# SUPPORTING MATERIAL

Here is a link to a supporting web page for this document. It contains all of the code, audio samples and videos referred to within this document.

A supporting web page for the work described in this document can be found here. All code, audio sample and videos mentioned in the text can be viewed/heard by clicking on the file name provided.

All relevant data produced as a result of this project, including that found on the supporting webpage can be found within the appropriate sub file within the following file structure:

```
Dissertation.pdf
└─ 0) SupportingFiles
    ├─ 1) Audio
    │   ├─ 1.1) Anechoic Recordings
    │   ├─ 1.2) Calibrated RIRs
    │   ├─ 1.3) Odeon Settings
    │   └─ 1.4) Odeon Materials
    ├─ 2) Matlab
    ├─ 3) MaxPatch
    │   ├─ 3.1) Javascript files
    │   └─ 3.2) Max patch file
    └─ 4) Room Model
        ├─ 4.1) SketchUp
        ├─ 4.2) Odeon
        └─ 4.3) Material List
```

For example, if the text were to read:

*...the relavant code can be found in file 3.1...*

the file in question can be found in the `Javascript files` folder.

# Contents

# Acronyms

**VSS** Virtual Singing Studio . . . 14
**RIR** Room Impulse Response . . . 7
**ISM** Image Source Method . . . 12
**TO** Transition Order . . . 11
**VBS** Vector Based Scattering Method . . . 11
**LR's** Late Rays . . . 12

...to avoid this problem it is recommended that distances to surfaces are kept greater than say 03 and 0.5 meters”

## 2.1 - ODEON

### 2.1.1) Water Tightness Test

Once the SketchUp model had been exported as a .par file using the SU2Odeon plug-in [1] it could be opened in Odeon and checked to ensure that there were no gaps in the model for which rays to escape. If this was the case, the model would have to be fixed in SketchUp and reimported into Odeon. Odeon makes checking the model easy by running a ‘**water tightness**’ check where a large number of rays are traced around the room, seeing whether any of them manage to escape. Figure 1 shows the Hendrix Hall model undergoing such a test. Once it was ensured that the model was fit for use, the surface materials could be assigned.

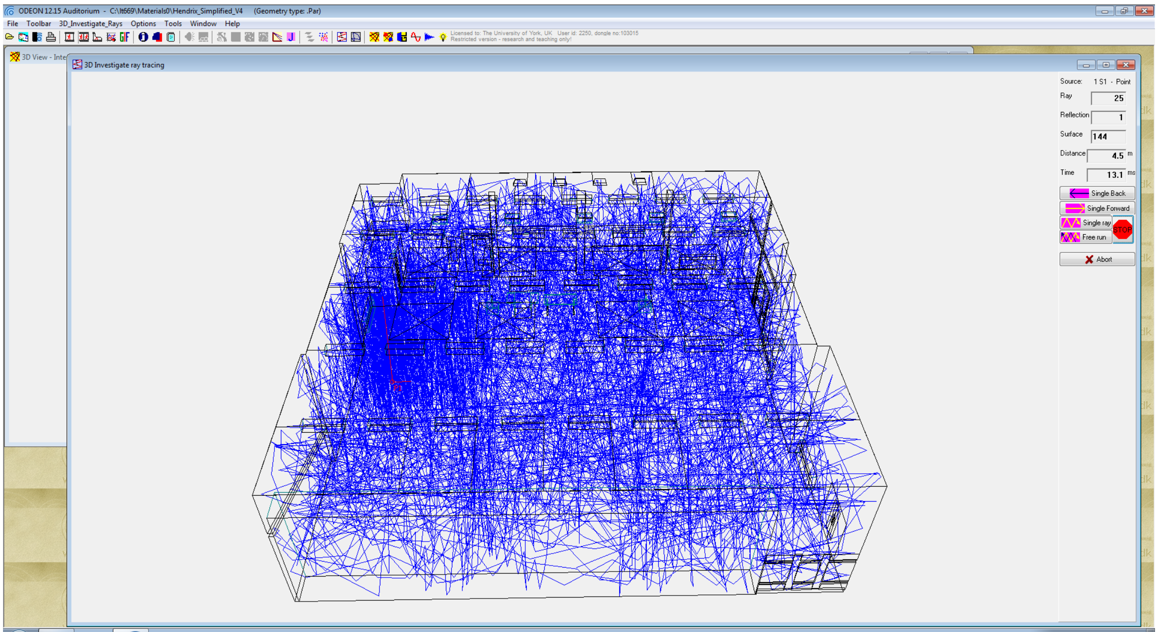



Figure 1: Hendrix Hall model undergoing a water tightness test in Odeon.

### 2.1.2) Material Selection

Surface materials greatly influence how sound reflects around a room, effecting reverberation time and the rate at which frequency bands are absorbed. It is therefore imperative to assign materials that closely match those within the real room in order to produce an accurate representation of the acoustic environment. For this, Odeon provides a list of common materials often found when constructing buildings.

#### 2.1.2.1 Initial Materials

Odeon's material list appears to be designed for the auralisation of structures with very basic interior. Due to a lack of choice, exact materials in the room could not be modelled accurately, however the closest match to what was thought to be the true material were selected as an approximation. In some cases, appropriate replacement materials were not available, therefore new materials had to be added to the material list. This can be done by finding a materials absorption coefficients and selecting '**Edit an existing material**' where a new materials absorption coefficients can be entered and saved, as shown in figure 2. Absorption coefficient values range from 0 - 1 (0% - 100%), indicating the percentage of attenuation applied to the selected frequency band upon a contact with the surface. Required materials for which an appropriate replacement could not be found in the material list are listed in table 2.

| Material | Surface applied to |
|---|---|
| Hard Plastic [2] | Roof lights and projector covers |
| Mineral fibre [3] | Ceiling Tiles |
| Slate [1] [4] | Blackboard |

Table 1: Table of materials for which absorption coefficients were sought and added to Odeon's material list.

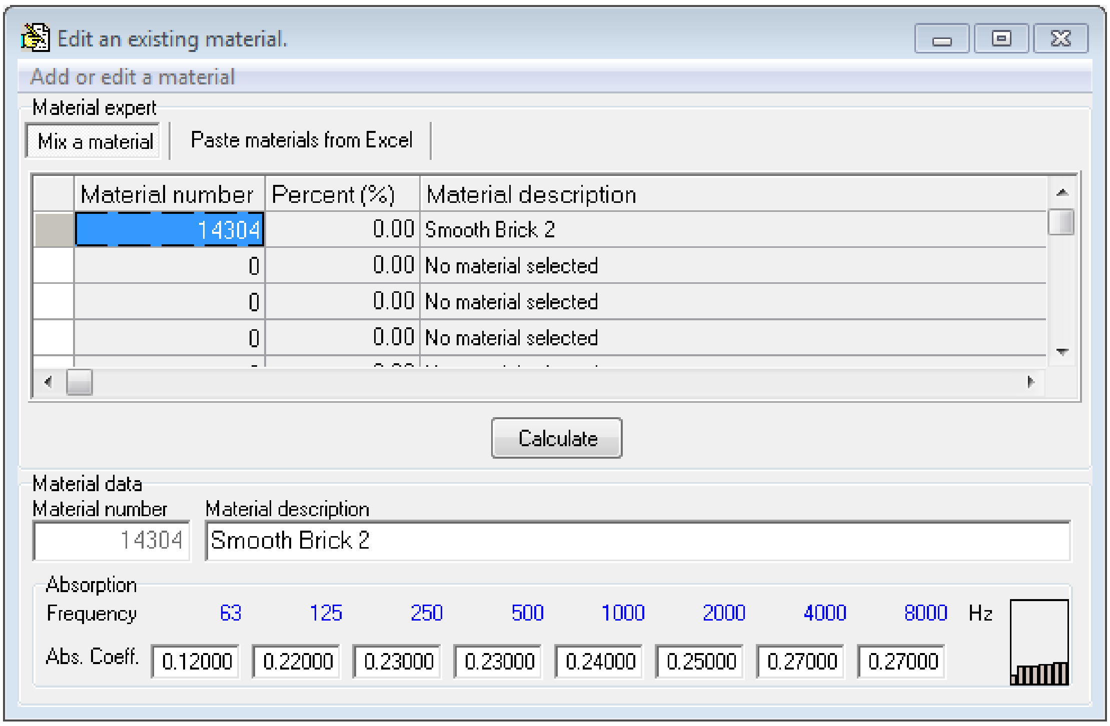



Figure 2: Absorption coefficient editing window in Odeon used to add unavailable materials.

[1]The absorption coefficients provided were only available for the 125Hz to 4kHz octave bands, therefore absorption coefficients for the 63Hz and 8kHz octave bands were given a value of 0.1

#### 2.1.2.2 Surface Types

For a number of surfaces it is appropriate to edit properties other than just their absorption coefficients. As explained in section **??**, the roof hangings are constructed of four joint slanting surfaces. By changing the individual surface types from **'Normal'** to **'Fractional'**, Odeon avoids erroneously calculating diffraction due to each of the individual surfaces and treats them as a whole surface. For surfaces such as the contracted seating shown in figure 3 where gaps are present in the over all structure, it was possible to model this as one solid object and to set a **transparency** value. A transparency value of 0 means the surface is a solid, whereas a value of 1 makes a surface totally transparent, allowing rays pass straight through it. A transparency value of 0.3 was chosen as a reasonable estimate, the effects of which are shown in figure 4, showing how a ray can pass through the front of the object, reflect around the inside and eventually reflect back out again.

Figure 3: Contracted seating in Hendrix Hall showing gaps in the structure.

As previously mentioned in section **??**, the lights were modelled as simple rectangles as opposed to the more complex objects made from a large number of surfaces with the intention of more accurately modelling the scattering effect (described in section **??**) due to their shape by altering the objects **Scattering Coefficient**. Odeon provides a table of initial indicators for possible scattering coefficients, shown in figure 5 which can be used to select the scattering coefficient for a surface that fits a similar description to the one given. Given the vague descriptions, a scattering coefficient of 0.2 was selected for the lights.

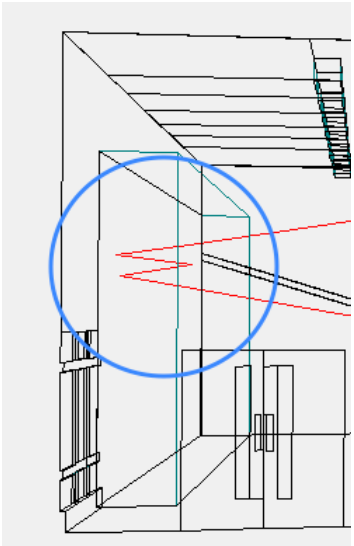


Figure 4: Blue circle highlights a ray penetrating the modelled seating area, reflecting 3 times and escaping the seating area due to the transparency value set.

| Material | Scattering coefficient at mid-frequency (707 Hz) |
|---|---|
| Audience area | 0.6–0.7 |
| Rough building structures, 0.3–0.5 m deep | 0.4–0.5 |
| Bookshelf, with some books | 0.3 |
| Brickwork with open joints | 0.1–0.2 |
| Brickwork, filled joints but not plastered | 0.05–0.1 |
| Smooth surfaces, general | 0.02–0.05 |
| Smooth painted concrete | 0.005–0.02 |

Figure 5: A table provided in the Odeon Manual [5] that can be used to set an approximate scattering coefficient for surfaces similar to those described.

#### 2.1.2.3 RIR Comparison

Once the initial materials had been selected, the authenticity of the Room Impulse Response (RIR)'s were checked by comparing their frequency content against that of a real measured RIR from Hendrix Hall (see section **??**). This was done by using Matlab to plotting the spectrograms of the omni-directional W channel of both RIR's, as can be seen in figure 6. The spectrogram shows the high frequency content in the real RIR attenuating in a smooth roll off fashion from 20kHz to about 8kHz starting from approximately 0.2 seconds in, whereas the full audible spectrum appears to attenuate almost evenly in the RIR produced in Odeon.

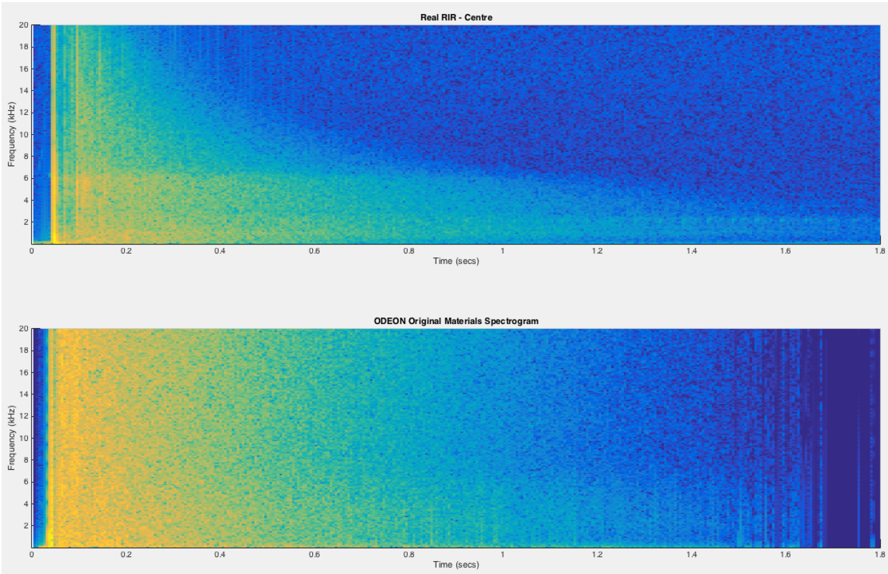



Figure 6: Spectrograms of real RIR (top) against spectrogram of first rendered RIR from Odeon.

As the most obvious difference between the two is the difference in high frequency attenuation, large surfaces with low valued high frequency absorption coefficients, such as the walls and ceiling, were edited to absorb more of the high frequency content. Several iterations can be seen in figure 7, showing the spectrogram of the real RIR and several iterations of the RIR's rendered from Odeon. The Spectrogram at the bottom is that of an RIR produced in Odeon with the final material selections. The difference between the real RIR and the final Odeon RIR can be seen more clearly in figure 8. This was done by adding 10% onto the existing absorption coefficients for the 63Hz - 2000Hz octave bands, 20% to the 4000Hz band and 30% to the 8000Hz band. This also included edited the ceiling material coefficients by adding an extra 10% to the 8000Hz octave band.

Table 2 contains a list of audio samples (can be found here or in file 1.4) indicating the edits made to the material list in order to produce each one, with their corresponding spectrogram graph name which can be found in figure 7.

By analysing figure 8 it can be seen that it was possible to produce a much more fitting RIR compared to the one produced using the original materials. By listening to the audio samples, it is clear that though the final Odeon RIR (Odeon7RIR.wav) is much more similar to the real RIR (RealRIR.wav) than the one produced using the original surface materials (OdeonOriginal.wav) in terms of reverberation time, they still sound very different in terms of their frequency content, where the RIR's produced in Odeon lack what the author considers *depth*. This is most likely due to the fact that the geometrical acoustic modelling methods used to produce the RIR do not accurately reproduce the low-frequency content, the reasons for which were discussed in section **??**. However, it can be said that the final Odeon RIR sounds much more similar to the real one than the other three RIR's produced using earlier iterations of the material list.

| Audio file | Graph Name | Absorption Coefficient Edits | | |
|---|---|---|---|---|
| | | Material | Addition | Octave band |
| RealRIR.wav | Real RIR | NA | | |
| OdeonOriginal.wav | ODEON Original | NA | | |
| Odeon3RIR.wav | ODEON 3 | Ceiling (Mineral Fibre) | +10% | 8kHz |
| Odeon6RIR.wav | ODEON 6 | Walls (Solid Brick) | +10% | All |
| Odeon7RIR.wav | ODEON 7 | Walls (Solid Brick) | +10% | 4kHz |
| | | | +20% | 8kHz |
| | | Ceiling (Mineral Fibre) | +10% | 8kHz |

Table 2: List of audio samples with corresponding spectrogram title in figure 7 and the edits made for each one. (Only Odeon(3)(6)(7) audio sample are shown here as the other audio samples produced as a result of other material list iterations provided no relevant information).

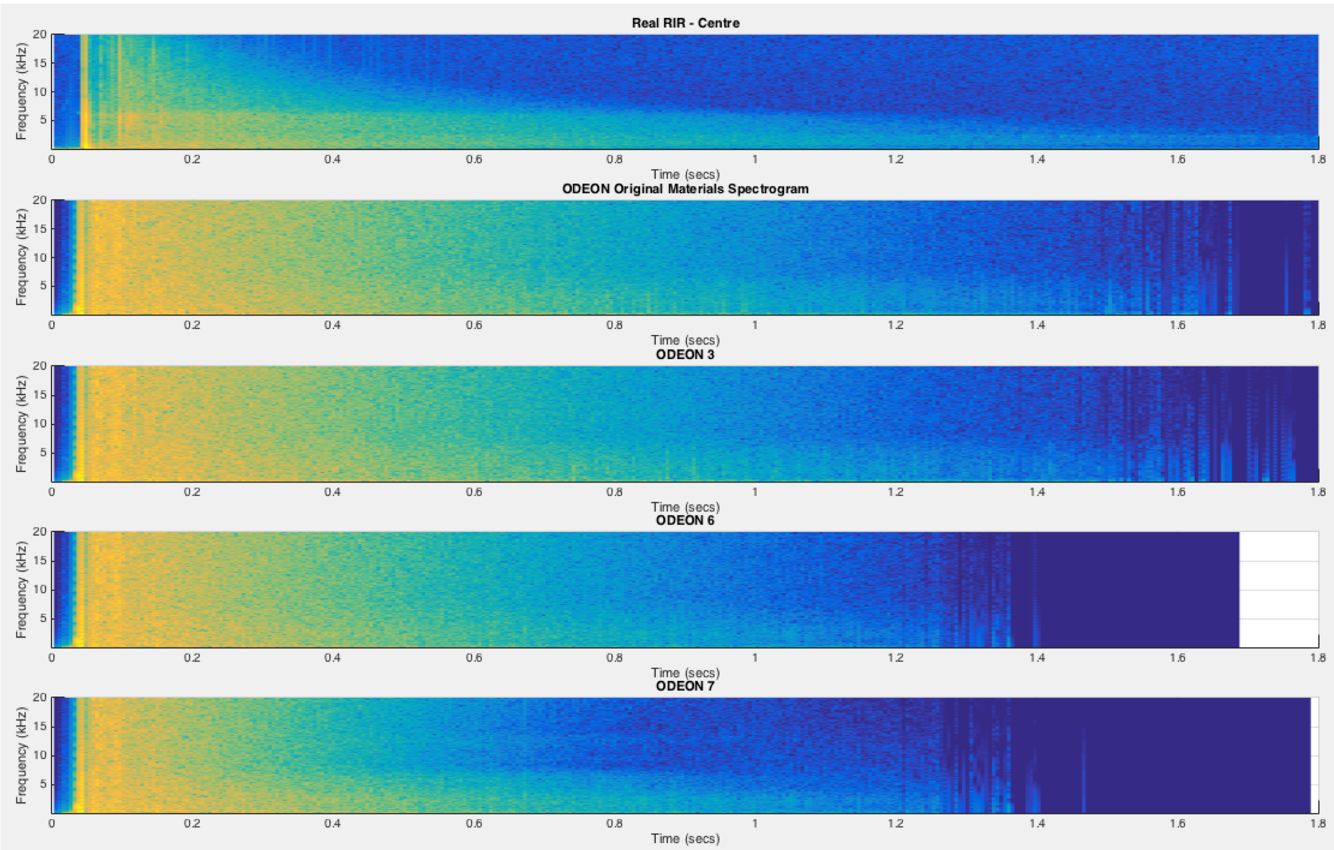



Figure 7: Spectrograms of real RIR (top) against spectrogram of several rendered RIR from Odeon with different material absorption coefficients where ODEON (bottom) shows the final RIR used. Spectrogram titles and corresponding material edits can be found in table 2

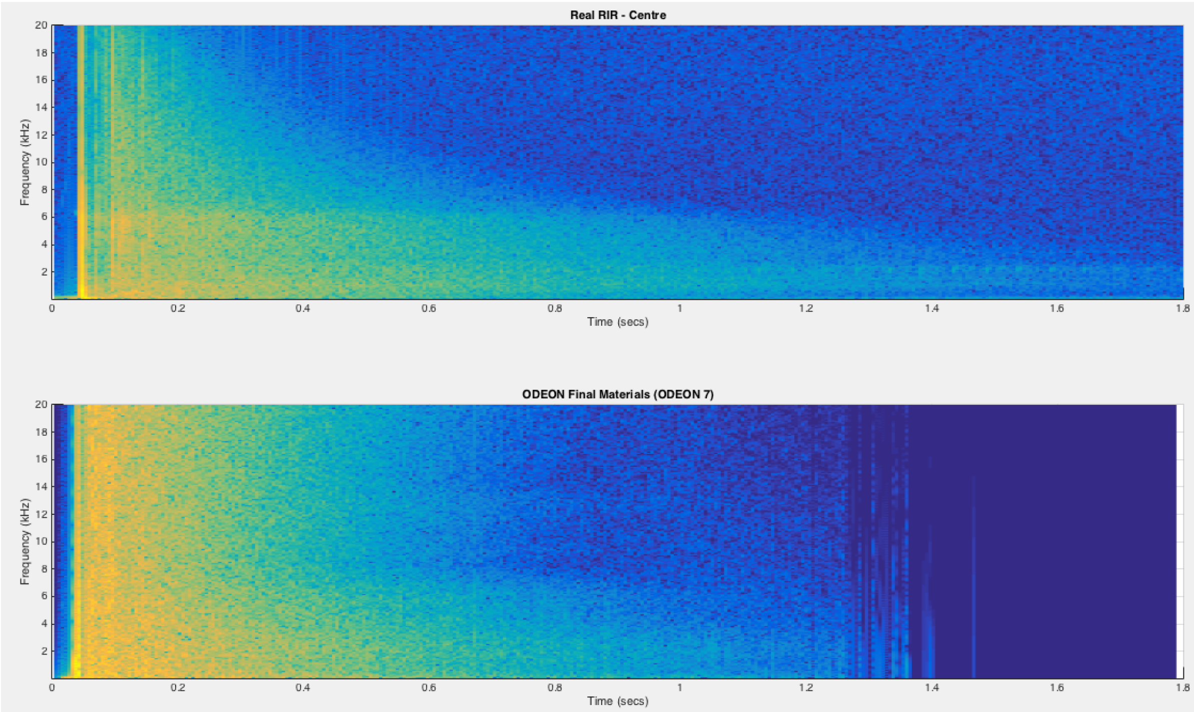



Figure 8: Spectrograms of real RIR (top) against Odeon RIR from Odeon with final material absorption coefficients showing a more convincing replica of the real RIR by editing the rooms material absorption coefficients.

A full material list is available in '`Material_List_V2.xlsx`', showing which materials are applied to each surface within the Odeon model, as well as a list of absorption coefficients iterations that were applied to the brick walls and ceiling can be found in `Absorption Coefficients.xlsx`. Both can also be found in file 4.3.

It was later discovered (see section RIR Topology Problems) that the RIR's used to decide the final materials were incorrect, a result of which caused them to be much louder than they should have been. To ensure that the final surface materials that were selected when using the incorrect RIR's were still appropriate for the new, quieter ones, two new RIR's were rendered using both the original materials and the final materials. Both were plotted to compare against the same real RIR used before, the results of which are shown in figure 9. As the new Odeon RIR's were so much quieter, the real RIR was reduced in level to match those produced by Odeon. As it can be seen, the final materials used in the room model reduce the length of the reverb and attenuate high frequencies quicker than when the room model contained the original materials as they had shown previously.

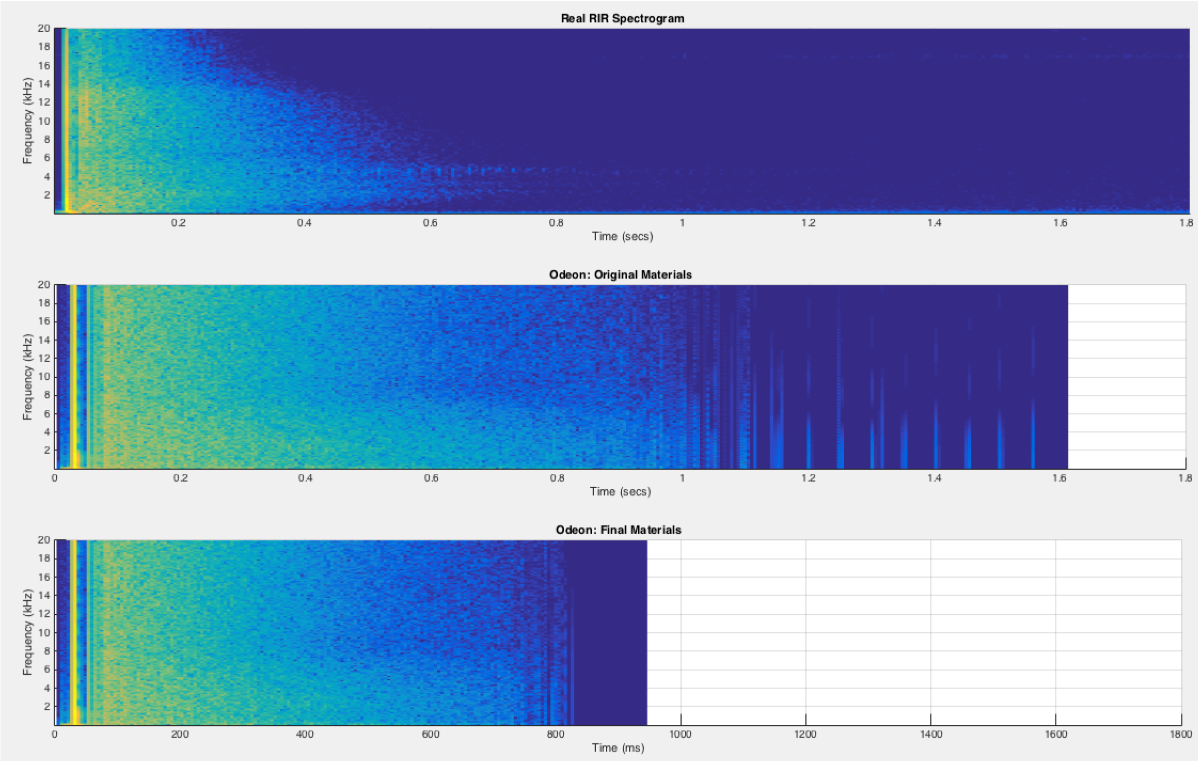



Figure 9: Spectrograms of correct Odeon RIR's with original materials (top) and final materials (centre) to compare against real RIR that has been level calibrated (bottom).

## 2.1.3) Odeon Output Settings

### 2.1.3.1 Ray Settings

Odeon provides the following options for RIR rendering:

**Astop** - The maximum possible attenuation of each octave band.

**Apass** - Ripple of octave band filters in dB.

**Band overlap** - 100% overlap gives a smooth transition between the FIR octave band filters as they are not completely rectangular.

**Maximum reflection order** - Maximum allowed is 2000 as is the default. This means a ray can only reflect around a room 2000 times before the simulation is stopped. This prevents trapped rays from prolonging the impulse response if it never reaches the receiver.

**Late Rays** - These are emitted from the source and reflected according to the Vector Based Scattering Method (VBS) (described in section **??**) taking into account scattering due to surface size and roughness.

**Transition Order** - As explained in section **??**, a Transition Order (TO) can be set to determine the number of early rays sent out to find a number of wall combinations for reflections using the

| Audio Sample | TO | LR's | Graph Name |
|---|---|---|---|
| TO_2_LR_1285.wav | 2 | 1285 | TO_2_LR_1285 |
| TO_2_LR_20560.wav | 2 | 20560 | TO_2_LR_20560 |
| TO_4_LR_1285.wav | 4 | 1285 | TO_4_LR_1285 |
| TO_4_LR_20560.wav | 4 | 20560 | TO_4_LR_20560 |

Table 3: Table of audio samples with the corresponding settings information and graph name. Samples can be listened to here or in file 1.3

Image Source Method (ISM). After this number of rays has been reached, the ray-tracing method is used.

All settings apart from Late Rays (LR's) and TO are left at their default values as they are sufficient for accuracy and prevent the increase of computation time which will be needed given the large number of RIRs required for free movement.

Odeon suggests two possible modes, **Engineer** and **Precision** which both suggest using a different number of LR's which are 1285 and 20560 respectively. As both the TO and LR's value increase both the accuracy and computation time of the RIR's, different combinations of the values were tried and RIR's were rendered for comparison. Figure 10 shows plots of the rendered RIR's with the name indicating the TO and LR's value. For example, 'TO_2_LR_1285' indicates the sample was produced where TO = 2 and LR's = 1285. Table 3 can be used to clarify the settings used for each RIR in figure 10 and indicates which audio sample can be listened to. For consistency, all RIR's were taken from the centre of the modelled room.

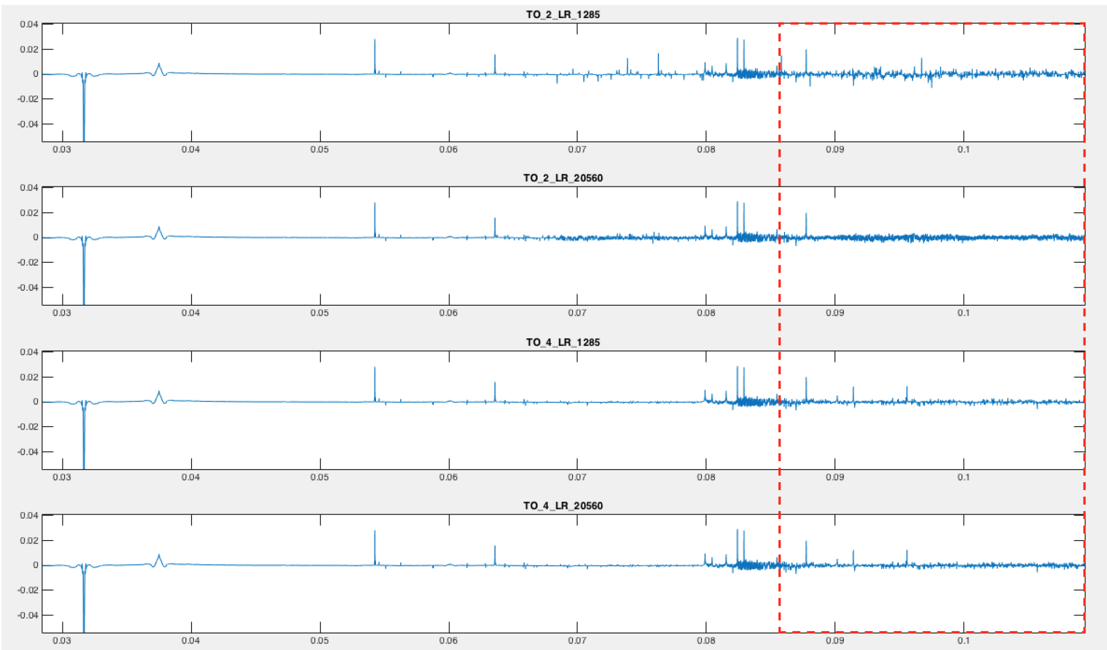



Figure 10: RIR's rendered using different TO and late ray values

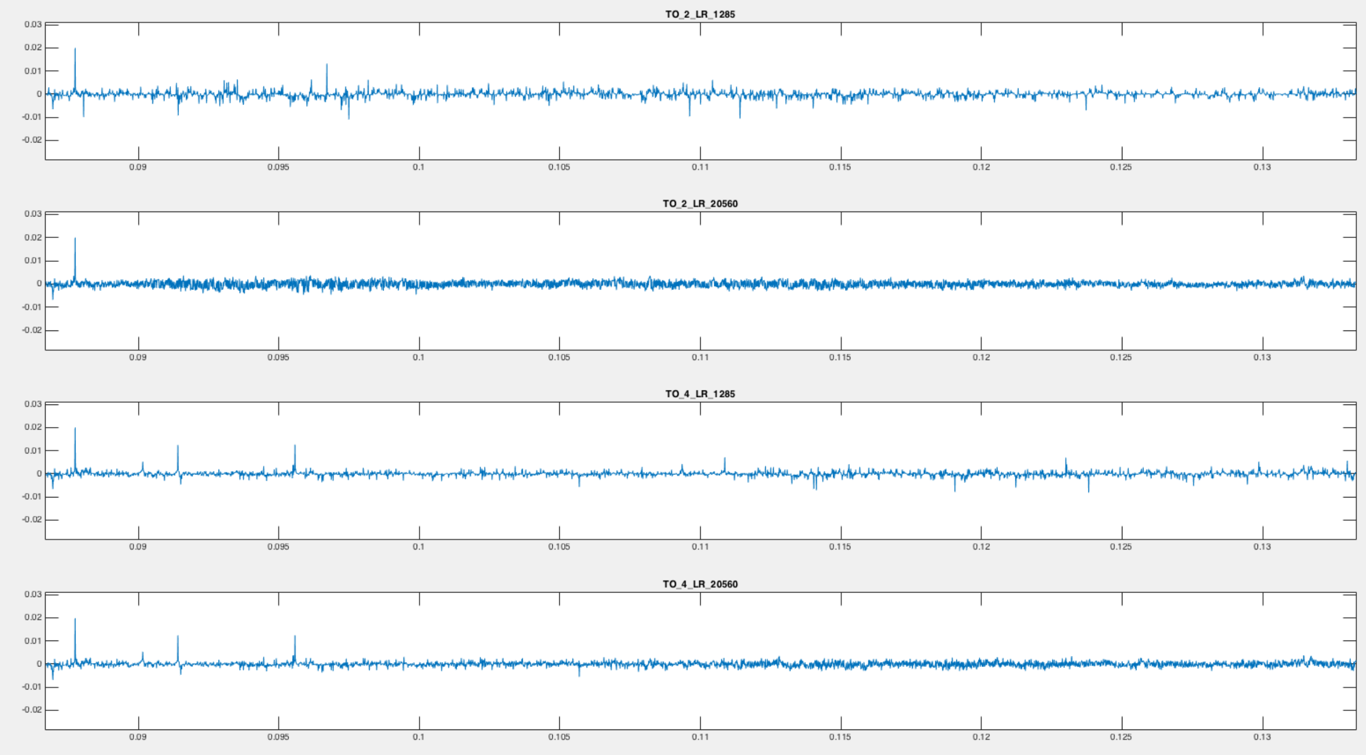



Figure 11: Zoomed in section indicated by the red dashed line in figure 10

Through aural analysis, it can be heard that the number of LR's used has an audible impact, where a lower LR's value produces an RIR that sounds *grainy*. This can be seen in figure 11, which is zoomed in on the section highlighted in red in figure 10 showing the RIR's from approximately 0.085s, where the difference caused by the different setting combinations can be seen. The two RIR's with a high LR's value (2nd and 4th plot) show more densely packed reflections due to the fact that more rays are used. This produces a 'smoother' more natural sounding reverb tail.

Figure 12 shows the same RIR's from the start of the plots. For the RIR's produced with a TO value of 2 (1st and 2nd plots), it can be seen that there are a lot more randomly occurring peaks than in the RIR's with a TO value of 4. It is assumed that this is caused by the fact that the ray-tracing method is used much earlier on in the top two RIR plots, thus the randomly reflected rays cause more randomly occurring peaks. Though the difference in using a TO of 2 or 4 can be seen, is not audible by listening to the RIR audio samples.

The time difference taken to render the RIR's is also an important factor when choosing which of these settings to use. Given the potentially large number of RIR's required, a slightly more accurate output may result in a much greater calculation time. Odeon uses information gathered when previously rendering RIR's to speed up calculations for other RIR's being rendered in the same room, however there initial time taken to calculate each of the RIR's was noted and compared for reference.

Initially it was found that increasing the TO value from 2 to 4 increased the time taken to render by

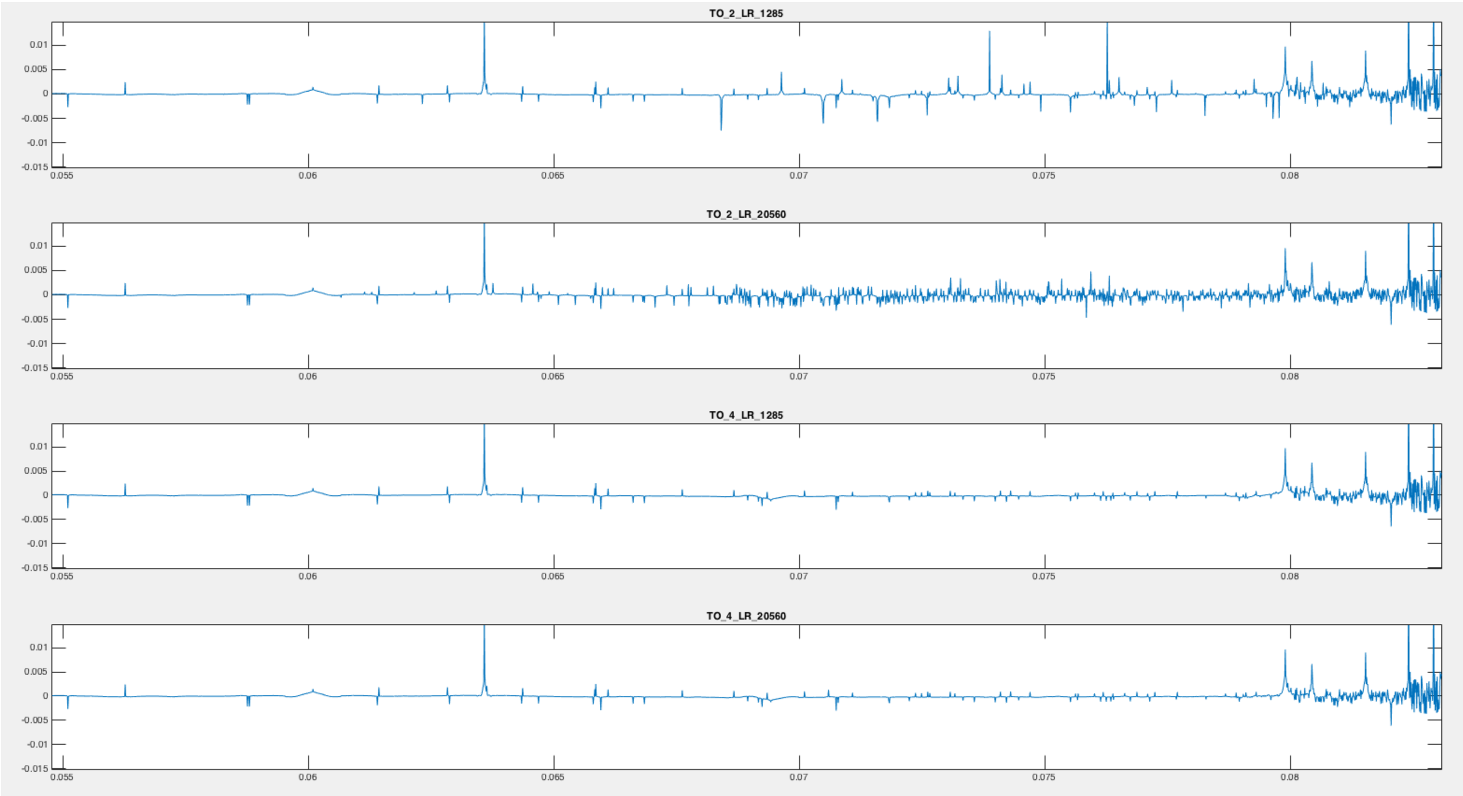



Figure 12: Zoomed in section showing the effects of using a high TO value

approximately 20 seconds, at which point it was decided that a TO value of 2 would be used, given that it made no perceptual difference. However, much later it realised that inaccurate calculation times may have been given due to the fact the project was being stored on the University network, thus adding variable delay times. The project was later stored on a local C drive, eliminating any network speed issues and inconsistencies. After doing so, the difference in TO values made almost no difference to the calculation times. In hindsight, a higher TO value could have been used to produce more accurate RIR's.

It was found that the LR's value made quite a difference to the calculation time:

| Transition Order | Late Ray Value | Calculation Time (s) |
|---|---|---|
| 2 | 1285 | 11 |
| 2 | 20560 | 28 |
| 4 | 1285 | 11 |
| 4 | 20560 | 28 |

Despite the increase in calculation time by 17 seconds, the inaccuracy of RIR's produced when using the lower LR's makes using the higher value justifiable.

As the Virtual Singing Studio (VSS) requires B-format audio files, the B-format option was selected in Odeons options menu.

#### 2.1.3.2 Directivity

Without selecting otherwise, Odeon uses an omni-directional point source as the sound source. This can be changed however by providing a .cf2 file. This file stores loudspeaker performance data and polar plots in what is known as a Common Loudspeaker Format, a standard used by loudspeaker manufacturers. By importing one of these files, Odeon simulates the directionality of the selected loudspeaker as the sound source. This can be used to attempt to more accurately recreate an RIR that would be taken in a real space, by finding the .cf2 file that corresponds to the loudspeaker used for the measurement.

It is also possible however, to input a custom directivity pattern to model a specific sound source. This provides an opportunity to create a more accurate sound source for a human head, something desirable when creating RIR's that will be convolved with the audio input from a singer. Techniques for recording the directivity of the human head have been reviewed and improved upon in [6], however only directivity data for the horizontal plane was recorded. Calum Armstrong, a student from the University of York has recently investigated the directivity of a human head in both the vertical and horizontal planes ranging from the 63Hz to 8kHz octave band [7] and generously provided said data to be used in this project. Figure 13 shows the directivity patten editor window in Odeon with the horizontal and vertical directivity patterns for the 8Khz octave band taken from [7].

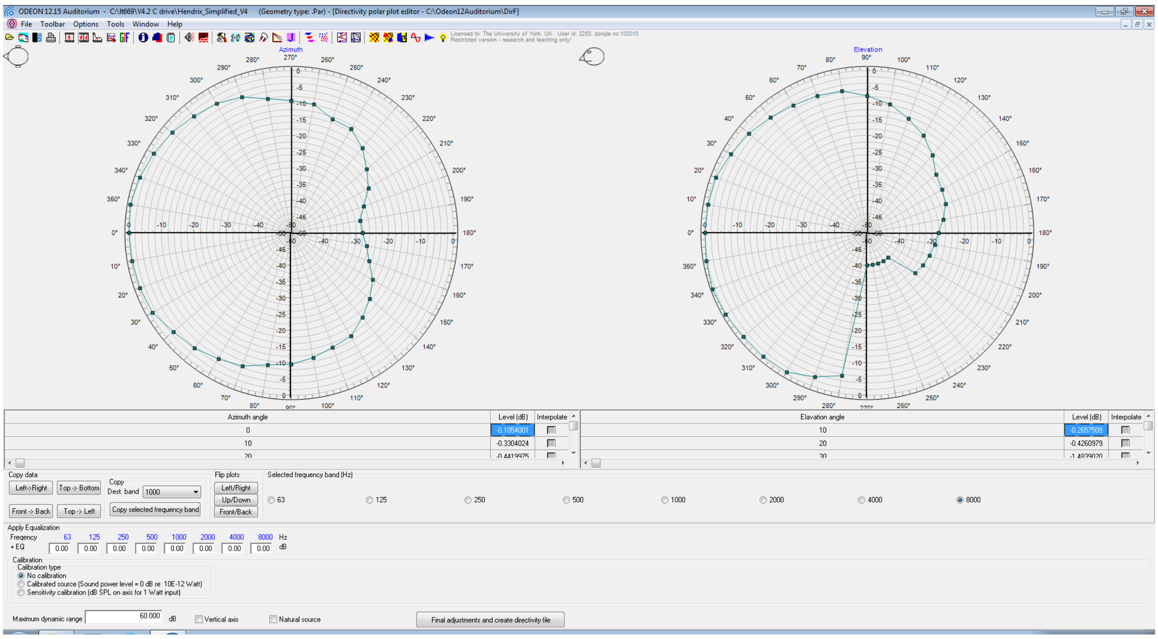


Figure 13: Directivity pattern editor window in Odeon showing the directivity pattern for the horizontal plane (left) and the vertical plane (right) for the 8Khz octave band using data from [7].

## 2.1.4) RIR Topology Problems

### 2.1.4.1 RIR Analysis

It has been noted that the RIR's produced will attempt to accurately resemble a human head. This involves placing the sound source below the receiver, to resemble the mouth (sound source) below the ears (receiver). When taking real RIR measurements it is often not possible to get the source and receiver close enough due to the physical dimensions of the equipment used and therefore result in an unnatural distance between them (see section **??**). However, with using software such as Odeon, where sound sources are calculated from a point source, it is possible to move the source and receiver much closer together. Initially, the sound source was placed 1m off the ground and the receiver was positioned 0.05m above the sound source. Three RIR's positioned along the centre of the room at varying distances from the left wall can be seen in figure 14. (These RIR's are taken from grid positions 76, 77 and 83 respectively which can be seen in figure 20 and is explained in section 2.1.5 RIR Locations)

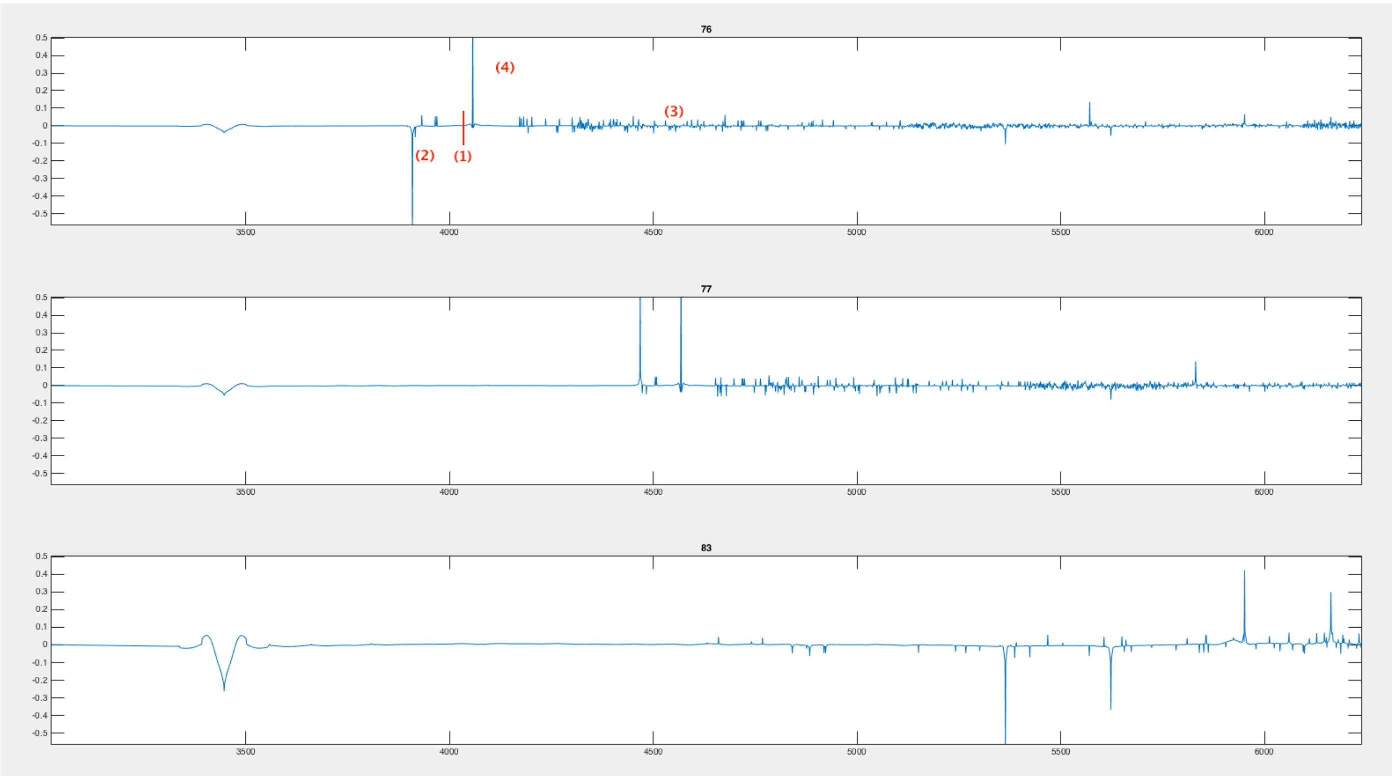



Figure 14: Initial RIR output with receiver 5cm above sound source showing RIR's from grid positions 76, 77 and 83 shown in figure 20

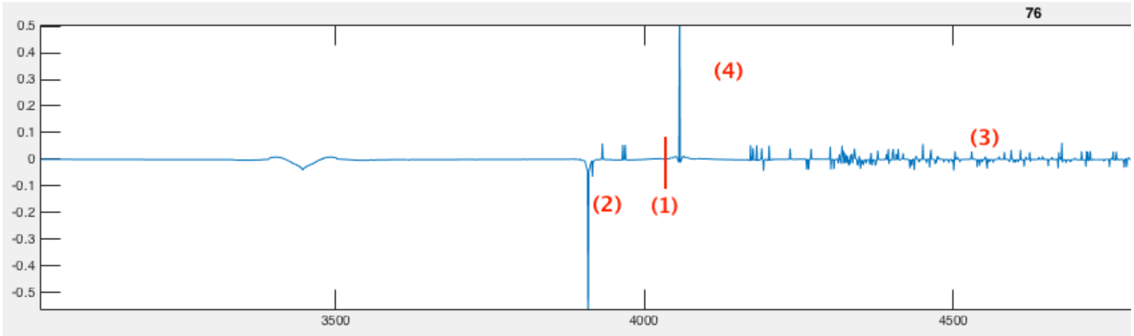



Figure 15: Initial RIR output with receiver 5cm above sound source from grid position 76

These RIR positions were chosen as the varying distance from the left wall can be used to estimate when to expect reflection other than that from the floor which will remain the same across each RIR. At the beginning of each plot, a dip of varying amplitude at the same instance in time can be seen. Without further inspection, this could be assumed as the direct sound from the source to receiver, which would be in the same place in time as the source and receiver are kept the same distance across all RIRs, though the varying amplitude is questionable.

Figure 15 shows the first RIR from figure 14 which is used the for analysis of early reflections. The RIR's can be validated by calculating where expected reflections should be, using the following equation:

$$t_s = t_p - \frac{d}{c} \tag{1}$$

where:

$t_s$ = Expected start of impulse (seconds)
$t_p$ = Time of first recorded reflection (first peak) (seconds)
$d$ = Distance between source and receiver (meters)
$c$ = Speed of sound ($344m/s$)

Therefore, if this dip is assumed to be the direct sound which occurs at $t_p = 0.03448s$ , where $d = 0.05m$, then the start of the impulse can be calculated to be:

$$0.03448s - \frac{0.05m}{344m/s} = 0.03433s = t_s \tag{2}$$

This is indistinguishable from the direct sound peak, which is understandable given the source and receiver are only separated by 0.05m. Now the expected start of this impulse is known, the expected first reflection due to the floor can be calculated:

$$\frac{2.05m}{344m/s} + t_s = 0.4029s \tag{3}$$

However, it can be seen that the first reflection occurs at (2), at a time of $0.039s$ suggesting that sound has travelled 1.6392m before reaching the receiver.

$$(0.039s - t_s) * c = 1.6382m \tag{4}$$

However, the shortest available distance a sound could travel other than directly from the source to the receiver is 2.05m (the first floor reflection).

As the source and receiver are placed 1.85m away from the left side wall (meaning sound has to travel slightly over twice this distance), a second reflection would be expected to occur at:

$$\frac{1.85m \times 2}{c} + t_s = 0.04509s \tag{5}$$

indicated by (3) in figure 15. However, it can be seen that no strong reflection occurs here.

After an email was sent to the technical department at Odeon [8], it was suggested that the small dip at the beginging of the RIR may be a bug in the version of Odeon being used. It could therefore be assumed that the direct sound is actually captured at the peak shown at (2) in figure 15 at a time of 0.03908. From this the new expected time for the start of the impulse could be calculated:

$$0.03908s - \frac{d}{c} = 0.03893s = t_s \tag{6}$$

It could then be assumed that the first reflection caused by the floor is the peak at (4) in figure 15 which occurs at a time of 0.04058s. However, this would suggest that sound has travelled 0.05532m:

$$(0.04058s - t_s) \times c = 0.05532m \tag{7}$$

meaning that it has reflected off a surface approximately 0.28m away, of which there is no surface.

It is also apparent that if the first peak other than the dip at the beginning of the file is taken to be the direct sound, it can be seen across the RIR plots shown in figure 14 that these would occur at different times depending on room position. This is also incorrect as it has been mentioned that the source and receiver distance is kept constant for all RIR simulations. The first strong reflection appears to be delayed in time the further away from the left side wall the RIRs are taken which would be accurate, however the timings are incorrect.

#### 2.1.4.2 Experiment

After some further investigation, it was found that the Odeon Manual [5] states in the section **Minimum distance from receiver to closest surface**:

> *"If a receiver is placed very close to a surface then results will be sensitive to the actual position of the secondary sources generated by ODEON's late ray method. If such a secondary source happens to be very close to the receiver, e.g. 1 to 10 centimetres, this may produce a spurious spike on the decay curve, resulting in unreliable predictions of the reverberation time [...] to avoid this problem it is recommended that distances to surfaces are kept greater than say 0.3 and 0.5 meters"*

Though this is a problem said to be caused by secondary sources, it can be assumed that this issue could also be caused by primary sources. This was investigated by producing RIR's with the source in a stationary location with receivers varying in distance above the source. Figure 16 shows six RIR's produced using different distances between the source and receiver where the numbers 1 - 6 indicate the following:

| Graph Number | Distance from floor | |
|---|---|---|
| | Source | Receiver |
| (1) | 1m | 1.05m |
| (2) | 1m | 1.15m |
| (3) | 1m | 1.30m |
| (4) | 1m | 1.50m |
| (5) | 1m | 1.60m |
| (6) | 1m | 1.70m |

The following numbers (1) - (6) refer to the individual plots in figure 16.

A 5cm distance between the source and receiver shown in (1) was originally used. This plot can be seen to be greatly different from the rest (***Note:*** *The reason the RIR shown in (1) looks different to previously investigated RIR's such as those in figure 14 is due to the plot being zoomed in, causing it to look greater in amplitude. This was done so it could be compared to (2) - (6) whilst being on the same scale*). It is in this RIR that the 'dip' found in the previously investigated RIR's is contained, whereas the RIR's with a greater distance between source and receiver (2) - (6) do not.

#### 2.1.4.3 Results

It is suggested in the Odeon manual (quoted above) that source and receiver be at least between 30cm and 50cm apart. Though a distance of 15cm (2) shows an RIR similar to the rest of the assumed correct ones, it is obvious that there is one strong direct sound, however it is difficult to see other reflections, whereas when the receiver is moved further away to a distance suggested by the Odeon manual, the early reflections become more obvious, with clarity of amplitude being more consistent through (4) - (6). The peaks in the RIR (5) (60cm) were investigated to assure they correctness, shown in figure 17. This RIR was chosen as the distance between the source and receiver for the real RIR's taken in Hendrix Hall was also 60cm. As the aim of achieving a more accurate human head topology is not possible due to this issue caused by Odeon, it was decided that for now, consistency is the next best thing.

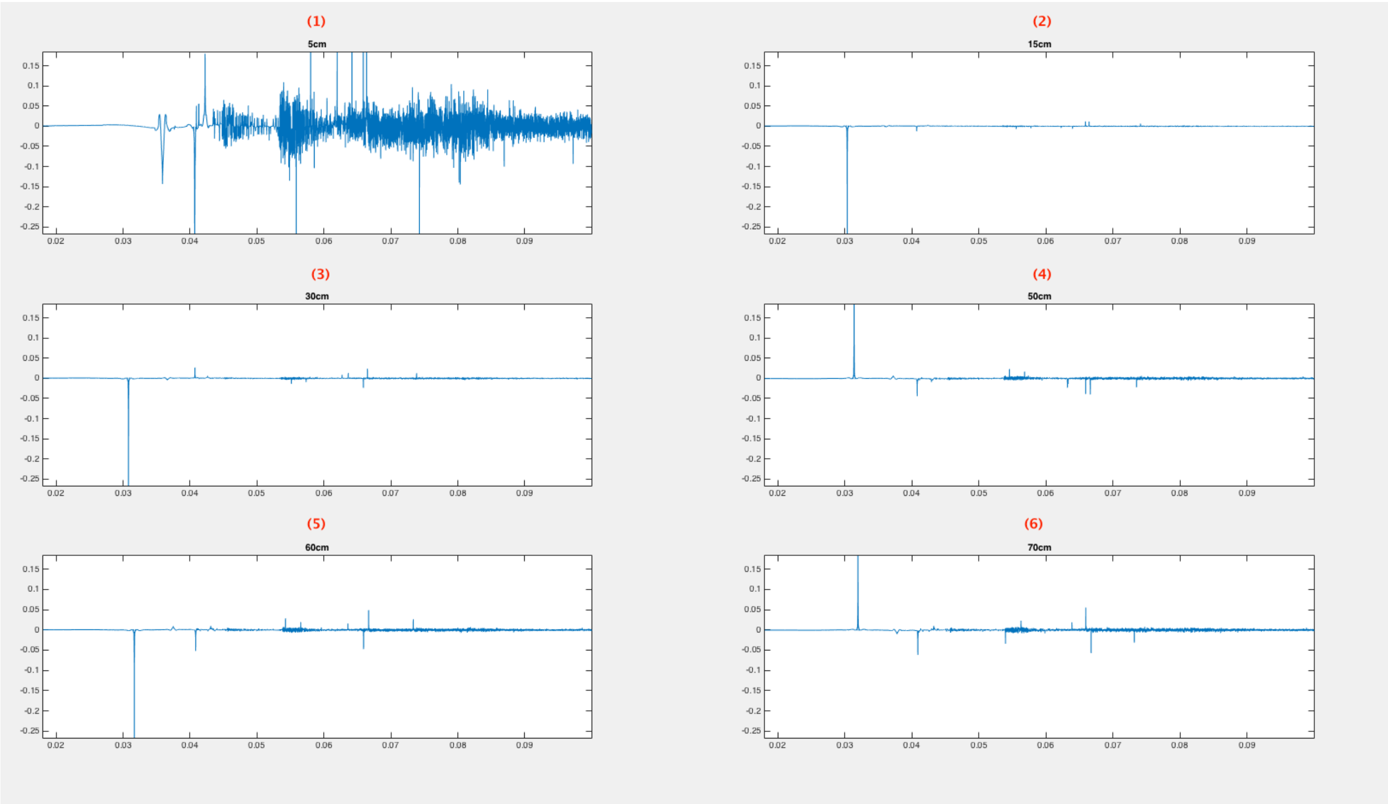



Figure 16: Plots of RIR's produced with the receiver placed at varying distances from the source, indicated by the title of each plot where **(1)**: 5cm **(2)**: 15cm **(3)**: 30cm **(4)**: 50cm **(5)**: 60cm **(6)**: 70cm

As can be seen, one the source has been moved away from the receiver by the amount suggested in the Odeon manual,

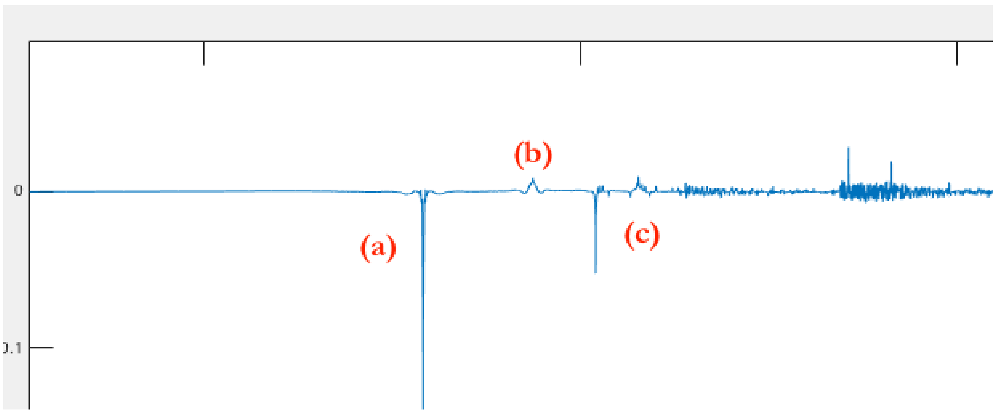



Figure 17: RIR produced when the receiver is places 60cm above the source.

The first peak (a) can be assumed to be the direct sound, present at $0.03168s$ meaning the start of the impulse can be calculated using equation 1, where now $d = 0.6$:

$$0.03168s - \frac{0.6m}{c} = 0.02994s = t_s \tag{8}$$

From this, it can be calculated that the peak at (b), which occurs at 0.0375$s$ had been captured after sound has travelled 2.6m:

$$(0.0375s - t_s) \times c = 2.6m \tag{9}$$

This is the reflection from the floor which is 1.6m below the receiver, causing the sound to travel 1m to the floor and then a further 1.6m to reach the receiver. This is shown in **Path 1** in figure 18. The peak at (c), occuring at 0.04084$s$ can be calculated as having travelled 3.75m which is approximately the distance travelled by a reflection caused by the left side wall shown as **Path 2** in figure 18.

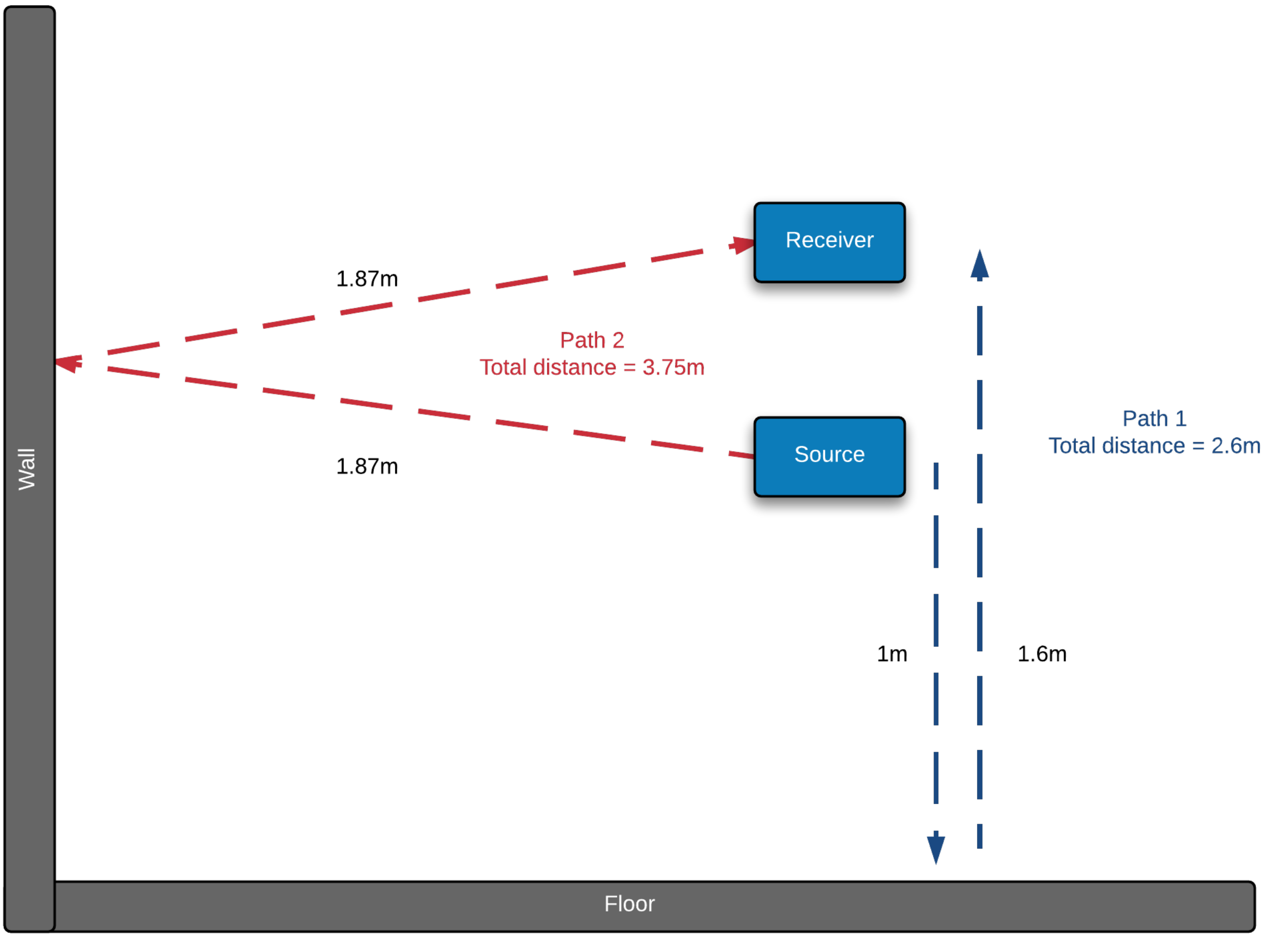



Figure 18: Illustration of the two early reflection paths due to the floor and left side wall when the receiver is placed 60cm above the source.

It can therefore be noted that the RIR's with a distance of 60cm between the source and receiver are accurate and can be used to produce a mass of RIR's.

### 2.1.5) RIR Locations

As discussed in section **??**, RIR's were to be taken every 1m starting from the centre of the room and should form a grid (rectangular or square). However, there were some positions within the room that would cause some RIR positions to be invalid. The location of these invalid positions could be found anywhere the sound source and receiver were placed closer than 0.75m next to a surface. This is because the loudspeakers in the VSS are placed 1.5m away from the centre, where the user will be stood meaning that any sound that has to travel less than 1.5m the reach the receiver will be impossible to reproduce at the correct time within the VSS. If a source and receiver were to be placed placed closer than 0.75m from a wall, sound would have to travel less than 1.5m before reaching the receiver. When taking this into account, figure 19 shows the positions that would be invalid due to them being placed too close to a surface, marked as U, V, W and Z. W represents a large space of invalid RIR positions and they would all be too close to the table surfaces.

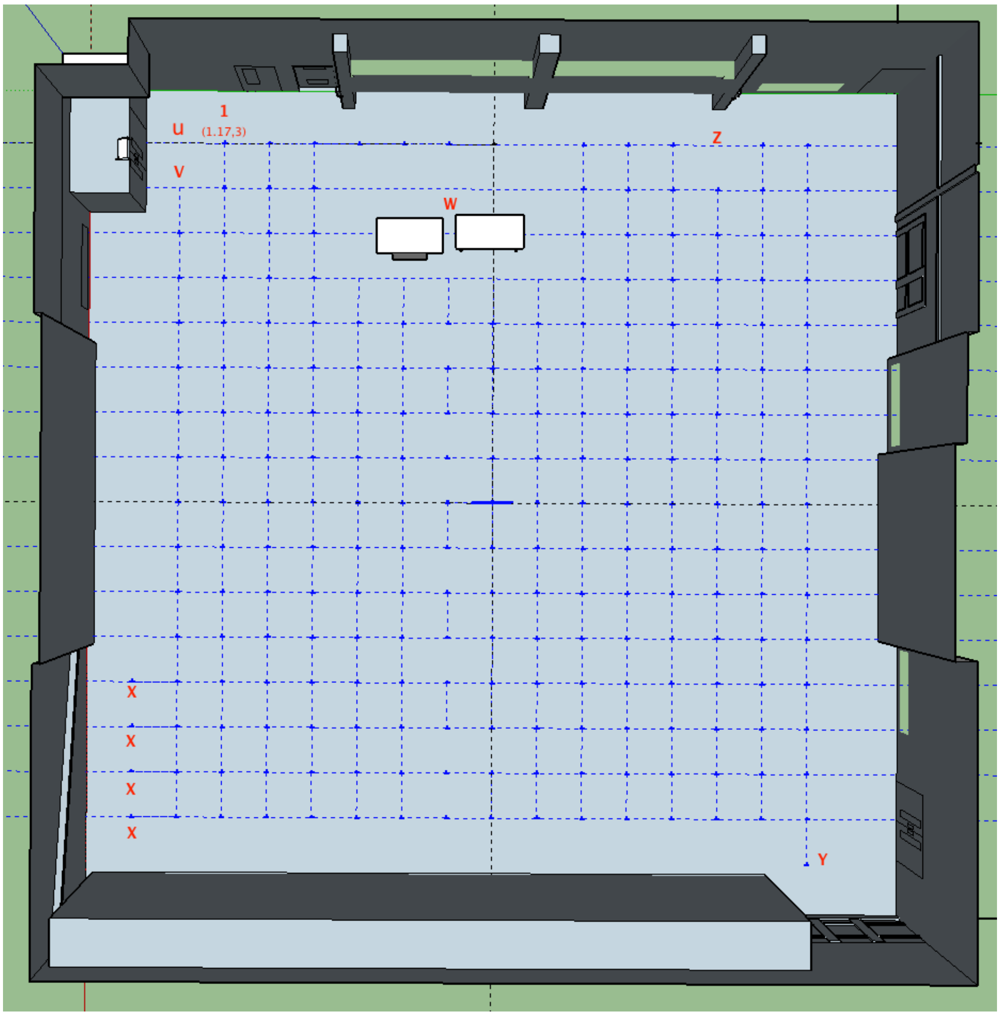



Figure 19: Top down view of the Hendrix Hall model indicating the RIR positions that are invalid due to being too close to a surface (U, V, W and Z), where other RIR's that are valid are not included as it would cause inconsistency (X and Y)

It was decided however that these RIR's would still be produced and used as part of the final system. If they were not included, there would be large gaps in the space where the user would have to jump between, thus be unable to produce what what feel like an consistent path.

It was also later discovered (see section **??**) that the latency due to the software used would mean that the beginning of the RIR's would have to be trimmed anyway, avoiding this issue.

There are several RIR's in figure 19 labeled X and Y. These positions would still produce valid RIR's, however they were not included in the final grid as they would cause inconsistencies in the positions available to the user. All other blue crosses in figure 19 were also included in the final RIR grid, producing a grid 15x16, totalling 240 available positions. Figure 20 shows the final maximum avilable RIR's locations.

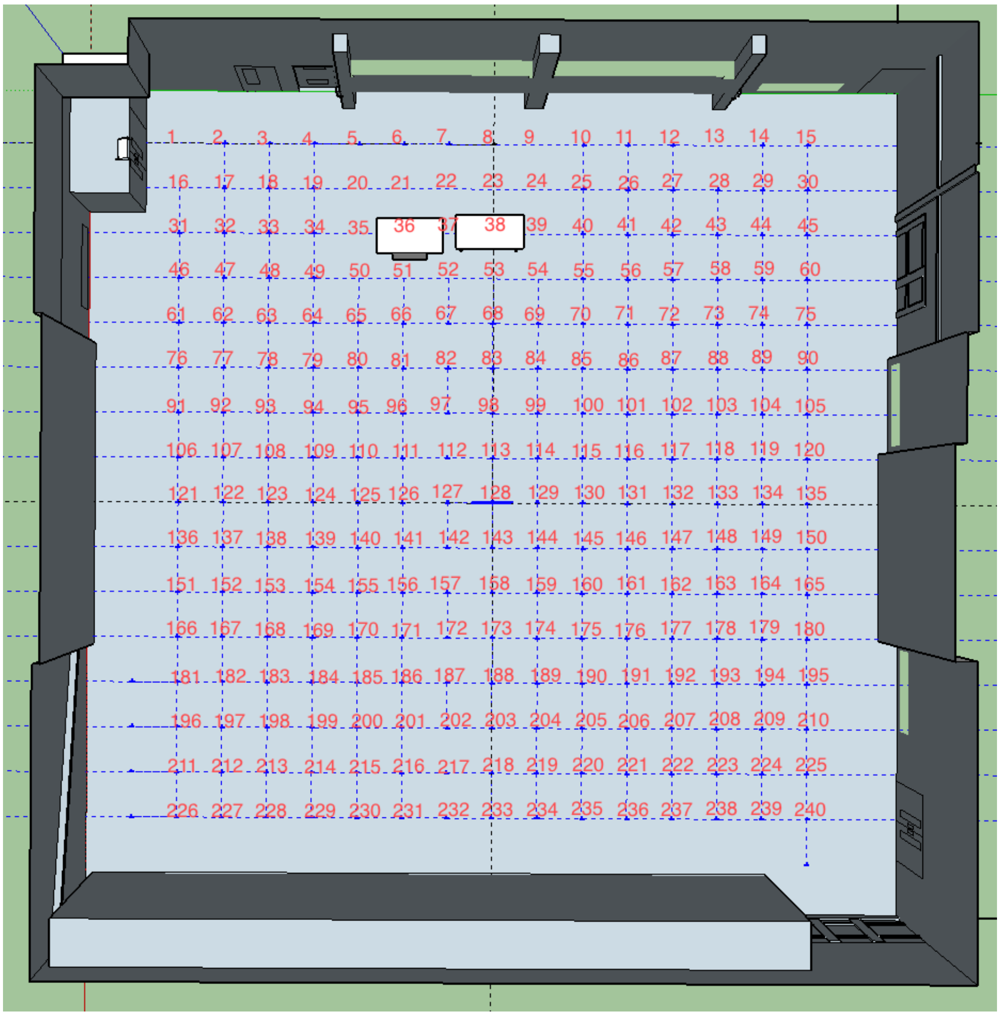


Figure 20: Top down view of the Hendrix Hall model showing the maximum available RIR locations

The tables present within the room so not cause an issue

## 2.1.6) RIR Rendering

The source and receiver positions can be set using a .SouRecScript file, which contains information regarding the Cartesian coordinates of both sources and receivers within the room as well as the sound source directivity patter and the direction in which the source is facing. This made producing 240 RIR's almost effortless. However, once the script had set to locations of the source and receivers, each source and receiver had to be paired up individually. This is because Odeon provides a **Job List**, where any number of RIR's can be rendered one after another automatically, allowing the process to run without user interaction. For this to work, the user must tell Odeon which receiver to use for each sound source. Figure 21 shows the job list menu where each source has to be connected to each receiver, using a series of drop down menus and scroll bars. This part of Odeon is clumsy and in no way designed to make the process fast and simple. This had to be done for all 240 RIR's and then again for the other three sets of RIR's used for facing different direction within the room, totalling 960 RIR's.

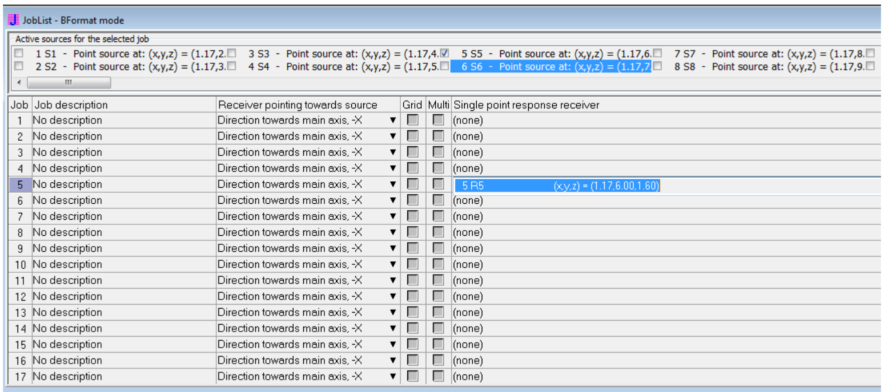



Figure 21: Source and receiver pairing window in Odeon

A naming system was used for the RIR files to indicate which grid positions they were taken from and which direction they are facing. For example: '008_0.wav' is is an RIR from grid position 8 with 0° rotation (facing the blackboards) where '128_90' is taken from grid position 128 and has a 90° rotation (facing the left side wall).

**Files available: SupportingFiles/sourecscripts/**

## 2.2 - File Sorting (Extra Grids)

As described in section **?? ??**, user test #3 requires there to be a number of available RIR grids to choose from, each containing RIR locations separated by a fixed distance. It was decided that there would be a total of 5 grids to choose from, the first (the one already produced) containing RIR's separated by 1m, the second grid containing RIR's separated by 2m and so on. Four Matlab algorithms were written (fully commented code available: renameSortFiles.s) to extract the appropriate RIR's required for each new gird, and rename them starting from 001 up to the maximum number of RIR's in each grid. This made it easily integrable with the software described in section **?? ??**.

The four grids can be seen in figures 31, 32, 33 and 34 respectively in Appendix B, showing which RIR locations from the original grid they contain.

## 2.3 - RIR Calibration

As both the real and synthetic RIR's were to be used in the user test, it was imperative that they were calibrated to be the same level to prevent a difference in perception when convolving with an audio signal. As the real RIR's were louder than the synthetic ones and the fact that there were far fewer real RIR measurements than synthetic ones, the real measurements were decreased in level to match those produced by Odeon.

### 2.3.1) Matlab

In an attempt to perfectly reduce the level of the real **RIRs!** (**RIRs!**) to match those of the synthetic ones, using Matlab, a multiplier was calculated by taking the RMS level of both the real and synthetic **RIRs!**:

```
Multiplier = RMS(real)/RMS(synthetic)

```

RIR's from the centre of the room were taken and used to calculated the multiplier. Figure 22 shows the plots of both of the original RIR's (top = real, centre = Odeon) and the calibrated version of the real RIR (bottom). As it can be seen, the peaks of the real RIR fall within the same range as the synthetic RIR. However, when these RIR's were convolved with an audio sample taken from OpenAir [9] (sing44.wav) **[LINK TO SAMPLE]**, it still sounded significantly louder than that of the synthetic RIR convolved with the same audio sample.

This can be heard in audio examples: **calRMSConv.wav and odeonConv.wav [LINK]**

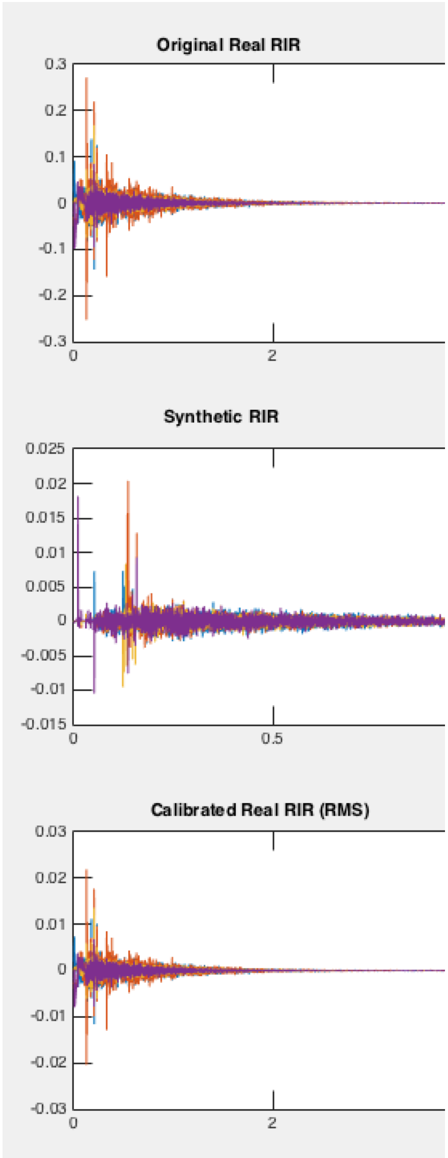



Figure 22: Plots of the original real RIR (top), the synthetic RIR (centre) and the calibrated real RIR (bottom) using the multiplier calculated using the RMS ratio value showing that the peaks of the calibrated RIR sit approximately within the same range as the synthetic RIR

As can be seen in figure 23, the peaks in the signal convolved with the calibrated RIR are approximately twice as large as those in the synthetic RIR convolved with the sample thus is perceived as louder.

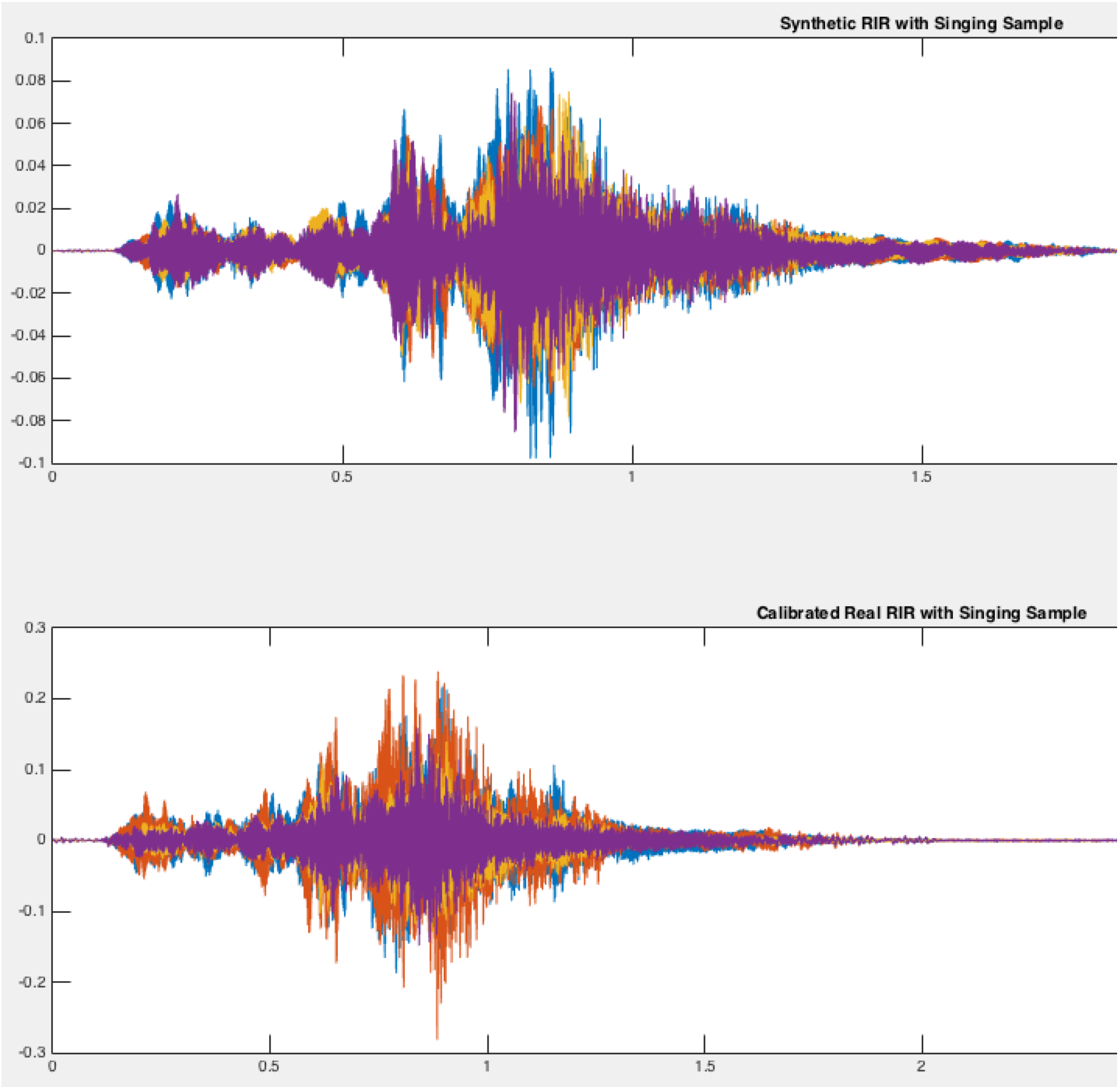



Figure 23: Plots of the synthetic RIR (top) and the RMS calibrated real RIR (bottom) convolved with an anechoic audio sample showing that the speaks of the calibrated RIR convolved sample sit approximately in the range double that of the synthetic RIR convolved sample, thus is perceived as louder.

Several manual values for the multiplier were tried starting from 0.7 down to 0.2. It was found that a value of 0.3 can the closest level similarity. Figure 24 shows that the level of the manually calibrated RIR falls within values half the magnitude of those found in the synthetic RIR. As the RMS calibrated audio sample shows that its magnitude sits within value twice as large as the synthetic RIR sample, now that the peaks of the manually calibrated RIR sit within half the synthetic RIR's range, it makes sense that the audio samples now approximately lay within the same range, as can be seen in figure 25.

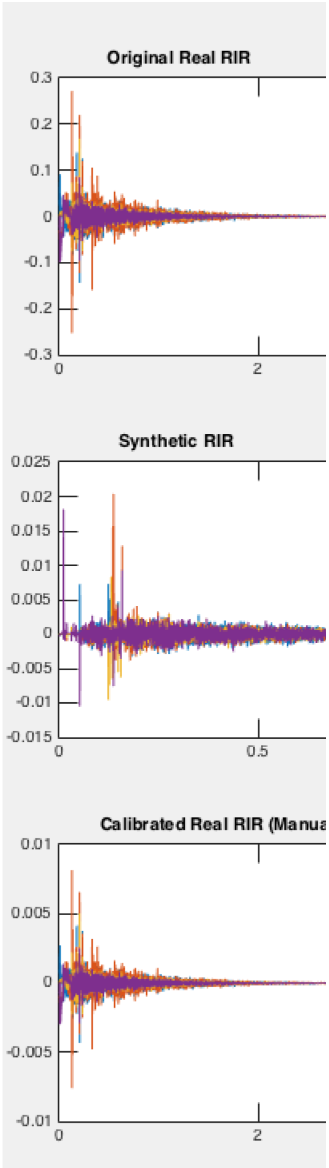



Figure 24: Plots of the original real RIR (top), the synthetic RIR (centre) and the *manually* calibrated real RIR (bottom) showing that the peaks of the manually calibrated RIR sit within half of the range of the synthetic RIR.

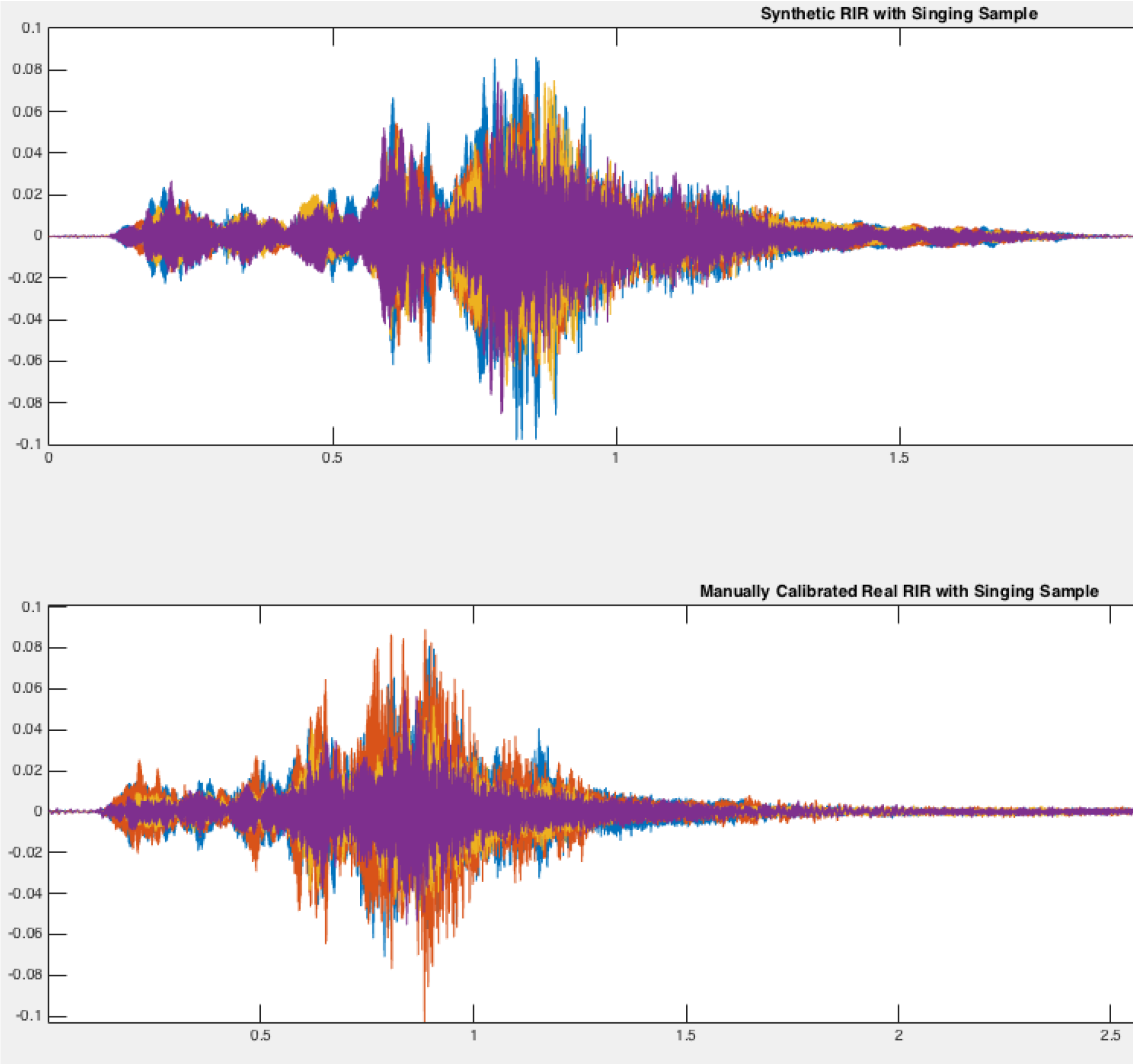



Figure 25: Plots of the synthetic RIR (top) and the RMS calibrated rel RIR (bottom) convolved with an anechoic audio sample showing that the speaks of the calibrated RIR convolved sample sit approximately in the range double that of the synthetic RIR convolved sample.

After listening to the convolved audio file more closely, it could be heard that the tail of the reverb seemed to increase in volume, producing what sounded like a 'swelling' effect. Figure 26 shows the reverb tail of the manually calibrated convolved audio samples, there small pulses can be seen (mostly in the channel shown in yellow). Further investigation did not reveal the cause of the pulsing, though it was found that it occurred every 0.05s, making it a 20Hz pulse. The effect of the pulse can be heard in the audio sample: pulseExample.wav **[LINK HERE]**.

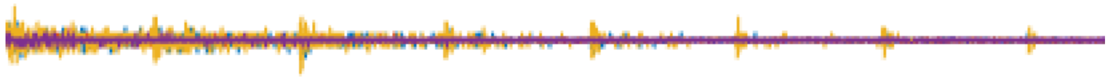


Figure 26: Reverb tail of the audio sample convolved with the manually calibrated RIR showing a pulse like pattern.

As the cause of the problem could not be uncovered within the time allocated, a solution was found.

**Matlab file used: calibrateALL.m**

### 2.3.2) Reaper

All of the real RIR files were imported into Reaper and send through a bus to a master channel in order to normalise all RIR's evenly. These were then compared by ear to a synthetic RIR and manually level matched. The closest match was by reducing the level of the real RIR's by 28.6dB. Though this method is less accurate, it prevented any unwanted errors in the signal such as those caused by using Matlab.

## 2.4 - Synthetic RIR Rendering Summary

As stated in section **?? ??**, one benefit of using the direct RIR rendering method is that it is simple. The process of obtaining the desired RIR's has shown that though time consuming, no complex procedures were undertaken therefore the statement is true. The fact that this process was time consuming can be put down to two points:

1. In experience
2. Starting from scratch

Though in this case, the dimensions of the building in question had to be obtained, it would be much faster to work from existing blueprints.

A benefit of using Odeon the obtain these RIR's is that it has allowed for the directivity of a human head has been modelled, thus providing a more accurate sound source directivity pattern than would have otherwise been achieved and was no more hassle to do than rendering RIR's from a point source.

Issues faced during this process were caused by using the software Odeon itself. Though the material list could be edited and built upon, the fact that its initial available options are lacking instantly reduced the possibility for an accurate simulation of room acoustics. It has been shown that it is possible to achieve a close to desired result through the manipulation of the material list, however this issue coupled with the already stated issues of geometrical acoustic modelling methods regarding their accuracy diminishes the plausibility of the system. In addition, the extra time spend on tweaking the resultant RIR's is not desirable, however this can be said for any modelling technique that involves simulation, thus is not restricted to this project alone.

Though it has been shown that producing 960 RIR's can be done sat in front of a computer, Odeon is not optimised for such a task. As stated in 2.1.6 RIR Rendering, though it is possible to import a siple script to set all of the source and receiver locations, having to join them up one at a time takes the best part of an hour and for this project has to be done four times over. This undesirable software flaw would be simply fixed by adding the option to pair them up within the source and receiver script.

In summary, a large grid of 960 RIR's in 240 locations within a virtual space of choosing have been relatively easily produced, providing a good approximation of its room acoustics. The only real issue faced was software that is not optimised for such a task.

# Appendices

## APPENDIX A

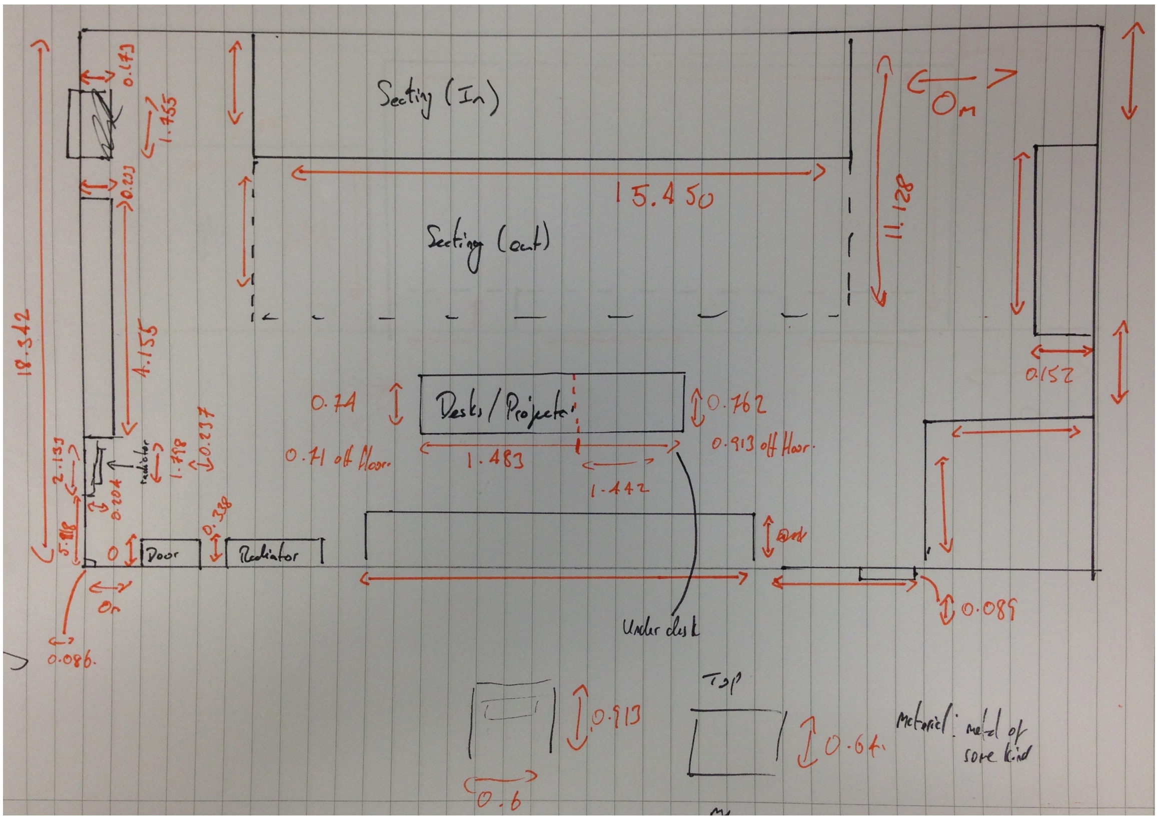



Figure 27: Annotated blueprint of Hendrix Hall from a birds-eye view

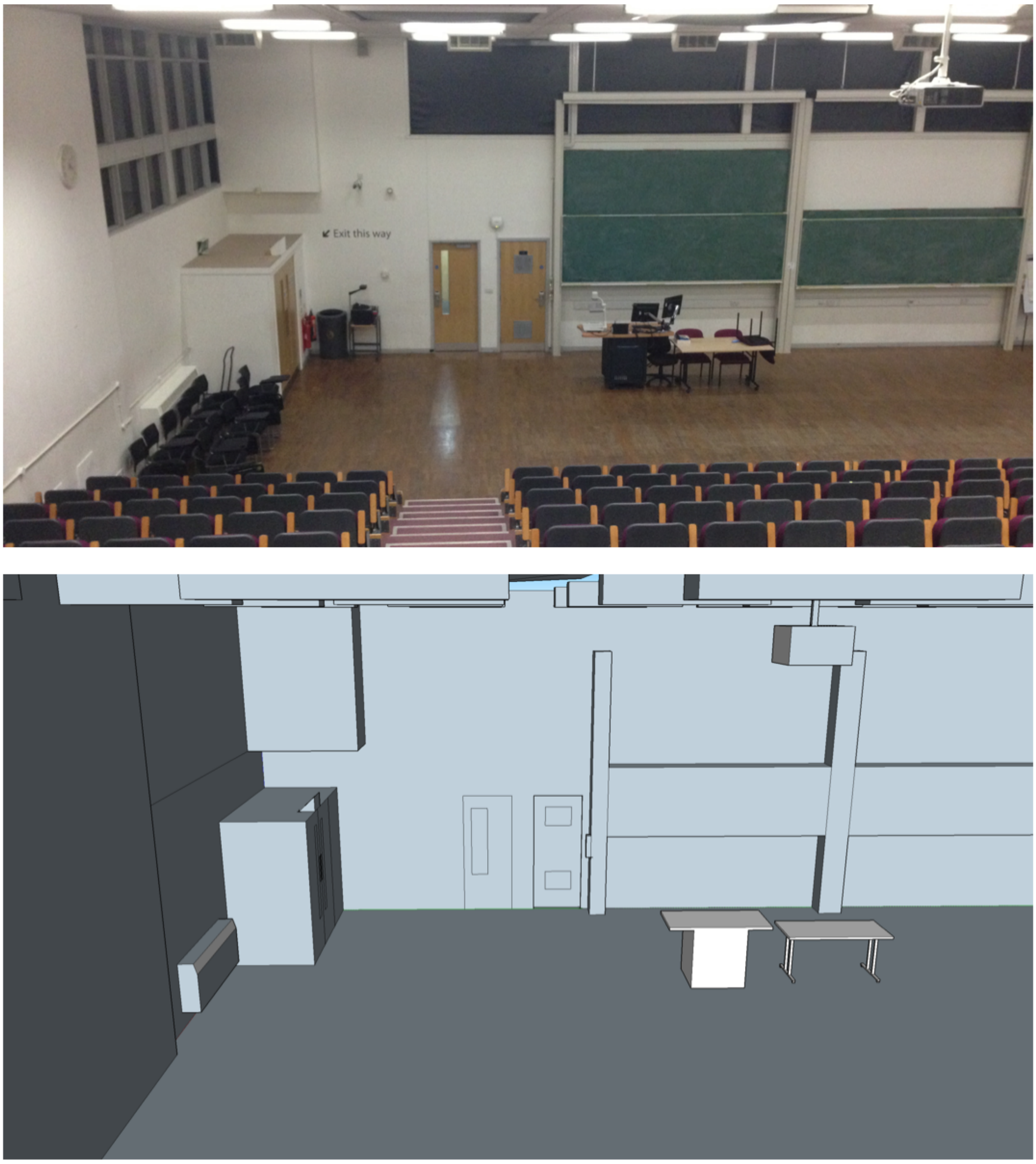



Figure 28: Image comparing a picture taken in Hendrix Hall (top) Compared to an image taken from the Google SetchUp Model (bottom)

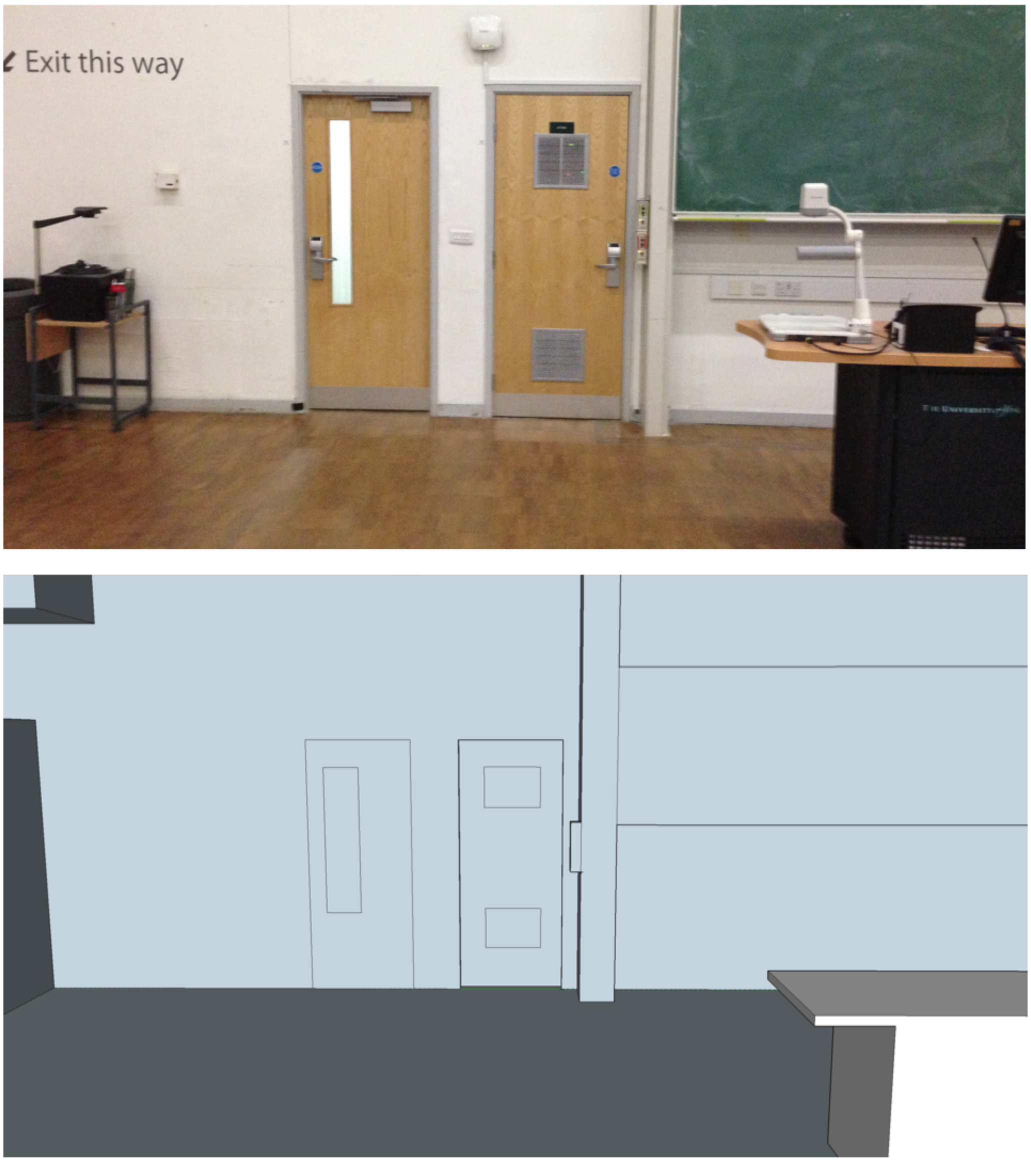



Figure 29: Image comparing a picture taken in Hendrix Hall (top) Compared to an image taken from the Google SetchUp Model (bottom)

Figure 30: Image comparing a picture taken in Hendrix Hall (top) Compared to an image taken from the Google SetchUp Model (bottom)

# Appendix B

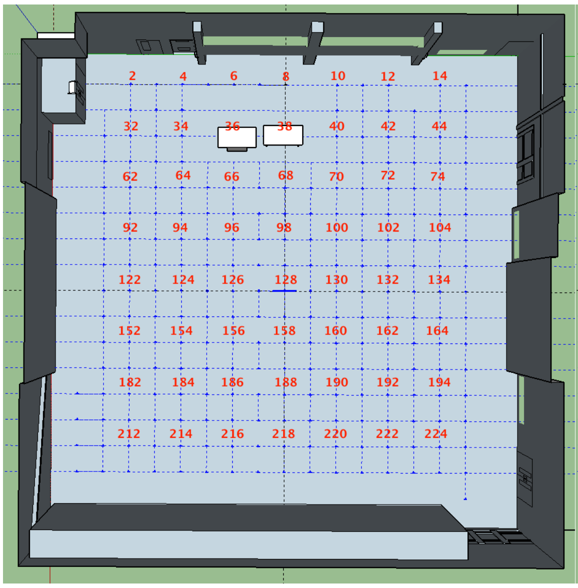



Figure 31: RIR grid with 2m separation

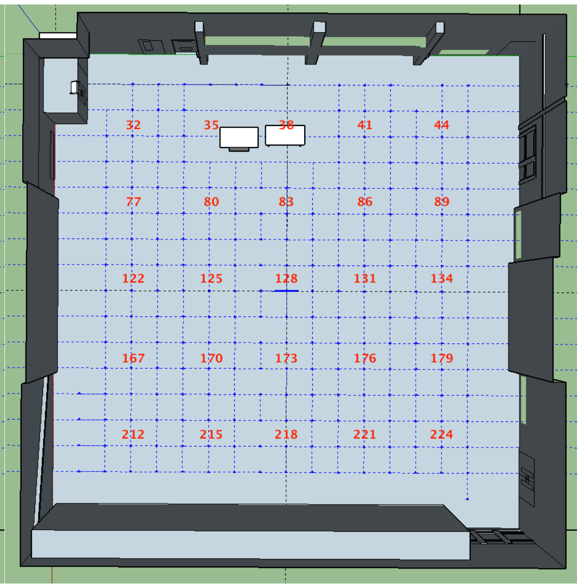



Figure 32: RIR grid with 3m separation

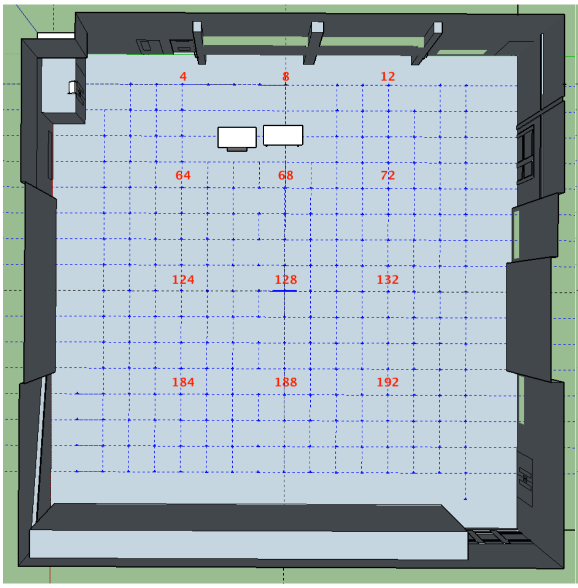



Figure 33: RIR grid with 4m separation

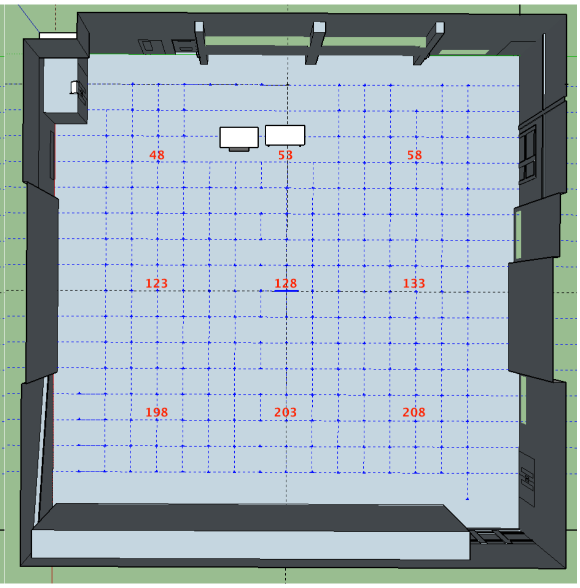



Figure 34: RIR grid with 5m separation

# APPENDIX C

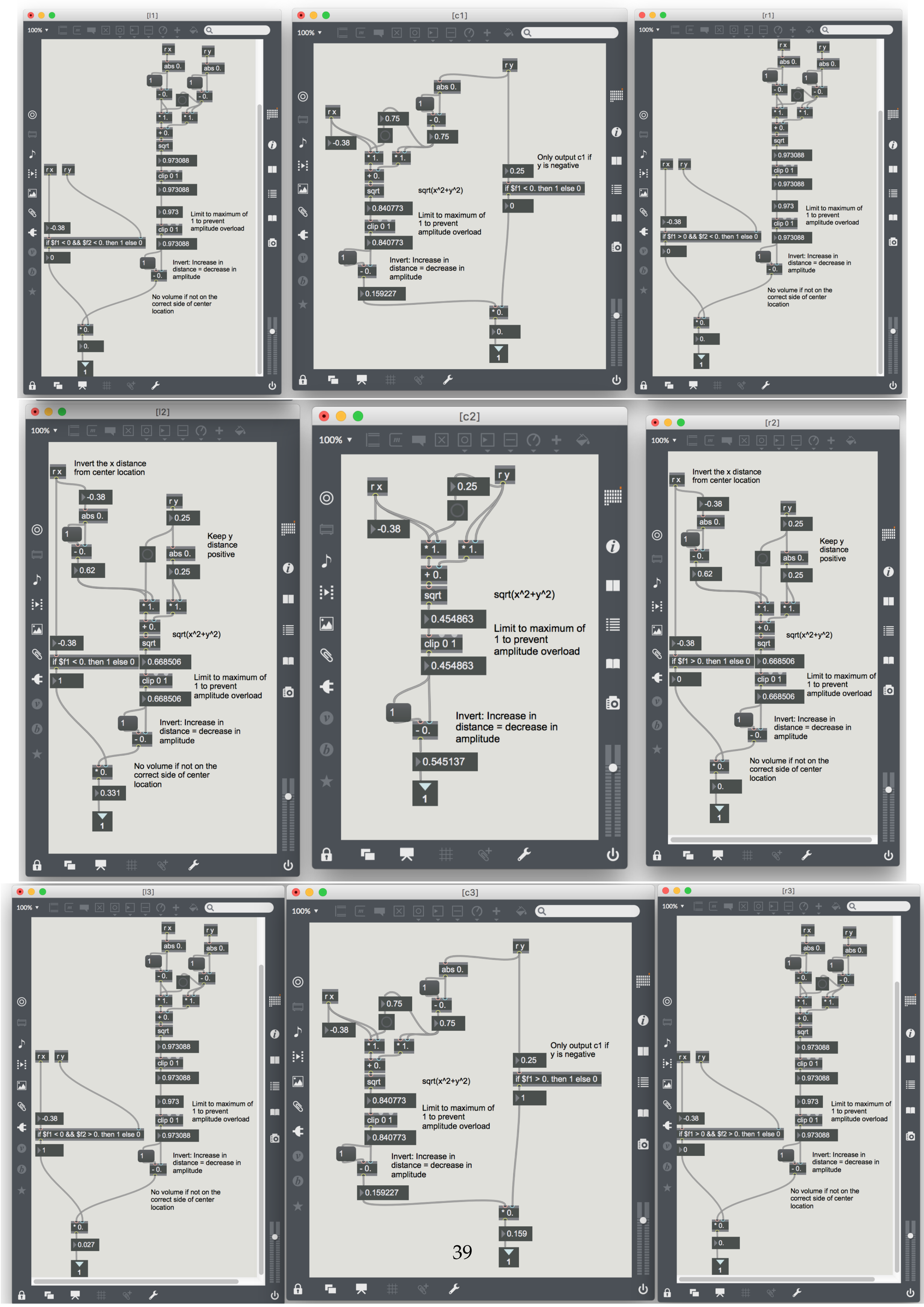



Figure 35: Overview of the individual panning algorithms used in iteration 1.

```
inlets = 5;
outlets = 5;

//Create arrays to store previous positions
var xArray = new Array(2);
var yArray = new Array(2);

var windowSize = new Array(2);

//Variables to use for file searching
var fileX, fileY, search;

//Defines how to split up the grid
var numberOfMeters;

//Loads appropriate files given users finger coordinates
function msg_int(input){
 if(inlet == 0){
  xPos = input;
 } else if (inlet == 1){
  yPos = input;//Add off set to start at (0,1)
 } else if (inlet == 2){
  windowSize [0] = input;
 } else if (inlet == 3){
  windowSize [1] = input;
 } else if(inlet==4){
  numberOfMeters = input;
 }

 //Split into sections
 if(numberOfMeters == 3 || numberOfMeters == 5){
  //Even grid for 3m and 5m
  xPosition = (xPos/windowSize[0])*(numberOfMeters);
  yPosition = (yPos/windowSize[1])*(numberOfMeters);
 } else if (numberOfMeters == 4 || numberOfMeters == 8){
  //4m separation requires different x,y coordinate scaling
  xPosition = (xPos/windowSize[0])*(numberOfMeters-1);
  yPosition = (yPos/windowSize[1])*(numberOfMeters);
 } else{
  //Extra row for others
  xPosition = (xPos/windowSize[0])*(numberOfMeters);
  yPosition = (yPos/windowSize[1])*(numberOfMeters+1);
 }

 //Round to nearest value
 xSection = Math.round(xPosition);
 ySection = Math.round(yPosition);

 //Start the lcd grid sections from column 1 row 1 instead of column 0 row 0
 if(xSection == 0){
  xSection = 1;
 }
```

```
 if(ySection == 0){
  ySection = 1;
 }

 //Distance in % away from center of section
 xBetween = 2*(xPosition - xSection); //x2 to get 100%
 yBetween = 2*(yPosition - ySection);

 //Which RIR to load in centre location
 outlet(0,xSection);
 outlet(1,ySection);

 //Output panning values
 outlet(2,xBetween);
 outlet(3,yBetween);

 //Store current location
 xArray[0] = xSection;
 yArray[0] = ySection;

 //If either coordinate is changed search for new files
 if(xArray[0] != xArray[1] || yArray[0] != yArray[1]){

  if(xArray[0] != xArray[1]){
   //Store previous value
   xArray[1] = xArray[0];
   X = xArray[0];
  }

  if(yArray[0] != yArray[1]){
   yArray[1] = yArray[0];
   Y = yArray[0]
  }

  //Output user location within grid
  if(numberOfMeters == 4 || numberOfMeters == 8){
   fileNumber = X + ((numberOfMeters-1)*(Y-1)); //Requires different algorithm for 4m
     due to different grid shape
  } else {
   fileNumber = X + ((numberOfMeters)*(Y-1));
  }
  outlet(4,fileNumber);
 }
}
```

Sections/Appendix/AppendixA/Code/loadFilesLogic.js

# Appendix D

# Test Participant Form

You have volunteered to partake in two user tests that should take no longer than 30 minutes to complete.

## Test Descriptions

The VSS (virtual singing studio) is a system that is used to simulate the acoustics of another room. The system can be used by standing in the centre of the speaker array and singing into a head mounted microphone. By wearing the provided head-tracking device, you can turn in the virtual space by turning your head/body.

### Test #1

This test aims to investigate the perception of movement within the virtual acoustic environment when using two different methods: Method **A** and Method **B**. You will be asked to step inside the VSS and say the word “Bob”. Your location within the virtual space will then be changed and you will be asked to produce another sound. This process will then be repeated a second time but this time using method **B**. You will then be asked to state whether method **B** felt like you had:

Moved a **shorter** distance than I had in **A**
Moved the **same** distance as I had in **A**
Moved a **further** than I had in **A**
I don’t know

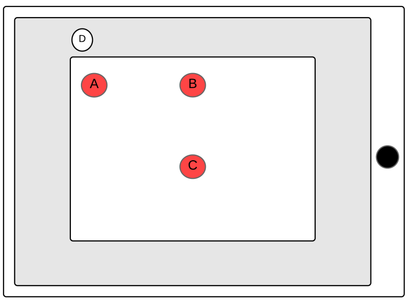
This process will be repeated 5 times in total.

---

### Test #2

**Part 1:** You will be asked to step inside the VSS and to sing or produce a noise. After a short amount of time you will be asked to do the same again. You will then be asked whether you feel you have changed location or not with a simple Yes/No answer. This will be repeated 8 times.

**Part 2:** In this part of the test, you will be asked to change your location within the virtual space yourself by tapping on a location or dragging your finger around the iPad provided for you. You will be asked to rate on a scale of **1 - 10** how free you feel you can move about the room with **1** being a *jumpy movement* and **10** being *complete freedom to move without limitations.* You will also be given the opportunity to add comments to further explain you score if you wish.




To the left is a diagram of an iPad. **A**, **B**, and **C** indicate where parts of the room can be located. When situated in the VSS, you will start in the center of the room **(C)** facing towards the front of the room **(B)**.

A = Top left corner of the room
B = Front of the room
C = Centre of room
D = Button to calibrate head tracker

## Answering Question

Note that when you’re within the VSS it will be difficult to write down your answers to the questions asked. Therefore you will be asked to answer verbally and your answers will be taken down for you. You will be asked at the end of the test to check that your answers have been taken down truthfully.

## Information and Consent

Experimenter: _______________

Please read the following statements and tick the boxes on the right hand side to indicate that you understand and agree.

| | |
|---|---|
| I understand that at any point I may choose to withdraw from the experiment | ☐ |
| I understand that I may omit answers to any questions | ☐ |
| I agree that I am here voluntarily | ☐ |
| I understand and agree that the experimenter conductor will be observing the experiment | ☐ |
| I agree that the system being used has been explained to me | ☐ |
| I agree that the point of this experiment has been explained to me | ☐ |

Participant Signature: _______________

## Answer Sheet

Participant Number: _______________ Date: _______________

**Test #1**

**Question 1:** Please state whether you feel you have:

Moved a **shorter** distance than I had in **A**
Moved the **same** distance I had in **A**
Moved a **further** distance than I had in **A**
I don't know

| **Trial** | **Score** | | | |
|---|---|---|---|---|
| | Shorter | Same | Further | Don't Know |
| 1 | ○ | ○ | ○ | ○ |
| 2 | ○ | ○ | ○ | ○ |
| 3 | ○ | ○ | ○ | ○ |
| 4 | ○ | ○ | ○ | ○ |
| 5 | ○ | ○ | ○ | ○ |

I agree that the answers that have been taken down on my behalf are correct ☐

Participant Signature: _______________

Participant Number: _______________ Date: _______________

**Test #2 - Part 1**

**Question 2:** Do you feel you have changed location within the room?

| Trial | Answer |
|---|---|
| 1 | YES/NO |
| 2 | YES/NO |
| 3 | YES/NO |
| 4 | YES/NO |
| 5 | YES/NO |
| 6 | YES/NO |
| 7 | YES/NO |

**Test #2 - Part 2**

**Question 3:** Please rate on a scale of **1 - 10** the mobility within the virtual space where **1** = Extremely "jumpy" movement and **10** = Completely smooth movement or please select "N/A" if you can not tell you are moving.

| Trial | Score | | | | | | | | | | N/A |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |
| 2 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |
| 3 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |
| 4 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |
| 5 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |

**Comments:**

I agree that the answers that have been taken down on my behalf are correct ☐

Participant Signature: _______________

**Question 4:** Please rate on a scale of **1 - 10** the mobility within the virtual space where **1** = Extremely staggered movement and **10** = Completely smooth movement or please select "N/A" if you can not tell you are moving.

| **Trial** | **Score** | | | | | | | | | | **N/A** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |
| 2 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |
| 3 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |
| 4 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |
| 5 | 1 ○ | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○ | ○ |

**Comments:**

I agree that the answers that have been taken down on my behalf are correct □

Participant Signature: _______________

pagebreak

# APPENDIX E

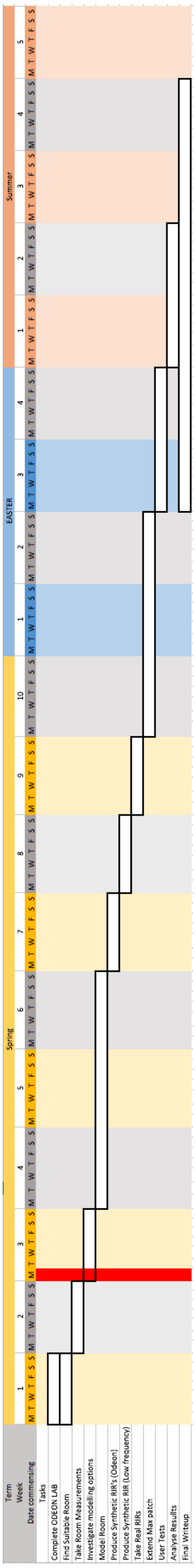



Figure 36: Initial Gantt chart showing the approximate time allocated to each of the tasks required to complete the project. The red line shows the time at which the Gantt chart was abandoned for a more appropriate planning method

# References


[1] (2016). Su2odeon, Odeon, [Online]. Available: http://www.odeon.dk/su2odeon-plugin-trimble-sketchup (visited on Apr. 25, 2016).

[2] (2016). Absorption coefficients, AcousticTraffic, [Online]. Available: http://www.acoustic.ua/st/web_absorption_data_eng.pdf (visited on Apr. 25, 2016).

[3] (2016), BemBook, [Online]. Available: http://www.bembook.ibpsa.us/index.php?title=Absorption_Coefficient&action=edit (visited on Apr. 25, 2016).

[4] P. Kovalchik, R. Matetic, S. Peng, and G. Cole, "Measurement method for determining absorption coefficients for underground mines," pp. 1–4, [Online]. Available: http://www.cdc.gov/niosh/mining/userfiles/works/pdfs/ammfd.pdf (visited on Apr. 25, 2016).

[5] C. L. Christensen and G. Koutsouris, *Odeon user's manual*, Scion DTU Diplomvej Kgs. Lyngby Denmark, 2015. [Online]. Available: http://www.odeon.dk/pdf/ODEONManual12.pdf (visited on Apr. 25, 2016).

[6] B. B. Monson, E. J. Hunter, and B. H. Story, "Horizontal directivity of low- and high-frequency energy in speech and singing.," *The Journal of the Acoustical Society of America*, vol. 132, no. 1, pp. 433–41, 2012, ISSN: 1520-8524. DOI: 10.1121/1.4725963.

[7] C. Armstrong, "A filter based approached to simulating acoustic radiation patters through a spherical array of loudspeakers," (Unpublished Masters Dissertation, University of York, May, 2016).

[8] L. Thresh, *Impulse response analysis question*, Personal Email, Odeon Support, Mar. 2016.

[9] A. F. Marina Theodoropoulou. (2016). Operatic voice, OpenAir, [Online]. Available: http://www.openairlib.net/anechoicdb?keys=singing&field_sound_cat_value=All (visited on May 5, 2016).