ELE00120M VASA Report

An Exploration of Computerised Acoustic Modelling

E3853335

April 21, 2021

Introduction

Acoustic modelling is a process involving complex computational algorithms in the aim to improve or explore a given modelled space. Often this is to discover problem areas before a retrofit or build of a musical venue[1]. concert halls and cathedrals are common explorations using acoustic modelling however, this is not where the applications are limited. Within this report the effectiveness of these techniques are explored when analysing the difference between the real space and modelled when the model is given constraints such as a low poly count.

Background

Acoustic modelling is commonly achieved through two different methods, calculating the parameters via solving the wave equation or by using geometrical acoustics[1]. The later is often used due to its lower computational cost and quicker results. In this case the sound waves are cast from the source as rays much like the way which light is cast from an emitter.

> "This assumption is valid at high frequencies, where the wavelength of sound is short compared to surface dimensions and the overall dimension of the space, but at lower frequencies the approximation errors increase as wave phenomena play a larger role." [1]

The first reported instance of an acoustic response with multiple receiver positions was by Krokstad in 1968 [2].

The primary objective of geometric room acoustics is to output a time-energy or impulse response(IR) which can then be used to calculate acoustic parameters or create auralisations. A term which is defined by:

> The process of rendering audible, by physical or mathematical modelling, the sound field of a source in a space, in such a way as to simulate the binaural listening experience at a given position in the modelled space."[3]

The program which allows this capability using ray/cone tracing with auralisation and IR output is CATT-A9[4]. This program sits along side other industry standards such as ODEON [5] and EASE [6], although these programs differ in their approach to modelling they all follow a similar Geometric Acoustic utilising ray tracing technologies.

Reflection and Absorption

The topic of pure reflection on a infinite, perfectly flat wall of an incoming signal acts much like a laser would, where the angle of incidence to the normal is equal to the angle of the reflector to the normal.

$$\theta_i = \theta_r$$

[1]

This is an example of a 'specular' reflection, one which appears visually as metallic or mirror like. However, this does not represent the vast majority of surfaces in the real world. Depending on the surface, not all the energy received will be reflected, some may be absorbed. In most cases some of this energy is transferred from kinetic energy of the particles in the air to heat and most is transmitted through the surface. This absorption in CATT-A9 is defined by what percentage of the energy will be absorbed during interaction with that surface. In addition to most surfaces not being perfectly reflective, surfaces are also not perfectly flat either.

Scattering

The vast majority of surfaces in everyday environments include some element of displacement, be it a rough texture like dirt or smooth bumps such as those found on painted brick and concrete. This disturbance in a material, when are inhabiting a similar order to that of the wavelength of the sound can cause audible deviations from an otherwise 'flat' wall[1]. The application CATT-A9 allows the user to specify the probability that a ray would turn from a specular reflection to a diffuse reflection where the ray cast changes direction. This is done as a percentage chance of the ray being changed, although this is not what happens in reality it is a close enough approximation if enough rays are cast from the source. In the real world a proportion of that sound is diffused and there remains a component which is then reflected specularly, however, computationally this would require every ray to double each time it hits a surface. If one was casting 10 rays out of a source, after only five interactions with surfaces there would be atleast 320 different rays to calculate. This would make it very difficult to complete a calculation in a reasonable length of time. Due to this the probability changes that ray from specular to diffuse to save on computation time/energy.

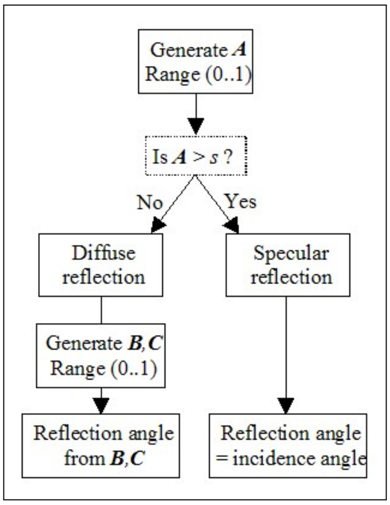


Figure 1: A flow chart to demonstrate the scattering process[7].

Diffusion

The process of a ray diverging from its reflector angle is called scattering, the width and uniformity of this scattering is labelled as diffusion, in this process the reflection of the wave is such that it does not follow its intended path. This process is often done by random or according to a distribution function, CATT-A9 does not model the individual material's diffusion parameters but instead relies on Lambert diffusion.[7] This process takes the Cosine function and makes energy it a function of the cosine of the angle of incident. This means that the most energy would be present at $\theta_d = \theta_i$ where θ_d is the angle of diffusion and the energy would fall off approaching $\theta_d = 180°$.

This method has been used in computer graphics since early ray tracing examples.[1] However, there is debate as to whether this diffusion system should be used for acoustic simulations. The concerns surrounding the appropriateness of this system originate from the wavelengths that this model targets. Due to the visual medium which this system has been predominantly developed, its primary wavelength range is far lower, and thus its frequency higher, than what humans are able to perceive auditorily. This also means that the interactions with the surface can be fundamentally different due to the far higher wavelength, leading to a distinct low end inaccuracy. [1]

3D Modelling

The process of choosing and modelling a space came from the aim to compare and analyse the differences between and real and modelled space in order to evaluate the effectiveness of the modelling software. Due to the author's personal experience recording and attending concerts in the Jack Lyons Concert Hall (JLCH) it was chosen to be modelled. When analysis the concert hall it appears to be a hexagon with different lengths on parallel faces. This results in three short walls and three long walls. Despite the simple shape it holds complexity with the seating arrangement, eight rows slanted up by 10-15cm each row and holding 330 people when fully booked. Another interesting aspect of the space is the presence of an organ, fitted in 1969 by Grant, Degens and Bradbeer and overhauled by J W Walker and Sons in 1983, the organ features a maximum pipe length of 16ft and 56 notes.

The program used to model was Google Sketchup 2016 [8] as it is highly accessible due to the low barrier for entry and highly customisable and complex. An advantage is the interaction between the physical modelling program and the acoustic modelling application CATT-Acoustic V9 as a trail version of SU2CATT or SketchUp to CATT is available for up to 50 faces of a model.[9] This means that it is possible to create a simplistic room and add small details allowing for a semi accurate depiction of the room itself. The process of referencing photographs and estimating the dimensions from personal experience allowed creative freedom in Google Sketchup. It was chosen to start with the main stage area, the floor is hexagonal but due to the angle of the stairs has two extra faces in addition to the cut-out at the back for the organ.

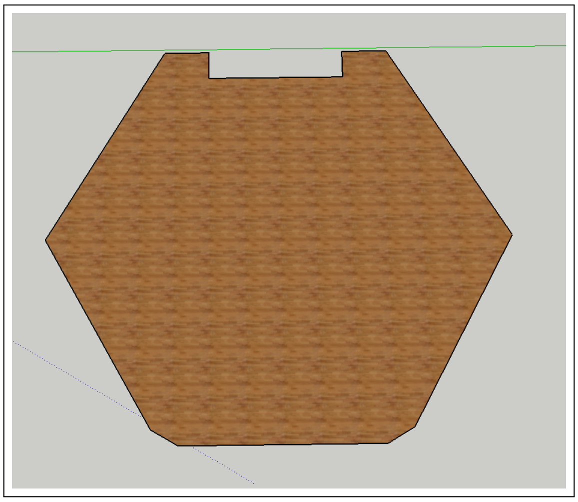


Figure 2: The floor space for the JLCH.

Next was the walls, when referencing photos the concrete wall only extended half of the length of the wall where a wooden cover was present. There is also an acoustic diffuser part way up the concrete wall, though physically modelling this surface would eat into the 50 face count it is possible to assign a different material to it and later assign diffusion to that material in CATT Acoustic. In order for the absorber panel to be set inside of the wall each face was required to be connected to the top and bottom face, creating a third, deducting one face for each side from the limit of 50 faces.

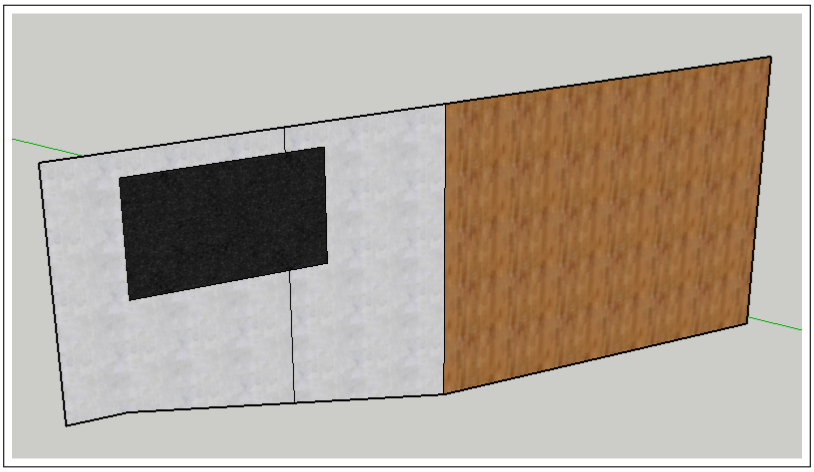


Figure 3: One of the walls including the acoustic diffusion panel (black) and the wooden cover.

The primary acoustic absorber in this room is the audience and the seating. It was decided to provide more than a surface to act as an absorbent face, instead a block would absorb from all sides mimicking the three dimensional nature of seating and the height that an audience would sit at. This was also done due to the examples on the official website which would place the audience plane elevated in a chunk, much like what is shown here. There was a simplification from stairs to ramps for the walkways, considerations were made regarding the 50 face count and where they would best be used. Different materials were used for the carpeted ground to the stage which was wooden.

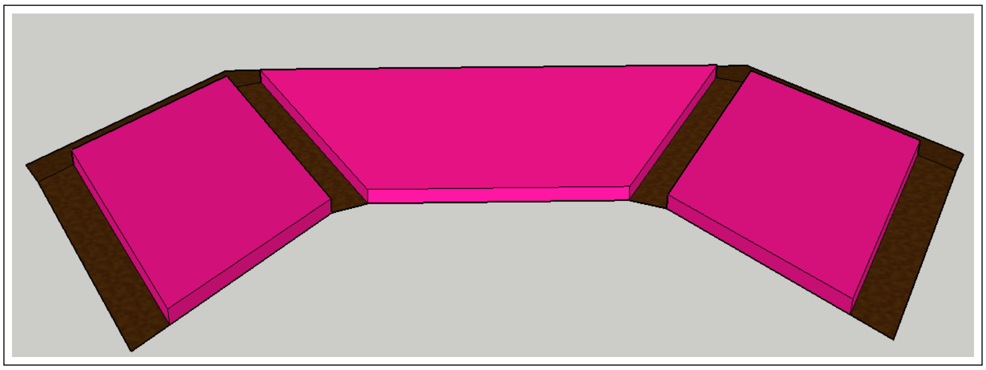


Figure 4: The seating format for JLCH.

The roof was the last component to the space. Initially it was chosen to simplify the roof into a single slanted face due to initial intricacies modelling the open pit and AV booth. However, a more complex roof was opted for once faces exceeded 50 in other cases. This roof was estimated at 45°angle with a 2m instep. This provided a smaller irregular hexagon to join faces to. There were always some faces which would not draw correctly regardless of how many attempts were made. This meant that the roof would require more faces than intended. However, considerations were made to the rest of the room in order to allow for these discrepancies including removing the small orchestra pit which had previously been modelled.

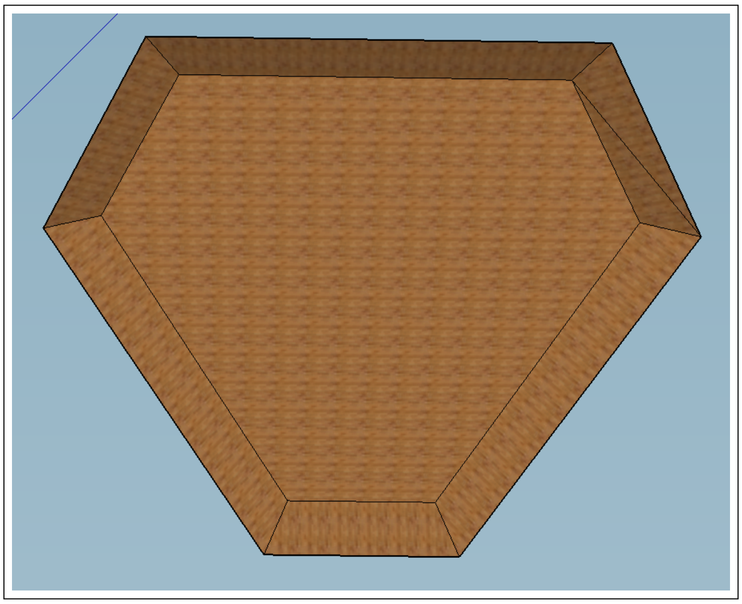


Figure 5: The roof of the JLCH.

The final outcome of the modelling process was a simplified but an arguably faithful reconstruction of the intended space. The primary walls and inclusion of the organ provide a good space for sound to be reflected off of. The seating allows for a highly absorbent surface meaning less reflection from the back of the room. The complexity could have been increased if given a larger number of faces to work with. However, it could also be argued that the acoustic modelling software would not benefit from these extra details as it would lead to increased time of render and possibly diminishing returns.

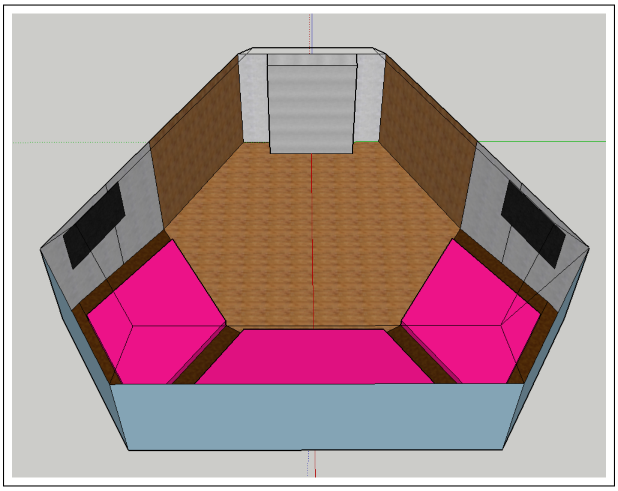
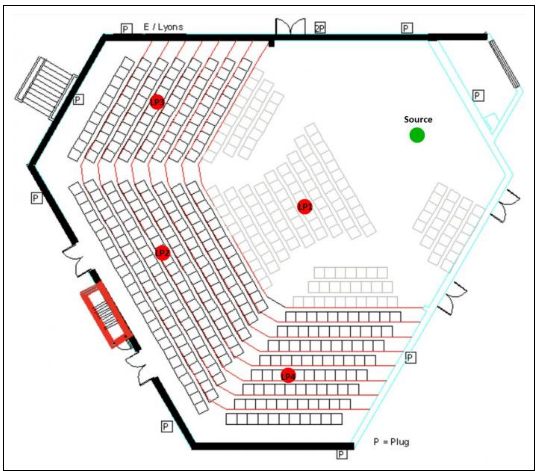


Figure 6: The model used during the acoustic modelling process with the roof faces hidden for clarity.

Finally to finish off the model and check scale an extension called VR Sketch and an Oculus Quest 2 [10] was used to look around the space in Virtual Reality and check scale factor for how big the space really should be.

CATT Acoustic

The process of exciting the room followed a similar design to a real IR captured in the room. It was chosen to mirror similar position of the real life IR with one placed in front of the source and one placed in each of the three audience stalls.

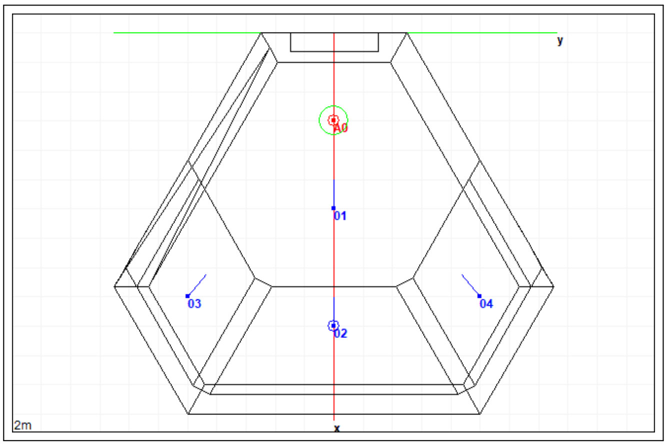



```
RECEIVERS
 1 12 0 1.7
 2 20 0 2
 3 18 -10 2
 4 18 10 2
```

Figure 7: The microphone positions placed for the real IR which the modelled will be compared against[11].

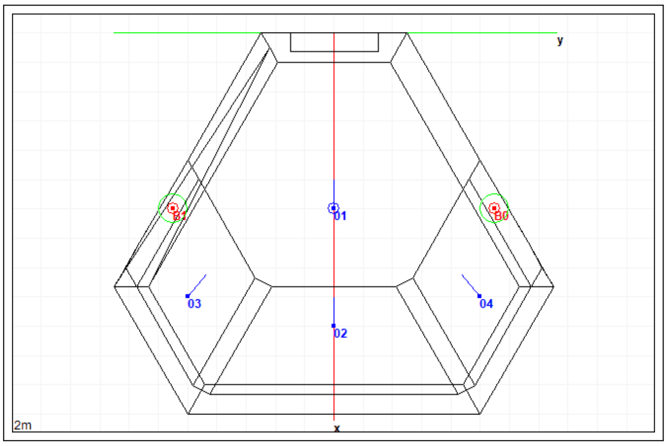
Tests

Each test conducted followed a scenario which could be experienced at this venue, a single point source at the back of the stage and a stereo point source from the speakers currently placed on each of the long walls.




```
a0  6 0 1.7    OMNI      0 0 1.7
Lp1m_a    = Lp_white 94 ;
white spectrum, 94 dB at 1kHz
```

Figure 8: Scenario 1.




```
b0  12 11 3    OMNI      0 0 1.7
Lp1m_a    = Lp_white 94 ;
white spectrum, 94 dB at 1kHz
b1  12 -11 3    OMNI      0 0 1.7
Lp1m_a    = Lp_white 94 ;
white spectrum, 94 dB at 1kHz
```

Figure 9: Scenario 2.

Absorbtion and Scattering

Data used in the ABS Defs file were obtained from the inbuilt coefficients after referencing several other sources. Starting by taking the options which seemed most relevant. Absorption is how much energy from each ray is destroyed in the software, thus high values create a dryer sounding environment.

CATT-A9 defaults to using 10 percent scatter for each surface, although this is useful for plain rooms, upon reading the CATT manual it recommends different coefficients for an audience surface along with any surfaces which you presume to have higher scattering capabilities. An interesting feature of the Jack Lyons Concert Hall is the use of acoustic diffusers on each side of the audience. These diffusers are in the format of Maximum Length Sequence meaning that they do not absorb any sound but instead scatter it around the room. Due to the lack of information regarding the measurements of these diffusers it was estimated. The MLS diffusers work to half the length of the wavelength of the frequency, this means that anything below roughly 500hz would not be affected by this panel as the wavelength. As half of the wavelength fits a rough estimation of the size, although this is highly speculative, if given the proper measurements it would be a good aspect to model.

```
ABS Metal_PERF_Organ = <76 76 90 99 85 70> L <10 10 10 10 10 10>
ABS AcousticPannel = <4.0 4.0 7.0 6.0 6.0 7.0> L <40 40 40 90 90 90>
ABS AUDIENCE51_Seeting = <44 60 77 89 82 70> L <40 40 50 50 60 70>
ABS CARPET_5MM_Carpet = <2.0 3.0 5.0 10 30 50> L <10 10 10 10 10 10>
ABS BRICK_WALL3_ConcreteWalls = <1.0 2.0 2.0 2.0 3.0 3.0> L <5 5 5 5 5 5>
ABS WOOD_FLOOR1_WoodFloor = <4.0 4.0 7.0 6.0 6.0 7.0> L <10 10 10 10 10 10>
ABS no_name = <10 10 10 10 10 10>
```

Test Settings

Each receiver was set to point towards the first source at (6, 0, 1.7) mimicking the normal direction when seated. The audience was mapped to the top sections of each plane much like the manual demonstrates.

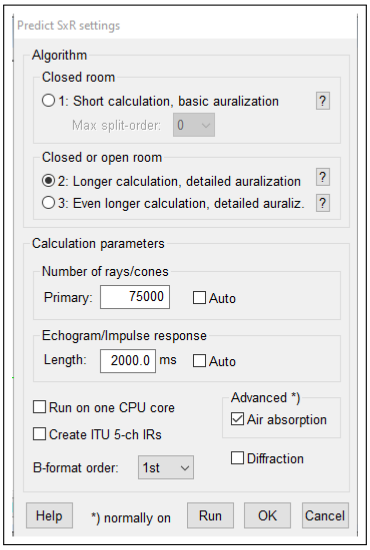


Figure 10: The settings panel for SxR predictions

The decision to chose the longer calculation with detailed auralisation meant that a longer render time would result in better results and a more accurate model of the space. Upon listening a comparing a significant amount of 'graining' was produced when using the 'Short Calculation' algorithm so the 2nd 'Longer Calculation' algorithm was chosen. It was speculated whether choosing the 3rd option was appropriate, however, it was deemed to be beyond the point of diminishing returns for this purpose. This choice would have been relevant if the space was modelled more accurately and with more than 50 faces. It was chosen that 75,000 rays was sufficient to create a highly detailed auralization due to the recommended balance between time and accuracy was roughly 56,000 rays. Due to the time of the render not being a limiting factor, it was chosen to calculate about one and a half that amount.

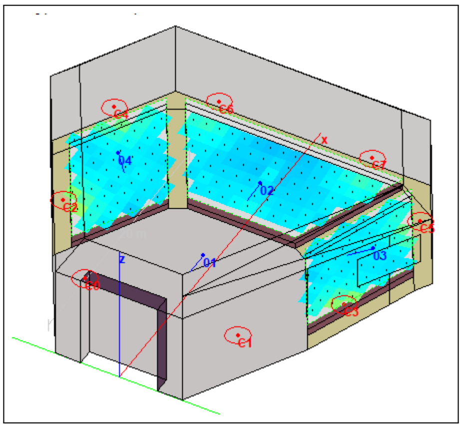


Figure 11: An example of audience mapping showing results

Although audience mapping is not a consideration in the intended research it remains within the expanded scope looking into concert hall design. CATT-A9 allows for the ability to map points within your audience to measure several aspects of the sound.

Critical Listening

For analytic and critical listening the Matthias Kronlachner - Ambisonic Suite was used inside of Reaper. After exporting individual B-Format channels, Stereo Binaural and Stereo Blumlein IRs from TUCT they were then brought into Reaper and converted to 4 channel B-Format. The ordering and normalisation output from TUCT is Furse-Malham which does not work with the Ambix binaural decoder. This is important as the critical listening with a different channel set could change the orientation that the listener would perceive the impulse. As an example, listening position 3 and 4 both are placed next to large walls. The reflection of these walls are audible and the precense of the wall can be felt, however, with the wrong channel ordering this could result in the wall being felt infront or behind the listener. The Furse-Malham is converted to ACN formating which shifts the traditional W-X-Y-Z ordering to fit with the Ambix Binaural Decoder.

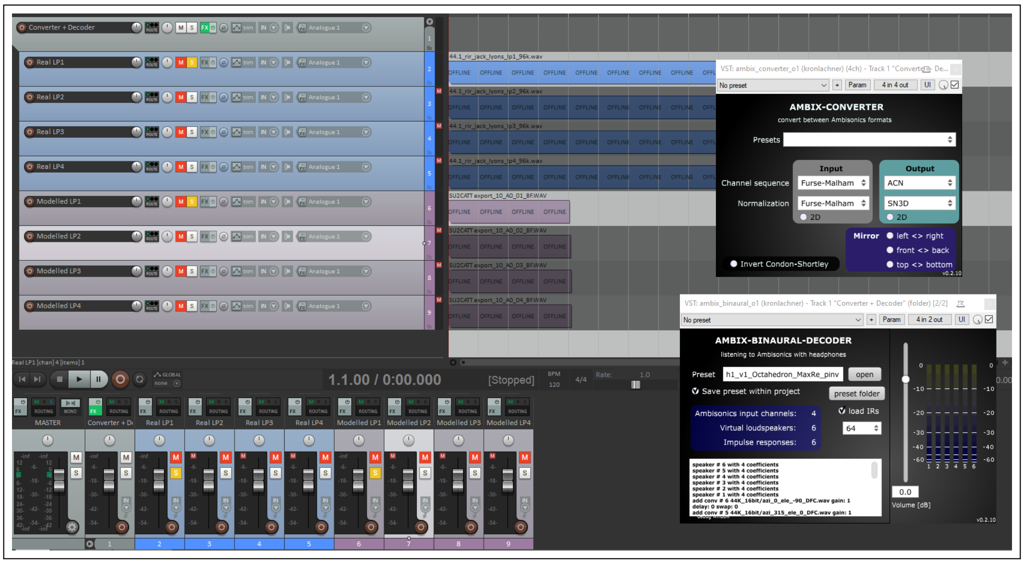


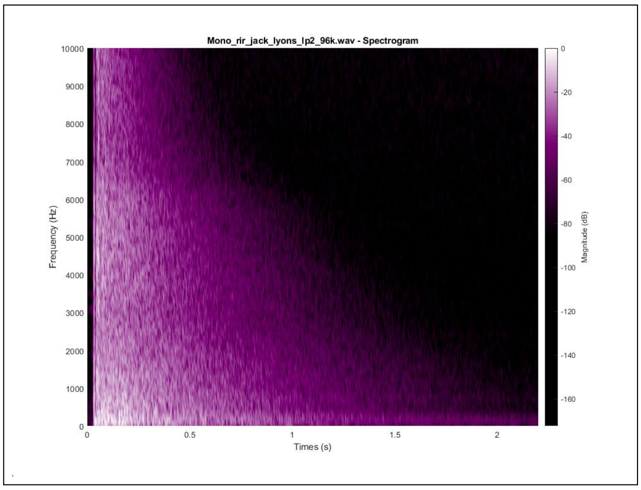
Figure 12: The project file which listening was done

Results Discussion

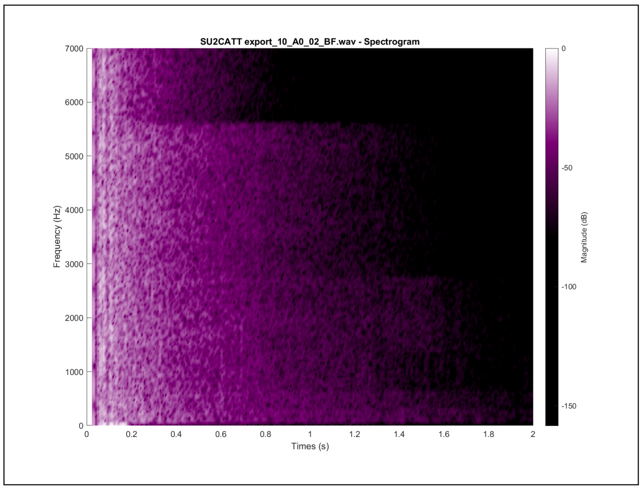
Impulse Responses

All results and files labelled as "Real" are created using the IRs from openair.hosted.york.ac.uk by Alexander Duffell, Zhong Li and, Aishwarya Sridhar. These are redistributed under Creative Commons Attribution 4.0.[11]

Test 1



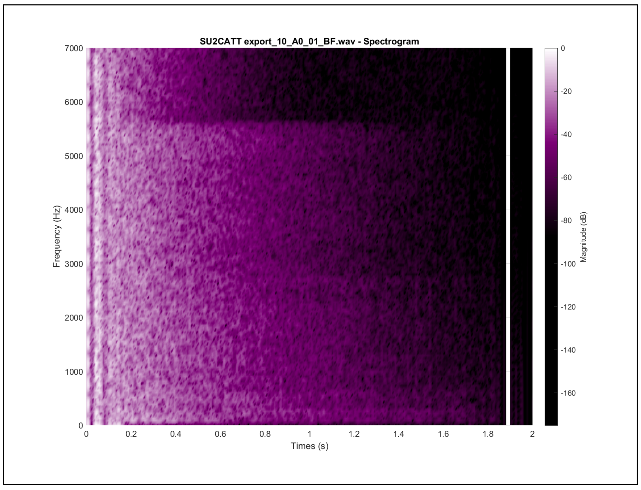
(a) IR_Real_LP2



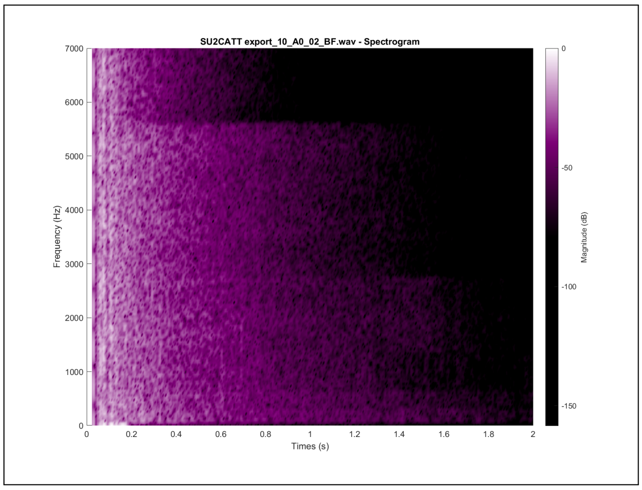
(b) IR_Test2_LP2

Figure 13: A comparison between modelled and real spectrograms

A notable discovery is the distinct frequency banding which occurs in all the modelled IRs. The modelled IR (15b) appears to have a clear bands at roughly 5600hz, 2900hz and 700hz. The real IR (15a) can be seen to have a higher concentration of low frequency during the length of the impulse in comparison to the modelled IR which only lasts 1.8 seconds in comparison to the real's over 2 seconds. It can also be seen that the decay rate is quicker in the real IR compared to the modelled one. This leads to a higher concentration of lower frequencies during auralisation, resulting in a muffled tone with speech and a soothed tone with musical instruments. As the purpose of the hall remains primarily for music this follows what is expected.



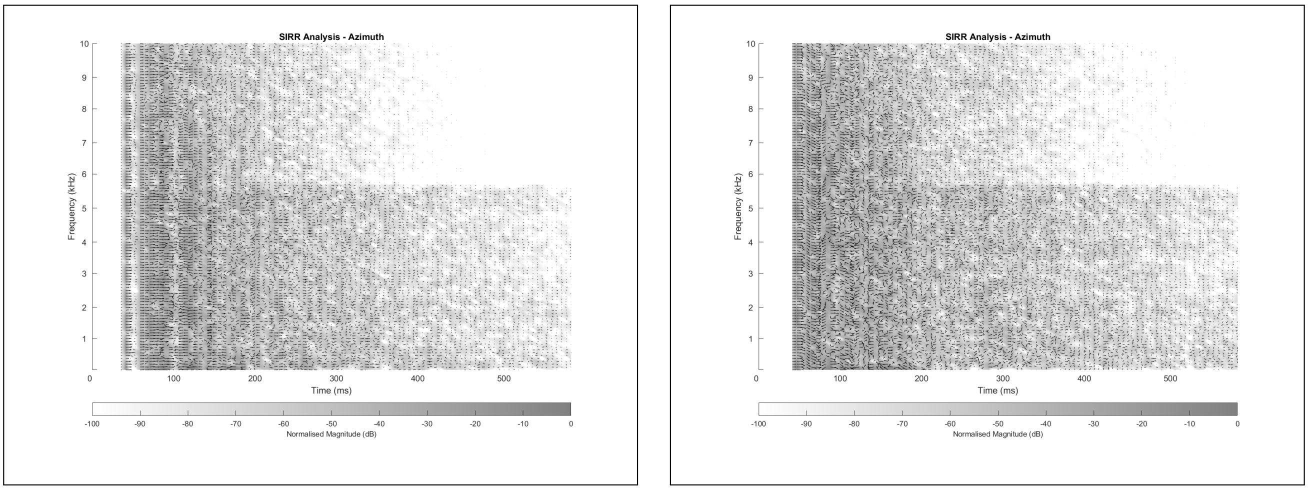
(a) IR_Test1_LP1



(b) IR_Test1_LP2

Figure 14: A spectrogram comparison between two modelled receivers

The comparison of the decay on these two examples shows that the closer the receiver is to the source, not only the louder the initial spike is but also the longer the decay time due to a more central position. This means that the reflected sound waves are more likely to bounce against less absorbent and less diffuse surfaces. Another important distinction is that Listening Position 2 is placed in the audience, among a very absorbent and highly diffuse setting. The magnitude of this change is not apparent until the scales are considered. 15(a) is using a less condensed scale for intensity of energy, this means that the differences are even more severe in 15(b) where the intensity is far more condensed.



(a) IR_Test1_LP2

(b) IR_Test1_LP4

Figure 15: An SIRR comparison between two modelled receivers

The SIRR analysis allows a B-Format signal to be analysed and allows a user to localise the space that they are in. In this case 16(a) the initial impulse at 40ms is the initial wave from the source as calculated by a 14m distance, the next wave roughly at 75ms is the source reflecting off the wall in front of the receiver and travelling back. This then gets more complicated as the reflection from that off the back wall causes an impulse from behind the direction of the listener at 110ms. Another distinction to make is the lack of directivity about 7800hz, this appears to be the result of scattering appearing more in higher frequencies, as indicated by the higher scattering probability in the ABS DEFs. This is in stark contrast to the complexity of 15(b) at all frequencies where after the initial impulse at ≈40ms there is a complex web of directions and scattering. A clear direction which is interpreted is one to the right of the listener between ≈55ms to ≈65ms, this is a reflection from the right hand wall to the listener. The position of the receiver shows that the right hand wall is close enough to only make a small time deviation from the initial response.

Due to the more complex placement of the receiver the reflection impulses are less defined. In 15(a) there are identifiable reflections at ≈55ms, 110ms, 140ms, 200ms and 270ms. This contrasts to the more diffuse and dense 15(b) where the only identifiable impulses are at ≈40ms and 60ms

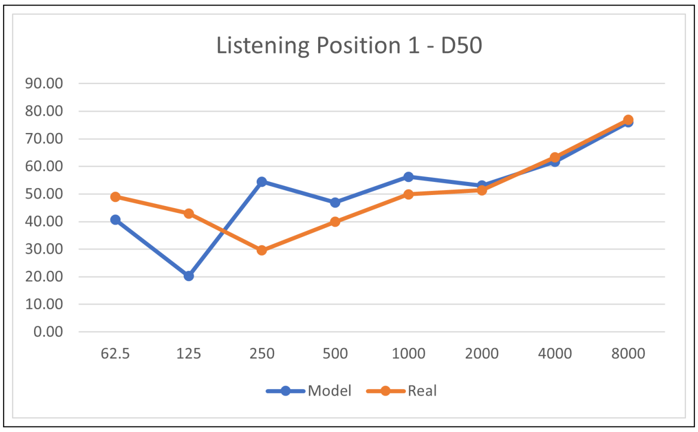
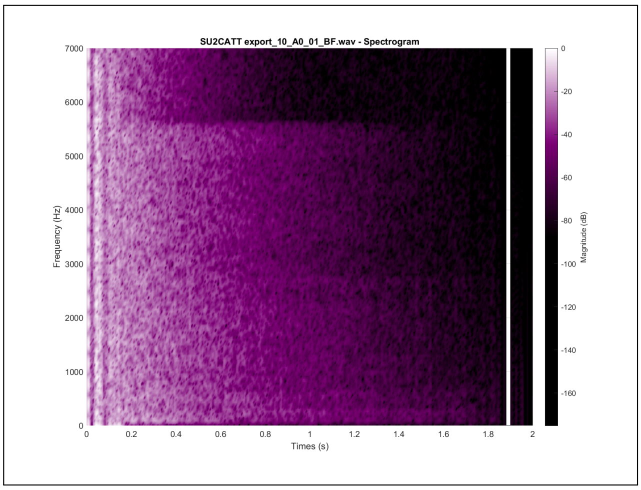


Figure 16: A line graph to show the differences in definition in a modelled acoustic space, data taken from Table 1,2

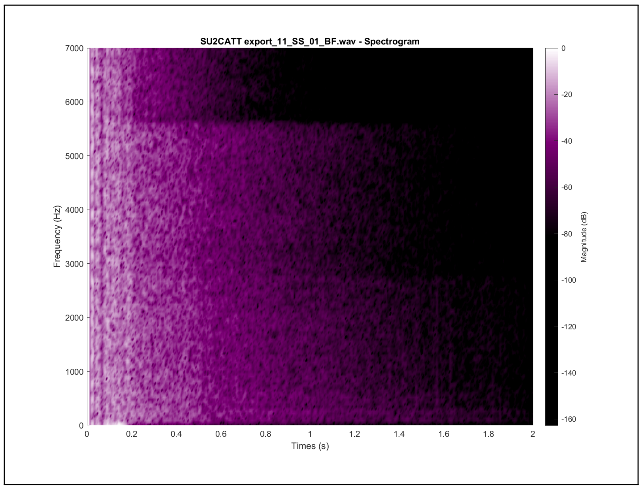
Looking to the acoustic parameters it is identified that the measurements only start to look similar at higher frequencies, this is possibly a result of the ray-tracing and Lambert diffusion techniques. Looking at listening position number 3 we can see that the results converge as the limits of the test are reached, towards 8000hz the results are closer than at the lower frequencies. This trend holds true in most acoustic parameters.

Test 2

Taking a look at the 2nd test conducted with two speakers, we can see a far more diffuse scenario where reflections are more difficult to resolve.



(a) IR_Test1_LP1



(b) IR_Test2_LP1

Figure 17: A spectrogram analysis of a single point source versus two point sources

This shows the sources interacting with each other and causing a richer more diffuse sound oppose to the first test. This may also be due to the source's proximity to the wall. The tests were kept as point sources emitting in all directions, this is not how speakers would emit straight to the audience in a cone shape. As a result more of the audio is given the opportunity to become diffuse reflection from the wall due to its proximity. One of the downsides to multiple sources is the interaction between the two sources, due to the use of the speakers in this case is vocal re-enforcement or announcements it is chosen to compare the D50 parameter.

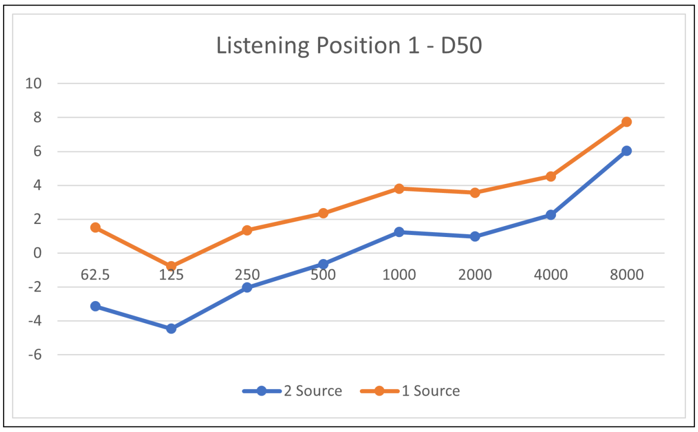
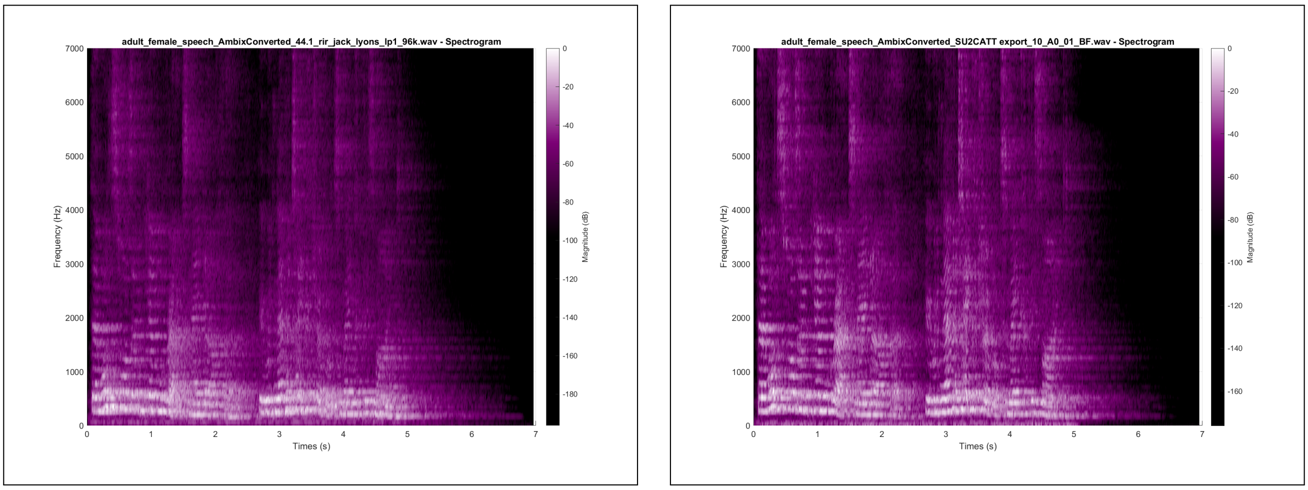


Figure 18: A line graph compare the definition of one source against two sources in a modelled space, data taken from Table 2,3.

Auralisations

Looking towards auralisations is a method which allows us to cast a different context and discover traits which would otherwise go unnoticed. The auralisations used here are from the "Acoustics and Psychoacoustics" book where it supplies a CD with anechoic material to use. This was chosen due to the reputation of the book and the accessibility to measure against other IRs given the opportunity[12].

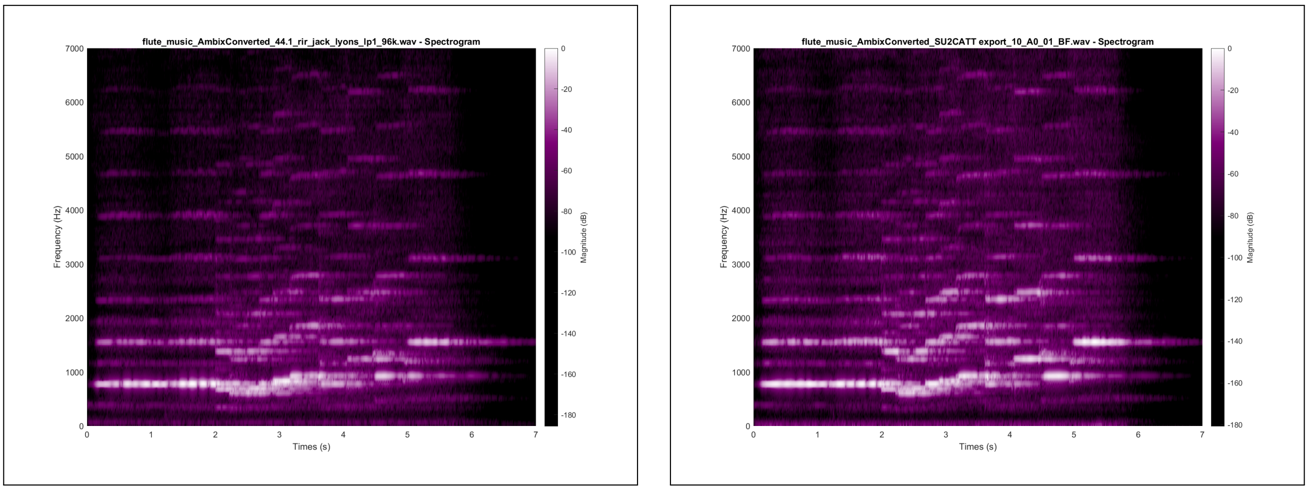
Test 1



(a) adult_female_speech_Real_LP1 (b) adult_female_speech_Test1_LP1

Figure 19: A spectrogram comparison between a modelled and real receiver in the application of speech

In the case of listening position one, it becomes clear after listening to the auralisations that the high frequency content is very different, this gives the speech a muffled tone and makes it hard to distinguish the intelligibility of speech. Another noticeable feature of this is the more pronounced reflections in the modelled result, this cause the speech to become masked by previous parts. The application of the subjectively 'smoother' sounding real IR would be better suited for music, however the modelled IR is better at speech but also the reflections could be useful in a musical application, depending on genre and intended effect.

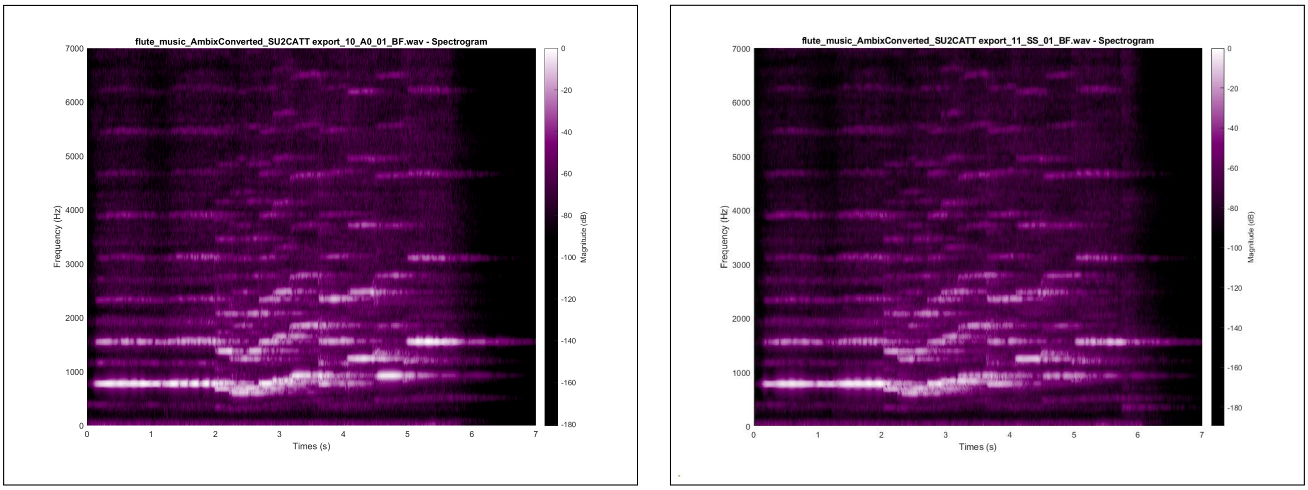


(a) flute_music_Real_LP1 (b) flute_music_Test1_LP1

Figure 20: A spectrogram comparison between a modelled and real receiver in the application of music

For this example the real impulse response obtains a subjectively better result out of the flute sample, although the extended delays of the modelled impulse allows for a more rhythmic and dynamic articulation, the brightness and high frequency content results in the flute sounding piercing and grating.

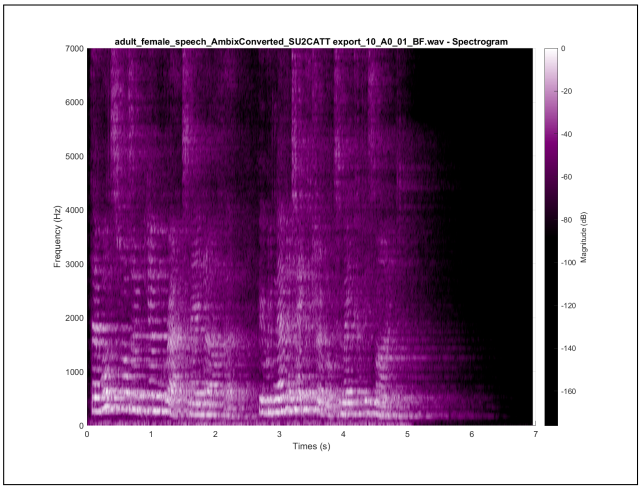
Test 2



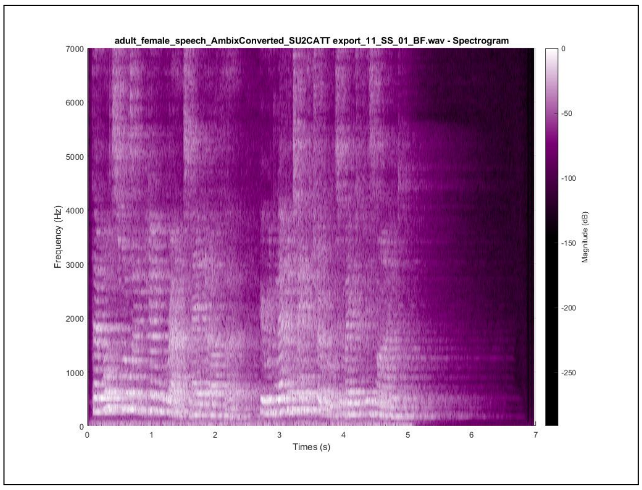
(a) flute_music_Test1_LP1 (b) flute_music_Test2_LP1

Figure 21: A spectrogram comparison between the two tests conducted in the application of music

The flute does not benefit much from the reinforcement given by two sources, a noticeable difference is the width of the perceived signal. Depending on the signal and piece this may work as an advantage or to the detriment of the artistic intent. Due to the sources being further back and directly to each side of the listener, the width is increased and the reverberation and clarity is replaced with more diffusion. This more diffuse sound makes the content smoother with its attacks allowing for a more pleasurable listening experience depending on the desired effect of the piece being played. This is due to the increased distance from the sources and the proximity to the walls allowing for more scattering.



(a) adult_female_speech_Test1_LP1



(b) adult_female_speech_Test2_LP1

Figure 22: A spectrogram comparison between the two tests conducted in the application of speech

Much like the flute the difference here is the clarity of speech, the interaction between the two sources means that they often mask or cancel each other, this means that a lot of detail is lost in the intelligibility of the speech. This is common in the room and something which can be experienced in person inside the room during announcements made through the PA.

Conclusion

The software managed to produce a noticeable recreation of the space even with the lacking low frequency accuracy. It was possible to make out the features of the room and localise when next to a wall and the SIRRs display. It was also possible to sense more than one source when placed at the perimeter walls with listening position three or four. This was due to in Test 2 the inter aural time difference and the inter aural level/energy difference.

The application of these technologies has proven their worth in many cases relating to audio and room design. An area which has recently become of interest is video game audio rendering. With the industry moving to a real time ray tracing technique for visuals, more and more developers are doing the same with audio. Nvidia's VR-Works Audio[13] SDK allows for such process in real time. Taking into account the same material properties that describe the way light interacts with a face, these can be used for calculations creating a real time impulse response for the in-game audio to be auralized with. Although in this case auralisation is either inaccurate or no computationally feasible due to the lack of computing power for most machines. However, the Nvidia "turing" architecture provides dedicated processing cores created for ray traced calculations which are supported and can be utilised with the SDK allowing for a accelerated Ray Traced audio which can create a more immersive environment and greater retention and realism.

Evaluation

The process of generating and analysing the model seemed to be easy due to my experience with the modelling process on Google Sketchup. However, the problem came from finding exact measurements of the space as these are not published online, nor are official floor plans. This meant that the process of creating the structure came from memory of the location and approximations. Due to my intimate knowledge of the space and having recorded in the space I had a good idea as to how the space looked and sounded. Add to this the ability for me to transport myself inside the space in VR to get a sense of scale allowed a more comprehensive approach to sizing the object.

The process of utilising CATT's various features was primarily trial and error, in total there were 18 different exports of the model into TUCT before I was able to get a result which I was pleased with, I attempted to model the pit of the Lyons but due to the original IR having the pit closed I chose not to. After every export I listened and found a new setting or a new parameter to help better explore the space. This is the process which took the most time, the exploration of the software and its capabilities made up most of the time.

An improvement which would allow for a better result would be using the inbuilt 'Electro-Acoustic' sources which emit from a single source with a direction oppose to in all directions. This means I would get a more accurate reading from a 'speaker' source such as in test number two.

An additional test was run to test the limits of the software and the usability of its output, a surround or 'Choral Spezatti' formation which was used in 2018 during the Music Department's Practical Project module. This is an opportunity to see how well running eight different sources sounded. Although this was for personal enquiry it was interesting to listen to the summed source to check for diffusion. These renders took over an hour for all four receivers.

Overall I feel that the renders and output were successful in comparison to what I was expecting, before starting this module I was not aware that we had the computing power not that the software would be so accessible such that it could be picked up so quickly. During the research phase of this module I found myself reading more than I needed to as I simply found the reading stimulating and interesting. I have come out of this module wanting to know and learn more about the various technologies that allow us to model a space to a certain degree of accuracy.

References

[1] L. Savioja and U. P. Svensson, "Overview of geometrical room acoustic modeling techniques," *Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 708–730, 2015. DOI: https://doi.org/10.1121/1.4926438.

[2] A. Krokstad, S. Strom, and S. Sorsdal, "Calculating the acoustical room response by the use of a ray tracing technique," *Journal of Sound and Vibration*, vol. 8, no. 1, pp. 118–125, 1968. DOI: https://doi.org/10.1016/0022-460X(68)90198-3.

[3] M. Kleiner, B.-I. Dalenback, and P. Svensson, "Auralization-an overview," vol. 41, pp. 861–875, 1993.

[4] CATT, *Catt-a9*, Online; accessed 18 April 2021, 2020. [Online]. Available: https://www.catt.se/.

[5] O. A/S, *Odeon room acoustic*, Online; accessed 18 April 2021, 2021. [Online]. Available: https://odeon.dk/.

[6] A. F. M. Group, *Enhanced acoustic simulator for engineers - ease*, Online; accessed 18 April 2021, 2021. [Online]. Available: https://ease.afmg.eu/index.php/software-new.html.

[7] CATT and B.-I. Dalenbäck, *Reverberation time, diffuse reflection, sabine, and computerized prediction - part i.* Online; accessed 18 April 2021, 1999. [Online]. Available: https://www.catt.se/diffseries/index.htm.

[8] Google and Trimble, *Sketchup make*, Online; accessed 18 April 2021, 2016. [Online]. Available: https://www.sketchup.com/.

[9] A. K5, *Su2catt*, Online; accessed 18 April 2021, 2021. [Online]. Available: https://www.k5-akustik.de/en/software/su2catt/.

[10] Oculus, *Oculus quest 2*, Online; accessed 18 April 2021, 2020. [Online]. Available: https://www.oculus.com/quest-2/.

[11] Audiolab, D. Murphy, and J. Rees-Jones. (2017). "Jack lyons concert hall (university of york)." Online; accessed 18 April 2021, [Online]. Available: https://www.openair.hosted.york.ac.uk/?page_id=571.

[12] D. M. Howard and J. Angus, *Acoustics and Psychoacoustics*, 5th. Routledge, 2017.

[13] Nvidia, *Vr works - audio*, Online; accessed 18 April 2021, 2021. [Online]. Available: https://developer.nvidia.com/vrworks/vrworks-audio.

Appendix 1 - Table of Acoustic Parameters

		Real										
LP	Param	62.5	125	250	500	1000	2000	4000	8000	A	C	L
	D50	48.97	42.95	29.53	39.90	49.92	51.36	63.32	76.86	54.24	46.69	52.25
	C50	-0.18	-1.23	-3.78	-1.78	-0.01	0.24	2.37	5.21	0.74	-0.58	0.39
1	C80	2.22	0.81	-2.61	-0.18	1.26	1.72	4.23	7.60	2.30	0.92	1.84
	EDT	1.43	1.75	1.76	1.72	1.76	1.70	1.24	0.62	1.55	1.68	1.63
	T30	2.18	2.12	2.03	1.82	1.80	1.69	1.26	0.92	1.72	1.89	1.89
	D50	26.36	20.68	21.42	25.67	29.90	31.82	55.33	56.23	39.43	30.92	34.49
	C50	-4.46	-5.84	-5.65	-4.62	-3.70	-3.31	0.93	1.09	-1.87	-3.49	-2.79
2	C80	-0.65	-3.06	-1.59	-1.38	-1.27	-1.11	3.02	4.44	0.42	-0.79	-0.18
	EDT	1.66	1.74	1.71	1.66	1.76	1.77	1.24	0.89	1.55	1.62	1.57
	T30	2.34	2.27	2.26	1.89	1.83	1.70	1.27	0.91	1.78	2.05	2.06
	D50	7.20	32.28	7.94	17.21	24.50	36.77	58.84	67.16	41.06	28.87	33.67
	C50	-11.10	-3.22	-10.64	-6.82	-4.89	-2.35	1.55	3.11	-1.57	-3.92	-2.94
3	C80	-7.48	-0.84	-5.77	-4.37	-2.21	-0.21	3.74	5.84	0.50	-1.64	-0.83
	EDT	1.84	2.09	2.08	2.01	1.90	1.84	1.12	0.72	1.63	1.82	1.76
	T30	2.01	2.34	2.07	1.81	1.86	1.65	1.27	0.93	1.75	1.97	1.98
	D50	22.53	49.44	18.69	18.68	23.47	35.02	57.44	64.74	39.86	35.52	39.65
	C50	-5.36	-0.10	-6.38	-6.39	-5.13	-2.69	1.30	2.64	-1.79	-2.59	-1.83
4	C80	-0.75	1.66	-3.65	-2.82	-2.88	-0.61	3.42	5.64	0.31	-0.40	0.23
	EDT	1.68	1.55	1.86	1.85	2.05	1.77	1.27	0.83	1.69	1.75	1.71
	T30	2.05	2.13	2.21	1.79	1.81	1.70	1.27	0.91	1.76	1.99	2.00

Table 1: A table to display the acoustic parameters of the real IR taken in the Jack Lyons Concert Hall

		Modelled Test 1										
LP	Param	62.5	125	250	500	1000	2000	4000	8000	A	C	L
	D50	40.65	20.23	54.42	46.96	56.22	53.05	61.71	75.96	62.14	60.13	68.66
	C50	-1.64	-5.96	0.77	-0.53	1.09	0.53	2.07	5.00	2.15	1.78	3.41
1	C80	1.50	-0.78	1.35	2.35	3.81	3.57	4.53	7.74	4.79	4.34	5.89
	EDT	1.95	2.03	1.60	1.74	1.17	1.31	1.13	0.69	1.09	1.19	0.99
	T30	1.78	1.98	2.26	1.88	1.66	1.56	1.38	0.97	1.50	1.62	1.56
	D50	21.73	20.59	28.32	27.68	48.86	40.62	43.87	60.38	47.48	45.14	48.65
	C50	-5.57	-5.86	-4.03	-4.17	-0.20	-1.66	-1.09	1.83	-0.44	-0.85	-0.24
2	C80	0.79	-3.04	-0.82	-1.27	2.19	1.75	1.68	4.82	2.24	1.82	2.31
	EDT	1.41	1.88	1.48	1.48	1.24	1.08	1.06	0.70	1.03	1.12	0.99
	T30	1.78	2.39	2.15	1.89	1.64	1.51	1.36	0.95	1.47	1.60	1.53
	D50	31.13	30.75	46.65	29.29	34.02	40.20	42.68	58.84	44.75	43.75	47.92
	C50	-3.45	-3.53	-0.58	-3.83	-2.88	-1.72	-1.28	1.55	-0.91	-1.09	-0.36
3	C80	1.51	0.39	1.80	0.01	-0.09	1.37	2.21	5.17	2.42	2.13	2.88
	EDT	0.95	1.49	1.79	1.98	1.53	1.44	1.25	0.78	1.23	1.30	1.07
	T30	2.62	2.46	2.45	2.12	1.82	1.68	1.45	1.00	1.61	1.80	1.72
	D50	38.19	32.71	49.62	31.48	33.49	38.36	41.64	57.55	43.48	42.81	46.72
	C50	-2.09	-3.14	-0.07	-3.38	-2.98	-2.06	-1.47	1.32	-1.14	-1.26	-0.57
4	C80	2.91	0.63	2.41	1.19	0.13	0.95	1.99	5.10	2.19	2.00	2.71
	EDT	0.93	1.71	1.80	1.76	1.62	1.43	1.27	0.77	1.24	1.32	1.06
	T30	2.70	2.43	2.45	2.04	1.71	1.67	1.43	1.03	1.58	1.78	1.71

Table 2: A table to display the acoustic parameters of the Test 1 IR in the modelled Jack Lyons Concert Hall

LP	Param	Modelled Test 2 62.5	125	250	500	1000	2000	4000	8000	A	C	L
1	D50	22.73	19.23	26.39	35.98	39.72	39.12	47.00	67.13	47.52	44.96	50.58
	C50	-5.32	-6.23	-4.45	-2.50	-1.81	-1.92	-0.52	3.10	-0.43	-0.88	0.10
	C80	-3.15	-4.46	-2.03	-0.65	1.24	0.98	2.25	6.04	2.34	1.91	2.72
	EDT	0.99	1.71	1.78	1.52	1.34	1.27	1.17	0.89	1.22	1.24	1.16
	T30	1.73	2.16	2.05	1.83	1.66	1.52	1.34	0.93	1.47	1.60	1.53
2	D50	29.51	42.60	48.70	52.92	50.63	57.23	58.41	73.99	60.20	58.36	63.83
	C50	-3.78	-1.30	-0.23	0.51	0.11	1.27	1.48	4.54	1.80	1.47	2.47
	C80	-1.15	-0.86	0.65	1.67	1.30	2.54	2.91	6.49	3.25	2.86	3.84
	EDT	1.37	2.11	1.59	1.65	1.40	1.35	1.21	0.79	1.20	1.27	1.11
	T30	1.79	2.09	2.15	1.80	1.63	1.50	1.33	0.94	1.45	1.56	1.49
3	D50	64.53	19.17	39.57	32.45	39.68	34.98	45.26	60.65	46.16	44.62	51.54
	C50	2.60	-6.25	-1.84	-3.18	-1.82	-2.69	-0.83	1.88	-0.67	-0.94	0.27
	C80	3.51	-4.88	-0.59	-0.61	0.77	-0.54	1.27	4.96	1.75	1.42	2.44
	EDT	1.21	1.93	2.32	2.18	1.65	1.58	1.42	0.91	1.36	1.45	1.12
	T30	2.14	2.31	2.21	2.04	1.80	1.71	1.40	1.02	1.60	1.74	1.65
4	D50	64.64	20.76	42.21	31.01	41.68	34.38	42.63	60.95	45.01	43.67	51.14
	C50	2.62	-5.82	-1.36	-3.47	-1.46	-2.81	-1.29	1.93	-0.87	-1.11	0.20
	C80	3.43	-4.86	-0.54	-1.45	0.83	-0.17	0.79	4.70	1.51	1.19	2.30
	EDT	1.15	1.93	2.18	1.91	1.55	1.52	1.42	0.95	1.35	1.42	1.13
	T30	2.12	2.33	2.25	1.88	1.75	1.60	1.40	0.99	1.55	1.68	1.61

Table 3: A table to display the acoustic parameters of the Test 2 IR in the modelled Jack Lyons Concert Hall

Appendix 2 - List of Audio Files

```
Source_Scenario_ListeningPosition_ListeningFormat
- adult_female_speech.wav
- adult_female_speech_Real_LP1_BFormat.wav
- adult_female_speech_Real_LP2_BFormat.wav
- adult_female_speech_Real_LP3_BFormat.wav
- adult_female_speech_Real_LP4_BFormat.wav
- adult_female_speech_Test1_LP1_BFormat.wav
- adult_female_speech_Test1_LP1_Binarual.wav
- adult_female_speech_Test1_LP2_BFormat.wav
- adult_female_speech_Test1_LP2_Binarual.wav
- adult_female_speech_Test1_LP3_BFormat.wav
- adult_female_speech_Test1_LP3_Binarual.wav
- adult_female_speech_Test1_LP4_BFormat.wav
- adult_female_speech_Test1_LP4_Binarual.wav
- adult_female_speech_Test2_LP1_BFormat.wav
- adult_female_speech_Test2_LP1_Binarual.wav
- adult_female_speech_Test2_LP2_BFormat.wav
- adult_female_speech_Test2_LP2_Binarual.wav
- adult_female_speech_Test2_LP3_BFormat.wav
- adult_female_speech_Test2_LP3_Binarual.wav
- adult_female_speech_Test2_LP4_BFormat.wav
- adult_female_speech_Test2_LP4_Binarual.wav
- flute_music.wav
- flute_music_Real_LP1_BFormat.wav
- flute_music_Real_LP2_BFormat.wav
- flute_music_Real_LP3_BFormat.wav
```

- flute_music_Real_LP4_BFormat.wav
- flute_music_Test1_LP1_BFormat.wav
- flute_music_Test1_LP1_Binarual.wav
- flute_music_Test1_LP2_BFormat.wav
- flute_music_Test1_LP2_Binarual.wav
- flute_music_Test1_LP3_BFormat.wav
- flute_music_Test1_LP3_Binarual.wav
- flute_music_Test1_LP4_BFormat.wav
- flute_music_Test1_LP4_Binarual.wav
- flute_music_Test2_LP1_BFormat.wav
- flute_music_Test2_LP1_Binarual.wav
- flute_music_Test2_LP2_BFormat.wav
- flute_music_Test2_LP2_Binarual.wav
- flute_music_Test2_LP3_BFormat.wav
- flute_music_Test2_LP3_Binarual.wav
- flute_music_Test2_LP4_BFormat.wav
- flute_music_Test2_LP4_Binarual.wav
- IR_Real_LP1_BFormat.wav
- IR_Real_LP2_BFormat.wav
- IR_Real_LP3_BFormat.wav
- IR_Real_LP4_BFormat.wav
- IR_Test1_LP1_BFormat.wav
- IR_Test1_LP1_Binarual.wav
- IR_Test1_LP2_BFormat.wav
- IR_Test1_LP2_Binarual.wav
- IR_Test1_LP3_BFormat.wav
- IR_Test1_LP3_Binarual.wav
- IR_Test1_LP4_BFormat.wav
- IR_Test1_LP4_Binarual.wav
- IR_Test2_LP1_BFormat.wav
- IR_Test2_LP1_Binarual.wav
- IR_Test2_LP2_BFormat.wav
- IR_Test2_LP2_Binarual.wav
- IR_Test2_LP3_BFormat.wav
- IR_Test2_LP3_Binarual.wav
- IR_Test2_LP4_BFormat.wav
- IR_Test2_LP4_Binarual.wav