Room Acoustics and Virtual Reality: An implementation of auralisation and 360 degree image techniques to create virtual...

Alexander Vilkaitis
University of Derby

Adam J. Hill
University of Derby

Charlie Middlicott
University of Derby

Bruce J Wiggins
University of Derby

ROOM ACOUSTICS AND VIRTUAL REALITY: AN IMPLEMENTATION OF AURALISATION AND 360 DEGREE IMAGE TECHNIQUES TO CREATE VIRTUAL REPRESENTATIONS OF SPACES

A Vilkaitis Department of Electronics, Computing & Mathematics, University of Derby, UK
M Dring Department of Electronics, Computing & Mathematics, University of Derby, UK
C Middlicott Department of Electronics, Computing & Mathematics, University of Derby, UK
B Wiggins Department of Electronics, Computing & Mathematics, University of Derby, UK
AJ Hill Department of Electronics, Computing & Mathematics, University of Derby, UK

1 INTRODUCTION

There has been a huge increase in enthusiasm for virtual reality in recent years. Spatial audio is of significant importance when creating virtual reality content if the experience is to be perceptually congruent. This project aims to intersect the worlds of virtual acoustic auralisation and virtual reality, creating a novel method of demonstrating room acoustic environments with maximal audio visual impact in a user friendly fashion. An open source library of 3D impulse responses together with 360° image/video capture using a variety of techniques will be created in different spaces (and positions within). Various spaces will be measured and analysed including classrooms, music venues, buildings of historical interest and theatres. As well as impulse response (IR) measurements, 360° images will be recorded using photospheres, captured on android smart phones [1] and the Ricoh theta S [2]. Future applications for these impulse responses will be the development of a virtual mixing tool, where the user will be able to experience mixing live performances within an auralised virtual environment, a method of allowing audiences to view and hear auralisations of different seating positions within a space for ticketing and marketing purposes, and the possibility of a real time auralised virtual concert.

2 LITERATURE

2.1 Impulse Responses, Auralisation and Virtual Reality

Impulse responses, auralisation and virtual acoustics, in their various forms and implementations, have been well documented in the literature. The use of auralisation and virtual reality can help to bridge the gap between acousticians, architects, and lay persons. Virtual acoustic demonstrations have an inherent ease of comprehension, clarity and impact that traditional, data-based methods may not when viewed by a multidisciplinary team working on a particular project or the general public. They also allow the combination of the architect's vision of a space with the acoustician's design, and bring them into a prototype that can be experienced during the planning stage. The methodology also allows users to experience existing spaces using standard impulse response capture methods and 360° images for a variety of applications and purposes.

2.1.1 Ambisonics

Ambisonics is a spatial audio format where the aim is to accurately reproduce the originally recorded sound field. The minimum requirement for full 3D ambisonic capture and reproduction is four channels, while more channels increase the spatial resolution. Encoding and decoding of ambisonics is separate which creates a flexible system that can be played back over a variety of speaker array sizes. In this case, first order ambisonics is utilised, however the use of higher order ambisonics for future applications is in the early development stages.

B-format ambisonics has four channels, WXYZ, where W contains omnidirectional mono information, X one horizontal plane, Y the perpendicular horizontal plane, and Z for height information. Channels

X, Y and Z are directional figure of eights, pointing forward, leftward and upward (Figure 1) with a maximal directional gain being 3dB above W, so the signals have approximately equal energy compared to the omnidirectional W channel[3].

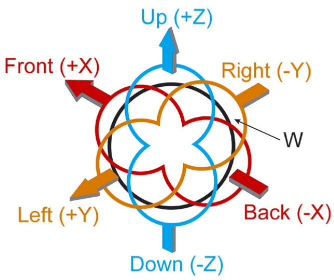

Figure 1 B-Format [4]

Traditional ambisonics uses large speaker arrays in order to provide the 3D listening experience which are cumbersome, and impractical for commercial and consumer use. Binaural audio uses Head related Transfer Functions, which are filters that describe inter-aural differences between both ears when a source is at a certain location, and create a full 3D sound field with the possibility of external localisation. Binaural decoding of ambisonics is a suitable solution, as it accurately represents and allows rotation of the sound field at the listener's ears corresponding to head movements whilst only requiring a pair of headphones [5] [6]. YouTube have implemented a 1st order ambisonic to binaural head tracking system into their 360° videos, which gives a convenient way of demoing virtual room auralisations in the beginning stages of this project. This implementation allows for first order B-format recordings to be uploaded and demonstrated on YouTube, provided they are manipulated in order to adhere to the required specification [7]. In this case, the processing also involves the convolution of anechoic content gathered from OpenAir [8], as well as B-format impulse responses specifically recorded for this project.

The ease of use of binaural ambisonic decoding has increased due to the implementation of 3D head tracked spatial audio on Google's YouTube platform [3] [6]. Both created and recorded spatial audio content can be shared online, without site visits to laboratory facilities or the use of specialist hardware/software systems. All that is required is a standard consumer Android phone/tablet, a pair of headphones and the YouTube Android application or browser. Access to spatial audio has never been easier and more inclusive.

2.2 Auralisation Applications

There are several applications for utilisation of the recorded impulse responses. The first is to develop a library of impulse responses within different spaces, and to also develop novel applications for auralisations in a virtual reality environment, which are discussed in section 4.2. The first stage is to use the impulse responses and auralised them using the YouTube spatial audio platform

3 METHOD

This work is chiefly concerned with the auralisation of impulse responses with corresponding images before working on real time applications. ISO3382 [10] describes the international standard for measuring impulse responses. The technique used in this investigation, while informed by ISO 3382, did not strictly conform to its guidelines as the end purpose was auralisation and not objective acoustical analysis.

The first task was to capture impulse responses in the various identified spaces. Three contrasting spaces were chosen: Derby Theatre, a large lecture hall, and a small multipurpose performance venue. The methodology was adapted to each space according to its primary usage and equipment availability. All spaces used a SoundField ST350 [11] as the receiver.

A logarithmic sine sweep [12] and its inverse filter were generated using MATLAB [13] with a sweep duration of 15 seconds and frequency range of 20 Hz – 20 kHz, with 3 seconds of silence prior to the sweep and a 0.1 second envelope with a sampling rate of 48 kHz/24 bit[14]. The sweep includes a series of locator samples (alternating between -0.1 and 0.1) 1 second before the start of the sweep which is used for automatic detection and truncation during deconvolution when recording using portable recorders.

3.1 Equipment

3.1.1 System 1 – Derby Theatre

Derby Theatre served as the first venue to be captured. In this initial stage, several methods were carried out in order to allow workflow optimisation when capturing subsequent venues. A passive omnidirectional dodecahedron speaker array powered by a QSC amplifier [14] and an active KRK Systems Rokit 8 (RP8) [15] were used as sources to allow for preliminary analysis to shape the method. An Apple Mac-based RedNet system [16] and a SoundField ST50 were used to capture impulse responses. The system consisted of a single RedNet 4 unit (8 channel A/D at either instrument, microphone or line level), and a RedNet 1 unit (8 channel line level A/D D/A), which were both connected to a Gigabit Ethernet Switch, and in turn, a MacBook Pro. The MackBook Pro ran Dante Virtual Soundcard [17] and Dante Controller [18] to facilitate routing into ProTools HD [19] for recording and playback. A ¼" patch bay was also utilised to facilitate termination from DB25 to ¼" jack, thus allowing the playback to be sent to the sound reproduction system (active speaker or amplifier).

3.1.2 System 2 – University Auditorium and The VoiceBox

For measurements of an auditorium at the University of Derby's Markeaton Street campus and the multipurpose venue (The VoiceBox) the method had to be adjusted due to equipment changes and availability. Whilst not ideal, it was an opportunity for the method to be further streamlined, including the introduction of a Focusrite Saffire Pro 40 [20] interface instead of a fully racked and flightcased RedNet system, and the Ricoh Theta S 360° camera [2] as well as a KRK Rokit 8 as a directive source.

The Focusrite Saffire Pro 40 necessitated the use of their mix control software in place of Dante Virtual Soundcard and Dante Controller to facilitate routing in/out of Reaper for recording/playback. Due to the Saffire Pro40 having dedicated ¼" line outputs, there was no need for an external patch bay – thus reducing the size of the system substantially, aiding portability.

3.2 Measurement Scenarios

Outlined below are the measurement procedures for the chosen spaces. All sources were placed at a height of 1.4 m, and all receivers at 1.2 m, in order to adhere closely to ISO 3382.

3.2.1 Derby Theatre

Derby Theatre is a 530-seat capacity theatre that regularly produces their own shows, as well as accepting tours and other productions into its main house. It has many curved walls inside the auditorium, which cause problematic focused reflections. It has an orchestra pit that can be raised to stage height, as it was during recording, in order to give extra stage space in front of the proscenium arch. The floors are carpeted and the seats are upholstered with foam cushioning and a cloth covers. It is a single tier venue, and the production rooms overhang the row of seats at the rear of the auditorium.

Three source positions were used on stage, 1m behind the proscenium arch; stage left, centre and right. The three positions were chosen in order to give an accurate representation of possible actor positions on stage, and to enable comparison of localisation at different source locations. The receiver locations were chosen to represent the auditorium as comprehensively as possible without measuring at each one of the 530 seating positions. The auditorium is not symmetrical, and therefore measuring only one half of the space was not appropriate. An impulse response was also taken at the mix position as several staff at the theatre had reported interesting acoustic effects around the area. This will feed into a later project developing a virtual sound mix trainer, where students would be able to mix in an auralised virtual venue.

The receiver positions were spread around the auditorium, however only a limited number have accompanying photospheres due to the lengthy process of capturing them on a smartphone. Specific seating locations were chosen to reproduce on YouTube that best demonstrate the different areas of the theatre, as can be seen in Figure 2 - Plan of Derby Theatre's main house and receiver locations. This virtual auralisation of different seating positions could be continued in other performance spaces and implemented by tickets sites to allow the users to choose a seat not only based on a seating chart, but experiencing both the visuals and auralisation in 360°.

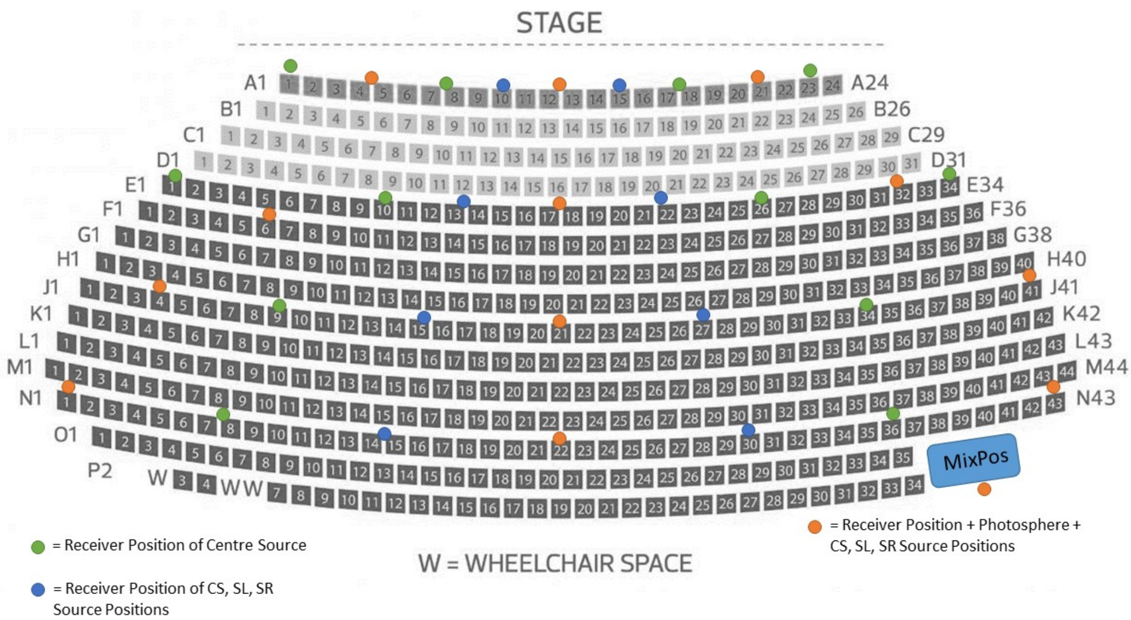

Figure 2 - Plan of Derby Theatre's main house and receiver locations (not to scale)

Due to the limitations on equipment availability at that time, further impulse responses and images need to be captured to benefit from the improved methodology. Since the recording of the impulse responses, the theatre auditorium has also undergone extensive refurbishment due to the installation of a new air management system, which may have altered the acoustic response of the space. For the virtual mixing application, it would also be necessary to record impulse responses in various positions using the house PA system as the source to simulate a realistic use-case.

3.2.2 The VoiceBox

The multipurpose performance venue is used for a variety of events and concerts, usually choral or solo singing, and has a standing capacity of 110 persons. It is a converted malting factory and has a pitched ceiling in two dimensions made of wood cladding, carpeted floor, and bare brick walls, as well as three large and recessed windows. The hall is 10.95 m long, 8.68 m wide, with the beginning and apex of the pitched ceiling being 4.77 m and 7.49 m high, respectively, giving an approximate volume of 582m^3.

Figure 3 - 360° image of Voices Hall, The VoiceBox

The source was placed at 3 m from the rear wall, as that is the usual performance location, and 2.15 m from each of the side walls. The six receiver positions were distributed to represent an audience with three rows of three positions: front, middle and rear. All of the front positions were 1.3 m from the source, with subsequent positions at a 2 m behind them.

3.2.3 Auditorium at Markeaton Street

The large lecture hall is situated at the University of Derby's Markeaton Street campus, and is used to give lectures and seminars to a cohort of up to 128 students. During the measurements, it was being used as an 'empty' space, with all the tiered seating in its storage position. The space is 13.08 m long at floor level, with a balcony level at 3.9 m height extending an extra 2.16 m in length, 9.83 m wide and a ceiling height of 6.5 m. The approximate volume of the room is 1037 m^3. It has some acoustic treatment in the form of perforated walls and ceiling tiles. The source location was chosen to simulate a lecture scenario, with the talker standing in front of the desk at 2.85 m from the projection screen and white board. The receiver positions began at 2 m from the source, and continued in 2 m increments away from the source, up to a distance of 8 m. This was repeated to either side of centre, 2.45 m from each of the side walls.

Figure 4 - 360 degree image of Auditorium 4

3.3 Impulse Response Processing

Once the sweeps were captured in a given space and matrixed into B-Format, the four channel sweeps (WXYZ) were catalogued and exported from ProTools/Reaper in preparation for deconvolution. A MATLAB script was written to automatically search the directory folders and find the source logarithmic sine sweep, the inverse filter, and the WXYZ audio files so that they are all automatically loaded into the workspace. The locator samples are identified in the sine sweep files which allow appropriate truncation to ensure all recordings are time-aligned. The truncated sweeps are then deconvolved using the inverse filter to give one B-Format file and four separate WXYZ files for auralisation purposes.

Acoustic parameter analysis was undertaken, with the aim to give a general indication of the reverberation within the space. To accomplish this, the W channel of each impulses response was truncated to two seconds in length and RT20 values in octave bands were generated using Aurora Acoustical Parameters software [21].

3.4 YouTube Preparation

At this stage, the use of YouTube is as a platform for auditioning the virtual auralisations, and to evaluate their effectiveness, before implementing the further aims of this project. Preparing the auralisations and images for YouTube requires a number of processes and libraries: FFmpeg [22], Python [23] and the Google Spatial Media Python Tools [24]. These external libraries can be easily installed using Homebrew [25].
Firstly, the impulse responses are loaded into a four channel bus for convolution with anechoic recordings in Reaper using ReaVerb by Cockos [26], which was then reordered into ambiX format and rendered into a four channel file. The four channel auralisation and photosphere are then loaded into terminal using the FFmpeg library, and are converted into a .mov video file:

> *ffmpeg -loop 1 -i **PS.jpg** -i **ambiX.wav** -map 1:a -map 0:v -c:a copy -channel_layout quad -c:v libx264 -b:v 40000k -bufsize 40000k -shortest **PSandAmbiX.mov***

Where ***PS.jpg*** is the 360° photosphere, ***ambiX.wav*** is the four channel aurailsed .wav file, and ***PSandAmbiX.mov*** is the output as a .mov video file. The next stage is to tag the video with spatial audio metadata by changing the terminal working directory to the location of Google Spatial Media Pyton Tools, and then launching the script with the command:

> *cd <location of your spatial media tools folder>*
> *python gui.py*

This command launches a GUI which requires the user to select the image projection type, in this case equirectangular, and then load the previously generated .mov file. The python script then injects spatial audio metadata and outputs this injected .mov as a separate file, which can be uploaded to YouTube and viewed as a 360° auralised video.

4 RESULTS

The measurements at Derby theatre will be used as an analysis case study. Preliminary analysis of RT20 values from the measurements captured at Derby Theatre showed that this particular omnidirectional dodecahedral source was not suitable for auralisation purposes, as it could not reach sufficient levels at low frequency to fully excite the space, resulting in distortion and therefore the source was changed to a directive Rokit 8 monitor for subsequent IR capture sessions.

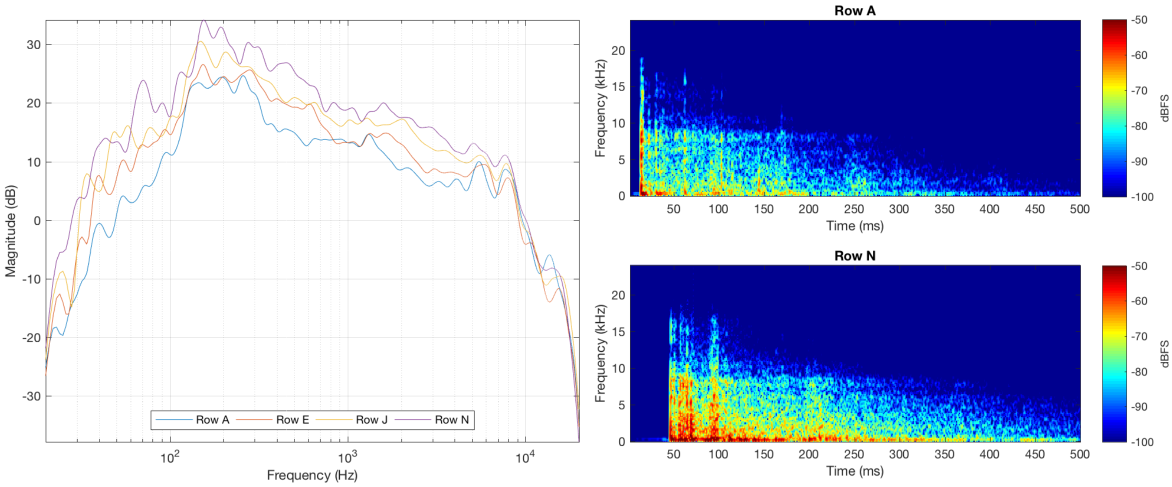

Figure 5 - Frequency analysis of the centre of each measurement row

Figure 5 shows both a spectrogram of the centre of row A and N (right plot) and a frequency domain plot of the impulse responses recorded in the centre of row A, E, J and N (left plot) with the source being centre stage. All frequency plots have been smoothed using a 1/6th octave Savitzky-Golay [28] filter. There is a variation in spectral characteristics and a dramatic change in overall energy over time between the two positions which may be due to the orchestral riser being risen to stage level, creating an acoustic obstruction for row A. At row N, a low ceiling due to the production room, as well as greater reverberant energy may have led to the greater spread of energy. There is a reduction in energy at around 12kHz-14kHz at row N that may be caused by high frequency absorption from the seating. The arrival of numerous reflections and strength of low frequency energy at row N is likely to give rise to a sense of envelopment and warmth which could be lacking in row A.

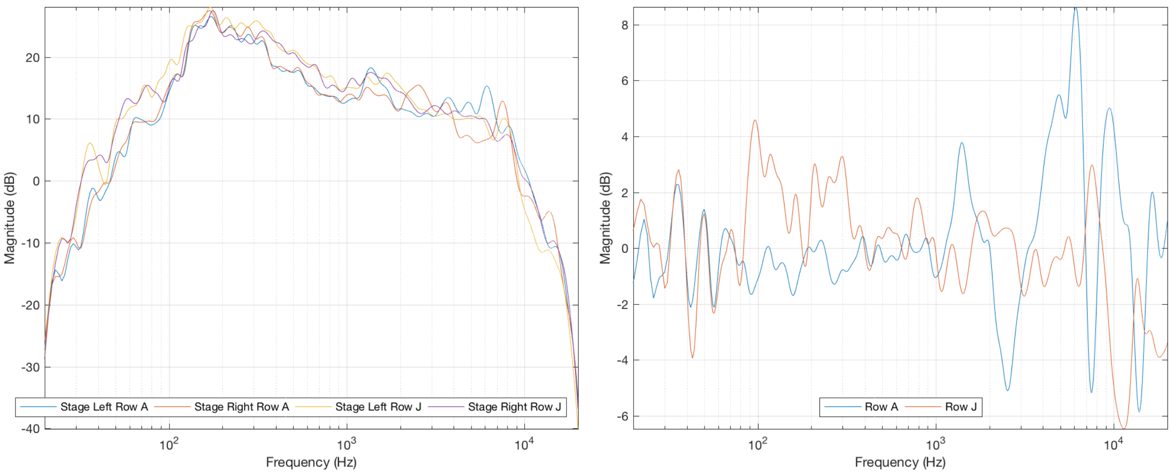

Figure 6 – Frequency analysis at the centre of row A & J with SL and SR source positions

Figure 6 shows two different source positions at the centre of row A and J and the difference between them, Stage Left (SL) and Stage Right (SR). There is a difference in spectral characteristics at 1kHz and above, with a large variance between 7kHz-8kHz, and between 1-3kHz. There is enough spectral difference in the IR at the same receiver position when moving the source to warrant auralisation of multiple source positions in order to accurately audition a space.

There is an immersive effect when using multiple source locations to auralise three different sources simultaneously, for instance a viola trio, giving the impression that the sources are spread across the stage. This technique may be built upon when for both the virtual mixing tool, and the virtual concert experience, where sound sources are placed in their real-world performance positions inside a virtual auralised venue.

The large variation in the spectral characteristics of the room at different seating locations shows that the auditory experience is hugely dependent on seating position, which justifies the use of virtual acoustic auralisations in order for the public to choose their preferred seating location.

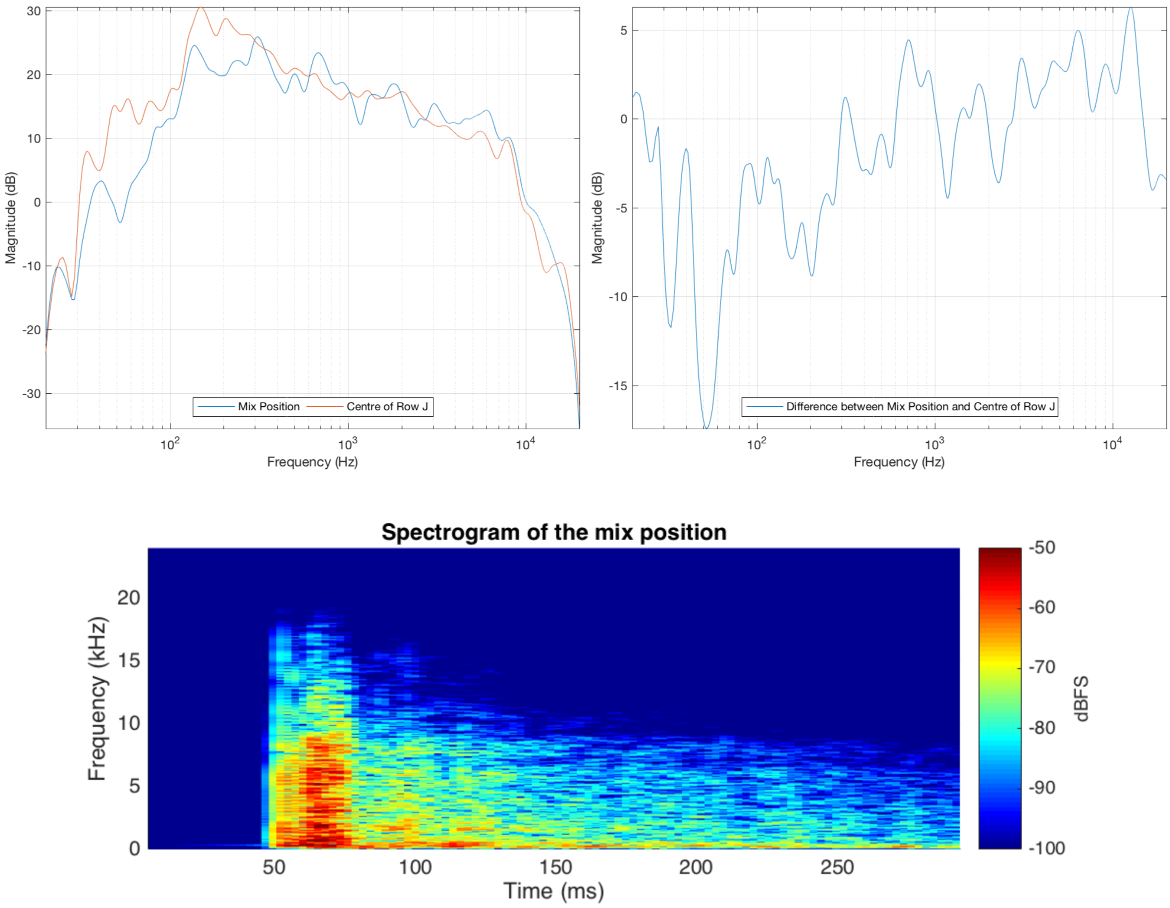

Figure 7 - Frequency analysis of the mix position compared to the centre of row J

Fig 7 shows the frequency analysis of the mixing desk position. Staff at the theatre had reported that it was an acoustically difficult position to mix in, and it would therefore make a good first application situation for the virtual mixing tool. The frequency analysis shows that there is an almost uniform arrival of acoustic energy up to around 8kHz, with much more energy in that frequency range compared to positions in the auditorium closer to the stage. This also correlates with positions on row N (Fig 5), which also demonstrates an increase in energy below 8kHz compared to row A. This could possibly be due to the theatre's curved walls creating a focusing effect or because of using an omnidirectional source on the stage, instead of the house PA system. This would need investigating further, with subsequent IR's captured using the PA. Once again, there is a significant variation in spectral characteristics at the mix position compared to other positions in the auditorium to warrant real time auralisation to aid mixing.

There may also be differences in the arrival directions of the first reflections which would alter the perception of the reverberation of the room, however further direction of arrival analysis would be needed.

4.1 Perceptual Evaluation of YouTube Virtual Acoustic Experience

It was necessary to perceptually evaluate the YouTube auralisations in order to gauge their success. The effect was convincing with a good HRTF match for the author, with good head tracking. Spatial resolution and localisation of sources was not particularly convincing, but this will improve as support for higher order ambisonics is added. First order ambisonics implementations have received some criticism in online forums, however 3rd order examples seem to perform better. As a method of demonstrating room acoustics and auralised spaces, it provides an acceptable sense of place and immersion for it to be a worthwhile endeavour as it gives a plausible rendition of the original space. More rigorous listener testing needs to be carried out in further work, however this paper focuses on the application and explanation of the method.

One reason why the accuracy of localisation for sources isn't particularly good may either be due to the use first order ambisonics instead of higher order, or the lack of coherent visual cues. The Derby Theatre photospheres were taken with an empty stage, however the VoiceBox and Auditorium 4 both had photos taken with and without the source in place from all the receiver positions. This may lead to listening tests where the subject is asked to localise sources with and without the visual representation of the source speaker. The implications of visual representation of sources on localisation in general also needs investigating, as visual cues often dominate over auditory stimulus due to multimodal perception [29].

The variety and size of a space also affects the YouTube auralisations. As the receiver distance from a source increases, the differences in the experience due to head tracking decreases since a given change in head position yields a smaller shift in perceived source position. The reverberation of the room also varies the experience, with the localisation effect being more apparent in smaller rooms. In the author's opinion, the auralisations of the VoiceBox are the most immersive and realistic, possibly due to its small size (and relatively high RT20), ensuring the receiver is always proximal to the source.

4.2 Further Work

Further work is planned for the continuation of this project, including the development of the three applications for the auralisations. The first is the development of a virtual reality auralised demonstration of seating positions within performance venues. This tool would allow the customer to be able to choose their seating position based not only on a seating diagram, on an easily accessible virtual reality experience with head tracked audio. In the initial stages of development, the YouTube spatial audio platform will be utilised due to its ease of use for both content producer and consumer. However, with the development of online systems that support higher order ambisonics, the methodology may change in the future in order to take advantage of the improved spatial resolution.

Using the impulse responses from Derby Theatre, a virtual mixing tool will also be developed. This will allow the user to experience mixing in different, and often difficult acoustic environments. A multitrack Reaper session would be loaded with convolution reverb using the impulse response recorded from the mix position, which in turn would be connected to a mixing desk and headphones. Further development to include head tracking will also be undertaken. This tool would also allow the user, when mixing in an actual venue, to be able to load impulse responses from different audience positions as they are mixing in order to experience the mix from an audience perspective. This may be especially useful if the mix position is located in an undesirable location, which it often is, and has issues with certain parts of the frequency band due to balconies and other structures.

The final application using the impulse responses would be a virtual real time auralised concert. Real time auralisation of singing has been carried out to great success, however it has mostly been only from the performer's perspective, without considering an audience [30]. At each venue, a coincidental impulse response is recorded, with the source and receiver as close together as possible. This mimics the acoustic the performer experiences as the source and receiver are in the same position. More impulse responses of audience positions will be captured and auralised, allowing both the performer and audience to be in the same virtual space simultaneously.

More spaces will be recorded, including historical properties including Sudbury Hall and the Devonshire Dome in Buxton, which is known for its interesting acoustical properties derived from the large freestanding domed roof, as well as another visit to Derby Theatre to carry out recording with the improved methodology. The use of higher order ambisonics platforms to improve localisation as discussed previously will also be investigated. Listening tests to thoroughly evaluate these platforms from an auralisation perspective, in terms of localisation, perceived sense of space, and spatiality etc., also need investigating.

5 CONCLUSIONS

The consumption of spatial media and dissemination methods have never been simpler and more user friendly for the consumer. YouTube, and the other platforms, have enabled a wave of development of 360° content with spatial audio to be created and shared. These advancements have allowed spatial auralisation to be available to the general public using nothing but a pair of headphones and a web browser or mobile app, allowing the user to experience the space from within a virtual reality experience, and not as numbers on a page.

YouTube playlists of the auralisations can be found at:

Derby Theatre – bit.ly/DerbyTheatreAuralisation
The VoiceBox – bit.ly/VoiceBoxAualisation
Auditorium 4 – bit.ly/Aud4Auralisation

6 REFERENCES

[1] Google Inc., “Google Street View Photosphere,” [Online]. Available: https://www.google.com/maps/streetview/publish/. [Accessed 04 08 2016].

[2] Ricoh Company Ltd., “Ricoh Theta S,” 2016. [Online]. Available: https://theta360.com/uk/about/theta/s.html. [Accessed 4 08 2016].

[3] R. K. Furness, “Ambisonics - An Overview,” in *AES 8th International Conference*, 1990.

[4] Sound on Sound, “livesoundfielddiag.l.jpg,” 2016. [Online]. Available: http://media.soundonsound.com/sos/jun04/images/livesoundfielddiag.l.jpg. [Accessed 5 10 2016].

[5] G. Kearney and T. Doyle, “Height Perception in Ambisonic Based Binaural Decoding,” in *139th AES Convention*, New York, 2015.

[6] R. Höldrich, S. Leitner and A. Sontacchi, “Multichannel sound reproduction system for binaural signals - The ambisonic approach,” in *DAFX-00*, Verona, 2000.

[7] Google, Inc., “Use spatial audio in 360-degree and VR videos,” 2016. [Online]. Available: https://support.google.com/youtube/answer/6395969?hl=en. [Accessed 13 09 2016].

[8] The Univesity of York, “OpenAir,” 2016. [Online]. Available: http://www.openairlib.net/. [Accessed 29 09 2016].

[9] Google, Inc. & GitHub, Inc., “spatial-audio-rfc.md,” Google, Inc., 27 07 2016. [Online]. Available: https://github.com/google/spatial-media/blob/master/docs/spatial-audio-rfc.md. [Accessed 13 09 2016].

[10] ISO, “ISO 3382-1:2009 Acoustics -- Measurement of room acoustic parameters,” 2009. [Online]. Available: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=40979. [Accessed 29 09 2016].

[11] ManualsLib, “Manuals Lib - Manual Database,” [Online]. Available: http://www.manualslib.com/manual/777936/Soundfield-St350.html#manual. [Accessed 16 09 2016].

[12] A. Farina, “Advancements in impulse response measurements by sine sweeps,” *Audio Engineering Society,* May 2007.

[13] The MathWorks Inc., “MathWorks,” 2016. [Online]. Available: http://uk.mathworks.com/. [Accessed 30 09 2016].

[14] A. Chadwick, “GenerateSweep.m,” University of York, York, 2015.

[15] QSC LLC, “QSC PLX Amplifier Manual,” [Online]. Available: http://www.qsc.com/resource-files/productresources/amp/discontinued/plx/q_amp_plxspec_specs.pdf. [Accessed 29 09 2016].

[16] KRK Systems LLC, “KRK Systems Manual for KRK RP8 Gen3,” [Online]. Available: http://www.krksys.com/manuals/rokit/RokitG3-Manual.pdf. [Accessed 29 09 2016].

[17] Focusrite PLC, 2016. [Online]. Available: https://us.focusrite.com/ethernet-audio-interfaces/rednet. [Accessed 29 09 2016].

[18] Audinate Pty Ltd, “Dante Virtual Soundcard,” 2016. [Online]. Available: https://www.audinate.com/products/software/dante-virtual-soundcard. [Accessed 29 09 2016].

[19] Audinate Pty Ltd, “Dante Controller,” 2016. [Online]. Available: http://dev.audinate.com/GA/dante-controller/userguide/pdf/latest/AUD-MAN-DanteController-3.10.x-v1.0.pdf. [Accessed 29 09 2016].

[20] Avid Technologies Inc, “Avid - ProTools HD,” 2016. [Online]. Available: http://www.avid.com/pro-tools-hd. [Accessed 29 09 2016].

[21] Focusrite PLC, “Saffire Pro 40 - Manual,” 2016. [Online]. Available: https://d3se566zfvnmhf.cloudfront.net/sites/default/files/focusrite/downloads/7046/userguidepro40english04_0.pdf. [Accessed 29 09 2016].

[22] A. Farina, “Aurora Plug-ins,” 2007. [Online]. Available: http://www.aurora-plugins.com/. [Accessed 25 09 2016].

[23] FFmpeg, “Download FFmpeg,” 26 08 2016. [Online]. Available: https://ffmpeg.org/download.html. [Accessed 27 09 2016].

[24] Python Software Foundation, “Download Python,” 27 06 2016. [Online]. Available: https://www.python.org/downloads/. [Accessed 27 09 2016].

[25] Google inc., “Github - google/spatial-media,” 2016. [Online]. Available: https://github.com/google/spatial-media. [Accessed 27 09 2016].

[26] M. Howell, “Homebrew,” 2016. [Online]. Available: http://brew.sh/. [Accessed 27 09 2016].

[27] Cockos, “ReaVerb,” Cockos, 2016. [Online]. Available: http://wiki.cockos.com/wiki/index.php/ReaVerb. [Accessed 27 09 2016].

[28] M. Gorzel, D. Corrigan and G. Kearney, “Distance Perception in Virtual Audio-Visual Environments,” in *25th Conference: Spatial Audio in Today's 3D World*, 2012.

[29] J. S. Brereton, D. T. Murphy and D. M. Howard, “A loudspeaker-based room acoustics simulation for rea-time musical performance,” in *UK 25th Conference: Spatial Audio in Today's 3D World*, 2012.

[30] B. Wiggins, “YouTube 360 VR Ambisonics Teardown!,” 28 04 2016. [Online]. Available: http://www.brucewiggins.co.uk/?p=700#content. [Accessed 16 09 2016].