DOI: 10.61782/fa.2025.0508

FORUM ACUSTICUM EURONOISE 2025

IR64-CAR: A DATASET OF IMPULSE RESPONSES CAPTURED WITH A SPHERICAL MICROPHONE ARRAY FOR AUTOMOTIVE ACOUSTIC RESEARCH

Jesus Lopez-Ballester[1*] **Gioele Greco**[2] **Sandra Roger**[3]
Jose J. Lopez[4] **Francesc J. Ferri**[3] **Maximo Cobos**[3]
[1] Department of Electronic Engineering, Universitat de València, Spain
[2] Department of Electronic Informatics and Bioinformatics, Politenico di Milano, Milan, Italy
[3] Computer Science Department, Universitat de València, Spain
[4] iTEAM Institute, Universitat Politècnica de València, Spain

ABSTRACT

This work presents the creation of a novel dataset of impulse responses captured with a 64-capsule microphone array (Eigenmike 64) positioned at the center of a car cabin. The dataset was generated to address the growing demand for realistic acoustic data in automotive audio research, spatial audio systems, and machine learning applications. Impulse responses were recorded with sound sources positioned at 8 angular locations around the car (every 45°), simulating real-world external noise conditions. All measurements were conducted with the car windows closed, providing a controlled acoustic environment representative of standard driving conditions. This configuration enables the study of how sound propagates, reflects, and penetrates the vehicle's interior under typical isolation settings. The primary applications of this dataset include the simulation of in-vehicle recordings, enabling realistic testing and validation of audio systems; source localization, to determine the direction of external sounds entering the vehicle; spatial filtering, supporting beamforming and noise suppression techniques; and advanced driver-assistance systems (ADAS), to improve external sound detection and classification for safety-critical scenarios. Calibration and validation procedures were applied to ensure the accuracy and reliability of the dataset. This resource provides significant potential for advancing both academic research and industrial innovation in automotive acoustics.

**Corresponding author*: *jesus.lopez-ballester@uv.es.*

Keywords: *dataset, spherical microphone, ambisonics, impulse responses, vehicles.*

1. INTRODUCTION

The increasing integration of advanced audio and sensing technologies in modern vehicles has amplified the need for realistic, high-resolution acoustic datasets. Applications such as in-cabin sound simulation, source localization, spatial filtering, and acoustic scene analysis require precise modeling of how external and internal sounds interact within the vehicle interior [1–3]. In particular, the development of machine learning models and spatial audio systems for automotive use relies heavily on representative datasets that capture the complex acoustic behavior of real-world environments. One of the key challenges in automotive acoustic research is characterizing how sound propagates from the exterior into the cabin [4, 5]. Realistic impulse response (IR) data is essential for a wide range of applications, from training models for direction-of-arrival (DoA) estimation and beamforming to evaluating signal enhancement methods and designing perceptually optimized spatial audio systems. Such data also underpins advanced driver-assistance systems (ADAS) that must reliably detect, localize, and classify external auditory cues under varying acoustic conditions [6, 7].

Several existing datasets have made strides in providing acoustic data for automotive contexts. For example, the ANIR corpus dataset [8] offers valuable binaural recordings of background noises and impulse responses in various driving situations, while other works provide datasets for audio-visual events in automotive scenarios [9–11]. However, these resources are often limited either in spatial resolution, in the number of source directions, or in the availability of multichannel measurements suitable for high-order spatial processing. Furthermore, while outside the automotive context, several datasets of measured IRs in high-order ambisonics (HOA) format are available [12–15], there remains a gap for high-resolution, three-dimensional spatial audio captured in realistic automotive interiors.

To address these limitations, we present IR64-CAR, a novel dataset of automotive impulse responses recorded using the Eigenmike EM64, a 64-capsule spherical microphone array capable of high-order Ambisonics capture. The array was positioned at the geometric center of a standard passenger vehicle cabin, and impulse responses were recorded from eight equidistant source directions (every 45°) at ear-height level. All recordings were conducted with the car windows closed, ensuring a consistent acoustic environment representative of typical driving conditions. This experimental setup enables detailed analysis of how external sounds propagate and reflect within the vehicle's interior, supporting research into spatial audio processing, source localization, and in-cabin sound simulation.

The use of a spherical microphone array allows the dataset to be represented and processed in the spherical harmonic (SH) domain [16], a powerful mathematical framework for spatial audio. SH-based techniques enable efficient and compact representations of the sound field and are particularly well-suited for spatial filtering, beamforming, and source separation [17–19]. Moreover, the high-order ambisonics data supports virtual microphone synthesis and spatial rendering, making IR64-CAR a valuable resource for researchers and engineers in both audio signal processing and machine learning domains.

In summary, IR64-CAR offers the following key contributions:

- A comprehensive set of impulse responses captured in a real car cabin using a high-resolution spherical microphone array.
- Measurements from eight spatial directions simulating external noise sources around the vehicle, with full azimuthal coverage at 45° intervals.
- Recordings conducted with the car windows closed, ensuring consistent acoustic isolation representative of typical in-vehicle scenarios.
- Data suitable for spherical harmonic analysis, enabling advanced spatial audio processing, source localization, and beamforming applications.

This paper describes the dataset acquisition process, calibration procedures, and potential applications, and provides benchmarks for future work in automotive acoustic modeling, source localization, spatial filtering, and in-vehicle sound simulation.

2. DATASET DESIGN AND ACQUISITION

The IR64-CAR dataset was designed to provide a realistic and high-resolution representation of how external sounds interact with the interior of a passenger vehicle. This section details the hardware setup, recording conditions, and procedures used to capture the impulse responses.

2.1 Microphone Array and Recording System

Recordings were made using the Eigenmike EM64 spherical microphone array, which consists of 64 omnidirectional capsules mounted on the surface of a rigid sphere with a diameter of 8.4 cm. This array supports high-order Ambisonics (HOA) up to the 7^{th} order, allowing for precise spatial resolution in three dimensions.

The array was connected to a multichannel audio interface, and signals were recorded at a sampling rate of 48 kHz with a bit depth of 24 bits. The data was stored as 64-channel WAV files for each measurement, preserving the raw microphone signals prior to any Ambisonics encoding.

2.2 Vehicle Environment and Microphone Placement

Impulse responses were recorded inside a 2016 Peugeot 208, a common mid-sized passenger vehicle. All interior surfaces (dashboard, seats, windows, upholstery) were left in their standard configuration, and no passengers were present during recording.

The EM64 array was positioned at the geometric center of the cabin, roughly aligned with the midpoint between the front and rear seats and centered laterally between the driver and passenger sides. The height of the array was set at 1.2 m, approximating the ear level of a seated occupant. Figure 1 shows a photograph of the microphone positioning inside the car cabin.

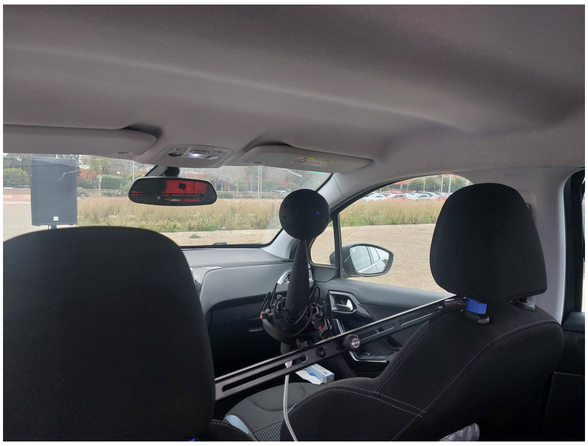

Figure 1. EM64 placement inside the car.

2.3 Recording Environment

The measurements were conducted in an open urban environment characterized by relatively low ambient noise and minimal acoustic reflections from surrounding structures. Specifically, the vehicle was parked in a large, paved courtyard adjacent to a university campus, with nearby buildings located at distances no closer than 90 meters. This setup provided a practical compromise between a fully anechoic outdoor setting and a typical urban context. Figure 2 shows the measurement environment.

Although the location is within a city, the selected area is notably clear, with unobstructed line-of-sight in all horizontal directions and reduced pedestrian or vehicular activity during recording hours. This was essential to minimize contamination of the impulse responses with environmental noise or spurious reflections, while still preserving the realistic boundary conditions that arise from open-air propagation. Additionally, recording in an outdoor urban space—as opposed to an anechoic chamber—helps preserve external sound behavior that is relevant to real-world automotive scenarios, such as mild ground reflections and natural spectral shaping. To further ensure signal quality, recordings were conducted during off-peak hours when ambient traffic and background noise levels were consistently below 35 dBA. Periodic background noise measurements confirmed that no significant interference occurred during the acquisition process.

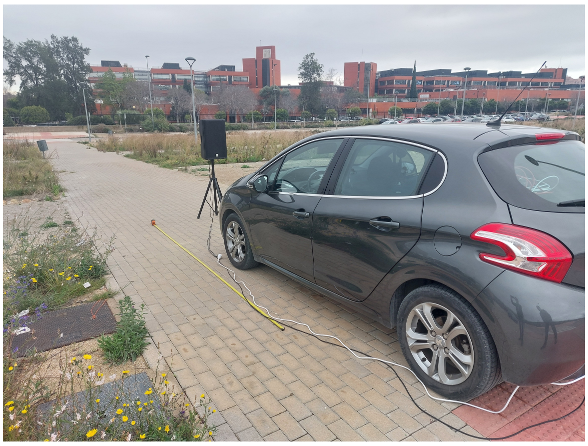

Figure 2. Recording environment.

2.4 Sound Source Configuration

To simulate realistic external noise sources, a single loudspeaker was used to emit excitation signals from eight distinct angular positions distributed uniformly around the vehicle at 45 ° azimuth increments (0 °, 45 °,..., 315 °), as shown in Figure 3. The elevation angle was fixed at 0°, corresponding to a horizontal plane relative to the car.

The loudspeaker was placed at a distance of 1.5 m from the front and rear vehicle boundary, following a circular shape and aligned such that it was facing the center of the car. A logarithmic sweep was used to excite the environment, enabling the extraction of the impulse response via deconvolution [20].

2.5 Acoustic Conditions

All impulse responses in the IR64-CAR dataset were recorded with the car windows and doors fully closed. This configuration represents the most common acoustic condition encountered during everyday driving and ensures a consistent and controlled environment for all measurements. Recording under sealed-cabin conditions allows for a focused study of how external sounds propagate into and interact with the interior space of a vehicle with standard acoustic isolation. This setup also facilitates the development and evaluation of audio processing algorithms—such as beamforming, source localization, and in-cabin sound simulation—under realistic and reproducible conditions.

2.6 Calibration and Data Processing

Prior to recording, a calibration procedure was performed to ensure consistent playback levels across all loudspeaker positions. The excitation signal used for the measurements was a logarithmic sine sweep, which is a widely adopted technique for obtaining high dynamic range impulse responses with robust separation of linear and nonlinear components. The sweep was pre-equalized to compensate for the loudspeaker's frequency response and was played back at a level ensuring a signal-to-noise ratio above 30 dB inside the cabin.

Impulse responses were extracted from the recorded signals using deconvolution of the sweep with its time-reversed version. This method produces a clean impulse response with minimal artifacts and allows for the separation of harmonic distortion components, which appear temporally after the linear impulse response. Each extracted IR was trimmed to a length of 1 second to preserve the full decay of reflections within the cabin.

To ensure the quality and consistency of the dataset, the frequency response of each impulse response was analyzed across all microphone channels immediately after each measurement. This verification step involved inspecting the spectral magnitude of the responses to confirm uniformity and to detect any anomalies such as spectral nulls, clipping artifacts, or unexpected deviations across the array. Measurements exhibiting irregularities—due to environmental noise, loudspeaker misalignment, or technical issues—were repeated. This procedure ensured that the dataset maintained a high standard of reliability suitable for spatial analysis and machine learning applications.

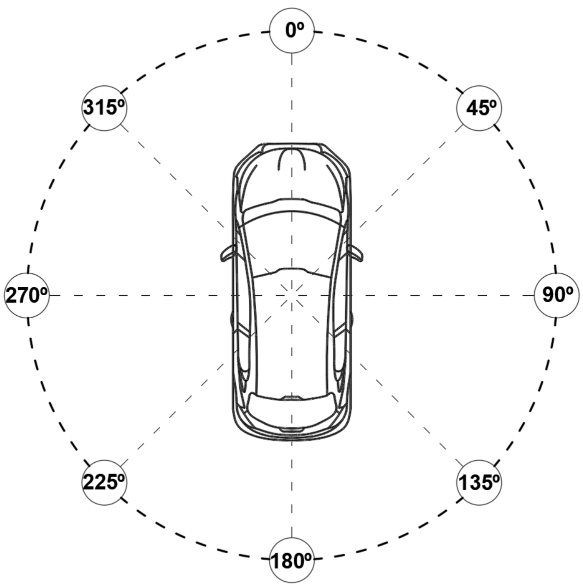

Figure 3. Measured source positions.

2.7 Data Format and Availability

The dataset is provided as raw 64-channel WAV files for each measurement, containing the impulse responses captured by the spherical microphone array. Each file corresponds to one source azimuth direction under consistent recording conditions. Metadata files accompany the audio recordings and specify the azimuth angle of the sound source, the recording parameters, and additional session information.

The complete dataset and accompanying documentation will be made publicly available at https://zenodo.org/records/15168578 upon publication.

3. VALIDATION AND ANALYSIS

To ensure the reliability and scientific utility of the IR64-CAR dataset, a series of validation and analysis procedures were conducted. These focused on assessing the acoustic characteristics of the recorded impulse responses, verifying measurement consistency across channels, and confirming the spatial diversity of the dataset. This section presents several representative analyses, including energy decay metrics, frequency-domain inspection, and spatial response patterns.

3.1 Energy Decay and Reverberation Analysis

Although the car cabin is a small and acoustically damped environment, time-domain energy decay profiles can provide valuable insight into the reflective characteristics of the space. Figure 4 shows T20 reverberation time estimates computed from octave-band-filtered impulse responses across several source directions.

The estimated T20 values exhibit short decay times, as expected for an enclosed vehicle cabin, typically below 300 ms across most frequency bands. Slight variations are observable between source directions, particularly at mid-frequencies (e.g., 500 Hz to 2 kHz), which are consistent with the dimensions of typical vehicle cabins and the differential absorption characteristics of materials such as upholstery, glass, and plastic surfaces.

These measurements serve as a quantitative description of the vehicle's acoustic response and may inform simulations or perceptual models.

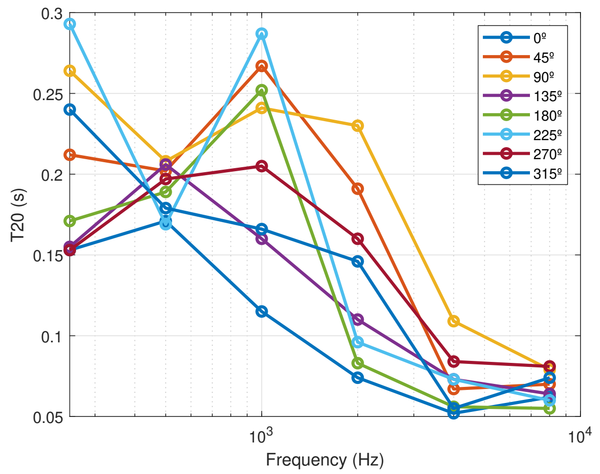

Figure 4. T20 reverberation time estimates for multiple source directions. Values were calculated using the backward integration method applied to octave-band filtered IRs.

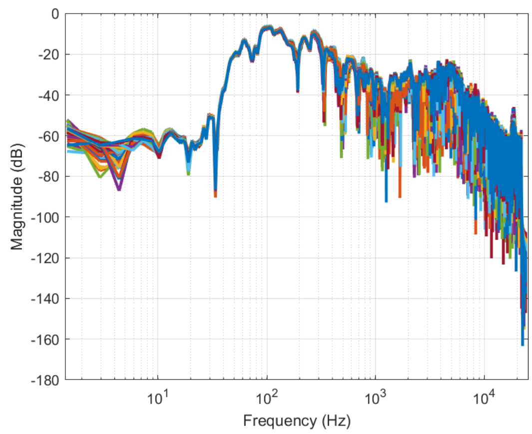

Figure 5. Frequency responses across all 64 channels for a representative source direction (90°). The responses show strong coherence, indicating a high-quality recording with consistent microphone behavior.

3.2 Channel Consistency and Frequency Response Uniformity

The Eigenmike EM64 array consists of 64 closely matched omnidirectional microphones. To validate the uniformity of the recordings across channels, we computed and plotted the magnitude response of each microphone channel in response to the same excitation. Figure 5 shows overlaid frequency responses for all 64 channels, derived from the Fourier transform of the impulse responses corresponding to a single source direction.

As shown, the responses exhibit strong consistency across the array, with only minor variations attributable to physical positioning on the sphere and small sensitivity differences among the microphones. No channels exhibited spectral anomalies or dropouts, confirming the overall calibration integrity and the reliability of the recorded data.

3.3 Impulse Response Shape for Lateral Directions

To further illustrate the spatial characteristics of the dataset, Figure 6 shows example time-domain impulse responses from lateral source directions at 90° and 270° azimuths for microphone 4, located almost zenith to the sphere. The direct-path peaks are clearly visible, followed by early reflections and decay patterns specific to each direction.

Differences in the timing and energy distribution of these responses highlight the spatial asymmetry of the car cabin and underscore the relevance of directionally resolved data for applications such as beamforming and source localization.

3.4 Directional Spectral Energy Distribution

Finally, to assess how energy varies across directions and frequencies, we computed a directional spectral energy map by integrating the energy of each impulse response in third-octave bands and plotting it as a function of azimuth. Figure 7 illustrates these energy distributions for selected frequency bands.

Some lateral symmetry can be seen with respect to the source positions, for example between the position at 90º and 270º, between 45º and 315º or between 135º and 225º. The observable differences between symmetrical directions are produced by the asymmetrical elements inside the vehicle.

The plots reveal meaningful direction-dependent spectral patterns. For instance, low-frequency energy tends to be more evenly distributed, while mid- and high-frequency bands exhibit stronger directionality, likely influenced by geometric occlusions and material absorption. Such insights are relevant for designing spatial filters, training machine learning models for directional per-

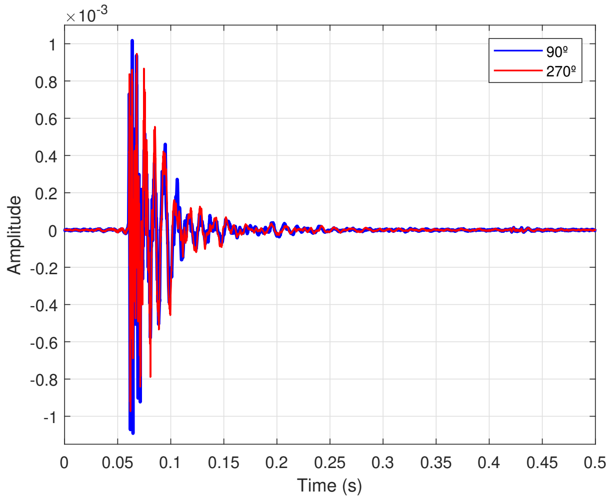

Figure 6. Impulse responses for two lateral source directions (90° and 270° azimuth). Differences in early reflection structure are visible, reflecting cabin asymmetry.

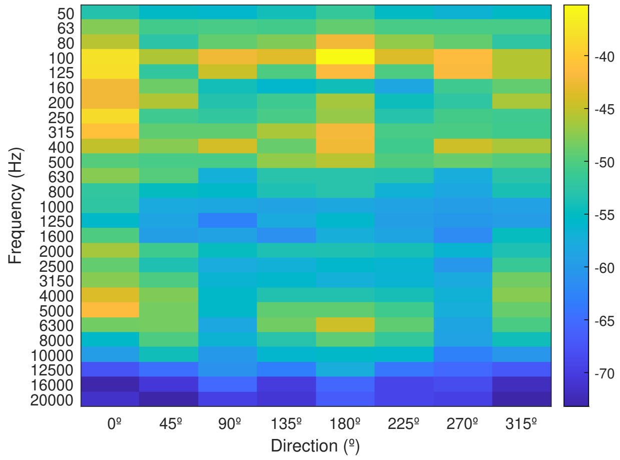

Figure 7. Spectral energy per direction across several frequency bands, averaged over microphone channels. Directionality increases with frequency, due to acoustic shadowing and cabin geometry.

ception, or evaluating in-car playback systems.

4. CONCLUSIONS

In this work, we presented IR64-CAR, a novel dataset of impulse responses captured with a high-order spherical microphone array inside a real passenger vehicle. The dataset fills an important gap in the field of automotive acoustics by providing high-resolution, spatially detailed recordings suitable for applications such as source localization, spatial filtering, in-cabin sound simulation, and machine learning for advanced driver-assistance systems (ADAS). Analyses of energy decay, channel consistency, and directional spectral energy distributions confirm the acoustic realism and spatial diversity of the collected data. However, several limitations must be acknowledged. First, the dataset was recorded under a single set of acoustic conditions — with the vehicle stationary, windows closed, and in a quiet urban environment — which may not capture the full variability encountered in real-world driving scenarios. Additionally, recordings were conducted using only one vehicle model and at a fixed source distance and elevation, limiting the diversity of cabin geometries and external noise conditions represented. Future work will focus on expanding the dataset to include multiple vehicles of different types, dynamic conditions with moving sources, and varying cabin configurations such as partially open windows. Further extensions may also involve recording under different environmental settings (e.g., rain, traffic) and adding more diverse source types and distances.

5. ACKNOWLEDGMENTS

This work has been supported by Grant TED2021-131003B-C21 funded by MCIN/AEI/10.13039/501100011033 and by the "EU Union NextGenerationEU/PRTR", as well as by Grant PID2022-137048OB-C41 funded by MICIU/AEI/10.13039/501100011033 and "ERDF A way of making Europe". Finally, the authors acknowledge as well the Artemisa computer resources funded by the EU ERDF and Comunitat Valenciana, and the technical support of IFIC (CSIC-UV).

6. REFERENCES

[1] G. Coelho, L. M. Matos, P. J. Pereira, A. Ferreira, A. Pilastri, and P. Cortez, "Deep autoencoders for acoustic anomaly detection: experiments with working machine and in-vehicle audio," *Neural Computing and Applications*, vol. 34, no. 22, pp. 19485–19499, 2022.

[2] S. Kim, J. J. E. Chang, H. H. Park, S. U. Song, C. B. Cha, J. W. Kim, and N. K. and, "Autonomous taxi service design and user experience," *International Journal of Human–Computer Interaction*, vol. 36, no. 5, pp. 429–448, 2020.

[3] S. Chu, S. Narayanan, and C.-C. J. Kuo, "Environmental sound recognition with time–frequency audio features," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 6, pp. 1142–1158, 2009.

[4] J. Cheer, *Active Sound Control in the Automotive Interior*, pp. 53–69. Cham: Springer International Publishing, 2021.

[5] *Robust Acoustic Head Orientation Estimation and Speech Enhancement for In-Car Communication Systems*, March 2019.

[6] A. A. Mehta, A. A. Padaria, D. J. Bavisi, V. Ukani, P. Thakkar, R. Geddam, K. Kotecha, and A. Abraham, "Securing the future: A comprehensive review of security challenges and solutions in advanced driver assistance systems," *IEEE Access*, vol. 12, pp. 643–678, 2024.

[7] M. Liu, E. Yurtsever, J. Fossaert, X. Zhou, W. Zimmer, Y. Cui, B. L. Zagar, and A. C. Knoll, "A survey on autonomous driving datasets: Statistics, annotation quality, and a future outlook," *IEEE Transactions on Intelligent Vehicles*, pp. 1–29, 2024.

[8] *A Background Noise and Impulse Response Corpus for Research in Automotive Speech and Audio Processing*, March 2022.

[9] M. Cantarini, L. Gabrielli, A. Mancini, S. Squartini, and R. Longo, "A3carscene: An audio-visual dataset for driving scene understanding," *Data in Brief*, vol. 48, p. 109146, 2023.

[10] P. de Tournemire, D. Nitti, E. Perot, D. Migliore, and A. Sironi, "A large scale event-based detection dataset for automotive," 2020.

[11] C. Castorena, M. Cobos, J. Lopez-Ballester, and F. J. Ferri, "A safety-oriented framework for sound event detection in driving scenarios," *Applied Acoustics*, vol. 215, p. 109719, 2024.

[12] T. Piquet, H. Teyssier, Émile Ouellet-Delorme, R. Duée, and N. Bouillot, "Two datasets of room impulse responses for navigation in 6 degrees of freedom: a symphonic concert hall and a former planetarium," in *25th International Conference on Digital Audio Effects DAFx20*, (Vienna University of Music and Performing Arts (mdw), Austria), pp. pp. 169–176, sep 2022.

[13] F. Miotello, P. Ostan, M. Pezzoli, L. Comanducci, A. Bernardini, F. Antonacci, and A. Sarti, "Homularir: A room impulse response dataset for teleconferencing and spatial audio applications acquired through higher-order microphones and uniform linear microphone arrays," 2024.

[14] G. Chesworth, A. Bastine, and T. Abhayapala, "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," *Applied Sciences*, vol. 14, no. 5, 2024.

[15] M. Wang, S. Clarke, J.-H. Wang, R. Gao, and J. Wu, "Soundcam: A dataset for finding humans using room acoustics," 2024.

[16] B. Rafaely, *Fundamentals of spherical array processing*, vol. 8. Springer, 2015.

[17] M. Cobos, M. Pezzoli, F. Antonacci, and A. Sarti, "Acoustic source localization in the spherical harmonics domain exploiting low-rank approximations," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, 2023.

[18] M. Pezzoli, M. Cobos, F. Antonacci, and A. Sarti, "Sparsity-based sound field separation in the spherical harmonics domain," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1051–1055, 2022.

[19] M. Cobos, J. Ahrens, K. Kowalczyk, and A. Politis, "An overview of machine learning and other data-based methods for spatial audio capture, processing, and reproduction," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2022, no. 1, p. 10, 2022.

[20] A. Farina, "Advancements in impulse response measurements by sine sweeps," in *Audio engineering society convention 122*, Audio Engineering Society, 2007.