

UNIVERSITY OF TARTU VILJANDI CULTURAL ACADEMY

Faculty of Arts and Humanities
Sound and Visual Technology

Qumrosh Abbas Khan

Soundscape of Estonia- **“A comparison of Soundscape of Tallinn and Viljandi”**

Masters project

Supervisor:
Christoph Felix Schulz

Viljandi 2023

Acknowledgment!

I would like to thank my Wife Aqsa Mahboob who supported me during my Master's Degree.

I would like to thank Professor Christoph Felix Schulz, who supported me in my education by providing me with the opportunity to explore new ideas, techniques, and equipment. I also want to Thank you (Christoph) for letting me use your equipment for the master's thesis.

I would like to thank the University of Tartu, for accepting me as their student and allowing me to express my desire to record the natural ambiance of Estonia in 1st order Ambisonics and Present it as my Master's Project.

Finally, I would like to Thank Viljandi Cultural Academy for their support in providing equipment and resources for the recording.

Abstract. Soundscape of Estonia- “A comparison of Soundscape of Tallinn and Viljandi”

The thesis aimed to record the Soundscape of two cities of Estonia, Tallinn, and Viljandi, and try to find how noise is affecting human life, the comparison is done on an aesthetical level. The locations recorded were selected on the basis of humans visiting that location. The resulting material was uploaded on Soundcloud in 1st order Ambisonics Audio format and Video was uploaded on YouTube. I hope that this study would help in contributing to and improving the noise pollution situation and how important it is to record the current Soundscape before we lose it.

Keywords: Ambisonics, VR, Adobe Premiere, Nuendo, Sennheiser Ambeo VR Mic.

1. Introduction:	6
1.1 Background and Motivation:	6
2. Methodology and Approach:	8
2.1 Research Questions and Objectives:	8
2.2 Understanding Ambisonics:	8
What is Ambisonics and why did I choose this as my audio recording format?	8
3. Literature Review	11
3.1 The Importance of Natural Soundscapes:	11
Reference 1:	11
Masking:	11
Physiological stress:	11
Behavioral Change:	11
3.2 The Impact of Noise Pollution on Humans:	11
Reference 1 (Hygge, 2003):	11
3.3 The Role of audio recording in documenting natural Soundscapes:	12
3.3.1 Biodiversity Monitoring	12
3.3.2 Non-Invasive Monitoring	12
3.3.3 Education and Monitoring:	12
3.3.4 Audio Library for films, games, and Metaverse:	12
4. Methodology:	14
4.1 Recording equipment and techniques:	14
4.1.1 Microphone:	14
4.1.2 Audio Recorder:	14
4.2 Selection of recording locations:	15
4.3 Data processing:	16
4.3.1 Audio Processing in DAW:	16
4.3.2 Converting Audio to Video	18
5. Results:	22
5.1 Analysis:	22
6. Discussion:	24
6.1 Implications of the Findings for environmental policy and Management:	24
6.1.2 Protecting natural soundscape:	24
6.1.3 Managing Noise Pollution:	24
6.1.4 Educating the public:	24
6.2 Future research directions:	24
6.2.1 Importance of the soundscape:	24
6.2.2 Application of Ambisonics:	25
6.2.3 Equipment cost:	25
6.2.4 Processing Ambisonics:	25
6.3 Limitations & Suggestions:	25
6.3.1 Field Recording Device:	25
6.3.2 Post-Production:	26
7. Hyper-Link to the album:	26
References:	27

1. Introduction:

1.1 Background and Motivation:

Humans tend to remember locations, emotions, and events not just visually but also with sounds that they hear. While growing up in my hometown I could hear different birds, animals, and natural ambiences, however with time the city changed and now there are huge buildings all over the city and you can hardly hear the sound of the birds, or the natural ambiance of the environment. This changed a lot of things, now all the beautiful sounds of nature are replaced with a very harsh sound spectrum of vehicles and human-generated noises. As a result, our ears don't get relaxed at all and we are losing our sense of listening as we are only focusing on specific frequencies which we hear all the time.

As a result of global warming and increasing temperature the glaciers are melting and we will soon lose the world as we know it. Last year 70% of Pakistan was drowned in flood due to global warming and scientist predict that it's just the beginning. Europe experienced a severe heat stroke and now the cold storms.

I came to Estonia- a country which is 52% forest and started doing my Master's in Sound and Visual Technology from the University of Tartu, Viljandi Campus, and started living in Viljandi. I started exploring the city and realized how beautiful and peaceful the city is. The natural ambiance, the wind in the forest, and the sound of birds were audible and very pleasant to the ears. After spending some time in this environment, I realized what the rest of the world is missing and is not paying attention to.

I believe we as humans have started to ignore our sense of listening, we have engaged our ears in so many different types of noises that we cannot focus on what we want to hear and how our listening can help us. Our ancestors were able to identify threats, and emotions just by listening to different sounds but now we cannot do that because we are occupied by different human-created noises, and with the passage of time our sense of listening has deteriorated.

I believe that if we don't recognize the importance of Geo-Phony (Sound created/made by earth) and Bio-Phony (sound created/made by all living beings other than humans) we will lose a very important part of this planet and humans will suffer because of it.

Almost all the organizations working towards a better environment don't realize the importance of sound, they consider noise pollution as an issue but what are we doing about it? We talk about noise pollution and discuss how it affects humans and this is very good but in the meantime are we recording or documenting the current natural sound that exists around us in any way?

If we look at recorded history then we can find some elements from the Prehistoric period and then we come to the initial civilization and then later on we are in the modern world. All of this historical information comes from books, paintings, sculptures, and remaining's buildings or civilizations. The only audio version of history we have is through written music or oral tradition which mostly has folk or religious context. Even after the industrial revolution and

the development of audio recording technology we still don't have enough audio recordings of the geophony or biophony. Most of the natural ambiance recordings are made for the sake of libraries for either films, video games, theater, and other visual media but not for the sake of audio archives so that the later generations can listen to the sound of nature that we can hear now or what our grandparents could hear and we can't.

If you search online and try to find natural ambiance you would be able to hear hundreds of albums with the sound of nature but none of them present the actual ambiance, these are all different sounds recorded separately on the location and later on, mixed in the studio according to the platform they were published for. If you listen to them, they don't sound natural at all. It just feels like it's a mockery of nature.

Concerning all this, I decided to record the natural ambiance of different locations in Estonia and compare their effects on humans to highlight the importance of how natural sounds affect us.

2. Methodology and Approach:

To record audio, I have decided to use the Ambeo VR microphone by Sennheiser (Sennheiser, 2023), this particular mic is chosen because it records audio in 1st order of Ambisonics (Furness, 1990), and if recorded properly then I don't need to apply any post-processing to the audio signal. Since the audio recorded is in Ambisonics format thus the user can experience the surround sound. The elements that would be recorded would remain exactly the way they were at the time of recording.

This master's project would contain 1 hour and 59 minutes of audio recordings of the different marked locations. The locations chosen are based on how frequently people visit those locations.

The written material of the thesis project would be the discussion on the technical elements like understanding Ambisonics and its applications, and how to implement the Ambisonics audio in 360* videos for YouTube. Understanding the software that would be used for this purpose. Future research and the result of the audio recording.

2.1 Research Questions and Objectives:

The research compares the audio soundscape between two cities of Estonia- Tallinn and Viljandi- I have chosen these cities because I have personally lived in these cities and I know how they sound and what is the major difference between the audio soundscape. So, my research question, in a nutshell, would be **“A comparison of the Soundscape of Tallinn and Viljandi”**.

2.2 Understanding Ambisonics:

What is Ambisonics and why did I choose this as my audio recording format?

Ambisonics is a 360 or full-sphere surround sound format, unlike other surround audio systems Ambisonics also records and presents the horizontal plane. Ambisonics is independent of the speaker setup and that is because it doesn't carry any speaker-oriented audio signal instead it carries an independent representation of the sound field called B-format. This feature allows the producer to work according to the audio source direction instead of the speaker layout.

This system was developed in the UK in 1970 with the support of the British National Research Development Corporation.

So 1st order Ambisonics would have 4 channels, W, X, Y, and Z.

- **W** is an omnidirectional polar pattern, which captures all the sound coming present in the sphere.
- **X** represents a bi-directional figure-8 polar pattern mic that points forward.
- **Y** represents a bi-directional figure-8 polar pattern mic that points left.
- **Z** represents a bi-directional figure-8 polar pattern mic that points up.

So, to understand the above patterns and their utility let's try to understand them. Let's choose channel X and since we know that it's a figure-of-8 bi-directional polar pattern mic with its positive side on the front and negative side on the rear. If an object enters this sound field it would be recorded with the highest gain when the object is exactly at the front or back side of the mic and since we know that the polar pattern is Figure 8 so the gain level will change with the movement of the object and it will become zero when the object reaches the exact 90-degree for this channel. Given that it is a 4-channel configuration the Y and Z mics would be recording the left-right and up-down signal. So technically by the time object would reach the 90-degrees for the X channel, it would be at the highest gain level for the Y channel, and if the object is also moving up and down then the Z channel would record this information, with all this information we still have a W channel which has an omnidirectional microphone which would be recording the whole sound sphere.

Ambisonics 360 audio is sometimes confused by people with other surround technologies. But they are not the same, they have major differences in between them. Traditional surround technologies are more immersive than a simple two-channel stereo setup, however, the principle is similar for all of these setups. The audio image is created by sending a specific channel to specific speakers; thus, the end product is predefined and you would always know how the sound is going to move in the space. A stereo setup channel would send audio to two channels, 5.1 would send it to 6, and 7.1 would send it to 8 channels. So, if you have a 7.1.4 speaker surround render file then you have to put 7 floor-standing speakers at the listener's ear level, 1 subwoofer, and 4 high speakers. Hence for this setup, you would have to route your audio signal for each channel accordingly. This is why Dolby Atmos uses bed and object-based configurations for their channels. The bed would represent the sound that needs to remain stationary and object sound would be the audio files that you want to move in this configuration. Let's say if you have managed to do all this configuration and the user doesn't have a 7.1.4 audio setup then s/he won't be able to experience the setup.

However, if you choose Ambisonics doesn't send audio to any specific speaker layout. Ambisonics can be decoded into any speaker array and this is because the end product in B-format (A-format refers to raw recordings from the mic and when they are decoded in surround format then it's called B-format). (Waves-blog, 2017)

Some important points to consider:

- Traditional surround format provides good information when the image is static, however, if they rotate the sound field they fail, but with Ambisonics you can create a more clear, smooth and precise soundscape.
- Traditional surround format is more focused on front-biased, information from side rare is not focused, whereas in Ambisonics the whole 360 audio spectrum is focused very clearly.
- Traditional surround format has difficulties in representing sounds beyond the horizontal dimension, whereas Ambisonics deliver full sphere information including elevation information.

Since I am preserving the audio history of Estonia so I wanted to make sure that the end product should have the least processing, so that maximum audio elements are preserved in their natural form. The mic I choose for my recording session was the Sennheiser Ambeo VR mic which provides high-quality 360-degree audio recordings. Once I am done with recording the ambiances, I can import the audio in my DAW (Nuendo) where I converted the mic audio into 1st order Ambisonic. Ambisonics has a better representation of how humans listen to sounds there my end product would in 1st order ambisonic audio track which renders in adobe premiere with VR monoscopic video (with still image) settings and 4.0 (1st-order ambisonic) audio channel, then the video would be injected using spatial injector so that YouTube reads the video as 360-degree video & audio. (Matthias, 2019)

3. Literature Review

3.1 The Importance of Natural Soundscapes:

Reference 1:

In the research paper titled "The Cost of Chronic Noise Exposure for Terrestrial Organisms," the authors focus on three major costs for chronic noise exposure. They refer to these three terms as masking, physiological stress, and behavioral changes. It's important to consider chronic noise in this report refers to continuous or long-term exposed noise created by traffic, industry, or construction. (Jesse R. Barber, March 2009)

They further explain how these three factors play their role.

Masking:

Masking is a very common term used by sound engineers and beauticians, the writer has used this term to highlight the fact that animals' ability to hear and interpret signals from their environment by the frequencies overlaps of technological waves, traffic, and industrial noise. This can lead to reduced reproductive success, change in social structure, and altered distribution patterns.

Physiological stress:

This occurs when animals are exposed to chronic noises which exceed their tolerance levels. This effect is different for different species of animals, the side effects commonly noticed are increased heart rate, suppressed immune function, and increased cortisol levels.

Behavioral Change:

This change occurs when animals change their behavior towards chronic sounds this can include changes in migration patterns and predator avoidance strategies. Since migration and predator avoidance play a very important role in the life of animal and plant composition, so this kind of change can affect the ecological system.

In conclusion, the authors highlight the importance of natural sound for biodiversity and emphasize the fact that noise created by humans has severe effects.

3.2 The Impact of Noise Pollution on Humans:

Reference 1 (Hygge, 2003):

"The Effect of road traffic noise on Reading and listening comprehension in School" was conducted by Hygge, Boman, and Enmarker in 2003. The researcher found that exposure to road traffic effects negatively to both the reading and listening capacity of children. The study also found that a frequency range between 2-8 kHz was more harmful to students than the lower frequency range of 0.125-2k. The age group of students was focused on 7 to 12 years and these

kinds were randomly assigned to different classrooms with different levels of noise exposure. The noise level in each classroom was measured standard sound level meter, and the same was done for the quiet room. Although the study doesn't specifically identify the issues the students faced it highlights the cognitive performance and annoyance in students.

Reference 2 (V Pathak, 2008):

The study showed that the traffic noise level has increased to an alarming level in modern cities if further indicates that 85% of the people are disturbed by the traffic noise and 90 percent of people have reported that traffic noise is the reason for their suffering from headache, high BP problems, dizziness, and fatigue.

Just as noise affects the reproductive cycle in animals, it was found in this research that marriages were significantly affected by traffic noise.

The study was mainly conducted in the Varanasi city of India, with special reference to noise levels. 20 sites were selected for sampling. This location selected included industrial, commercial, residential, and silent zones.

3.3 The Role of audio recording in documenting natural Soundscapes:

Audio recordings play a crucial role in documenting the soundscape of any location. It shows the ecological system of the location without interfering with the environment. Throughout history, we see several documents and visuals that show the history of any location during a certain period. However, we don't have such documents for audio.

3.3.1 Biodiversity Monitoring:

Audio recordings can be used to identify different breed of animals and birds or other living creature in a particular location. It also helps in identifying their behaviors.

3.3.2 Non-Invasive Monitoring:

Unlike the tapping and tagging method of documentation, with audio recording, you can simply place a mic with a recorder and power source to a location and record the soundscape. This removes the involvement of human beings in the environment.

3.3.3 Education and Monitoring:

Audio recording can be used to create an immersive experience for people who are away from a particular location and they want to listen to that place or it can also use as a memory book instead of pictures.

3.3.4 Audio Library for films, games, and Metaverse:

Any kind of field audio recording that captures nature can be used for multiple purposes in films, and games of metaverse.

In filmmaking, sound plays a very important role in making the scene for the film more interesting and immersive. These sounds used are not limited to natural ambiance like a forest, rain, waterfall, wind blowing, or bird sounds. In fact, in films the audio library can be added according to all the objects visible or non-visible in the scene. A good example of this can be the city traffic audible in a scene where actors are acting in a room with open windows.

In terms of Video Games, the ambiance becomes more player- and object-based. Each object in the game scene or location has its unique sound and multiple parameters are activated by different actions of the players. While in films, the recorded sounds and visuals remain the same so that background ambiances don't change however in video games the ambiances change with every action of the player.

4. Methodology:

4.1 Recording equipment and techniques:

4.1.1 Microphone:

The most important element for audio recording is the Microphone and, in this project, the Sennheiser Ambeo VR mic was chosen due to the following reasons. Ambeo VR Mic has 4 condenser capsules that deliver very high-quality 360-degree 1st order Ambisonic recording. The capsules of the Sennheiser Ambeo VR mic deliver A-format, a raw 4-channel output that has to be converted into a new set of 4 channels, the Ambisonics B format. This is done by the specifically designer Sennheiser Ambeo A-B converter Plugin. B-format is a W, X, Y, and Z reorientation of the sound field around the microphone. W is the sum of all 4 capsules, whereas X, Y, and Z are the three virtual bi-directional microphone patterns representing front/back, left/right, and up/down.

It is important to reduce all kinds of mechanical noise which can be generated by either the recorder, recording stand, or cables. It is important to make sure that when recording outdoors, as for this study the main location was the outdoor location which has more frequency of human moving, I had to use a Rycote with a dead-cat cover on a shock mount.

The microphone front upside (labeled on the mic) was always placed on the opposite side of me, which in other words meant that I was always on the back side of the front side. It was very important to make sure that I am aware of the direction of the mic and how it was placed, even a slight error can create serious confusion.

Another important element to be considered here is the fact that I had always placed the microphone in the upright, this decision was made based on how humans stand and listen to sound. It is important to consider that if you change the mic to upside down or in EndFire the audio orientation would change.

The height of the stand was always at my ear level and my height is 170cm. so the mic was always at my ear level.

The above configuration for mic placement was decided based on several tests, and eventually, I found my solution.

4.1.2 Audio Recorder:

Since the microphone has 4 channels and each channel should have an exact gain level so it is advised to use a modern field recording that can link channels and can record audio at 32bit float. This 32-bit factor allows an additional noise-to-signal ratio. Since the project was focused on recording natural soundscapes so the possibility of having a situation where someone would shout in the recording is common or there can be some random loud noise and if you have a 32-bit audio recording option in your recorder then you don't need to worry about this. However, due to a lack of equipment and financial equipment, the Roland R88 field recorder

was used and it has good pre-amps. Since the recorder doesn't link the audio channel if not connected to a computer so it's important to make sure that the gain level of each channel matches. Since the microphone has 4 phantom power mics so it's important to consider a good quality long-lasting power supply for the recorder. In this case, I was using a Swit Gold Mount battery pack. The battery can last one full day of recording with 4 phantom-powered channels.

4.2 Selection of recording locations:

Since the project focuses on the soundscape of two different cities, the locations were selected based on human activity in those locations. The following table shows similar locations in Tallinn and Viljandi

Tallinn.	Viljandi.
Tallinn Beach.	Viljandi Lake & Viljandi Lake 2.
Kadriorg Park.	Viljandi Ruins.
Public Play area for kids – Kids of Tallinn.	Public Play area for kids- Kids of Viljandi.
Birds in Tallinn.	Birds in Viljandi.
Café Maiasmokk.	Kohvik Fellin.
Stockmann Square Tallinn.	Freedom Square Viljandi.
Walk in Tallinn.	Walk in Viljandi.

4.3 Data processing:

4.3.1 Audio Processing in DAW:

The recorded data was processed in DAW Nuendo 11.

Since the microphone records audio in Ambisonics A format so it's necessary to convert this audio into B format Ambisonics. The process can be done in the following order

- Import audio in 4 different audio channels.
- Set a quadrophonic group track for these 4 audio channels.
- On the group track add the Ambeo A-B converter plugin.

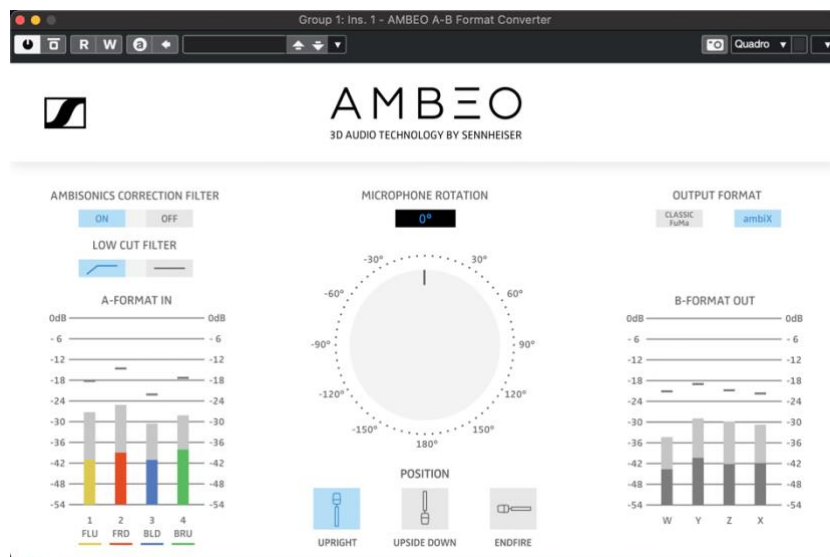


Figure 1- Ambeo A-B Format Converter (Sennheiser, 2023)

- You should now see the audio channel output in the plugin.
- Now add a 1st order Ambisonic audio bus in your outputs from your audio connection tab.
- Change the main mix for stereo out to 1st order Ambisonic track and remove the audio device connection from this bus, once you do this then you can monitor audio through the Nuendo control room.
- Now change the group track out from stereo mix to 1st order Ambisonic bus.
- As soon as you route the group track from stereo to 1st order Ambisonic the audio panner of the group track changes to Ambisonic.
- To monitor all four channels properly you have to change the group track panner from Vst MultiPanner to standard panner. If you don't do this, you would have one missing channel in your recording.

- Then 4 audio channels are panned in the following configuration- 1st Front left, 2nd Front Right, 3rd rear left, and 4th rear right



Figure 2- Nuendo 11 Channel Layout and Panning (GmbH, 2023)

- Once you are satisfied with the audio panning then you can export the audio by selecting the 1st order Ambisonic track from the Audio mixdown.

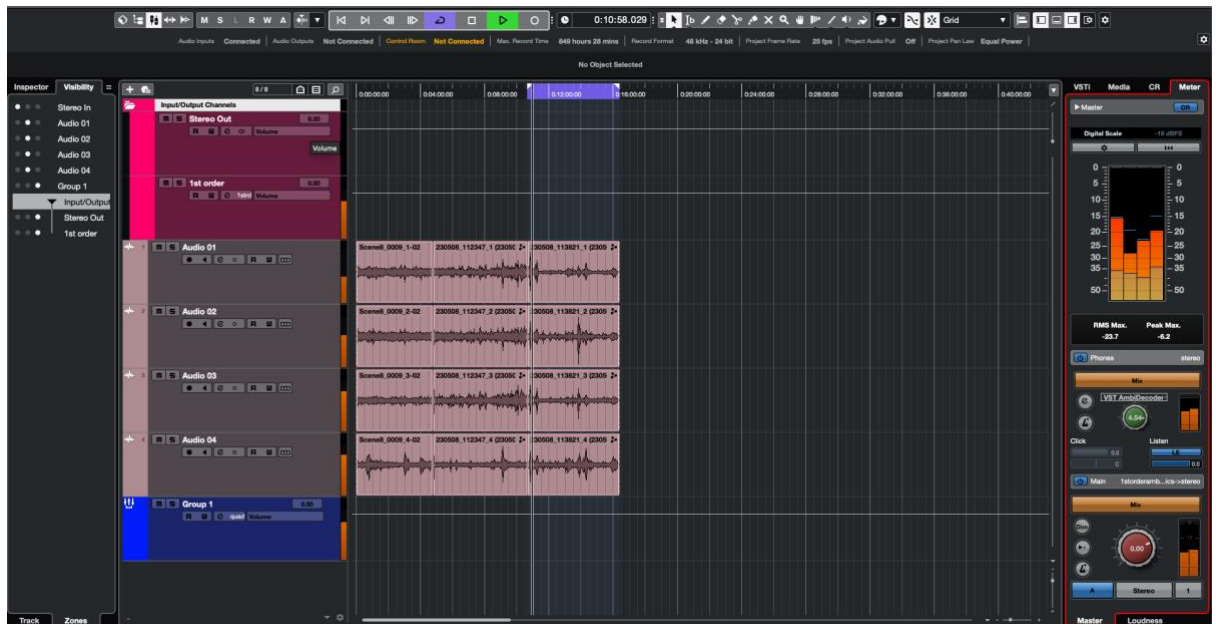


Figure 3- Nuendo 11 Project Layout (GmbH, 2023) (Adobe, 2023)

4.3.2 Converting Audio to Video:

Since there are very few online portals that support 1st order Ambisonic, so YouTube was selected for uploading the audio recording as it supports special audio. This can be achieved by the following process.

- Import your audio track in Adobe Premier along with a picture or video.
- Make sure that your timeline has the default audio track's Multichannel Mono Media and is adaptively selected.

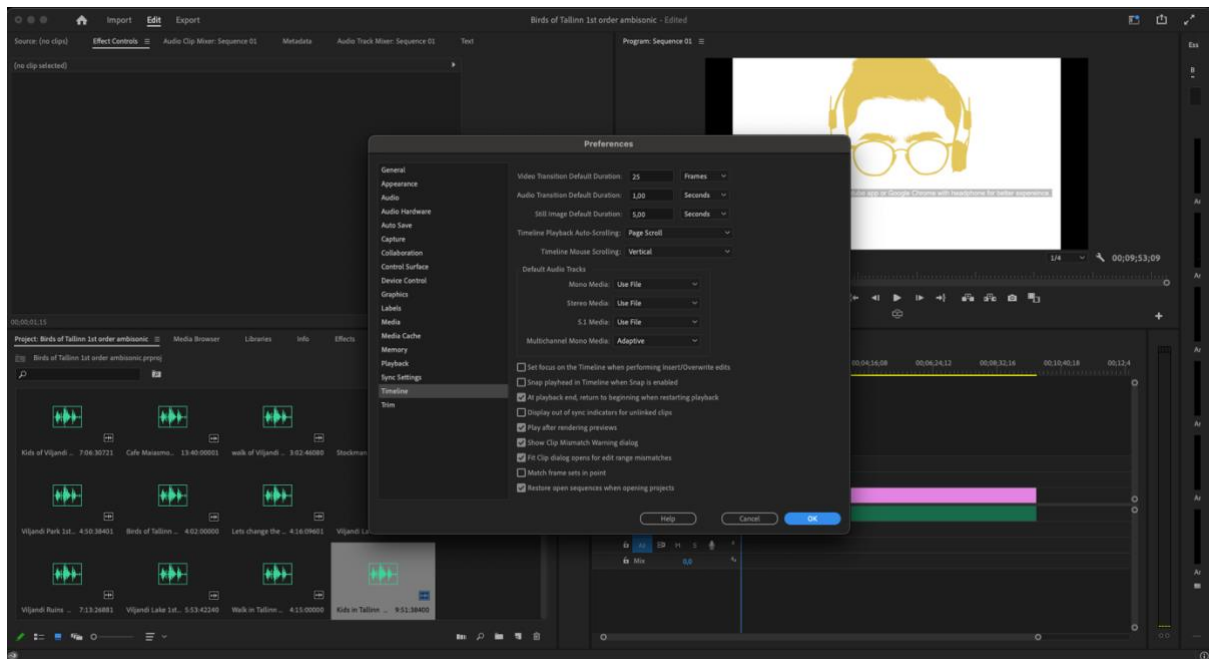


Figure 4- Adobe Premiere Multichannel Mono Media (Adobe, 2023)

- From your media, select the audio track, right-click on it, and go to modify from drop-down, then select Audio Channels.

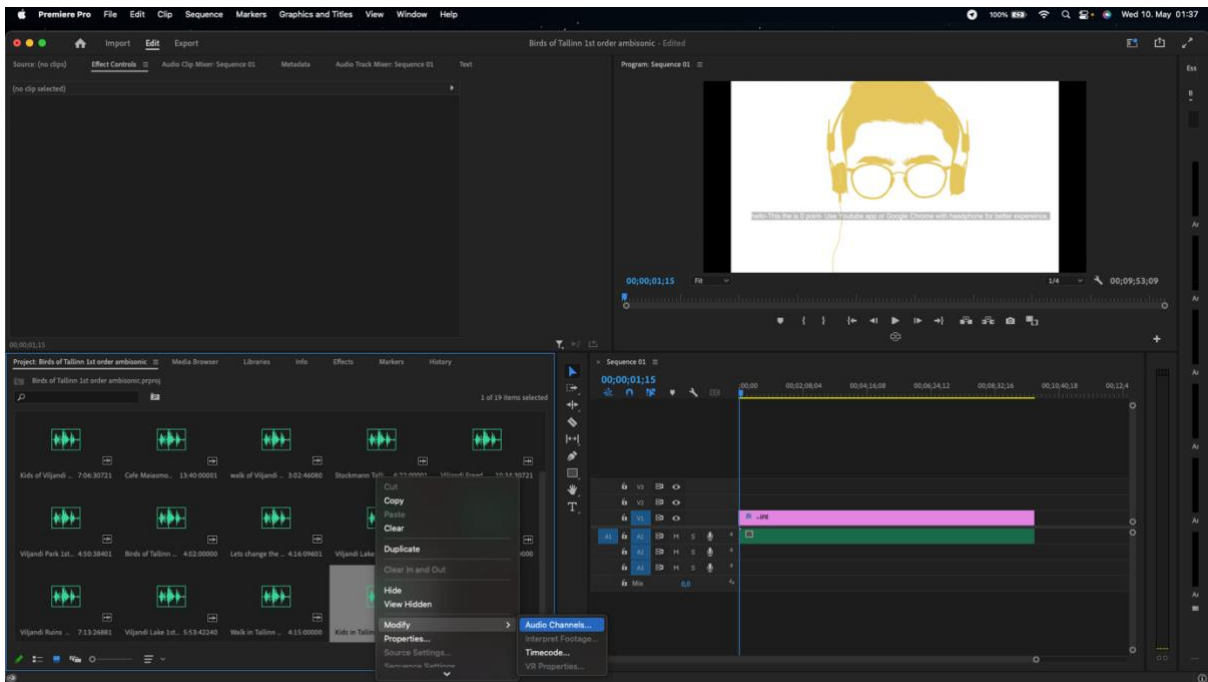


Figure 5 Adobe Premiere modify-Audio Channels (Adobe, 2023)

- In your audio channels make sure that the active channel per clip is 4 (for 1st order Ambisonic)

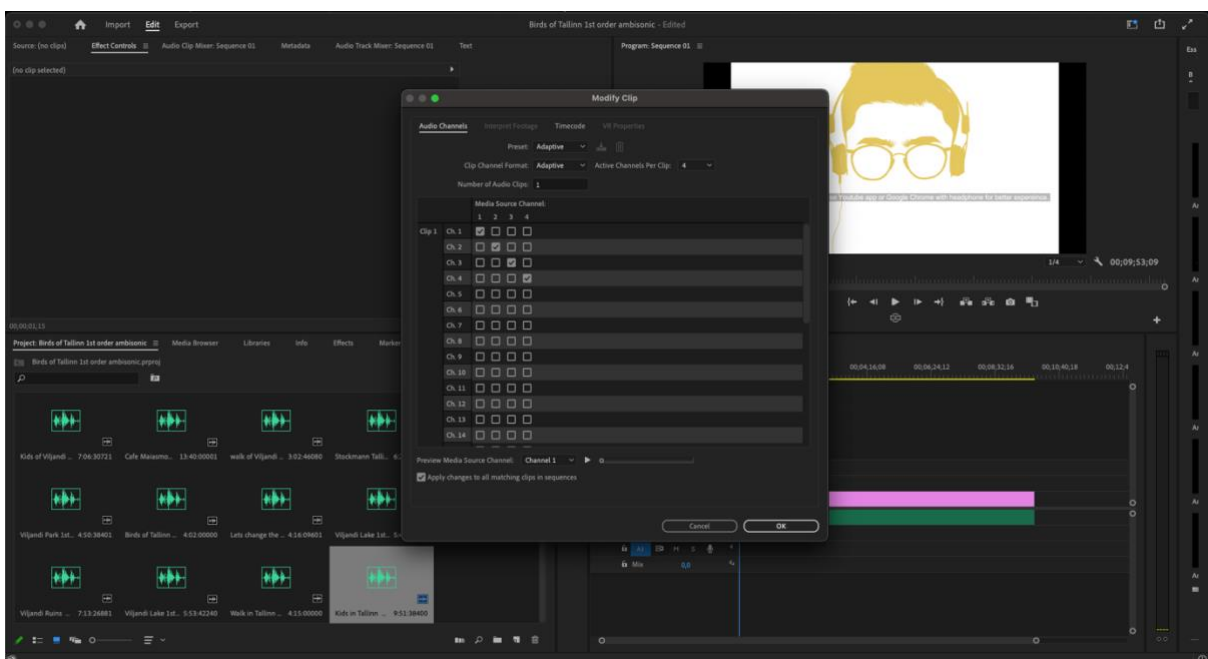


Figure 6 Adobe Premiere - Activate channel per clip (Adobe, 2023)

- Then create a new sequence- choose a VR preset
- Drag and drop your audio file (video or image if you want to use it) in the sequence.
- Click on Export from the File drop-down.
- In the export media tab click on the box that says video is VR

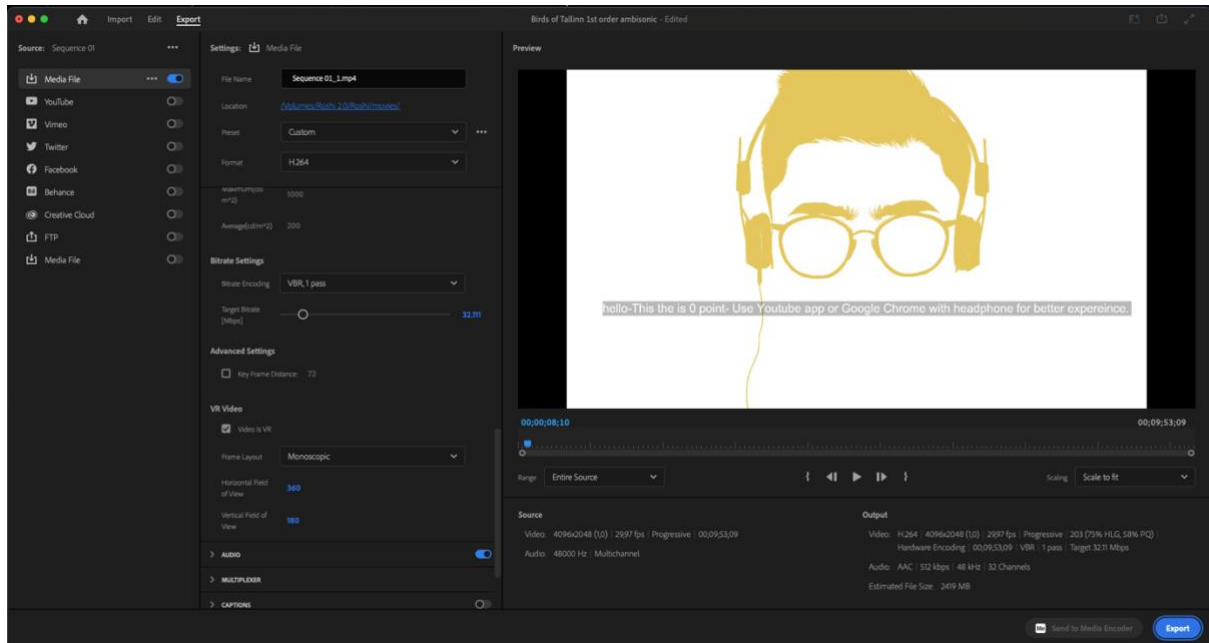


Figure 7 Adobe Premiere- Check the Box with VR (Adobe, 2023)

- Make sure that the rendered audio has 4 channels in it and Ambisonic audio is marked in the audio.

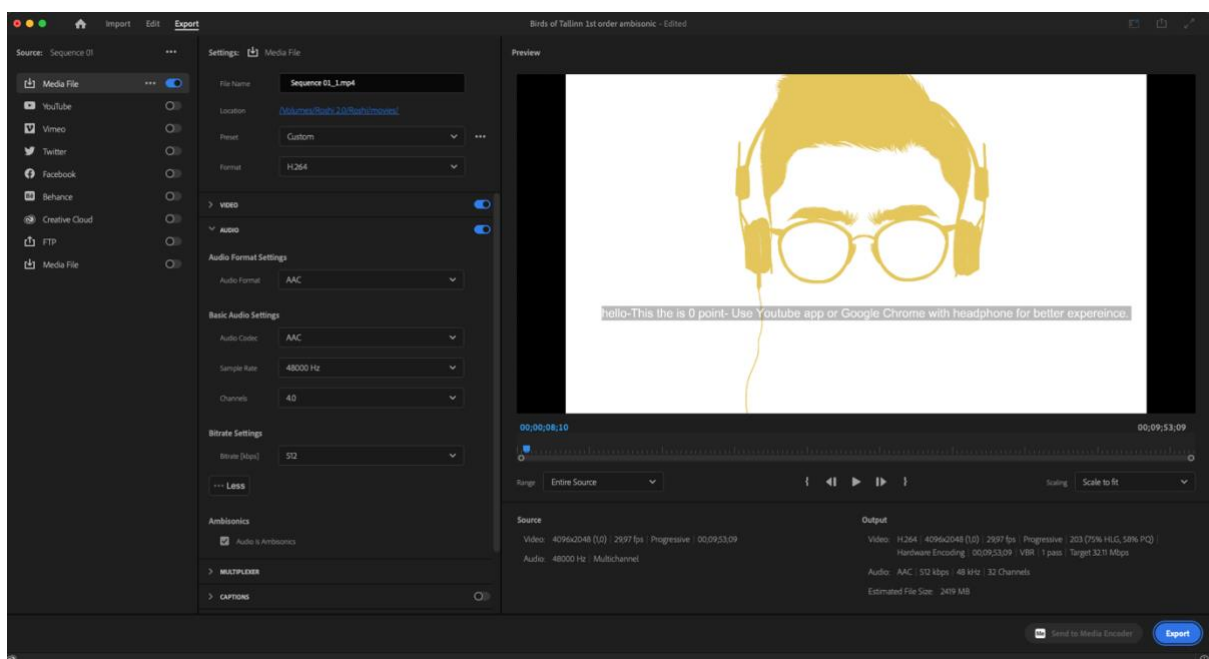


Figure 8 Adobe Premiere- Audio Channel 4.0 and Ambisonics check box (Adobe, 2023)

Then rendered product can be viewed on VLC and you can check the camera movement and see if it's working or not.

The next part is that you should use the spatial media metadata injector application on the video and then make sure you check the boxes.

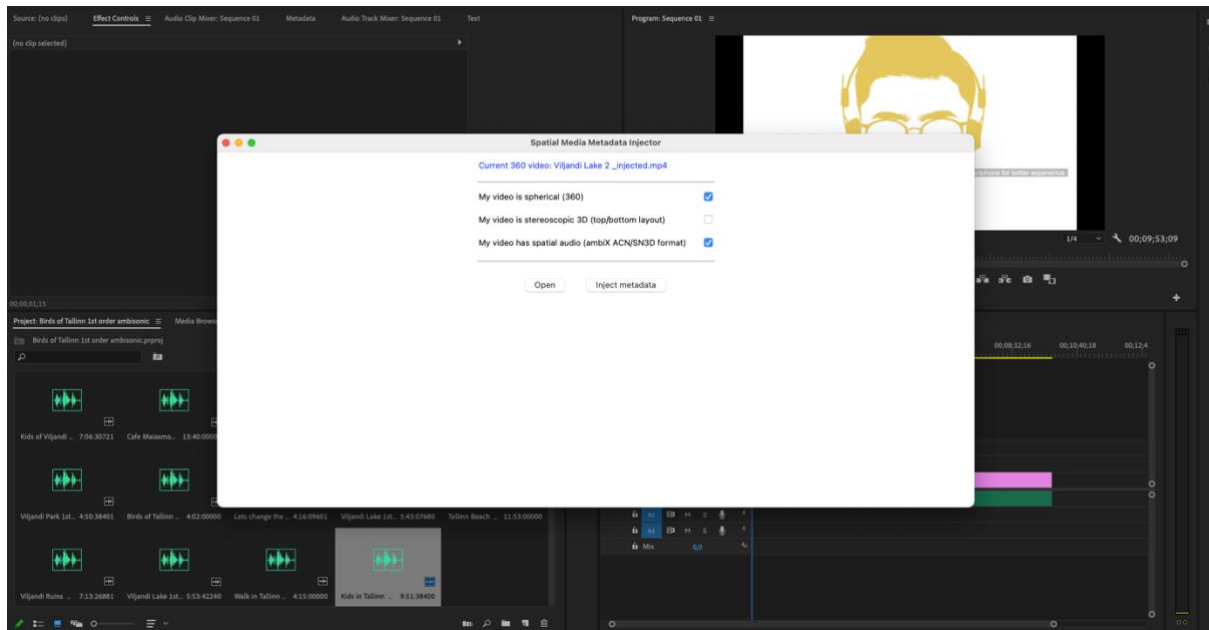


Figure 9 Spatial Media Metadata Injector (Spatial Media Metadata Injector, 2023)

Once done with this your video would be ready to be uploaded on YouTube.

5. Results:

The results of these recordings can't be classified because of different reasons like, microphones' responses to different frequency ranges, the environmental noise, and the pre-amps of the recorder. These are just general issues that I faced in my recording process. To present a scientific approach, the gain level of all 4 channels of the mic must remain the same. Since the audio recorder, Roland Professional AV R-88 field recorder doesn't have a 32-bit audio float recording feature and the audio level had to be readjusted for each location, there is no way to set a basic noise level. Therefore it is decided to approach the result from an aesthetical point of view.

5.1 Analysis:

Overall, after recording different locations in Tallinn and Viljandi, the project results show that different locations have different types of noise factors. Depending on the population of the cities and how the cities are designed with or without consideration to soundscape and impact of noise one, can hear that recording done in Tallinn have more traffic noise even when the recording is done in a park with a densely populated tree or when there are designated locations for kids to play their outdoor games.

It's important to understand how we as humans react to traffic noise, when we are on the road we are always conscious of our surroundings and always focused on the direction of sound to avoid any accident with any passing vehicle. What this does is that it makes sure that we stay focused hence there is more blood flow, which increases the heart rate,

Although Viljandi has a low population in comparison to Tallinn, the kid's playground areas have less public traffic, and even if there is any playground near roads, these roads have one-way traffic which is a really smart solution to reduce traffic noise.

Another important factor that needs to be highlighted is the fact that recording done in Tallinn has less wind noise, although similar recording equipment was used in both cities still the natural ambiance was more audible in Viljandi in comparison to Tallinn, which highlights the fact that most of the noise generated is created human-generated and this noise masks the natural ambiance of the landscape.

Another good example would be the comparison between the Bird sound recordings in Tallinn and Viljandi. It's quite evident that the Tallinn recording has more traffic noise in the background and it's consistent throughout the recording regardless of how dense the tree section of the park is. Whereas in Birds in Viljandi, we can hear that the traffics noise is replaced by the sound of the wind. Though there are vehicles in this recording too still the traffic noise doesn't mask the sound of the birds singing in nature.

If we compare the recording of Kids in Tallinn and Kids in Viljandi, it's easily audible that the playground area in Tallinn for kids has consistent traffic noise, whereas in Viljandi the traffic noise is present but it's not consistent.

We must start working against the traffic noise in our society, fortunately, we humans have grown in multiple directions and we have found solutions to a lot of our problems, still, the element of human health is ignored in almost every sector of life.

We step out of our home to walk in a garden and all we hear is the constant sound of traffic in our ears, to avoid this noise we have decided to use headphone or earphone which solves the noise problem by blocking the sound and increasing the power consumption and climate pollution with every new user.

People born before the 21st century still have a better lifestyle because they had fewer city noise issues but the kids born in the 21st century are forced into a world where they are constantly forced to listen to unnecessary noise. All this extra noise and constant sound vibrating our eardrums is not healthy for our life. Young kids playing on the grounds are exposed to additional noise reduces their listening capacity, although human ears are designed in a way that they adopt noise up to a certain level, if the ears cannot relax it causes ear fatigue, and this results in severe headaches.

6. Discussion:

6.1 Implications of the Findings for environmental policy and Management:

We can do the following things to improve our soundscape.

6.1.2 Protecting natural soundscape: Protecting natural soundscape should be the number one priority in the policy of environmental management. This includes protecting areas with minimum human noise polluted areas. Such as national parks or wildlife areas. The policies should focus on limiting or restricting the number of noise-generating factors.

6.1.3 Managing Noise Pollution: Noise pollution is a significant issue and we cannot ignore it. Protecting nature or wildlife from noise pollution is a good choice but we also have to implement the same policy in the city life. There should be noise measuring units on roads and vehicles with higher noise levels should be banned.

6.1.4 Educating the public: Local community should be made aware of this issue so that people can become a part of the process and change the environment for the better good.

Incorporating soundscape management into urban planning: urban areas can be a significant source of noise pollution however they also present a solution to the problem. The building construction material, the city map, and the facade of the building should help in reducing the noise levels.

6.2 Future research directions:

Documenting and recording natural soundscape in its current form is important. We don't have enough audio records that can show what kind of animals or birds were habitant of certain locations. The impact of man-made noise is still unknown to us and we don't see the actual problems related to it. Some of the important directions that can be highlighted for future research are as follows.

6.2.1 Importance of the soundscape: Humans are constantly progressing in the technological advancement and we are developing new and modern methods of listening to sounds, but what would happen if we would all start wearing headphones and never listen to natural sound on our own? We have started listening to audio recordings of nature through our headphones even when we are out for a walk in a park, or listening to natural ambiance while doing exercise. All of this led to extra consumption of natural resources and this can be easily adjusted by just listening to the sound of nature on its own without any electrical device.

We still need to write and present more documentation on how soundscapes play a role in our daily life. How sound helps in generating memories with loved ones? How do people remember location just by listening to the sound of it? How human brain creates a sketch of each individual just by listening to the sound and even if we don't see the person, we can still identify them just by listening to their sound.

6.2.2 Application of Ambisonics: Ambisonics is still a very complicated process of recording audio. The equipment is quite expensive and the whole audio gear for conducting an ambisonic recording could cost a lot more than a regular recording session. Though there are cheaper recording equipment available in the market but the audio quality is not up to the mark.

6.2.3 Equipment cost: We have to find a solution to make the equipment cheap in cost while maintaining the audio quality. Maybe introduce newer and better microphones, better recorders. The post-production process needs to be simpler so that more users can use it and implement it in their daily use.

6.2.4 Processing Ambisonics: Although several DAW can now easily read and bounce ambisonic tracks still the routing process takes a lot of time and effort.

6.3 Limitations & Suggestions:

6.3.1 Field Recording Device:

It is absolutely important to make sure that field recorded has 32-bit float recording in it. This helps, given the fact that all the recordings are conducted outdoors and locations where the general public would constantly visit so the noise level can change randomly, there are even situations when some drunk people would come and shout in the recording or there is suddenly an ambulance that passes by, the fact that all these factors are not controlled and they can happen randomly so it's not possible to have to change the gain level in each recording. With 32-bit float you don't need to worry about sound blast issues, this new method of recording records audio at higher resolution and you can avoid gain clipping errors. This feature is available in most of the new field recording devices, these devices can range from 500 US \$ to 2000 US \$ and more. Most commonly used are Zoom F8n pro, Tascam Portacapture X8, or from sound device MixPre-6 ii.

Another very important feature that is required for doing Ambisonic recording is linking each of the mic tracks. It doesn't matter what microphone you use, linking tracks is important and compulsory.

Another important feature required in the field recorder is the direct conversion of Ambisonic A-format to B-format, most of the new recorders have this feature. While Ambisonics field recording if you don't have a built-in converter for ambisonic then it gets difficult to judge the direction of the sound. Although this can be avoided by careful consideration during the setup and maybe marking the mic stand if your recorder has a built-in converter then you know how the sound be presented in the end.

Having a 32-bit float recording and gain link for the 4 channels audio and the A to B-format Ambisonics converter solves almost all the possible issues that you can face.

Apart from this if the recording is taking place outdoors then it's absolutely important to make sure that your mic is covered with windshield, this helps in avoiding unnecessary wind noise, however, you would have to consider the gain level with the windshield.

6.3.2 Post-Production:

Almost all the DAW support Ambisonics tracks, but the process of making it work is different for each DAW, even the routing system is different so it's very important to consider which DAW you would be using for this purpose. A general process would be to create a 4-mono channel (depending on the type of Ambisonic microphone you are using) routed into a bus or group channel that contains your Ambisonics A to B format convertor plugin and then you send this group tracks into your main out which can 1st, 2nd or 3rd Ambisonics.

Once you are done with rendering your ambisonic track in the desired order, you have to find a solution to upload it online. Very few portals music streaming portals support ambisonic tracks and Sound cloud only supports 1st order Ambisonics.

For video streaming portals like YouTube, Facebook, and Vimeo, you have to attach your audio with a video. The video can contain any image or visual, then you need to render it through premier pro or Davinci Resolve Pro with 4 channels out and make sure to check on the check box that says that the audio is ambisonic. Then you add this video to another software like FB360 or Spatial media metadata Injector.

This is a very long process, most of the music producers or mixing and mastering engineer are not even familiar with this process and some of the software doesn't work anymore. You will not find any tutorial on how to upload the Ambisonics track on YouTube. You would only find VR video tutorial which discusses ambisonic audio implementation in the video and all of these videos are at least 3 to 4 years old, which shows that there is less use of this technology.

7. Hyper-Link to the album:

YouTube Link

<https://youtube.com/playlist?list=PL4580s3ZxSQ3eZHBuMOIRIgSfMwfRTWA8>

SoundCloud Link

https://soundcloud.com/qumarosh/sets/soundscape-of-estonia/s-s7BTsSaDyFA?si=56150a7bd56945998b70d6c76e12236c&utm_source=clipboard&utm_medium=text&utm_campaign=social_sharing

References:

- Jesse R. Barber, K. R. (March 2009). The costs of chronic noise exposure for terrestrial organisms. *Trends in Ecology & Evolution*, 190-189.
- Maria Klatte*, K. B. (2013, August 30). *Does noise affect learning? A short review on noise effects on cognitive performance in children*. From [Frontiers.org: https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00578/full](https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00578/full)
- Hygge, S. B. (2003). The effects of road traffic noise on reading and listening comprehension in school. *Noise & Health*, 71-82.
- Sennheiser. (2023, March 12). *Ambeo VR Mic*. From <https://en-de.sennheiser.com/microphone-3d-audio-ambeo-vr-mic>
- Furness, R. K. (1990, May 1). *AES- E Library- Ambisonics an Overview*. From Audio Engineering society: <http://www.aes.org/e-lib/browse.cfm?elib=5417>
- Matthias, Z. F. (2019). *Ambisonics- A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*. Cham: Springer Nature.
- Waves-blog. (2017, April 12). *Ambisonics Explained: A Guide for Sound Engineers*. From [waves.com: https://www.waves.com/ambisonics-explained-guide-for-sound-engineers](https://www.waves.com/ambisonics-explained-guide-for-sound-engineers)
- V Pathak, B. T. (2008). Evaluation of traffic noise pollution and attitudes of exposed individuals in working place. In B. T. V Pathak, *Atmospheric Enviroment* (pp. 3892-3898). Varanasi: Banaras Hindu University.
- GmbH, S. M. (2023, April 12). *Steinberg-Nuendo*. From <https://www.steinberg.net/nuendo/>
- Adobe. (2023, April 20). *Adobe Premiere Pro*. From [Adobe: https://www.adobe.com/products/premiere.html](https://www.adobe.com/products/premiere.html)
- Spatial Media Metadata Injector*. (2023, April 25). From [Github: https://github.com/google/spatial-media/releases](https://github.com/google/spatial-media/releases)

Table of Figures

Figure 1- Ambeo A-B Format Converter (Sennheiser, 2023)	16
Figure 2- Nuendo 11 Channel Layout and Panning (GmbH, 2023)	17
Figure 3- Nuendo 11 Project Layout (GmbH, 2023) (Adobe, 2023).....	17
Figure 4- Adobe Premiere Multichannel Mono Media (Adobe, 2023)	18
Figure 5 Adobe Premiere modify-Audio Channels (Adobe, 2023).....	19
Figure 6 Adobe Premiere - Activate channel per clip (Adobe, 2023)	19
Figure 7 Adobe Premiere- Check the Box with VR (Adobe, 2023).....	20
Figure 8 Adobe Premiere- Audio Channel 4.0 and Ambisonics check box (Adobe, 2023)....	20
Figure 9 Spatial Media Metadata Injector (Spatial Media Metadata Injector, 2023)	21

Non-exclusive license to reproduce thesis and make thesis public

I, Qumrosh Abbas Khan,

1. herewith grant the University of Tartu a free permit (non-exclusive license) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

Puuluup Immersive,

supervised by Christoph Felix Schulz.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons license CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.
3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.
4. I certify that granting the non-exclusive license does not infringe on other persons' intellectual property rights or rights arising from the personal data protection legislation.

Qumrosh Abbas Khan

10/05/2023