

Article

Room Impulse Response Dataset of a Recording Studio with Variable Wall Paneling Measured Using a 32-Channel Spherical Microphone Array and a B-Format Microphone Array

Grace Chesworth, Amy Bastine * and Thushara Abhayapala *

Audio & Acoustic Signal Processing Group, The Australian National University, Canberra 2601, Australia;
grace.chesworth@anu.edu.au

* Correspondence: amy.bastine@anu.edu.au (A.B.); thushara.abhayapala@anu.edu.au (T.A.)

Abstract: This paper introduces RSoANU, a dataset of real multichannel room impulse responses (RIRs) obtained in a recording studio. Compared to the current publicly available datasets, RSoANU distinguishes itself by featuring RIRs captured using both a 32-channel spherical microphone array (mh acoustics em32 Eigenmike) and a B-format soundfield microphone array (Rode NT-SF1). The studio incorporates variable wall panels in felt and wood options, with measurements conducted for two configurations: all panels set to wood or felt. Three source positions that emulate typical performance locations were considered. RIRs were collected over a planar receiver grid spanning the room, with the microphone array centered at a height of 1.7 m. The paper includes an analysis of acoustic parameters derived from the dataset, revealing notable distinctions between felt and wood panel environments. Felt panels exhibit faster decay, higher clarity, and superior definition in mid-to-high frequencies. The analysis across the receiver grid emphasizes the impact of room geometry and source–receiver positions on reverberation time and clarity. The study also notes spatial variations in parameters obtained from the two microphone arrays, suggesting potential for future research into their specific capabilities for room acoustic characterization.



Citation: Chesworth, G.; Bastine, A.; Abhayapala, T. Room Impulse Response Dataset of a Recording Studio with Variable Wall Paneling Measured Using a 32-Channel Spherical Microphone Array and a B-Format Microphone Array. *Appl. Sci.* **2024**, *14*, 2095. <https://doi.org/10.3390/app14052095>

Received: 5 February 2024

Revised: 28 February 2024

Accepted: 29 February 2024

Published: 2 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: recording studio acoustics; ambisonics; room impulse response; virtual reality; dataset

1. Introduction

Sound waves emitted from a source within a room undergo intricate interactions with the surrounding environment, influenced by factors such as the room's geometry, surface materials, and furnishings. These elements collectively determine the acoustic response of the room and contribute to the auditory experience. Considering the room as a linear time-invariant system, the sound transformations between a source and a receiver are quantitatively encapsulated as a room impulse response (RIR), uniquely describing the room's acoustic fingerprint [1].

The RIRs can facilitate a comprehensive understanding of the room acoustics, which is pivotal in the design of performance venues [2,3], correction of acoustic anomalies [4], and selection of acoustic treatments [5]. It is also integral in the development of speech processing algorithms [6], the design of immersive audio in extended reality (XR) environments [7], and various soundfield analysis and synthesis methods. The spatial variations captured by RIRs across the room enable accurate listener translations, providing an enriched immersive experience. Therefore, the availability of diverse and authentic RIRs can greatly benefit the research and development of numerous audio applications.

RIRs can be obtained through two primary methods: simulation-based generation using wave-based and ray-based methods [8], or real-world measurements. While simulated RIRs are quicker to generate, they lack the authenticity of real acoustic environments due to challenges in modeling complex scattering effects and material compositions [9,10]. Hence,

algorithms validated using simulated data may malfunction when applied to real conditions. Even though resource-intensive, measured RIRs provide a realistic representation of real-world scenarios and preserve important perceptual features. Typically, a singular microphone or an array of microphones is employed to measure RIRs at spatially distributed receiver positions in a given room. Recently, higher-order ambisonic microphone arrays have been widely used to acquire directional characteristics of the acoustic field [11–13].

Several measured RIR datasets have been collected in the past, and they are framed to meet the demands of certain applications, like acoustic analysis, room geometry estimation, speech enhancement, automatic speech recognition, and VR environment design. These datasets differ in key attributes, like the type, size, and configurations of the room measured, the number and density of source and receiver locations, and the type and configuration of loudspeakers and microphones. Some of the publicly available datasets [12–20] are compared in Table 1. The OpenAIR database [11] contains impulse response data from more than 50 environments, mostly collected in the Ambisonic B-format.

Table 1. Comparison of publicly available RIR datasets. (#)RIRs indicate the number of RIRs times the number of channels. (#)Rooms indicate the number of rooms times the number of configurations. The last row corresponds to the introduced dataset.

Name	#RIRs	#Rooms	#Source Positions	#Receiver Positions	T ₆₀ (s)	Microphone
C4DM [14]	468 × 1, 468 × 4	3 × 1	1	130, 169	1.7–2.7	Omnidirectional, B-format array
Hadad [15]	234 × 8	1 × 3	26	3	0.16–0.61	8-element linear array
ACE [16]	14 × {2,3,5,8,32}	7 × 1	1	2	0.34–1.25	{2,3,5,8,32}-element arrays of different configurations
BUT ReverbDB [17]	1426 × 1	9 × 1	3–10	31	0.59–1.85	2–10 omnidirectional microphones of several mountings
dEchorate [18]	60 × 30	1 × 10	6	1	0.18–0.75	6 5-element linear array
MeshRIR [19]	3969 × 1, 441 × 32	1 × 1	1, 32	3969, 441	0.19–0.38	Omnidirectional
Motus [12]	3320 × 32	1 × 830	4	1	0.5–1.5	em32 Eigenmike
Arni [20]	26560 × 1	1 × 5312	1	5	0.25–1.5	Omnidirectional
BBC Maida Vale [13]	1586 × 32	2 × 1	3–7	4–45	0.7–0.9	em32 Eigenmike
RSoANU	396 × 4, 396 × 32	1 × 2	3	65–67	0.6–0.9	B-format array, em32 Eigenmike

Contribution

This paper introduces RSoANU, the RIR dataset of a recording studio [21] at the School of Music, Australian National University (ANU). This studio features variable wall paneling with felt and wood options. The RIRs are recorded using both a 32-channel spherical microphone array and a B-format soundfield microphone array for two wall configurations, three source positions, and a 1 m × 1 m receiver grid spanning a measurement region of dimensions 7 m × 7 m. Additional receiver points with a finer grid spacing of 50 cm around each active source are also included. This dataset comprises 396 4-channel and 396 32-channel RIRs with metadata of Cartesian coordinates of source and receiver positions, wall panel setting, microphone device, temperature, and other signal parameters.

To the best of our knowledge, no existing datasets offer RIRs of a room recorded with both a higher-order spherical microphone array and a B-format array, especially across an extensive receiver grid. RSoANU dataset is particularly intended to facilitate the analysis of spatial soundfields and room acoustics with both higher-order and first-order ambisonics, taking into account the differences in spatial resolution, frequency response, directivity, and economic considerations of the two types of microphone arrays. Additionally, the receiver grid is designed to support the validation of RIR interpolation/extrapolation and soundfield translation methods for VR applications. Since the room environment and source–receiver locations are closely controlled during the measurement, direct comparisons between different acoustic settings can also be performed.

This paper also presents an analysis of room acoustic parameters using the RSoANU dataset. The obtained results enable the identification of potential acoustic issues in the ANU Recording Studio, providing insights for necessary treatments and aiding in the selection of optimal recording locations. Given the scarcity of recording studio datasets [11,13], especially those with variable environments, the findings from this dataset offer valuable guidance for the design of similar studio settings.

2. Measurement Setup and Technique

2.1. Room Setup

The RSoANU dataset consists of RIRs measured in the ANU Recording Studio [21] with a floor dimension of $8.5\text{ m} \times 9\text{ m}$. The studio features a double-height space with a floor-to-ceiling height of 5.6 m. Two adjacent walls are fitted with variable wall paneling, offering options in both felt and wood materials, whilst the other two walls have fixed paneling. The room also includes a window and a sliding door made of glass material providing access to the control room.

The variable wall panels set to felt and wooden configurations are shown in Figure 1a and Figure 1b, respectively. The structural details and dimensions of the felt and wooden panels can be seen in Figure 1c and Figure 1d, respectively. In the wooden configuration, angled plywood frames with a width of 0.603 m and a depth of 0.398 m alternate with slat resonators measuring 0.617 m in width and 0.137 m in depth. Note that these depths are measured from the middle of the angled surface to the wall. Behind the angled plywood surface, three layers of 50 mm insulation separated by air cavities are affixed, while one insulation layer is attached behind the slat surface. In the felt configuration, each panel has a depth of 0.268 m with two layers of 50 mm insulation behind the surface. More details on similar variable wall panel structures can be found in [22].

2.2. Equipment

During the data collection process, a Tannoy System 600 loudspeaker (Tannoy Ltd.: Coatbridge, Scotland) served as the sound source. It has a frequency response with $\pm 3\text{ dB}$ tolerance bandlimited to 52–20 kHz with cross-over frequency at 1.8 kHz [23].

The RIRs were measured using two microphone arrays: the Rode NT-SF1 microphone (Røde: Sydney, Australia) and the mh acoustics em32 Eigenmike (mh acoustics: Summit, NJ, USA). The NT-SF1 comprises four matched condenser cardioid microphone capsules arranged in a tetrahedral array. It is capable of capturing spatial sound in first-order ambisonics with a frequency response of $\pm 3\text{ dB}$ tolerance in the range of 30–20 kHz [24]. The em32 consists of thirty-two omnidirectional electret microphones mounted on a rigid sphere of 8.4 cm diameter. It can capture spatial sound up to fourth-order ambisonics with a frequency response bandlimited to 30–20 kHz, offering a nominally flat response up to 8 kHz [25,26].

The source and the receivers are shown in Figure 2. The Eigenmike and NT-SF1 were connected to a computer using the Eigenmike Interface Box (EMIB) and the Zoom F6 [27], respectively. Note that the Zoom F6 interface was configured to record in Ambisonic B-format with a Furse–Malham (FuMa) normalization scheme.



Figure 1. ANU Recording Studio: (a,b) panorama view; (c,d) zoomed view of panels.

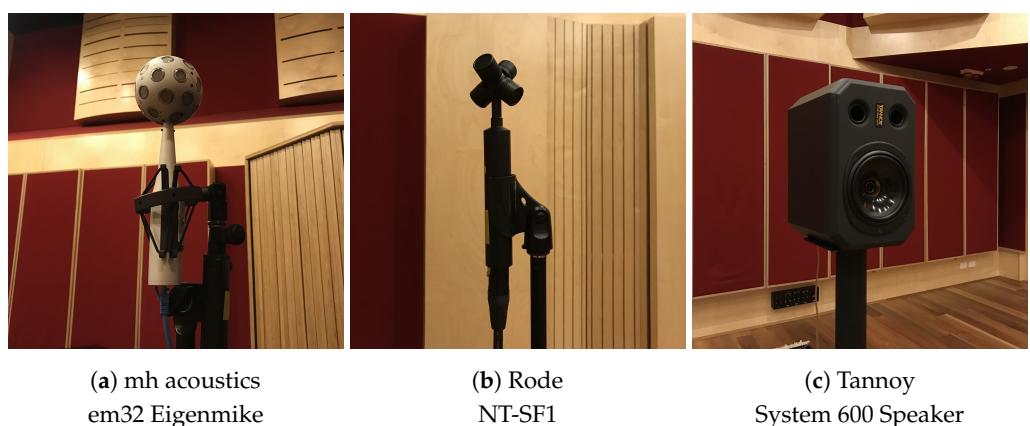


Figure 2. Receivers and source used in data collection.

2.3. Recording Setup

The source–receiver arrangement for the RIR measurement is depicted in Figure 3, with the room's bottom-left corner considered as the coordinate origin.

The RIRs were measured for three distinct source locations, with only one source being active during each measurement. The source locations are marked as S1, S2, and S3 in Figure 3, and their Cartesian coordinates are mentioned in Table A1. The height and

direction of the sound source were set to reflect the placement of typical performers or ensembles in the recording studio. Source S1 was positioned closest to the front door, at a height of 1.2 m, reflecting a typical drum kit height. Source S2 in the middle of the room was positioned at a height of 1.4 m to reflect a singer's projection height [28]. Source S3 was positioned at a height of 0.9 m, corresponding to the height of the strings on a grand piano. In terms of source orientation, both S1 and S2 faced the control room, while S3 was pointed to the left wall. Note that these source heights are measured from the studio floor to the base of the loudspeaker. Users can refer to the loudspeaker dimensions [23] to adjust the source height with reference to the loudspeaker cone.

Receiver locations were set up in an 8×8 planar grid pattern with $1\text{ m} \times 1\text{ m}$ spacing. These positions are marked as circles in Figure 3. A few points in the square grid were excluded since they were inaccessible due to the placement of furnishings and the grand piano (blank space to the right side of S3). In addition to these receiver points, a denser grid with 50 cm spacing was also included in front of each active source. These are indicated by colored triangles in Figure 3 corresponding to their respective source. During the measurement, the microphone array was centered at these positions at a consistent height of 1.7 m, reflecting the average height of an Australian [28]. This height was measured from the floor to the center of the array.

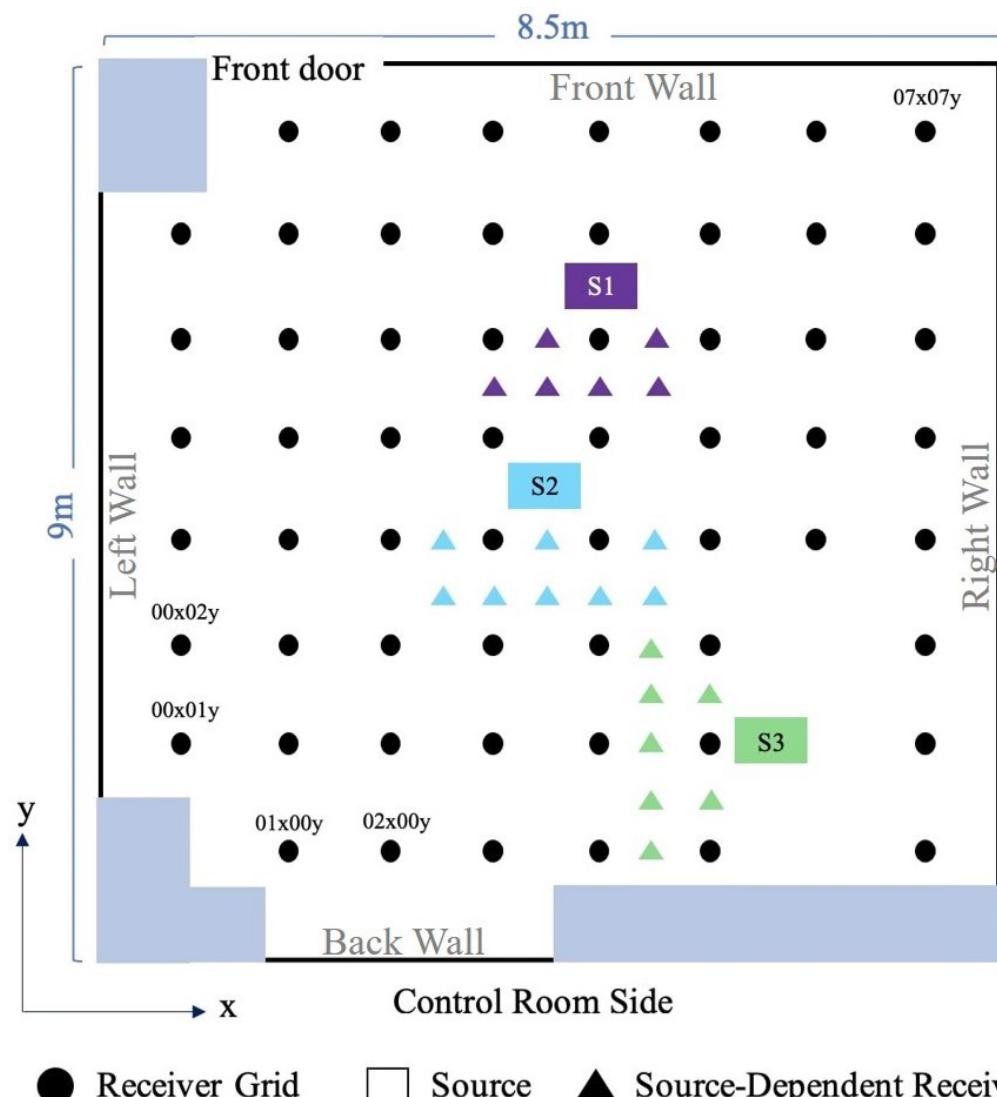


Figure 3. Diagram of the source and receiver positions. Light blue rectangular shapes indicate locations of furnishings.

Each receiver position is assigned a code name based on its coordinates along the x and y axes. Code names of some of the receiver positions are indicated in Figure 3. For instance, $01x00y$ is the code name of the receiver location in the second column from the left and the first row from the bottom. The Cartesian coordinates and code names of the receiver positions on the regular grid are listed in Table A2. The additional receiver points surrounding the sources are labeled differently; for example, $e30x45y$ corresponds to the receiver associated with S1 in the fourth column, positioned between the fifth and sixth rows from the bottom. The Cartesian coordinates and code names of these source-dependent receiver positions are listed in Table A3.

2.4. Recording Method

An exponential sine sweep (ESS) signal of 2 s duration and 20 Hz to 25 kHz frequency range generated at a sampling rate of 48 kHz was used as the excitation signal. We chose an ESS as it is widely used for RIR measurements due to its large dynamic range, excellent signal-to-noise ratio (SNR), and ability to reject harmonic distortions and non-linear effects [13,17,29]. The sweep signal was preceded and followed by 1 s silence periods to avoid time-aliasing problems.

In the initial acoustic setting for the RIR measurement, all variable wall panels were set to felt, and the loudspeaker was positioned at S1. The loudspeaker volume was set to receive around 65 dB of loudness at 1 m distance. The ESS signal was played through the loudspeaker, and recordings were made at each receiver position using the em32 Eigenmike. Next, all the variable wall panels were switched to wood, and the measurements were repeated using the em32 Eigenmike. The process was repeated for source locations S2 and S3. The entire measurement procedure was then repeated using the Rode NT-SF1 microphone. Following each recording, the RIRs were estimated through deconvolution of the recorded sweeps with the inverse sweep signal. MATLABTM R2021a was employed for controlling the source and receiver signals, as well as for performing the deconvolution process.

Before each measurement for every room configuration and source position, microphones underwent calibration, and latency adjustments were performed. The room temperature was also recorded before each run and was very stable, with variations of less than 0.5 °C across all measurements.

3. Dataset Description

The RSoANU dataset comprises RIRs captured in the ANU Recording studio, considering two room configurations and three source positions (S1, S2, and S3). RIRs were recorded for 65, 67, and 66 receiver positions for S1, S2, and S3, respectively. Recordings conducted using both the em32 Eigenmike and Rode NT-SF1 microphone generated a total of 396 32-channel and 396 4-channel RIRs, respectively. Note that the RIRs recorded by Rode NT-SF1 microphones are in Ambisonic B-format with FuMa normalization. All RIRs are saved with a sampling frequency of 48 kHz and a duration of 1.5 s.

The RIRs in the RSoANU dataset are organized in folders with the structure shown in Figure 4. Each multichannel RIR is saved in two file formats: MATLABTM struct (*.mat) and 16-bit audio (*.wav). The filenames follow a convention based on the source code names and receiver code names outlined in Tables A1–A3. For the em32 Eigenmike, the 32-channel RIR is saved under the *IR_Data* field name in the (*.mat) struct file. For the Rode NT-SF1 microphone, the RIR is saved under the *Bformat* field name with separate child fields for *W*, *X*, *Y*, *Z* channels. The room temperature, sampling rate, microphone array type, source and receiver coordinates, excitation signal, and raw recording are recorded as metadata. The (*.mat) struct files for NT-SF1 also include the normalization standard under the *Bformat_standard* metadata field.

The RSoANU dataset is publicly accessible online at <https://doi.org/10.5281/zenodo.10720345> [30] under the Creative Commons Attribution 4.0 International license.

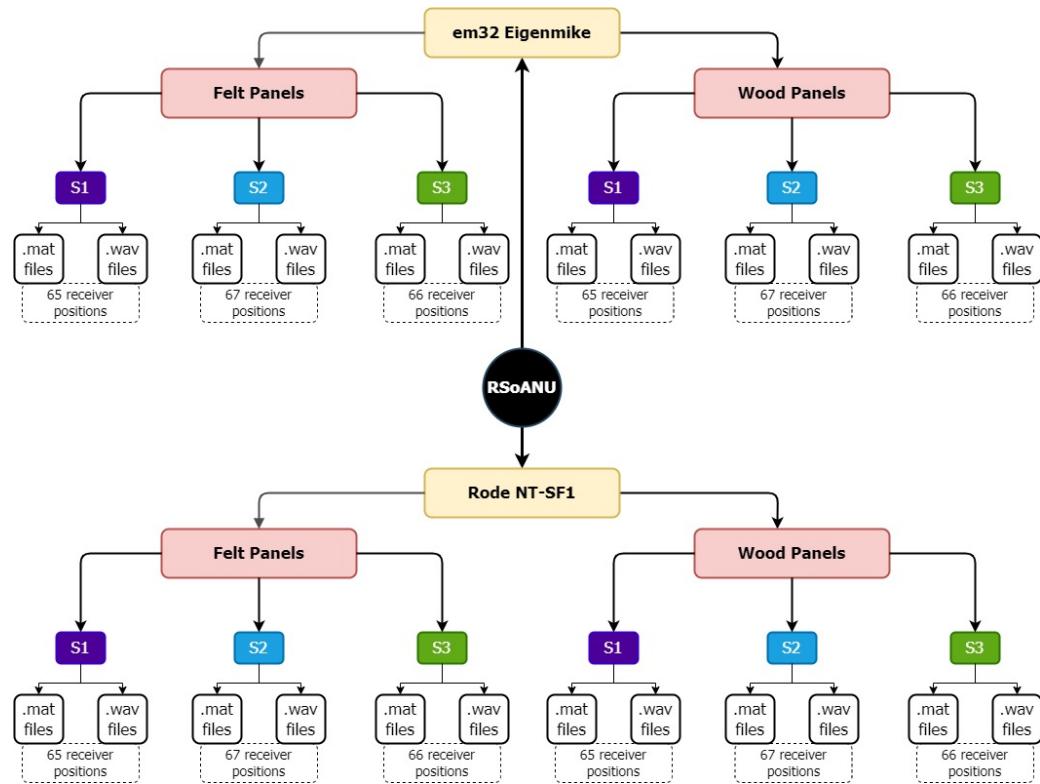


Figure 4. Organization of RSoANU dataset.

4. Room Acoustic Parameter Analysis

In this section, we utilize the RSoANU dataset to analyze the acoustics of the ANU Recording Studio by examining the parameters: reverberation time (T_{20}), clarity (C_{80}), definition (D_{50}), and center time (T_s) [31]. The parameters are derived from the W channel for NT-SF1 microphone recordings and from the average of 32 channels for the em32 Eigenmike recordings.

Table 2 presents the mean and standard deviation of all parameters for each source position in both felt and wood panel environments, calculated from the RIRs collected at all receiver positions. The difference in acoustics offered by the felt and wooden panel environments is evident from the parameter values in Table 2. Additionally, we can observe high standard deviations in C_{80} due to its strong reliance on the receiver positioning relative to the sound source.

In the following subsections, we delve into the variations in each parameter across frequencies in octave bands ranging from 125 Hz to 8 kHz. For T_{20} and C_{80} , we further investigate variations across different receiver locations using parameter values averaged across the octave bands.

Table 2. Mean and standard deviation of reverberation time (T_{20}), clarity (C_{80}), definition (D_{50}), and center time (T_s) calculated using RIRs across the receiver region.

Parameter	Felt Panels			Wood Panels		
	S1	S2	S3	S1	S2	S3
T_{20} (s)	0.74 ± 0.03	0.74 ± 0.04	0.73 ± 0.03	0.84 ± 0.03	0.84 ± 0.04	0.83 ± 0.04
C_{80} (dB)	8.20 ± 8.64	7.91 ± 6.6	7.94 ± 5.65	6.83 ± 5.62	6.88 ± 5.43	6.84 ± 4.27
D_{50} (%)	70.40 ± 8.45	69.98 ± 8.98	71.27 ± 8.04	66.39 ± 8.48	66.17 ± 9.39	67.39 ± 8.24
T_s (ms)	40.47 ± 10.73	41.58 ± 11.91	40.56 ± 11.23	46.44 ± 11.08	47.31 ± 12.72	46.13 ± 11.66

4.1. Reverberation Time

Reverberation time is the duration it takes for the sound pressure level to decay by 60 dB after the source ceases. From the Schroeder curve generated through backward integration of the squared RIR, we calculate the reverberation time (T_{20}) by extrapolating the least-squares fit between -5 dB and -25 dB following the ISO 3382 recommendation [31,32].

Figure 5a,b present the T_{20} values across frequency in octave bands, derived from RIRs measured by the Rode NT-SF1 microphone and em32 Eigenmike, respectively. The two environments with felt and wood panels show similar T_{20} values from 125 to 250 Hz. However, from 500 Hz to 8 kHz, the wood panel environment consistently demonstrates higher T_{20} compared to the felt environment. The average T_{20} for the felt and wood panel environments is approximately 0.74 s and 0.84 s, respectively. The faster sound decay in the felt panels aligns with their higher sound absorption capabilities compared to the wood panels. The overall trend remains consistent between the NT-SF1 and Eigenmike measurements, but the latter shows slight variations in T_{20} values among different source positions beyond 500 Hz.

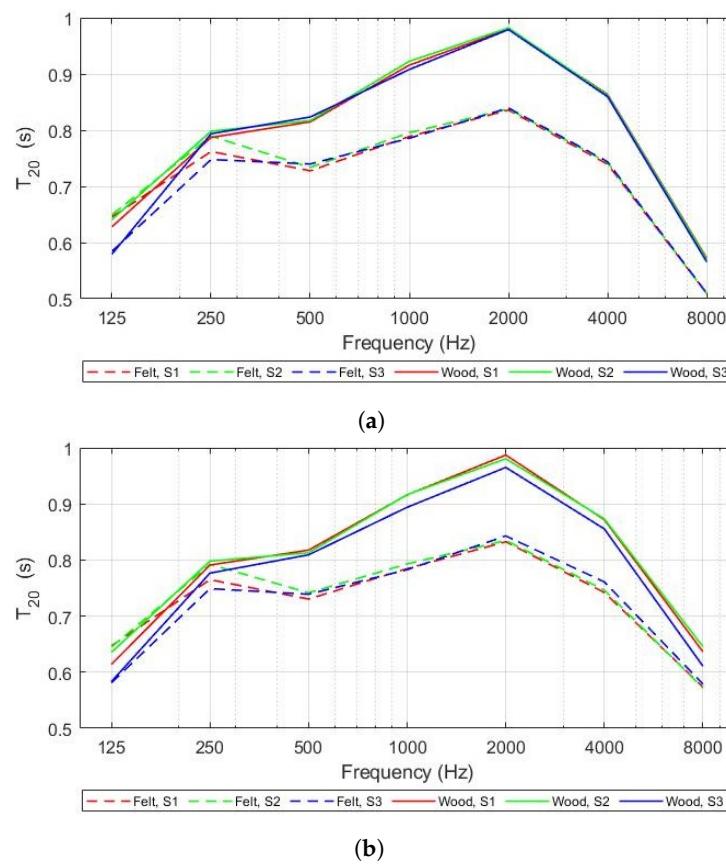


Figure 5. Reverberation time (T_{20}) across frequency in octave bands for the felt and wood environments calculated from the RIRs measured by (a) Rode NT-SF1 microphone, (b) em32 Eigenmike.

Figure 6a,b illustrate 2D heatmaps of T_{20} values across receiver locations for both paneling environments and all three source locations, calculated from the RIRs measured by the NT-SF1 microphone and em32 Eigenmike, respectively. In these visualizations, higher T_{20} values are noticeably closer to the front and right walls, particularly in the front-right corner of the room (refer to Figures 1 and 3 for orientation). This observation suggests the influence of the overhang in this corner, a phenomenon consistent with similar effects observed in other rooms with such structures [33,34]. Additional acoustic treatment around this corner can help to achieve a more uniform distribution of reverb effect across the room.

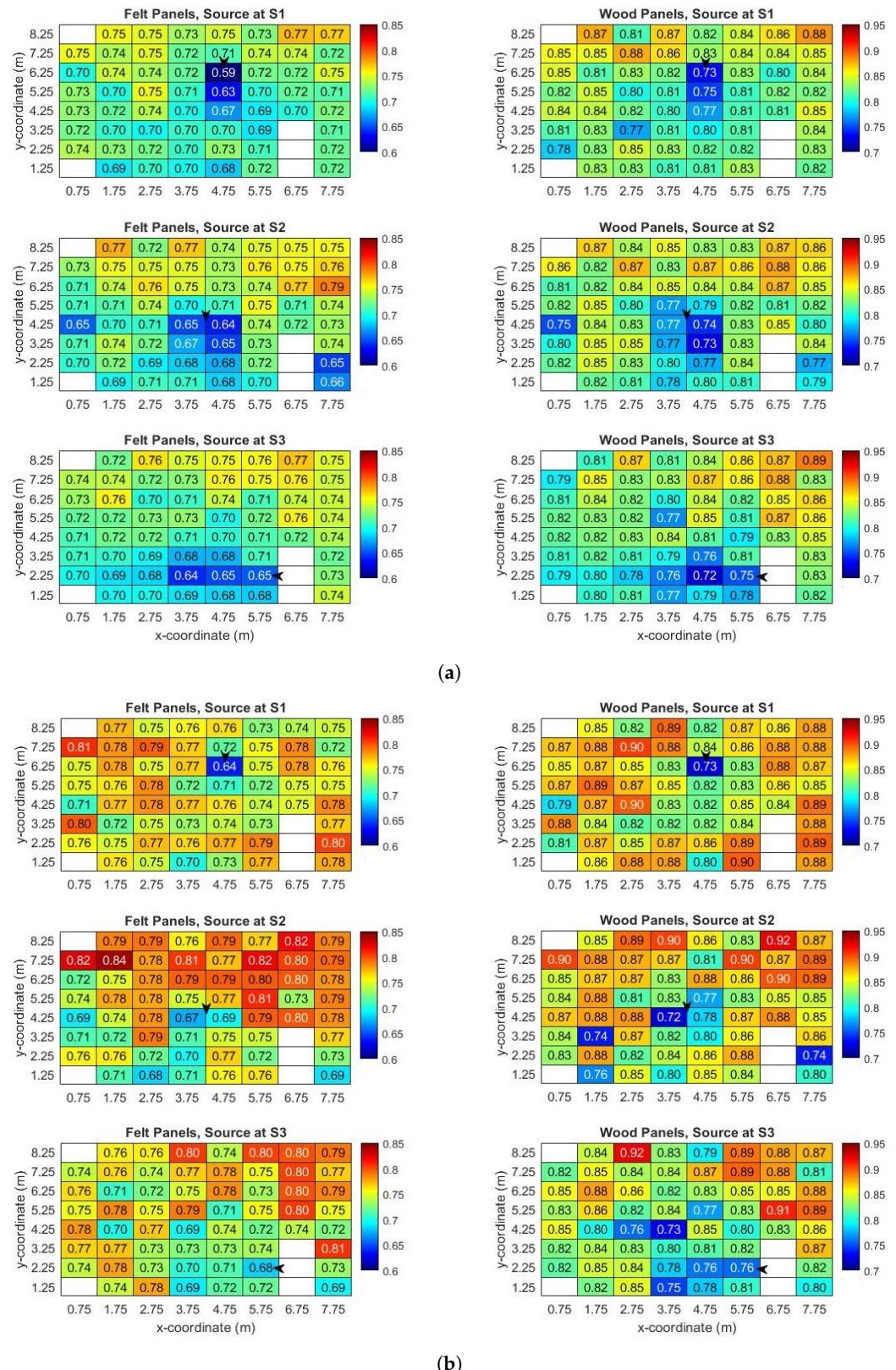


Figure 6. Reverberation time (T_{20}) in seconds at different receiver locations for each of the three source locations and the two wall paneling types calculated from the RIRs measured by (a) Rode NT-SF1 microphone, (b) em32 Eigenmike. The black arrowhead indicates the source location and orientation. The white areas represent spaces where no receivers were positioned. Note that colormap limits are different for felt and wood panels.

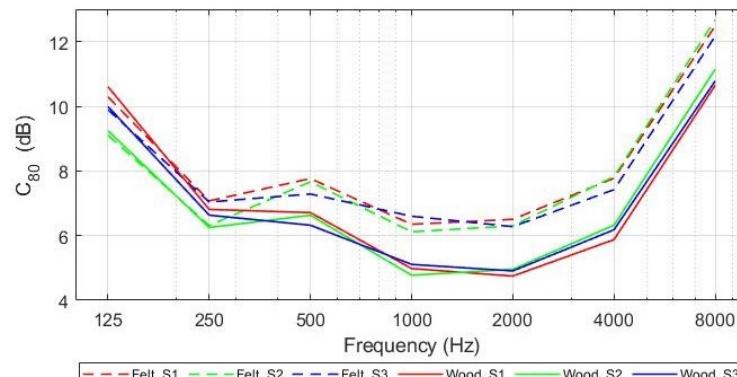
In Figure 6a, we can see lower T_{20} values for a few receiver points directly in front of the source. However, this observation is moderated in Figure 6b for the Eigenmike measurements. Moreover, the local variations in T_{20} are distinct between NT-SF1 (Figure 6a) and Eigenmike (Figure 6b). This discrepancy warrants deeper analysis, considering that the Eigenmike captures finer spatial details of the reflected sound compared to NT-SF1.

4.2. Clarity

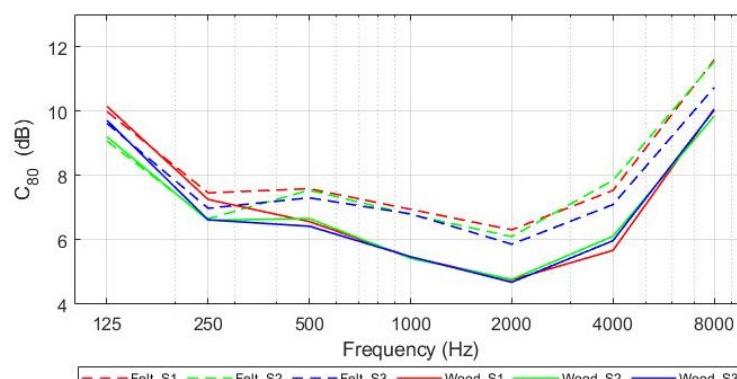
Clarity (C_{80}) characterizes the transparency of the sound and is computed from the RIR $h(t)$ as follows [1,31]:

$$C_{80} = 10 \log\left(\frac{\left(\int_0^{80 \text{ ms}} h^2(t) dt\right)}{\left(\int_{80 \text{ ms}}^{\infty} h^2(t) dt\right)}\right) \text{ dB}$$

Figure 7a,b present the C_{80} values across frequency in octave bands, derived from RIRs measured by the Rode NT-SF1 microphone and em32 Eigenmike, respectively. C_{80} has an inverse relation to T_{20} [35] and we can observe this between Figures 5 and 7. Similar to the T_{20} plots in Figure 5, the distinction between the felt and wood panels becomes apparent after 500 Hz, where the felt panel environment exhibits higher clarity. This aligns with expectations, as felt panels absorb more late reflections, resulting in enhanced clarity.



(a)



(b)

Figure 7. Clarity (C_{80}) across frequency in octave bands for the felt and wood environments calculated from the RIRs measured by (a) Rode NT-SF1 microphone, (b) em32 Eigenmike.

Figure 8a,b illustrate 2D heatmaps of C_{80} values across receiver locations for both room paneling environments and all three source locations, calculated from the RIRs measured by the NT-SF1 microphone and em32 Eigenmike, respectively. Higher C_{80} values are concentrated directly in front of the source location, with lower values for receivers situated further away. This underscores the influence of loudspeaker directivity and interference of reflections from the walls. Source S3 comparatively has the lowest levels of clarity. This

may be potentially due to its proximity to glass doors and windows along the control panel wall, leading to increased interference from reflections. In both felt and wood environments, the highest C_{80} values were observed for source S1. These observations are crucial for recording singers and speech sources, guiding the selection of optimal source–receiver placements.

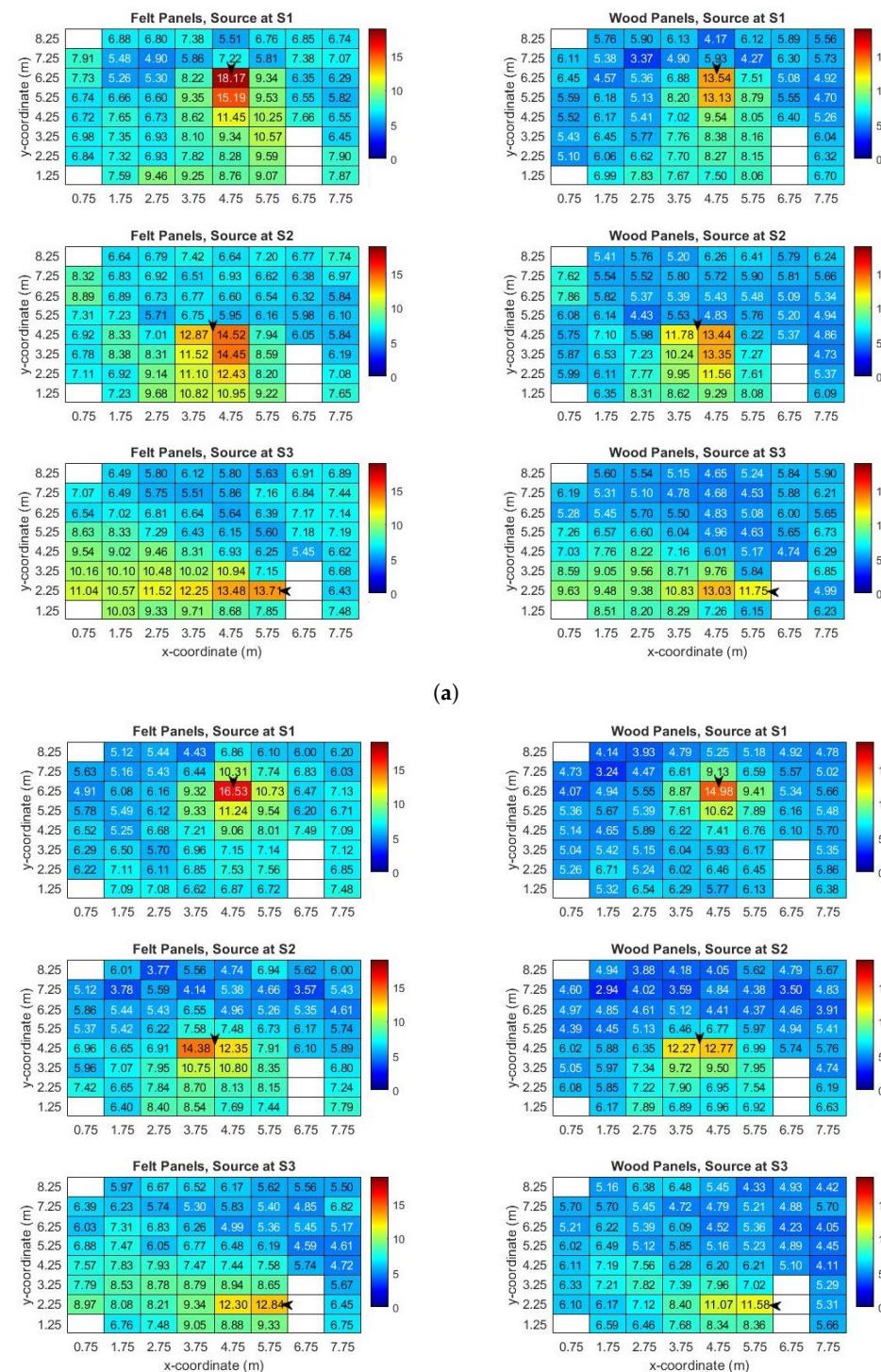


Figure 8. Clarity (C_{80}) in dB at different receiver locations for each of the three source locations and the two wall paneling types calculated from the RIRs measured by (a) Rode NT-SF1 microphone, (b) em32 Eigenmike. The black arrowhead indicates the source location and orientation. The white areas represent spaces where no receivers were positioned.

The spatial spread of high clarity around the sources is more restricted in Figure 8b compared to Figure 8a. This again corroborates the differences in spatial capturing capabilities of the two microphones, necessitating further analysis as mentioned in Section 4.1.

4.3. Definition

The acoustic definition (D_{50}) serves as a valuable metric for speech intelligibility, quantifying the distinctness of sound [1]. It is calculated from the RIR $h(t)$ as follows [1,31]:

$$D_{50} = \frac{\int_0^{50 \text{ ms}} h^2(t) dt}{\int_0^{\infty} h^2(t) dt} \times 100\%$$

Figure 9a,b present the D_{50} values across frequency in octave bands, derived from RIRs measured by the Rode NT-SF1 microphone and em32 Eigenmike, respectively. The overall trend aligns with C_{80} as they are closely correlated parameters [1]. Felt panels yield a higher level of definition, with an average of 71%, compared to wooden panels, with an average of 67%. The average D_{50} suggests that this recording studio can ensure a minimum of 90% syllable intelligibility in both environments [1].

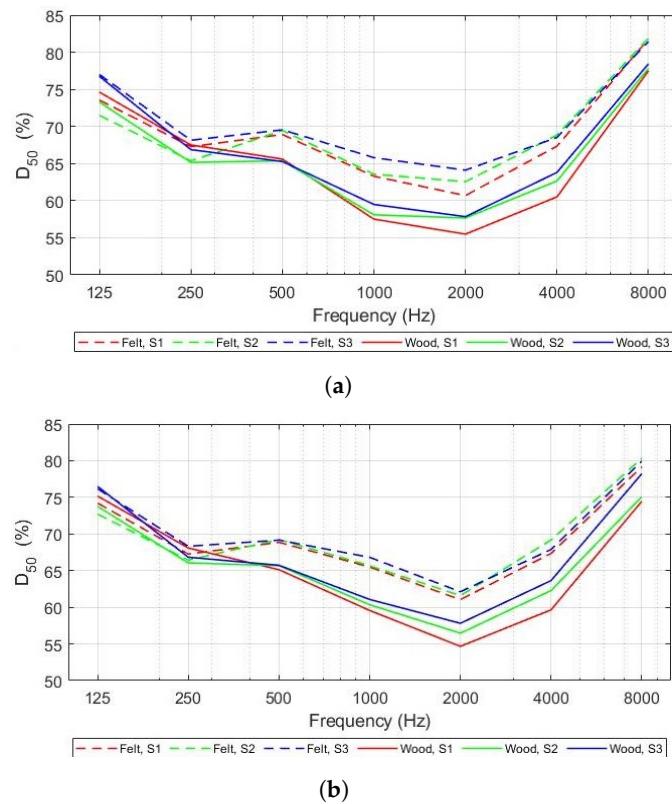


Figure 9. Definition (D_{50}) across frequency in octave bands for the felt and wood environments calculated from the RIRs measured by (a) Rode NT-SF1 microphone, (b) em32 Eigenmike.

4.4. Center Time

Center time (T_s) signifies the center of gravity or first moment of the squared RIR $h^2(t)$ and is computed as follows [1,31]:

$$T_s = \frac{\int_0^{\infty} t h^2(t) dt}{\int_0^{\infty} h^2(t) dt}$$

Figure 10a,b present the T_s values across frequency in octave bands, derived from RIRs measured by the Rode NT-SF1 microphone and em32 Eigenmike, respectively. The overall trend aligns with T_{20} shown in Figure 5. Similar to other parameters, the distinction

between the felt and wood panels becomes evident after 500 Hz. The felt and wood panels yield an average T_s of 0.04 s and 0.05 s, respectively. These values indicate that the felt environment can offer slightly higher music transparency and better speech intelligibility than the wood environment [1].

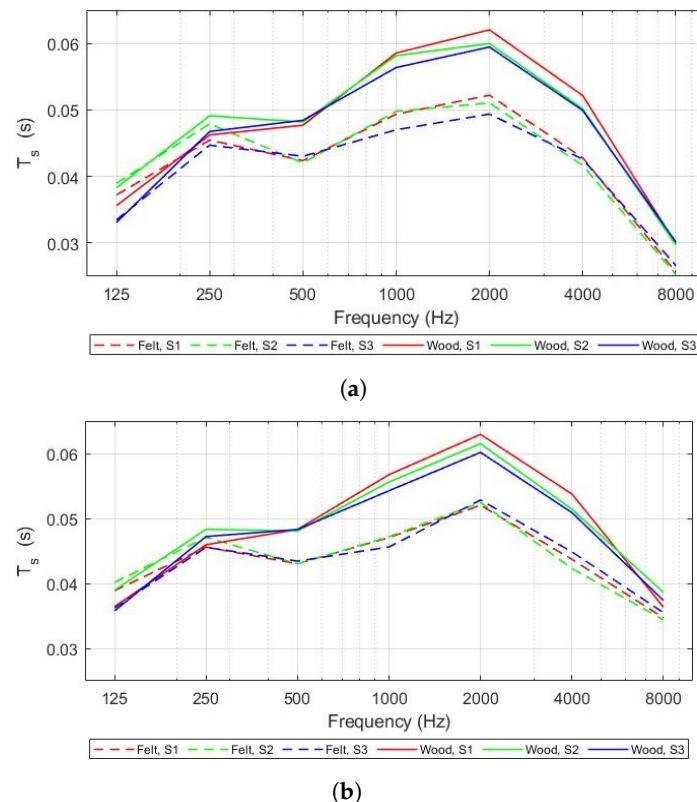


Figure 10. Center time (T_s) across frequency in octave bands for the felt and wood environments calculated from the RIRs measured by (a) Rode NT-SF1 microphone, (b) em32 Eigenmike.

5. Noise Interference and Outlier Measurements

During the organization of the dataset, an outlier measurement by the NT-SF1 microphone was identified at the receiver coordinate ($x = 0.75, y = 5.25, z = 1.7$) m labeled as 00x04y. The irregularity in the W-channel RIR at 00x04y, compared to the neighboring receiver point at 00x03y, is shown in Figure 11.

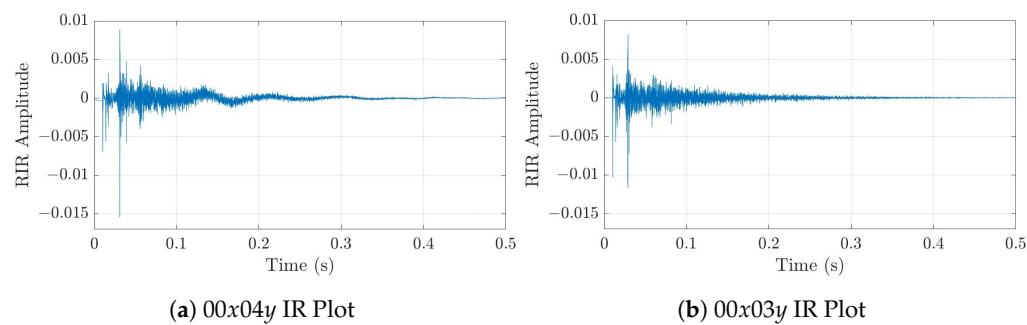


Figure 11. RIRs measured by the NT-SF1 microphone in the wood panel environment with source at S2 and receiver at the coordinates (a) 00x04y and (b) 00x03y.

Upon closer investigation, it became apparent that the anomaly was likely caused by air conditioner noise, given the proximity of a duct directly above this receiver point. Measurements repeated at and around this receiver point confirmed that the issue persisted at different heights at ($x = 0.75, y = 5.25$) m. However, the distortion diminished when the receiver was shifted by approximately 50 cm in the x - y plane.

In the recording studio, there are four air conditioner ducts, but only the duct above $04x00y$ appeared to influence the measurements. The impact of the other ducts was either obscured by furnishings or mitigated by felt surfaces. It is noteworthy that the measurements taken by the em32 Eigenmike at this position did not exhibit any noticeable distortions.

Increasing the SNR with respect to the receiver at $00x04y$ can mitigate the effects of noise interference. However, for this dataset, the loudspeaker volume was set to emulate typical studio performance levels and was kept unchanged while measuring at all receiver positions, reflecting real performance scenarios. Moreover, the large receiver region considered in our dataset makes SNR optimization impractical.

Despite the acknowledged impact of air conditioner noise, it is essential to highlight that the RIR measurements were intentionally conducted without switching off the air conditioners to capture real recording scenarios. Therefore, the RIRs corresponding to the receiver code $00x04y$ measured by the NT-SF1 microphone are still included in the dataset but separated into an ‘Outlier’ folder, warranting consideration for their specific nature in any application.

6. Conclusions

This paper introduced the RSoANU dataset, a collection of RIRs recorded in the Australian National University (ANU) Recording Studio. The dataset encompasses diverse acoustic scenarios featuring two wall panel environments, three source positions, and a planar grid of receiver locations. Recorded with both a 32-channel spherical microphone array and a B-format soundfield microphone array, the dataset includes 396 32-channel and 396 4-channel RIRs.

We utilized the RSoANU dataset to analyze the acoustics within the ANU Recording Studio by examining key parameters such as reverberation time (T_{20}), clarity (C_{80}), definition (D_{50}), and center time (T_s). The results highlighted distinct differences between felt and wood panel environments, emphasizing the impact of panel material on acoustic properties. Felt panels consistently exhibited faster sound decay, higher clarity, and better definition compared to wood panels, particularly in the mid-to-high frequency range. The variations in the parameter values across the receiver grid for different source positions revealed the influence of room environment and source–receiver positioning on acoustic features. Notably, C_{80} exhibited higher values directly in front of the source location, while T_{20} showed elevated values near the front and right walls.

The RSoANU dataset holds the potential to contribute to the research and development of various audio applications, facilitating their validation and enhancement in authentic acoustic environments. It can aid in the designing of recording studios and guide optimal source and receiver placements across different recording scenarios. Future work will involve a more in-depth acoustic analysis, focusing on examining the trade-offs between information captured by a higher-order spherical microphone array and a commercially viable B-format soundfield microphone array.

Author Contributions: Conceptualization, G.C., A.B. and T.A.; data curation, G.C. and A.B.; formal analysis, G.C. and A.B.; investigation, G.C.; methodology, G.C. and A.B.; project administration, G.C., A.B. and T.A.; resources, A.B. and T.A.; supervision, A.B. and T.A.; validation, G.C., A.B. and T.A.; visualization, G.C. and A.B.; writing—original draft, G.C. and A.B.; writing—review and editing, G.C., A.B. and T.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: RSoANU dataset is openly available in Zenodo at <https://doi.org/10.5281/zenodo.10720345> under the Creative Commons Attribution 4.0 International license.

Acknowledgments: Authors would like to thank Matthew Barnes and Craig Greening of the ANU School of Music for providing access to the recording studio and granting permission to conduct the measurements.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANU	Australian National University
RIR	Room Impulse Response
VR	Virtual Reality
XR	eXtended Reality
ESS	Exponential Swept Sine
SNR	Signal-to-Noise Ratio

Appendix A

Tables A1–A3 display the Cartesian coordinates and code names corresponding to the source, receiver, and source-dependent receiver locations, respectively.

Table A1. Source coordinates (in meters) and code names.

Code Name	X	Y	Z
S1	4.75	6.75	1.2
S2	4.25	4.75	1.384
S3	6	2.25	0.93

Table A2. Receiver coordinates (in meters) and code names.

Code Name	X	Y	Z	Code Name	X	Y	Z
00x00y	0.75	1.25	1.7	04x00y	4.75	1.25	1.7
00x01y	0.75	2.25	1.7	04x01y	4.75	2.25	1.7
00x02y	0.75	3.25	1.7	04x02y	4.75	3.25	1.7
00x03y	0.75	4.25	1.7	04x03y	4.75	4.25	1.7
00x04y	0.75	5.25	1.7	04x04y	4.75	5.25	1.7
00x05y	0.75	6.25	1.7	04x05y	4.75	6.25	1.7
00x06y	0.75	7.25	1.7	04x06y	4.75	7.25	1.7
00x07y	0.75	8.25	1.7	04x07y	4.75	8.25	1.7
01x00y	1.75	1.25	1.7	03x00y	5.75	1.25	1.7
01x01y	1.75	2.25	1.7	05x01y	5.75	2.25	1.7
01x02y	1.75	3.25	1.7	05x02y	5.75	3.25	1.7
01x03y	1.75	4.25	1.7	05x03y	5.75	4.25	1.7
01x04y	1.75	5.25	1.7	05x04y	5.75	5.25	1.7
01x05y	1.75	6.25	1.7	05x05y	5.75	6.25	1.7
01x06y	1.75	7.25	1.7	05x06y	5.75	7.25	1.7
01x07y	1.75	8.25	1.7	05x07y	5.75	8.25	1.7
02x00y	2.75	1.25	1.7	06x00y	6.75	1.25	1.7
02x01y	2.75	2.25	1.7	06x01y	6.75	2.25	1.7
02x02y	2.75	3.25	1.7	06x02y	6.75	3.25	1.7
02x03y	2.75	4.25	1.7	06x03y	6.75	4.25	1.7
02x04y	2.75	5.25	1.7	06x04y	6.75	5.25	1.7
02x05y	2.75	6.25	1.7	06x05y	6.75	6.25	1.7
02x06y	2.75	7.25	1.7	06x06y	6.75	7.25	1.7
02x07y	2.75	8.25	1.7	06x07y	6.75	8.25	1.7
03x00y	3.75	1.25	1.7	07x00y	7.75	1.25	1.7
03x01y	3.75	2.25	1.7	07x01y	7.75	2.25	1.7
03x02y	3.75	3.25	1.7	07x02y	7.75	3.25	1.7

Table A2. Cont.

Code Name	X	Y	Z	Code Name	X	Y	Z
03x03y	3.75	4.25	1.7	07x03y	7.75	4.25	1.7
03x04y	3.75	5.25	1.7	07x04y	7.75	5.25	1.7
03x05y	3.75	6.25	1.7	07x05y	7.75	6.25	1.7
03x06y	3.75	7.25	1.7	07x06y	7.75	7.25	1.7
03x07y	3.75	8.25	1.7	07x07y	7.75	8.25	1.7

Table A3. Coordinates (in meters) and code names of additional receiver locations corresponding to each source.

Source Code Name	Code Name	X	Receiver Y	Z
S1	e30x45y	3.75	5.75	1.7
	e35x45y	4.25	5.75	1.7
	e35x50y	4.25	6.25	1.7
	e40x45y	4.75	5.75	1.7
	e45x45y	5.25	5.75	1.7
	e45x50y	5.25	6.25	1.7
S2	e25x25y	3.25	3.75	1.7
	e25x30y	3.25	4.25	1.7
	e30x25y	3.75	3.75	1.7
	e35x25y	4.25	3.75	1.7
	e35x30y	4.25	4.25	1.7
	e40x25y	4.75	3.75	1.7
	e45x25y	5.25	3.75	1.7
	e45x30y	5.25	4.25	1.7
S3	e45x00y	5.25	1.25	1.7
	e45x05y	5.25	1.75	1.7
	e45x10y	5.25	2.25	1.7
	e45x15y	5.25	2.75	1.7
	e45x20y	5.25	3.25	1.7
	e50x05y	5.75	1.75	1.7
	e50x15y	5.75	2.75	1.7

References

- Kuttruff, H. *Room Acoustics*, 6th ed.; CRC Press: Boca Raton, FL, USA, 2016. [[CrossRef](#)]
- Ando, Y. Design Study. In *Concert Hall Acoustics*; Springer: Berlin/Heidelberg, Germany, 1985; Chapter 6, pp. 89–101. [[CrossRef](#)]
- Farina, A.; Amendola, A.; Capra, A.; Varani, C. Spatial Analysis of Room Impulse Responses Captured with a 32-Capsule Microphone Array. In Proceedings of the 130th Audio Engineering Society Convention, London, UK, 13–16 May 2011.
- Cox, T.J.; D’Antonio, P.; Avis, M.R. Room Sizing and Optimization at Low Frequencies. *J. Audio Eng. Soc.* **2004**, *52*, 640–651.
- Long, M. *Architectural Acoustics*, 2nd ed.; Academic Press: Cambridge, MA, USA, 2014.
- Dong, H.Y.; Lee, C.M. Speech intelligibility improvement in noisy reverberant environments based on speech enhancement and inverse filtering. *Eurasip J. Audio Speech Music Process.* **2018**, *2018*, 1–13. [[CrossRef](#)]
- Kim, H.; Remaggi, L.; Jackson, P.J.; Hilton, A. Immersive Spatial Audio Reproduction for VR/AR Using Room Acoustic Modelling from 360° Images. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces, Osaka, Japan, 23–27 March 2019; pp. 120–126. [[CrossRef](#)]
- Vorländer, M. Computer simulations in room acoustics: Concepts and uncertainties. *J. Acoust. Soc. Am.* **2013**, *133*, 1203–1213. [[CrossRef](#)] [[PubMed](#)]
- Schröder, D. Physically Based Real-Time Auralization of Interactive Virtual Environments. Ph.D. Thesis, RWTH Aachen University, Aachen, Germany, 2011.
- Brinkmann, F.; Aspöck, L.; Ackermann, D.; Lepa, S.; Vorländer, M.; Weinzierl, S. A round robin on room acoustical simulation and auralization. *J. Acoust. Soc. Am.* **2019**, *145*, 2746–2760. [[CrossRef](#)] [[PubMed](#)]
- Murphy, D.T.; Shelley, S. OpenAIR: An Interactive Auralization Web Resource and Database. In Proceedings of the 129th Audio Engineering Society Convention, San Francisco, CA, USA, 4–7 November 2010.

12. Götz, G.; Schlecht, S.J.; Pulkki, V. A Dataset of Higher-Order Ambisonic Room Impulse Responses and 3D Models Measured in a Room with Varying Furniture. In Proceedings of the 2021 Immersive and 3D Audio: From Architecture to Automotive (I3DA), Bologna, Italy, 8–10 September 2021; pp. 1–8. [[CrossRef](#)]
13. Kearney, G.; Daffern, H.; Cairns, P.; Hunt, A.; Lee, B.; Cooper, J.; Tsagkarakis, P.; Rudzki, T.; Johnston, D. Measuring the Acoustical Properties of the BBC Maida Vale Recording Studios for Virtual Reality. *Acoustics* **2022**, *4*, 783–799. [[CrossRef](#)]
14. Stewart, R.; Sandler, M. Database of omnidirectional and B-format room impulse responses. In Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, USA, 14–19 March 2010; pp. 165–168. [[CrossRef](#)]
15. Hadad, E.; Heese, F.; Vary, P.; Gannot, S. Multichannel audio database in various acoustic environments. In Proceedings of the 2014 14th International Workshop on Acoustic Signal Enhancement (IWAENC), Juan-les-Pins, France, 8–11 September 2014; pp. 313–317. [[CrossRef](#)]
16. Eaton, J.; Gaubitch, N.D.; Moore, A.H.; Naylor, P.A. Estimation of Room Acoustic Parameters: The ACE Challenge. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2016**, *24*, 1681–1693. [[CrossRef](#)]
17. Szöke, I.; Skácel, M.; Mošner, L.; Paliesek, J.; Černocký, J. Building and evaluation of a real room impulse response dataset. *IEEE J. SElect. Top. Signal Process.* **2019**, *13*, 863–876. [[CrossRef](#)]
18. Carlo, D.D.; Tandteitnik, P.; Foy, C.; Bertin, N.; Deleforge, A.; Gannot, S. dEchorate: A Calibrated Room Impulse Response Dataset For Echo-Aware Signal Processing. *EURASIP J. Audio Speech, Music Process.* **2021**, *2021*, 1–15. [[CrossRef](#)]
19. Koyama, S.; Nishida, T.; Kimura, K.; Abe, T.; Ueno, N.; Brunnström, J. MESHRR: A Dataset of Room Impulse Responses on Meshed Grid Points for Evaluating Sound Field Analysis and Synthesis Methods. In Proceedings of the 2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, 17–20 October 2021; pp. 1–5. [[CrossRef](#)]
20. Prawda, K.; Schlecht, S.J.; Välimäki, V. Calibrating the Sabine and Eyring formulas. *J. Acoust. Soc. Am.* **2022**, *152*, 1158–1169. [[CrossRef](#)] [[PubMed](#)]
21. Recording Studio—ANU School of Music Sound Recording & Music Technology Studios. Available online: <https://music.cass.anu.edu.au/services/studio-services/recording-studio> (accessed on 4 February 2024).
22. John Sayers Productions. The Recording Manual Part II—Studio Design. 2000. Available online: <https://www.tendolla.com/design/studio%20design.pdf> (accessed on 28 February 2024).
23. TANNOY. TANNOY SYSTEM 600 NEARFIELD MONITOR. Available online: <https://www.fullcompass.com/common/files/2709-System600SpecSheet.pdf> (accessed on 4 February 2024).
24. RODE. NT-SF1. Available online: https://edge.rode.com//pdf/products/94/NT-SF1_V02.pdf (accessed on 28 February 2024).
25. mh acoustics. Em32 Eigenmike® Microphone Array Release Notes (v18.0). 2014. Available online: <https://mhacoustics.com/sites/default/files/EigenmikeReleaseNotesV18.pdf> (accessed on 28 February 2024).
26. mh acoustics. Eigenbeam Data Specification for Eigenbeams. 2016. Available online: https://www.lesonbinaural.fr/EDIT/DOCS/mh_acoustics_eigenmike.PDF (accessed on 28 February 2024).
27. Zoom. The Zoom F6. Available online: <https://zoomcorp.com/en/us/field-recorders/field-recorders/f6/> (accessed on 1 March 2024).
28. 4338.0—Profiles of Health, Australia, 2011–13. Available online: www.abs.gov.au/ausstats/abs@.nsf/lookup/4338.0main+features212011-13 (accessed on 28 February 2024).
29. Guidorzi, P.; Barbaresi, L.; D’Orazio, D.; Garai, M. Impulse Responses Measured with MLS or Swept-Sine Signals Applied to Architectural Acoustics: An In-depth Analysis of the Two Methods and Some Case Studies of Measurements Inside Theaters. *Energy Procedia* **2015**, *78*, 1611–1616. [[CrossRef](#)]
30. Cheshire, G.; Bastine, A.; Abhayapala, T.D. RSoANU: RIR Dataset. Zenodo, Dataset. 2024. Available online: <https://zenodo.org/records/10720345> (accessed on 28 February 2024).
31. ISO 3382-1; Acoustics—Measurement of Room Acoustic Parameters—Part 1: Performance Spaces. International Organization for Standardization: Geneva, Switzerland, 2009.
32. Karjalainen, M.; Ansalo, P.; Mäkivirta, A.; Peltonen, T.; Välimäki, V. Estimation of Modal Decay Parameters from Noisy Response Measurements. *J. Audio Eng. Soc.* **2002**, *50*, 867–878.
33. Gulsrud, T.; Exton, P.; van der Harten, A.; Kirkegaard, L. Room Acoustics Investigations in Hamer Hall at the Arts Centre, Melbourne. In Proceedings of the International Symposium on Room Acoustics (ISRA), Melbourne, Australia, 29–31 August 2010.
34. Kimmich, J.M.; Schlesinger, A.; Ochmann, M.; Frank, S. Optimisation of the Orchestra Pit Acoustics in Opera Houses by Acoustical Simulations using the Finite Element Method. *MATEC Web Conf.* **2018**, *251*, 1–8. [[CrossRef](#)]
35. Adelman-Larsen, N.W.; Thompson, E.R.; Gade, A.C. Suitable reverberation times for halls for rock and pop music. *J. Acoust. Soc. Am.* **2010**, *127*, 247–255. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.