Towards rapid development of interactive acoustics virtual reality (AVR) application of existing spaces through the implementation of SLAM-based acoustics and visual mapping

Dhong Fhel Gom-os (1), Anthony Zander (1), Daniel Pitman (2) and William Robertson (1)

(1) School of Electrical and Mechanical Engineering, The University of Adelaide
(2) Elder Conservatorium of Music, The University of Adelaide

ABSTRACT
Interactive acoustics virtual reality (AVR) environments of existing spaces provide a means for users to enhance existing spaces and reconstruct historical acoustics of spaces which have undergone modifications over time. These applications allow users to auralise the acoustical changes in real-time as materials are modified in the VR environment. Currently, there are very limited studies and applications on interactive AVRs. This is due to the time-consuming and labour-intensive process of collecting and processing acoustics and visual data for the AVR application. This project is aimed towards addressing this issue by building a robotic system based on Simultaneous Localisation and Mapping (SLAM) algorithms that can autonomously and simultaneously collect visual and acoustic data of a space for rapid application development. This paper will discuss the motivations, overall framework, system architecture, hardware and software technologies, initial results on the characterisation and integration of robot platform, and data acquisition instrumentation. In addition, we will also discuss the future work aimed at collecting impulse responses autonomously.

1 INTRODUCTION
Acoustic virtual reality (AVR) (Pind et al. 2018) is an immersive technology that combines auditory and visual elements to produce realistic and interactive virtual environments. This technology provides a user interface to apply acoustic treatments on spaces and listen to their impact on room acoustics through the process called auralisation (Vorländer 2020). Auralisation in AVR is achieved through real-time simulation of sound propagation of the virtual model accounting for the geometry of the space and its associated materials and their acoustic properties. All these must be well-integrated with the visuals seen by the user. This kind of technology is particularly helpful in both design support and design review processes (Delgado et al. 2020), among others, in architecture, engineering, and construction.

Interactive AVR may be categorised into three applications: (a) enhancement of existing spaces (EES), (b) design of future spaces (DFS), and (c) reconstruction of historical acoustics (RHA). The first category involves the application of AVR to currently existing spaces to evaluate and implement acoustic treatments. It allows for virtual experimentation of various acoustic solutions without any physical alterations, thereby saving time and resources. The second category uses AVR to simulate and evaluate the acoustics of spaces that are yet to be built. This enables architects and acousticians to optimise the design of the space before construction begins. The last category can be further divided into two: (i) reconstruction of completely lost spaces and (ii) modification of existing spaces to historical configuration. In both cases, AVR is applied to recreate the acoustic environments of spaces that no longer exist or have undergone significant changes over time. This allows for the exploration and preservation of historical acoustic experiences. The interest of this research project is both categories (a) and (c). For (c), interest is given to (ii). Such types of AVR have existing spaces to begin with, where reference measurements can be taken for more accurate results.

There are not many AVR applications that fall under this category probably because of the tedious processing pipeline of producing the AVR environment, which is one of the aspects that this project aims to address. One recent study (Heimes et al. 2019) is the development of an AVR application for psychoacoustic experiments. In such application, the 3D CAD model was manually created using SketchUp software. The acoustic simulation is done using ray tracing. Unity software was used to build the application. The acoustic properties of materials were manually assigned in the software. Limited functions could be done by the user of the application, such as the activation of room acoustics, source/receiver characteristics, and choice of the methods for sound insulation filter construction. Despite the existence of the physical space, no field measurements were done which could have improved the performance of acoustic simulation. In contrast, other related studies (Naeemaee and Sü Gül 2024; Pérez-Aguilar et al. 2024) use field measurements to fine-tune (Postma and Katz 2015) their simulation environment. In another study (Castro et al. 2019), an AVR application was developed to demonstrate navigable auralisation for a classroom in the Queen Mary University of London. The visual model of the room was created using Unity3D where the application was integrated, while geometric based CATT-Acoustic software was used in the acoustic simulation where material absorption coefficients were manually configured from CATT's available database. With the press of a button, the user could change from three limited acoustic scenarios — measured IR benchmark, simulated IR from CATT, and treatment scenario where ceiling of the room is changed from plasterboard to mineral fibre tiles. Unlike the previous study (Heimes et al. 2019), this study leveraged measured IRs as benchmarks by which simulated IRs can be compared against. AVR application was employed in an attempt to recover historical acoustics of the Linlithgow palace in Scotland built in 1424 with only few structural remains through time (Cook et al. 2023). LIDAR scanning was used to build the structural basis of the 3D model which was then completed through archival records. Acoustic simulations were run in both ODEON and Google Resonance. Although not a full-fledged AVR application since users do not have the interactive capability, this case can benefit from an interactive feature. One study (Horvat et al. 2022) did a survey of the literature of VR applications in the area of design review where AVR could fall under. It is interesting to note that no particular AVR in any of the three categories described above is included. There is limited study in the area of interactive AVR of existing spaces, in part because of the tedious process of data acquisition itself on top of the already demanding development of the application.

The key measurable parameter for acoustic mapping is the room-impulse response (RIR), which is the response of the space when subjected to an impulsive sound. It captures the acoustics of the space based on its geometry and materials used, essentially providing insights into the behaviour of sound in the space and the means to reconstruct the space in a digital format. Recently, autonomous mobile robot platforms equipped with microphones have been developed (Götz et al. 2021; 2024; Stolz et al. 2023; 2024) to automate the process of RIR recording useful for AVR applications. However, they have not been equipped with cameras to collect both visuals and acoustic map of the space for rapid AVR application development.

The research goal of this project is to develop a new methodology combining geometry and acoustic mapping. For this paper, we discuss the motivations, overall framework, system architecture, hardware and software technologies, initial results on the characterisation and integration of robot platform, and data acquisition instrumentation. In addition, we will also discuss the future work aimed at collecting impulse responses autonomously and adaptively. This paper will present the methods in Section 2, the results in Section 3, a discussion in Section 4, and future work in Section 5.

2 METHODS

This project is motivated by the time-consuming and manual process of collecting acoustics and visual data of a space for AVR applications. It is divided into four different components: the SLAM-based room impulse response measurement, SLAM-based 3D modeling and material characterisation, coupling of the previous two components including the development of the VR environment, and user testing. This project seeks to develop a mobile robot platform that simultaneously and autonomously measures RIRs, creates a 3D model of the space, and characterises the acoustic properties of the materials in the space for the purpose of creating a VR environment for acoustic treatment applications with real-time auralisation.

The initial phase of the project is the configuration of hardware equipment which include the robot, field recorder, and microphone as well as the teleoperated RIR recording in a room – in this case a laboratory reverberation chamber as a controlled example. Figure 1 shows the system architecture of the measurement robot. The mobile robot base used is the ROSbot XL (Husarion, Poland), a universal robot for indoor applications, equipped with a single-board computer which is the ASUS NUC with Intel i3 processer, 16GB RAM, and 250GB SSD, which allows for fast data processing and sufficient storage of RIRs and visual assets. The robot is also equipped with a real-time STM32F407 microcontroller as an interface for peripherals such as the motors and sensors. The robot is installed with Ubuntu 24.04 LTS operating system with ROS2 Jazzy. An advantage of selecting this robot platform is the top plate which provides dozens of attachment points for sensors as well as required hardware such as the field recorder and microphone. In the future, simultaneous localisation and mapping (SLAM) and navigation algorithms will be installed to the SBC and a LiDAR sensor will be attached to the robot for autonomous operation. An 8-channel F8n pro multi-track field recorder (Zoom, Japan) was mounted on the robot and used to record RIRs from the first-order NT-SF1 ambisonic microphone (Rode, Australia). The recorded RIRs are currently stored in an SD card attached to the field recorder. Future work will interface the field recorder to the robot's SBC for processing and storage of RIRs and synchronisation of recording operations.

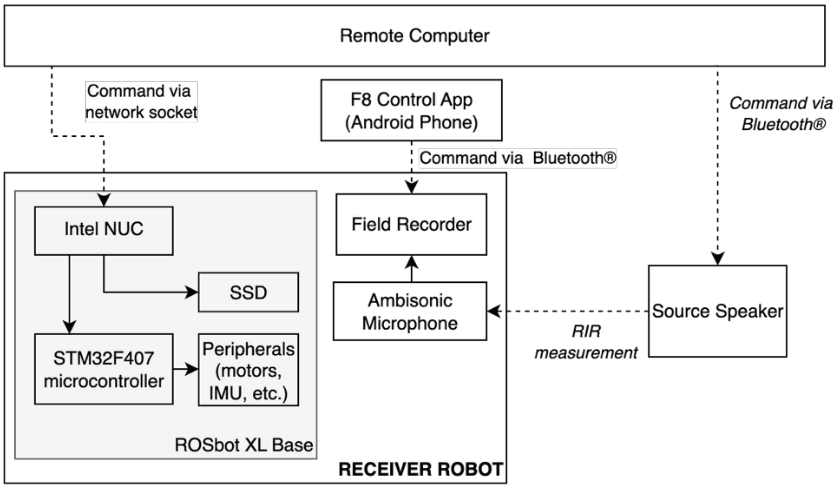

Figure 1: Overall system architecture of the RIR measurement robot

2.1 Characterisation of Ambisonic Microphone

RIRs are normally recorded through the use of microphones to record sound pressure levels of the sound. There are different ways by which RIRs can be recorded depending on the application. For applications that do not require directionality of sound such as the determination of reverberation time, omnidirectional recording approaches can be used permitting the use of omnidirectional microphones. For applications that require the directionality of sound to allow spatialisation and localisation, like in auralisation, directional microphones are desirable. In this project, the interest is in both global omnidirectional parameters like reverberation time and directional sound information for auralisation. Therefore, a recording device that is able to provide both directional and non-directional sound information is desirable.

Ambisonic microphones are a type of microphone that are able to provide both omnidirectional and directional sound information. They are designed to capture sound in all directions and separate components from these directions. They are also versatile in post-production, as they can capture sound that can be reproduced in virtually any format, including binaural reproduction for auralisation which is the interest of this project. This is why the ambisonic microphone will be used in this project, in particular, the first-order Rode NT-SF1 ambisonic microphone which has four cardioid polar pattern capsules (left-front: LF; right-front: RF; left-back: LB; right-back: RB) arranged in a tetrahedral pattern as shown in Figure 2.

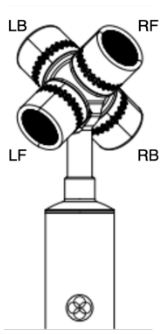

Figure 2: Schematic diagram of the Rode NT-SF1 ambisonic microphone with four capsules (Image: Rode)

The ambisonic microphone was successfully characterised inside the anechoic chamber at the University of Adelaide using a turntable to determine its polar response. It was placed on a tripod and has the same centre of rotation as that of the turntable. It was approximately 1 m away from the speaker and 1.6 m above the floor. Best effort was done to centre the microphone and the speaker. An omni-directional microphone (Brüel & Kjær Type 2669) was used as a reference microphone. Raw inputs of the ambisonic microphone were recorded using four channels of the field recorder. This recording yields what is known as A-format in ambisonics audio terms, which is the raw format from the four cardioid polar pattern capsules. Ambisonic's A-format inputs were successfully converted to FuMa B-format using Reaper DAW software with Rode's Soundfield plugin. The B-format is calculated from the A-format using the following equation (Zotter and Frank 2019), where W is an omnidirectional signal which is the summation of all four channels in the A-format. X, Y and Z are the directional signals which capture the sound direction longitudinal, lateral and vertical axes. The W channel may be used to calculate global acoustic parameters, and the rest may be used to decode binaural impulse response for auralisation.

$$\begin{bmatrix} W \\ X \\ Y \\ Z \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} LF \\ RF \\ LB \\ RB \end{bmatrix} \tag{1}$$

The ambisonic microphone's B-format W omnidirectional channel was also compared with a Brüel & Kjær type 2669 microphone as a reference signal using Brüel & Kjær 4 channel microphone power supply type 2829.

2.2 Teleoperated Recording in the Reverberation Chamber

The robot attached with the ambisonics microphone and field recorder was placed inside the reverberation chamber for a teleoperated recording of RIRs (see Figure 3). The floor area of the chamber was divided into 25 equidistant locations for the robot to take measurements. It was manually controlled to navigate to these locations. The field recorder was controlled remotely from a smartphone to enable recording while the speaker was remotely triggered with white noise from a computer via Bluetooth. A total of 25 recordings were taken. The recording session lasted for around an hour. The robot was also measured inside the anechoic chamber to determine its impact to the acoustic field.

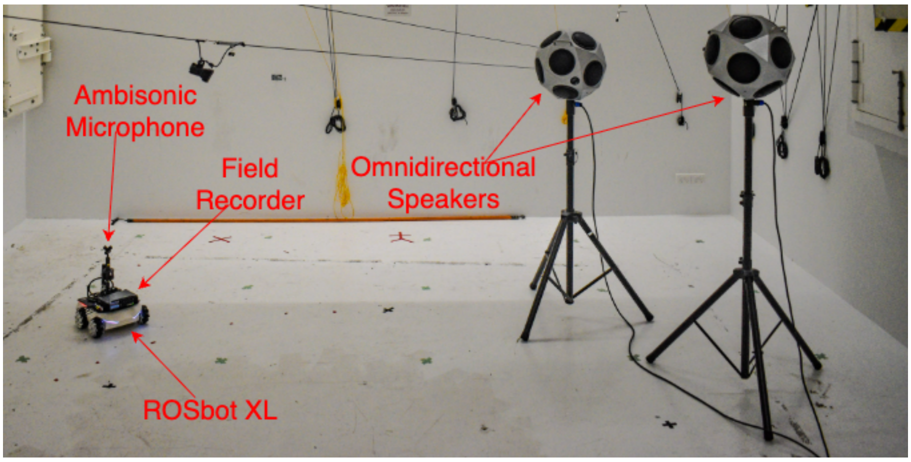

Figure 3: Receiver robot inside the reverberation chamber at the University of Adelaide

3 RESULTS AND DISCUSSION

3.1 Polar Responses of the Ambisonic Microphone

The polar response in 1/3 octave bands of the A-format recordings of the ambisonic microphone, normalised against peak power (dB), corresponding to the four input channels are shown in Figure 4. Each channel exhibits a cardioid heart-shaped polar pattern response which picks up sound primarily from the front and less at the sides and rear.

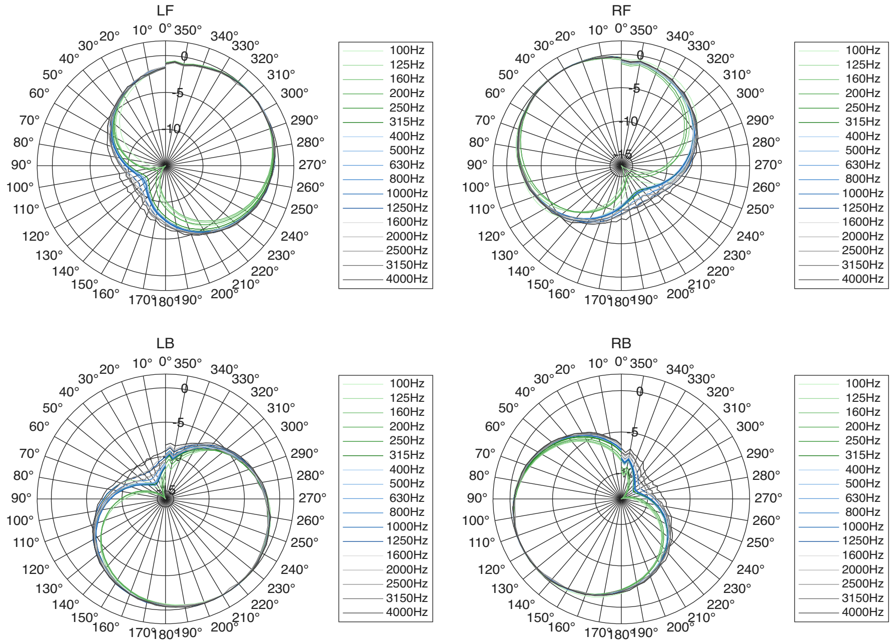

Figure 4: Cardioid polar response of the four ambisonics channels

Ambisonic A-format is of no interest to this project which is why it is necessary to convert A to B-format. The polar response in 1/3 octave bands of the B-format recordings of the ambisonic microphone, normalised against peak power (dB), corresponding to the four input channels are shown in Figure 5. As expected, the first channel (W) exhibits omnidirectional pattern while the second (X) and third (Y) channels exhibit a figure-of-8 pattern in front-back and left-right directions. The observed figure-of-8 pattern picks up sound from front and rear while it supresses sound at the sides. The last channel (Z) shows the up-down direction which is less useful in these tests because the speaker is positioned at fixed height during this measurement.

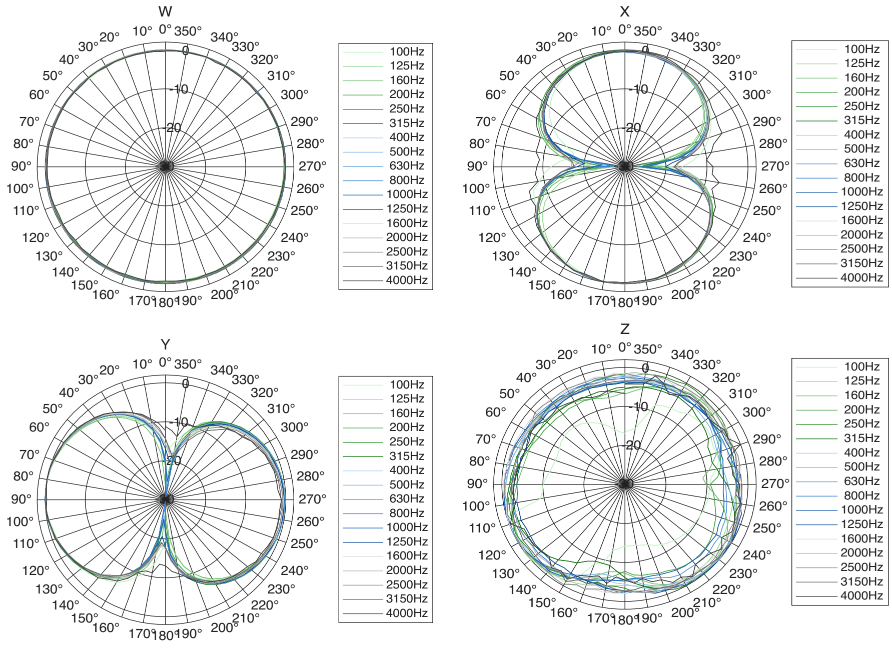

Figure 5: B-format polar response of the four directions of the ambisonic microphone.

3.2 Spectral Comparison of Ambisonic and Reference Microphones

The W channel of the Rode NT-SF1 microphone was compared with a calibrated reference microphone inside the anechoic chamber using a white noise input sound. Magnitude spectra were analysed with 1/3-octave band smoothing and expressed in dBFS to account for varying microphone sensitivities. Generally, as shown in Figure 6, the ambisonics microphone is consistent with the overall spectral shape of the reference microphone which indicates that it accurately captures the relative frequency content of the input signal, except at extremely low and high frequencies. However, a noticeable level offset of approximately 10 dBFS is registered between the two microphones reflecting the differences in their sensitivity and internal gain, including microphone noise characteristics and chamber effects. The difference plot highlights this offset. While absolute SPL measurements require separate calibration for each microphone, the W channel of the ambisonic microphone reliably preserves the spectral shape, making it suitable for applications emphasising relative spectral fidelity.

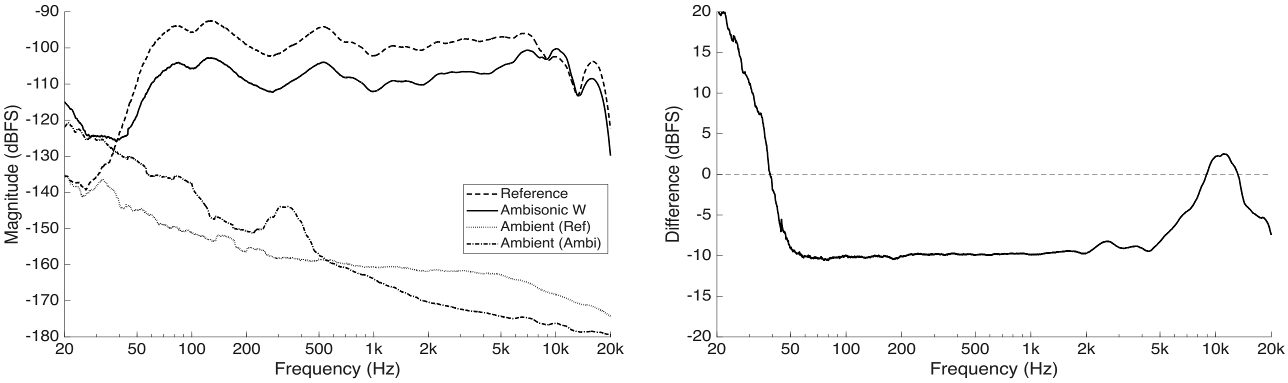

Figure 6: Spectra of reference and ambisonics (W channel) microphones with background noise (left) and difference between the reference and ambisonic microphones (right)

3.3 Anechoic Robot Measurement

Essentially, any physical objects introduced inside the acoustic field will alter its RIR. To assess the impact on the RIR from the presence of the robot itself, the frequency response of a speaker inside the anechoic chamber with and without the mobile robot was measured. Figure 7 shows the frequency response of the speaker with and without the robot, including the ambient noise inside the chamber. The frequency responses are relatively similar across frequencies except for extremely low and high frequencies. This suggests that the robot has minimal effect on the acoustic field, particularly in the main frequency range of interest (50 Hz – 10 kHz).

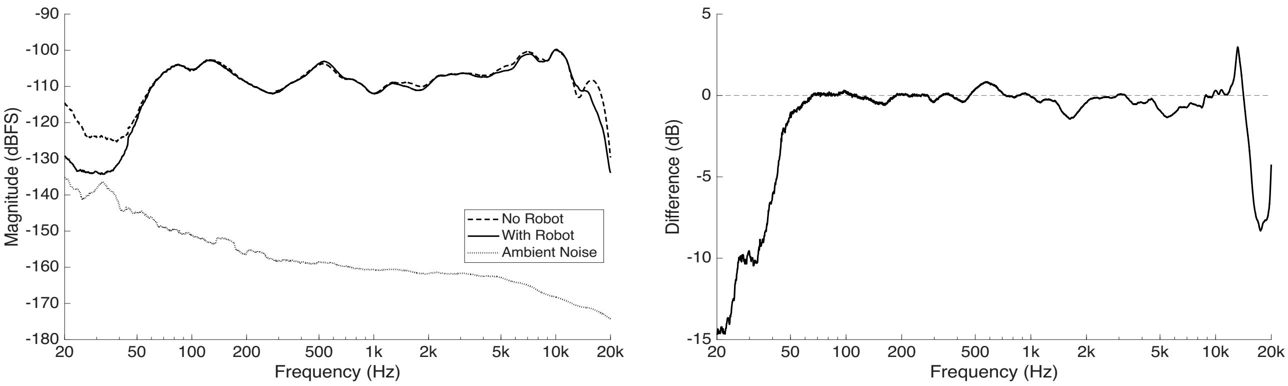

Figure 7: Spectra with and without the robot (left) and their difference (right)

3.4 RIR Recordings

A total of 25 RIR recordings were taken inside the reverberation chamber. The reverberation times (T60s) based on the slope of the Schroeder curve over a variable range selected to produce the best linear fit using the REW software are also plotted in Figure 8. Black line is average based on ISO 3382 Acoustics — Measurement of room acoustic parameters (ISO, n.d.). Heat map of T60s at 1 kHz, scaled between minimum and maximum T60 values, is also shown. Interpolation between points was done using MATLAB's interp2 function which interpolates for 2-D gridded data in meshgrid format. This heatmap illustrates the acoustic variance across the space (a reverberation chamber at that) and therefore supports the work of an adaptive sampling strategy. Future work will look into an adaptive sampling strategy to take dense RIR sampling in acoustically intricate areas such as near walls and corners, and sparse sampling in homogeneous regions like open spaces.

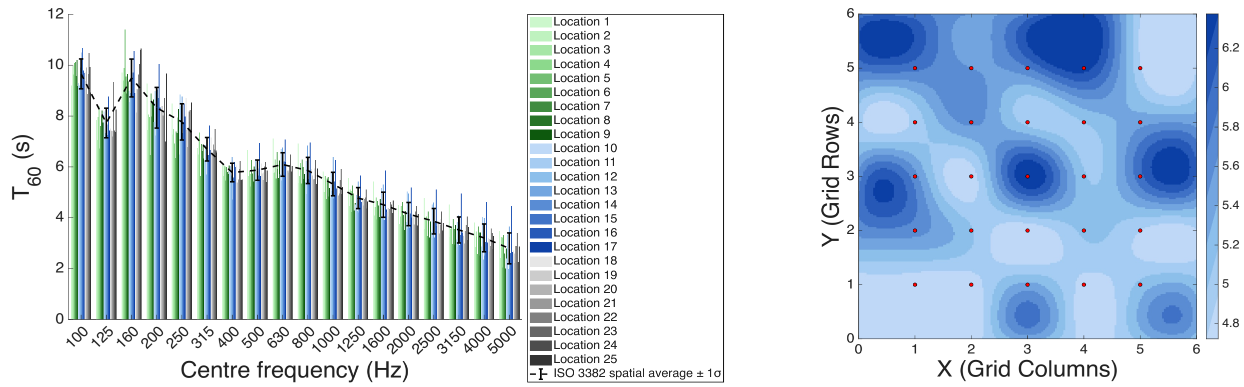

Figure 8: Reverberation times (T60s) at 25 locations and 1/3-octave band frequencies (left) and heatmap at 1kHz (right)

4 CONCLUSION AND FUTURE WORK

Recording RIRs and visual assets for AVR environments of existing spaces is time-consuming if done manually via the traditional tripod work. While some studies have explored the use of mobile platforms, none have explored the use of SLAM to autonomously and simultaneously record measurements. By performing teleoperated recording using the ROSbot XL inside a reverberation chamber, we have shown that such an objective is achievable and beneficial when incorporated with autonomous algorithms such as SLAM.

It was also shown that the ambisonic microphone used is comparable with the reference microphone in terms of its spectra. It was found that the presence of the robot has minimal effect on the acoustic field.

Future work will investigate incorporating SLAM and navigation algorithms for semi-autonomous operation and rapid AVR application development. Adaptive RIR sampling will also be introduced for more efficient data acquisition, towards the goal of autonomous three-dimensional RIR capture for six degree-of-freedom auralisation.

ACKNOWLEDGEMENTS

This work was supported in part by the Research Small Equipment Support Scheme 2024 (University of Adelaide); in part by the University of Adelaide Research Scholarship; and in part by the Faculty of SET, University of Adelaide.

REFERENCES

Castro, Daniel, André Verstappen, and Samuel Platt. 2019. *Walk-through Auralization Framework for Virtual Reality Environments Powered by Game Engine Architectures, Part II*.

Cook, James, Andrew Kirkman, Kenneth McAlpine, and Rod Selfridge. 2023. "Hearing Historic Scotland: Reflections on Recording in Virtually Reconstructed Acoustics." *Journal of the Alamire Foundation* 15 (1): 109–26. https://doi.org/10.1484/J.JAF.5.133797.

Delgado, Juan Manuel, Lukumon Oyedele, Peter Demian, and Thomas Beach. 2020. "A Research Agenda for Augmented and Virtual Reality in Architecture, Engineering and Construction." *Advanced Engineering Informatics* 45 (August): 101122. https://doi.org/10.1016/j.aei.2020.101122.

Götz, G., I. Ananthabhotla, S.V. Amengual Garí, and P. Calamia. 2024. "Autonomous Room Acoustic Measurements Using Rapidly-Exploring Random Trees and Gaussian Processes." *Proceedings of the 10th Convention of the European Acoustics Association Forum Acusticum 2023*, January 17, 1655–62. https://doi.org/10.61782/fa.2023.0796.

Götz, Georg, Sebastian J. Schlecht, Abraham Martinez Ornelas, and Ville Pulkki. 2021. “Autonomous Robot Twin System for Room Acoustic Measurements.” *Journal of the Audio Engineering Society* 69 (4): 261–72. WOS:000638270300003. https://doi.org/10.17743/jaes.2021.0002.

Heimes, A., I. Muhammad, and M. Vorländer. 2019. “A Real-Time Virtual Reality Building Acoustic Auralization Framework for Psychoacoustic Experiments with Contextual and Interactive Features.” 2019-September: 6417–23. Scopus. https://doi.org/10.18154/RWTH-CONV-239031.

Horvat, Nikola, Steffen Kunnen, Mario Štorga, Arun Nagarajah, and Stanko Škec. 2022. “Immersive Virtual Reality Applications for Design Reviews: Systematic Literature Review and Classification Scheme for Functionalities.” *Advanced Engineering Informatics* 54 (October): 101760. https://doi.org/10.1016/j.aei.2022.101760.

ISO. n.d. “ISO 3382-2:2008.” Accessed September 29, 2025. https://www.iso.org/standard/36201.html.

Naeemaee, Rozhin, and Zühre Sü Gül. 2024. “Assessment of Acoustical Comfort in Educational Atriums of Closed Central and Linear Typologies: Reverberation Perception and Preference.” *Building and Environment* 252 (March): 111230. https://doi.org/10.1016/j.buildenv.2024.111230.

Pérez-Aguilar, B., A. Quintana-Gallardo, J.L. Gasent-Blesa, and I. Guillén-Guillamón. 2024. “The Acoustic and Cultural Heritage of the Banda Primitiva de Llíria Theater: Objective and Subjective Evaluation.” *Buildings* 14 (8): 8. Scopus (WOS:001304981900001). https://doi.org/10.3390/buildings14082329.

Pind, Finnur, Cheol-Ho Jeong, Hermes Sampedro Llopis, Kacper Kosikowski, and Jakob Strømann-Andersen. 2018. “Acoustic Virtual Reality – Methods and Challenges.” *Proceedings of BNAM 2018*.

Postma, Barteld N. J., and Brian F. G. Katz. 2015. “Creation and Calibration Method of Acoustical Models for Historic Virtual Reality Auralizations.” In *Virtual Reality*, vol. 19. no. 3. Springer. Scopus (WOS:000363862800003). https://doi.org/10.1007/s10055-015-0275-3.

Stolz, Georg, Georg Götz, Lukas Treybig, Stephan Werner, and Florian Klein. 2024. “Spatial Room Impulse Response Dataset: A Robot’s Journey Through Coupled Rooms of a Reverberant University Building.” Zenodo, October 24. https://zenodo.org/records/13987509.

Stolz, Georg, Florian Klein, Stephan Werner, Lukas Treybig, Andreas Bley, and Christian Martin. 2023. “Discussion of Acoustic and Perceptual Optimization Methods for Measuring Spatial Room Impulse Responses with a Mobile Robotic Platform.” *2023 Immersive and 3D Audio: From Architecture to Automotive (I3DA)*, September, 1–7. https://doi.org/10.1109/I3DA57090.2023.10289338.

Vorländer, Michael. 2020. *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. RWTHedition. Springer International Publishing. https://doi.org/10.1007/978-3-030-51202-6.

Zotter, Franz, and Matthias Frank. 2019. *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*. Vol. 19. Springer Topics in Signal Processing. Springer International Publishing. https://doi.org/10.1007/978-3-030-17207-7.