions to improvements, give us insight if there was anything we had not thought about, and any other thing not covered by previous questions and so forth. Some testers noted that we should focus on the audio, both the quality and 3D audio perception, before focusing on other parts of the VR experience.

Examples

- "It felt a bit unnatural to sit so close to musicians, but this could be a preference for uneducated listeners. Some distant instruments were too close in sound. I could wish for a weighted listening for example when turning my focus."

- "Fix the audio before you fix the video. It will have a bigger impact on immersion. The audio might have been due to a sub-par headset."

- "It was exciting"

- "Would be nice, if possible, to have the score next to us, for better overview of who is playing next to the listening experience."

- "Great project."

- "It was very fun, recommended"

4.6   Performance on the Oculus Quest

In order to be approved for releasing an application on the Oculus Quest [16], you need to meet the requirements set by Oculus [114]. One of these requirements is stable performance, and for the VR application to run at a constant and stable 72 FPS, as 72Hz is the refresh rate of the displays in the Oculus Quest [16]. When we setup ONSP to use the 3D model for calculating reverb and reflections in real-time, we run into performance issues on the Oculus Quest version of our VR experience. We have used Oculus' own tool for performance metrics on Android devices, the OVR Metrics Tool [115], in order to gather performance data from the Oculus Quest version of our VR experience. As seen in 4.8, the CPU utilization is high and sporadic, and we see that the FPS drops below 72. The audio also cuts out and does not play correctly, as it only functions and play properly sporadically. The audio not working correctly when using the 3D model for calculation reverb and reflections, is not surprising as the Oculus documentation for ONSP says to not use reflections for Android applications. [50].
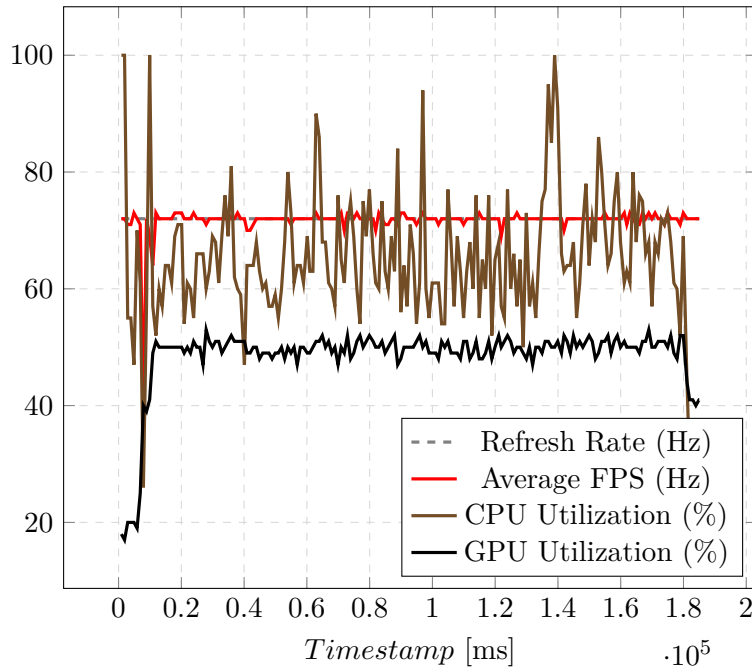
Figure 4.8: Reverb and reflections calculated with 3D model and materials.
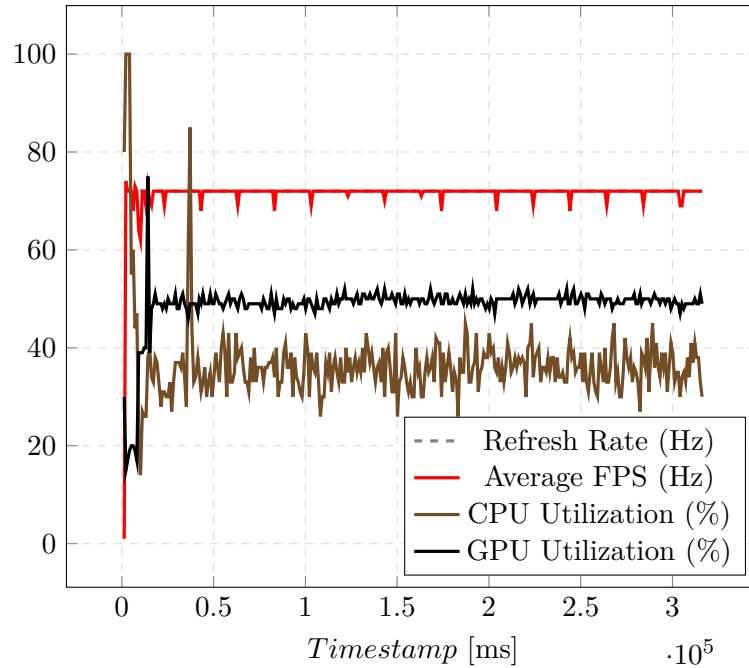
Figure 4.9: Performance when the box model is used.

When not using the 3D model to calculate reverb and reflections, and instead use the box model, CPU utilization is much lower and more consistent, as seen in figure 4.9. The GPU utilization remains consistent in both cases. Despite this, we are still seeing drops below 72 FPS. The reasons for this is not currently know, more profiling and performance testing is required. It may still be possible to achieve consisted performance when enabling ONSP to use 3D model for reverb and reflections calculation in real-time. When testing the VR experience with this feature enabled, the audio does work properly sometimes without any issues. Nevertheless, more development time is needed to look into this issue. The aforementioned performance issues are only related to the Oculus Quest version of our VR experience, we have not run into these issues when running the experience on PC VR.

Chapter 5

Conclusion

In our work, we decided to create a VR orchestral concert experience, and to figure out how to create it with the resources available too us. Furthermore, identifying how to create the 3D audio necessary for a VR experience, as well as the visual part of the VR experience. This chapter will conclude our thoughts on our method for creating a VR experience like ours, and whether or not we reached our initial goal and answers to our research questions.

5.1 Our VR Experience

With the available resources from Bergen Philharmonic Orchestra, such as access to professional grade recording equipment and a experienced sound engineer, we set out to figure out how we can create the 3D audio using the aforementioned resources. In order to use the audio from such a recording of the orchestra, as described in chapter 3.1, to create 3D audio we needed to use a spatial audio SDK. Our first research question was then defined as:

- What is the feasibility of using the audio recording from multiple micro-phones, along with a game engine and spatial audio SDK in order to create believable 3D audio of a orchestral concert?

From the works of Bergström et al. [31], we know that it is feasible on a smaller scale with a string quartet. Spatial audio SDKs are also commonplace in VR games and experiences, but they are typically used to spatialize several individual sounds that have little or no relation to one another. The question was then, does this work on a larger scale? Do we run into issues with synchronisation of the different audio sources?

Our second research question, defined as the following:

- How believable and well perceived by users of the experience is the resulting 3D audio and VR orchestral concert?

This research question was used as a complementary research question to the first one, in order to gauge whether or not our method actually provided a good 3D audio experience, and if the overall VR experience is any good and well perceived by the test users.

Looking at the results from our testing, as shown in chapter 4.4, we see that on questions about audio quality, video quality, perceived 3D audio and immersion; our experience scores above average on all questions. The answer then is yes - it is feasible to use this method in order to create an VR orchestral concert experience, and the resulting experience is considered to be good by our testers. There is definitely room for improvements on all aspects of the VR experience, and what these improvements are and how they can be developed will be discussed in chapter 6. Despite this, due to the COVID-19 pandemic, we are missing test data. Data from the planned tests with musicians from Bergen Philharmonic Orchestra and the longer in-depth tests with Stalheim and Sagstad would have provided valuable data for our conclusion.

Another point to touch on, is that our method is quite flexible and can be easily adopted to fit other microphone setups, concert halls and camera positions. We are not limited by any of our choices in technology or methods. Compared to other ways of creating 3D audio, as described in chapter 2, we can not argue that our method of creating 3D is *better* than others. Nevertheless, what we can argue is that our method is at least a feasible alternative, that provides some advantages in the overall production of a VR experience like ours when you have multiple camera positions or viewing spots.

We have, however, run into some performance issues with the Oculus Quest version of our VR experience, as discussed in chapter 6. Nevertheless, the performance issues is something we regard as fixable with further development.

Chapter 6

Further Work

This chapter will outline and discuss some potential additional features and improvements, that can be developed for our VR experience.

6.1   Including Full Length Symphonies

As mentioned earlier in chapter 3.9, due to file size limitations on the Oculus Quest and bugs within Unity, our prototype only has room for one movement in the symphony on the Oculus Quest version of our experience. Nevertheless, this is not an issue on the PC version of our experience, and adding the entire full length symphony is trivial. As is expanding the experience to include more symphonies performed by an orchestra, given the audio and video are recorded in a similar manner as was done for this project. In order to include the entire symphony, and have the ability to expand with other symphonies, on the Oculus Quest version, the 360° video will have to be streamed to the device from an external server. This is trivial to achieve, as the video player component in Unity has the ability to retrieve a video from an external source [99]. Unity also has integration with online video platforms, like Vimeo [116], which could be utilized to achieve this.

However, streaming the video somewhat defeats the purpose of the Oculus Quest being a standalone device free from any wires to be used anywhere, as one would need to be connected to the internet at all times when using our VR experience. It might also be trickier to achieve synchronisation of the audio and video, as these are handled separately in our experience and the audio is not bundled with the video like one normally would. Especially in

moments where the video might start buffering, or otherwise drop frames due to unstable internet.

6.2   Including Multiple Viewing Spots

An initial goal, and a request from Bergen Philharmonic, was to include multiple viewing spots in our VR experience. However, due the malfunctioning camera mentioned in chapter 3, we only ended up with one viewing position. If future orchestral concerts are recorded with multiple 360° cameras, we can include this functionality. This can be achieved by creating a interactive menu or otherwise creating an interactable object where that particular viewing spot is located. When the user selects a different viewing spot, we would move the player object to that particular viewing positions actual position in the model of Griegsalen in our Unity scene. The video would also be swapped out with the video from that position. The 3D audio would automatically begin to be calculated for the position the user was moved to. This can be done during the experience itself, allowing the user to change positions whenever they want during the VR experience.

6.3   Improving the Audio Quality

While the test users rated the audio quality to be above average, several of the test users noted that the audio sounded flat and that the quality was underwhelming. The test users either expressed this through the open ended question 8, or as a verbal comment after they had tested our VR experience. This is mainly due to us using the raw audio from the microphones, and that we have not done any audio mixing or producing on the audio files we received from NRK [26]. When ONSP uses the 3D model to calculate reverb and reflections, it is the authors opinion that the audio quality is improved. Depsite this, we did not test this with the test users due to the aforementioned performance issues described in chapter 4.6. Improving the audio quality through audio mixing or producing, will have to be done before importing the audio into Unity, by an experienced audio engineer or audio producer.

6.4   Live streaming

At the very start of the project, Bergen Philharmonic Orchestra was also interested in the possibility of live streaming their orchestral concerts in VR. Due to time constraints and the chosen scope of this thesis, live streaming a orchestral concert in VR was not included in the scope or goal of this thesis. Due to our chosen method for the 3D audio, we do not consider live streaming to be feasible with our method. As we create the 3D audio in real-time during our VR experience, and we require separate audio files from each microphone that is used for recording the audio. This would mean that each audio track would have to be streamed to the device, and this would require a lot of bandwidth. Nevertheless, live streaming is not impossible to achieve. As mentioned earlier in chapter 2.1.2, ONSP can be used with DAWs, and the 3D audio can be created there and be output as an ambisonics audio file format. The resulting audio file can be then streamed to the device and decoded by our VR experience. This would use less bandwidth. Otherwise, directly using ambisonics microphones for recording the 3D audio can be used to stream the audio, like how MelodyVR [32] and similar apps do.

Appendix A

Questionnaire Summary

Have you been to a classical concert in Grieghallen before?
15 responses

- ● Yes
- ● No

93.3%
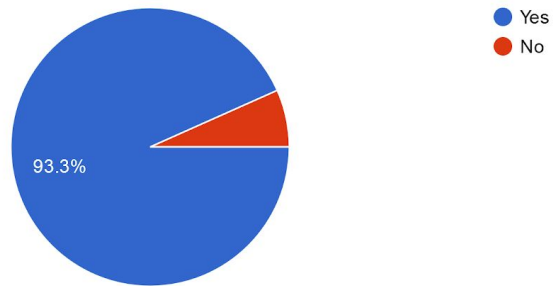
What did you think about the audio quality?
15 responses

0 (0%) — 1
0 (0%) — 2
0 (0%) — 3
4 (26.7%) — 4
6 (40%) — 5
5 (33.3%) — 6
0 (0%) — 7
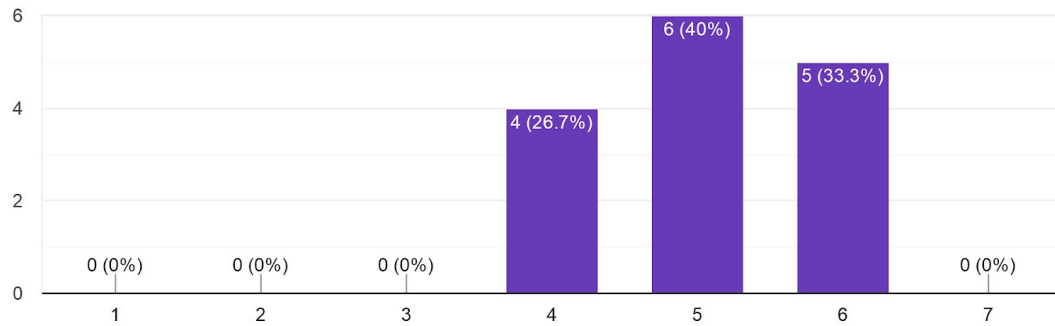
How well did you percieve the 3D audio? I.E Were you able to tell the direction the music from each instrument group came from.
15 responses

0 (0%) — 1
1 (6.7%) — 2
0 (0%) — 3
4 (26.7%) — 4
6 (40%) — 5
3 (20%) — 6
1 (6.7%) — 7
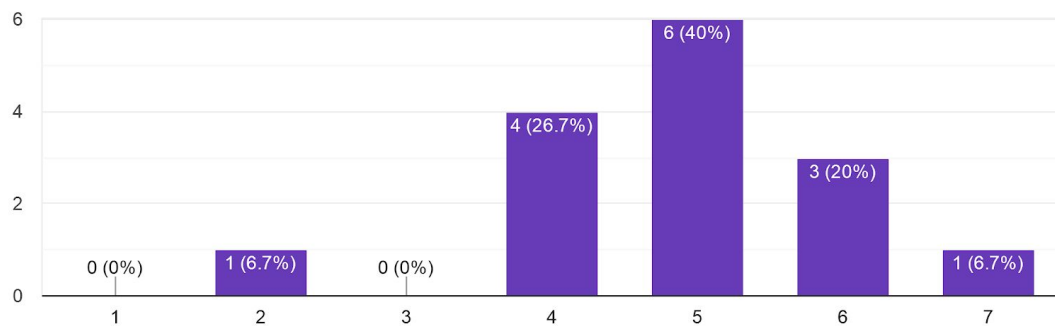
What did you think about the video quality?
15 responses
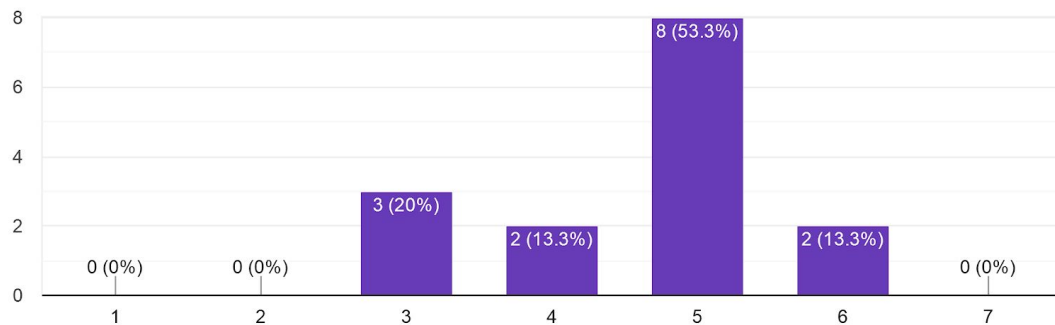
How immersed did you feel during the experience?
15 responses
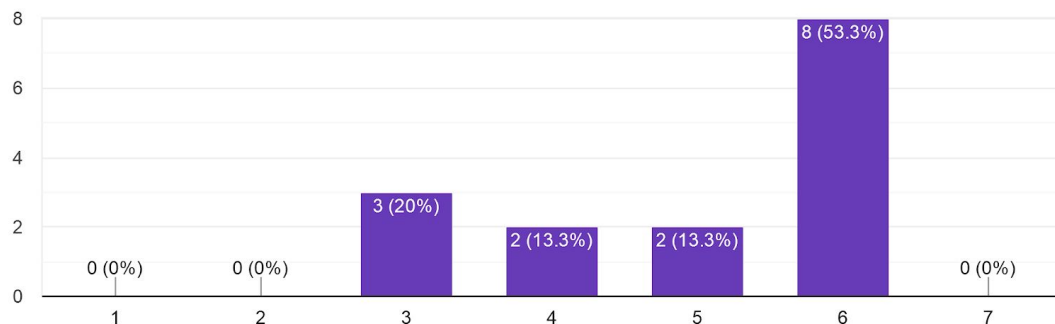
Did you like the viewing position? Would you have preferred a differentposition?

15 responses

1. I liked it. I would have preferred to see the conductor at the same time as the musicians, but I think this position worked very well.
2. Yes, I did. Cool to have the opportunity to see both te conductor and the orchestra.
3. I liked it, but it would also be cool to be able to change place
4. A bit scary, felt like a long way down to the floor
5. a bit further back maybe.
6. I would like the viewing position to be closer to the audience view, for a better concert experience and less need to rotate around the chair
7. Fin plass å sitte, men hadde vært interessant om man kunne bevege seg litt rundt.
8. den var grei. kunne gjerne hatt mulighet til å bevege meg rundt i orkesteret etter hva instrumentgruppe som var fremtredende der og da
9. Better to sit behind the conductor, in audience
10. quite good
11. I would prefer to be able to change to different positions during the performance.

12. Er veldig kult å kunne få sitte midt i ensemblet og lytte på hva som skjer, det gjør det jo definitivt enklere å se og høre hvor lyden kommer fra enn når du sitter fra salen og all lyden kommer forfra uansett. kunne kanskje vært litt høyere opp på samme posisjon, kanskje også gi muligheten for å bytte posisjon underveis? sette opp et par stasjoner man kan lytte fra.
13. It should be possible to switch places to sit in the audience, or to sit in the orchestra and watch the conductor.
14. Same position, but even higher up, to see everyone in the orchestra. Maybe 2 camera hights at the conductor, or just even higher up, but the angle was good to see what the conductor does and see the details in the closest string players. If possible, some cameras inside the orchestra would be cool.
15. Liked it.

Do you think experiences like this could be used in education? If yes, in what ways?
15 responses
1. Yes, it could be used to look at each individual instrument and how it is played with the others. Also for orchestral musiscians to look at the conductors gestures with the music.
2. Yes, it can help kids to see how an orchestra is put together, and experience how a philharmonic orchestra sounds like without actually being there. It's also easier to hear the details in a piece when you can see where the sound comes from.
3. Musical directing, learning about instrument groups, laerning about spesific pieces of music
4. Introduction to classical concerts
5. yes. this could be used to present music to students in a difrent way, were they get a visual representation of the instruments.
6. This experience could motivate younger listener to get a engaging look at classical music that can later inspire them to go to a concert in real life.
7. Absolutt. Første jeg tenker på er plasseringene på instrumentgruppene i en orkestral sammenheng.
8. ja. å reise på klassisk konsert kan fort bli dyrt. da kan dette være en god måte å oppleve klassisk musikk "live" i et klasserom
9. Yes, as an intro for deeper experience in a concert
10. Yes it would work. many ways
11. It could definately be used for conducting students.
12. ja, Definitivt, vet ikke helt hvordan, men for dirigentstudenter er det veldig bra å kunne observere en såpass prof dirigent på såpass nært hold.
13. I think the experience would be nice to use when learning to conduct or to play an instrument in the orchestra. The visual and audial cues would be easier to understand if you could be immersed in the experience of playing or conducting an orchestra.
14. This could be used in observating the conductors tecnique, and the communication between conductor and orchestra players. A other way is to see in detail how the individuals play. This is hard when you can se the once in the back.
15. Great for learning and experience the repertoare.

Do you have any other comments to the experience?

15 responses

1. I think it was a great experience.
2. No
3. it was exciting
4. The panning was ok, but the 3d experience was, lacking. Underwhelming sound quality.
5. if possible it would be nice to have the score in front of you when the music is played.
6. Would be nice, if possible, to have the score next to us, for better overview of who is playing next to the listening experience.
7. Fin opplevelse. Tenker dette er fint å videreutvikle.
8. var bra. videokvaliteten kunne vert bedre, og jeg ble sliten i øyner. men var kjekt å få et annen perspektiv i grieghallen, enn det du får som tilskuer i salen.
9. It felt a bit unnatural to sit so close to musicians, but this could be a preference for uneducated listeners. Some distant instruments were too close in sound. I could wish for a weigtened listening fex when turning my focus.
10. Interesting! Good potential. quality of sound very important to maximize experience.
11. Fix the audio before you fix the video. It will have a bigger impact on immersion. The audion might have been due to a sub-par headset.
12. det var veldig gøy, anbefales
13. The sound was kind of muted when you turned your head around and pointed your ear in that direction. It would be more immersive if the sound fitted more to the audio and positions.
14. The sound difference could be bigger, in the way that its more different from left to right, but not too big difference. You should hear an individual instrument more on the right, and slightly less in the left, for example. The video quality could be better.
15. Great project.

Appendix B

Statement from Musikkselskapet Harmonien

Til den det måtte angå.

Kandidaten hadde laget et fint og gjennomtenkt opplegg for å gjennomføre testing av VR-opplevelsen med musikere i Bergen Filharmoniske Orkester. Planlegging og påmelding var gjennomført, og tilbakemeldingsskjema var ferdig utarbeidet. Dato var lagt til onsdag 18. mars, i det som skulle vise seg å bli perioden Norge stengte ned for å hindre smitte av COVID-19.

Grunnet smitteverntiltak iverksatt av regjeringen og implementert av Musikkselskapet Harmonien, ble det i slutten av april 2020 bestemt at VR-testing med Harmoniens ansatte ikke lot seg utføre på en forsvarlig måte i inneværende semester.

Med vennlig hilsen,

Ingeborg Ekeland
Produsent barn og unge
Musikkselskapet Harmonien

Bibliography

[1] Philharmonia Orchestra, "Philharmonia vr | the virtual orchestra." `https://vimeo.com/214853600`. [Online, accessed 16-April-2020].

[2] 3DIO, "Omni binaural microphone." `https://3diosound.com/products/omni-binaural-microphone`. [Online, accessed 16-April-2020].

[3] 3DIO, "3dio." `https://3diosound.com/`. [Online, accessed 16-April-2020].

[4] Tchaikovsky Research, "Tchaikovsky." `http://en.tchaikovsky-research.net/pages/File:Photo050.gif`. [Online, accessed 10-May-2020].

[5] Resonance Audio, "Fundamental concepts." `https://resonance-audio.github.io/resonance-audio/discover/concepts.html`. [Online; accessed 8-March-2020].

[6] RØDE, "RØde nt-sf1." `https://www.rode.com/ntsf1`. [Online, accessed 23-January-2020].

[7] Sennheiser, "Ambeo vr mic." `https://en-us.sennheiser.com/microphone-3d-audio-ambeo-vr-mic`. [Online, accessed 8-March-2020].

[8] Humaneyes Technologies Ltd., "Vuze xr." `https://vuze.camera/camera/vuze-xr-camera`. [Online, accessed 06-January-2020].

[9] GoPro Inc, "Gopro max." `https://gopro.com/en/us/shop/cameras/max/CHDHZ-201-master.html`. [Online, accessed 26-March-2020].

[10] Insta 360, "Insta 360 pro 2." `https://www.insta360.com/product/insta360-pro2`. [Online, accessed 26-March-2020].

[11] Elaine.T, "01 research kickstarter: Images." `https://etsanggarp.blogspot.com/2016/03/01-research-kickstarter-images.html`. [Online, accessed 25-March-2020].

[12] I. E. Sutherland, "A head-mounted three dimensional display," in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pp. 757–764, 1968.

[13] Alex Vlachos, "Advanced vr rendering gdc 2015." `http://media.steampowered.com/apps/valve/2015/Alex_Vlachos_Advanced_VR_Rendering_GDC2015.pdf`. [Online, accessed 10-May-2020].

[14] Facebook Technologies, LLC., "Oculus rift." `https://www.oculus.com/rift/`. [Online, accessed 26-March-2020].

[15] HTC Corporation, "Htc vive." `https://www.vive.com/eu/product/vive/`. [Online, accessed 26-March-2020].

[16] Facebook Technologies, LLC., "Oculus quest." `https://www.oculus.com/quest/`. [Online, accessed 8-March-2020].

[17] The Khronos® Group Inc., "Openxr." `https://www.khronos.org/openxr/`. [Online, accessed 1-June-2020].

[18] Unity Technologies, "Xr plug-in framework." `https://docs.unity3d.com/Manual/XRPluginArchitecture.html`. [Online, accessed 20-April-2020].

[19] Grieghallen, "Grieghallen." `https://www.grieghallen.no/`. [Online, accessed 19-April-2020].

[20] COWI AS, "Cowi." `https://www.cowi.no/`. [Online, accessed 16-May-2020].

[21] Unity Technologies, "Audio mixer." `https://docs.unity3d.com/Manual/AudioMixer.html`. [Online, accessed 20-April-2020].

[22] Beyerdynamic, "Beyerdynamic dt-770 pro." `https://europe.beyerdynamic.com/dt-770-pro.html`. [Online, accessed 25-March-2020].

[23] Sony Electronics Inc., "Wh-1000xm2 wireless noise-canceling headphones." `https://www.sony.com/electronics/`

`headband-headphones/wh-1000xm2`. [Online, accessed 25-March-2020].

[24] Musikkselskapet Harmonien, "Bergen philharmonic orchestra." `https://harmonien.no/english/orchestra/`. [Online, accessed 19-April-2020].

[25] Jung - Woo Choi (Ed.) and Woon Seng Gan (Ed.), *Spatial Audio*. MDPI AG - Multidisciplinary Digital Publishing Institute, 2017.

[26] NRK, "Nrk." `https://www.nrk.no/`. [Online, accessed 19-April-2020].

[27] Philharmonia Orchestra, "Beethoven's fifth (full vr film)." `https://youtu.be/gFXWOOBUuW8`. [Online, accessed 8-March-2020].

[28] Philharmonia Orchestra, "Virtual reality: Sibelius 5." `https://www.philharmonia.co.uk/digital/virtual_reality_and_apps/sibelius_5`. [Online, accessed 8-March-2020].

[29] Philharmonia Orchestra, "Introducing the vr sound stage, from the philharmonia orchestra." `https://youtu.be/Un6HbDV8mQw`. [Online, accessed 22-October-2020].

[30] J. Janer, E. Gomez, A. Martorell, M. Miron, and B. de Wit, "Immersive orchestras: Audio processing for orchestral music vr content," in *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, pp. 1–2, Sep. 2016.

[31] I. Bergström, S. Azevedo, P. Papiotis, N. Saldanha, and M. Slater, "The plausibility of a string quartet performance in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, pp. 1352–1359, April 2017.

[32] MelodyVR Ltd, "Melodyvr." `https://melodyvr.com/`.

[33] Oculus, "Oculus venues." `https://www.oculus.com/experiences/go/1555304044520126`.

[34] Dolby Laboratories, Inc., "Dolby atmos." `https://www.dolby.com/us/en/brands/dolby-atmos.html`. [Online, accessed 28-October-2019].

[35] Sony Electronics INC., "360 reality audio." `https://www.sony.com/electronics/360-reality-audio`. [Online, accessed 23-March-2020].

[36] D. R. Begault and L. J. Trejo, *3-D sound for virtual reality and multimedia.* National Aeronautics and Space Administration, 2000.

[37] Alexander Poznansky, "Tchaikovsky: A life." `http://en.tchaikovsky-research.net/pages/Tchaikovsky:_A_Life`. [Online, accessed 10-May-2020].

[38] A. Poznansky, "Pyotr ilyich tchaikovsky." `https://www.britannica.com/biography/Pyotr-Ilyich-Tchaikovsky`, May 2020. [Online, accessed 10-May-2020].

[39] Tchaikovsky Research, "Swan lake." `http://en.tchaikovsky-research.net/pages/Swan_Lake`. [Online, accessed 10-May-2020].

[40] Tchaikovsky Research, "The nutcracker." `http://en.tchaikovsky-research.net/pages/The_Nutcracker`. [Online, accessed 10-May-2020].

[41] Tchaikovsky Research, "Violin concerto." `http://en.tchaikovsky-research.net/pages/Violin_Concerto`. [Online, accessed 10-May-2020].

[42] Tchaikovsky Research, "Symphony no. 5." `http://en.tchaikovsky-research.net/pages/Symphony_No._5`. [Online, accessed 10-May-2020].

[43] Musikkselskapet Harmonien, "Edward gardner." `https://harmonien.no/english/persons/conductors/edward-gardner/`. [Online, accessed 25-May-2020].

[44] J. Broderick, J. Duggan, and S. Redfern, "The importance of spatial audio in modern games and virtual environments," in *2018 IEEE Games, Entertainment, Media Conference (GEM)*, pp. 1–9, Aug 2018.

[45] K. Rogers, G. Ribeiro, R. R. Wehbe, M. Weber, and L. E. Nacke, "Vanishing importance: Studying immersive effects of game audio perception on player experiences in virtual reality," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, (New York, NY, USA), pp. 328:1–328:13, ACM, 2018.

[46] D. Hammershoi and H. Moller, "Binaural technique - basic methods for recording, synthesis, and reproduction," in *Communication Acoustics*, pp. 223–254, Springer Berlin Heidelberg, 2005.

[47] Soundfield, "Soundfield microphone basics." `https://www.soundfield.com/#/learn/microphone_basics`. [Online, accessed 8-March-2020].

[48] M. A. Gerzon, "Periphony: With-height sound reproduction," *Journal of the audio engineering society*, vol. 21, no. 1, pp. 2–10, 1973.

[49] S. Bertet, J. Daniel, E. Parizet, and O. Warusfel, "Investigation on localisation accuracy for first and higher order ambisonics reproduced sound sources," *Acta Acustica united with Acustica*, vol. 99, no. 4, pp. 642–657, 2013.

[50] Facebook Technologies, LLC., "Oculus audio sdk." `https://developer.oculus.com/design/audio-features/`. [Online, accessed 03-February-2020].

[51] Valve Corporation, "Steam audio." `https://valvesoftware.github.io/steam-audio/`. [Online, accessed 23-January-2020].

[52] Unity Technologies, "Unity." `https://unity.com/`. [Online, accessed 16-March-2020].

[53] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The cipic hrtf database," in *Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575)*, pp. 99–102, IEEE, 2001.

[54] W. G. Gardner, "3d audio and acoustic environment modeling," *Wave Arts, Inc*, vol. 99, 1999.

[55] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, "Localization using nonindividualized head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 94, no. 1, pp. 111–123, 1993.

[56] Michael Abrash, "Introducing michael abrash, oculus chief scientist." `https://www.oculus.com/blog/introducing-michael-abrash-oculus-chief-scientist/`. [Online, accessed 25-May-2020].

[57] Oculus, "Oculus connect 5 | keynote day 01." `https://youtu.be/o7OpS7pZ5ok?t=5605`. [Online, accessed 23-March-2020].

[58] Epic Games, "Unreal engine." `https://www.unrealengine.com/`. [Online, accessed 02-February-2020].

[59] Firelight Technologies Pty Ltd, "Fmod." `https://www.fmod.com/`. [Online, accessed 02-February-2020].

[60] Audiokinetic Inc., "Wwise." `https://www.audiokinetic.com/products/wwise/`. [Online, accessed 02-February-2020].

[61] Avid Technology, Inc., "Avid pro tools." `https://www.avid.com/pro-tools`. [Online, accessed 02-February-2020].

[62] MediaCollege.com, "What is reverb?." `https://www.mediacollege.com/audio/reverb/intro.html`. [Online, accessed 1-June-2020].

[63] M. M. A. Ermann, "Architectural acoustics illustrated," 2015.

[64] Facebook Technologies, LLC., "Simulating dynamic soundscapes at facebook reality labs." `https://www.oculus.com/blog/simulating-dynamic-soundscapes-at-facebook-reality-labs/`. [Online, accessed 1-June-2020].

[65] Epic Games Inc., "Unreal engine." `https://www.unrealengine.com/en-US/`. [Online, accessed 10-May-2020].

[66] Richard Gould, "Let's test: 3d audio spatialization plugins." `http://designingsound.org/2018/03/29/lets-test-3d-audio-spatialization-plugins/`. [Online, accessed 25-May-2020].

[67] Insta 360, "Insta 360 one x." `https://www.insta360.com/product/insta360-onex/`. [Online, accessed 26-March-2020].

[68] Ben Lang, "Analysis: Monthly-connected vr headsets on steam reach record high of 1.3 million." `https://www.roadtovr.com/analysis-monthly-connected-vr-headsets-on-steam-record-high/`. [Online, accessed 26-March-2020].

[69] T. Akenine-Möller, E. Haines, N. Hoffman, A. Pesce, M. Iwanicki, and S. Hillaire, *Real-Time Rendering 4th Edition.* Boca Raton, FL, USA: A K Peters/CRC Press, 2018.

[70] ARVI VR INC, "Combating vr sickness: Debunking myths and learning what really works." `https://vr.arvilab.com/blog/`

combating-vr-sickness-debunking-myths-and-learning-what-really-works.
[Online, accessed 10-May-2020].

[71] Valve Corporation, "Steamvr." `https://store.steampowered.com/app/250820/SteamVR/`. [Online, accessed 10-May-2020].

[72] Valve Corporation, "Valve." `https://www.valvesoftware.com/no/`. [Online, accessed 10-May-2020].

[73] Valve Corporation, "Steam." `https://store.steampowered.com/?l=norwegian`. [Online, accessed 10-May-2020].

[74] Ni-iX, "Unity vs. unreal: How to choose the best game engine." `https://medium.com/@N_iX/unity-vs-unreal-how-to-choose-the-best-game-engine-d3dbb4add73c`. [Online, accessed 1-June-2020].

[75] Facebook Technologies, LLC., "Oculus integration for unity." `https://developer.oculus.com/downloads/package/unity-integration/`. [Online, accessed 14-May-2020].

[76] Valve Corporation, "Steamvr unity plugin." `https://valvesoftware.github.io/steamvr_unity_plugin/`. [Online, accessed 23-January-2020].

[77] Unity Technologies, "Xr interaction toolkit." `https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@0.9/manual/index.html`. [Online, accessed 20-April-2020].

[78] Musikkselskapet Harmonien, "Tchaikovsky's fifth." `https://harmonien.no/english/programme/2019/11/tchaikovsky-s-fifth/`. [Online, accessed 1-February-2020].

[79] Harmonien, "Bergenphilive." `https://www.bergenphilive.no/`. [Online, accessed 19-April-2020].

[80] Musikkselskapet Harmonien, "Tchaikovsky: Symphony no. 5." `https://www.bergenphilive.no/video-konserter/2019/11/tchaikovsky-symphony-no-5/`. [Online, accessed 19-April-2020].

[81] Cockos Incorporated, "Reaper." `https://www.reaper.fm/`. [Online, accessed 19-April-2020].

[82] Senter for nye medier, "Western norway university of applied sciences." `http://www.hib.no/om-hogskolen/senter-for-nye-medier/`. [Online, accessed 18-November-2019].

[83] Western Norway University of Applied Sciences, "Western norway university of applied sciences." `https://hvl.no/en/`. [Online, accessed 25-March-2020].

[84] Humaneyes Technologies, "Vuze vr studio." `https://vuze.camera/support/vr-software/vuze-vr-studio`. [Online, accessed 02-February-2020].

[85] Facebook Technologies, LLC., "Encoding high-resolution 360 and 180 video for oculus quest and oculus go [updated to include h.265 support]." `https://creator.oculus.com/blog/encoding-high-resolution-360-and-180-video-for-oculus-go/`. [Online, accessed 20-April-2020].

[86] Blender Foundation, "Blender." `https://www.blender.org/`. [Online, accessed 22-January-2020].

[87] Autodesk Inc., "Autodesk." `https://www.autodesk.com/`. [Online, accessed 14-May-2020].

[88] Bart Veldhuizen, ""next gen" - blender production by tangent animation soon on netflix!." `https://www.blendernation.com/2018/08/20/next-gen-blender-production-by-tangent-animation-soon-on-netflix/`. [Online, accessed 14-May-2020].

[89] Facebook Technologies, LLC., "Testing and performance analysis." `https://developer.oculus.com/documentation/unity/unity-perf/`. [Online, accessed 14-May-2020].

[90] Unity Technologies, "Occlusion culling." `https://docs.unity3d.com/Manual/OcclusionCulling.html`. [Online, accessed 14-May-2020].

[91] Unity Technologies, "Audio normalize effect." `https://docs.unity3d.com/Manual/class-AudioNormalizeEffect.html`. [Online, accessed 20-April-2020].

[92] Facebook Technologies, LLC., "Audio propagation (beta)." `https://developer.oculus.com/documentation/unity/audio-osp-unity-propagation/`. [Online, accessed 21-May-2020].

[93] Unity Technologies, "Panoramic video." `https://docs.unity3d.com/Manual/VideoPanoramic.html`. [Online, accessed 26-May-2020].

[94] J. P. Snyder, *Flattening the earth: two thousand years of map projections.* University of Chicago Press, 1997.

[95] Unity Technologies, "Render texture." `https://docs.unity3d.com/Manual/class-RenderTexture.html`. [Online, accessed 20-February-2020].

[96] Unity Technologies, "Skyboxes." `https://docs.unity3d.com/Manual/skyboxes.html`. [Online, accessed 26-May-2020].

[97] Unity Technologies, "Layers." `https://docs.unity3d.com/Manual/Layers.html`. [Online, accessed 26-May-2020].

[98] Unity Technologies, "Unity ui: Unity user interface." `https://docs.unity3d.com/Packages/com.unity.ugui@1.0/manual/index.html`. [Online, accessed 26-May-2020].

[99] Unity Technologies, "Video player." `https://docs.unity3d.com/ScriptReference/Video.VideoPlayer.html`. [Online, accessed 20-April-2020].

[100] Unity Technologies, "Coroutines." `https://docs.unity3d.com/Manual/BestPracticeUnderstandingPerformanceInUnity3.html`. [Online, accessed 26-May-2020].

[101] Dragefjellets musikkorps, "Bjørn sagstad." `http://dragene.no/p/bjorn-sagstad/`. [Online, accessed 23-May-2020].

[102] Facebook Technologies, LLC., "Uploading android apps." `https://developer.oculus.com/distribute/publish-uploading-mobile/`. [Online, accessed 18-April-2020].

[103] Facebook Technologies, LLC., "Asset files to manage download size." `https://developer.oculus.com/documentation/unity/ps-assets/`. [Online, accessed 18-April-2020].

[104] Unity Technologies, "Assetbundles." `https://docs.unity3d.com/Manual/AssetBundlesIntro.html`. [Online, accessed 18-April-2020].

[105] Jostein Stalheim, "Jostein stalheim." `https://www.josteinstalheim.no/`. [Online, accessed 15-March-2020].

[106] R. Likert, "A technique for the measurement of attitudes.," *Archives of psychology*, 1932.

[107] HVL, "Fm1-avm2-10 music in use 2." `https://www.hvl.no/en/studies-at-hvl/study-programmes/course/fm1-avm2-10`. [Online, accessed 16-March-2020].

[108] HVL, "Study plan - teacher education - bachelor of music." `https://www.hvl.no/en/studies-at-hvl/study-programmes/2019h/fag-musikk/study-plan/`. [Online, accessed 16-March-2020].

[109] University of Bergen, "University of bergen." `https://www.uib.no/`. [Online, accessed 23-May-2020].

[110] University of Bergen, "Ensemble conducting." `https://www.uib.no/en/course/MUV211`. [Online, accessed 25-March-2020].

[111] World Health Organization, "Coronavirus disease (covid-19) pandemic." `https://www.who.int/emergencies/diseases/novel-coronavirus-2019`. [Online, accessed 23-May-2020].

[112] Christian Thomas, "Beyerdynamic dt 770 studio 80 ohm review." `https://www.soundguys.com/beyerdynamic-dt-770-studio-80ohm-review-15348/`. [Online, accessed 4-June-2020].

[113] Christian Thomas, "Sony wh-1000xm2 review." `https://www.soundguys.com/sony-wh-1000x-m2-review-14304/`. [Online, accessed 4-June-2020].

[114] Facebook Technologies, LLC., "Quest virtual reality check (vrc) guidelines." `https://developer.oculus.com/distribute/publish-quest-req/`. [Online, accessed 21-May-2020].

[115] Facebook Technologies, LLC., "Ovr metrics tool." `https://developer.oculus.com/documentation/native/android/mobile-ovrmetricstool`. [Online, accessed 20-April-2020].

[116] Vimeo Inc., "Vimeo." `https://vimeo.com/`. [Online, accessed 10-May-2020].