ChurchIR: A Dataset of Multichannel Church Impulse Responses for Spatial Audio Applications

Riccardo Giampiccolo, Sofia Parrinelli, and Fabio Antonacci
Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, Milan, Italy
E-mails: {riccardo.giampiccolo, sofia.parrinelli, fabio.antonacci}@polimi.it

***Abstract*—Spatial audio and virtual acoustics research heavily rely on measurement-based datasets of impulse responses to model and simulate real-world environments. Churches, in particular, are of significant interest due to their large, reverberant spaces and complex sound propagation characteristics. To this aim, we introduce ChurchIR, a dataset of church impulse responses acquired for spatial audio applications. The dataset has been captured using a 4th-order microphone and an omnidirectional microphone at thirty receiver positions for three different source locations. Key acoustic metrics and descriptors, including octave-band reverberation times and pseudo-spectra for source localization, are presented to characterize the dataset. ChurchIR offers researchers and developers a resource for spatial audio reproduction and analysis, providing them with data crucial for applications in immersive audio rendering, virtual reality, and architectural acoustics.**

***Index Terms*—church, impulse responses, spatial audio, virtual acoustics.**

Fig. 1: Measurement setup in the church of Saints Marcellino and Pietro, Cremona, Italy.

I. Introduction

Church acoustics is a particular field of room acoustics that has been gaining increasing attention over the past few decades [1]. Since the beginning of the first millennium, religious spaces have played a central role in the development of music as this was considered a means of spiritual expression, community bonding, and ritual enhancement [2]. Restricting our discussion to Christian churches, these are spread all over the world and amount to almost 3.4 million, covering a total area that is estimated to be almost that of Cyprus. The evolution of music and sacred performance were influenced by early Christian worship that took place in reverberant environments, such as catacombs and basilicas, and, with the passing of time, conditioned the architecture of churches to the point that the attention to acoustic details, such as the positioning of choirs and the integration of organs, became a fundamental constant [1].

The preservation and study of church acoustics are now recognized as part of cultural heritage efforts, with UNESCO being among the lead institutions in conservation projects [1], [2]. Modern acoustic research applies standardized measurement techniques and digital simulations to analyze sound behavior in said spaces. Applications like real-time auralization of church acoustics [3], Impulse Response (IR) compression [4], virtual acoustics [5], [6], and spatial upsampling [7] are just some of the many applications of interest for researchers working in this field. Lately, with the advent of machine/deep learning models, said applications have been requiring measurement data recorded in different environments and with particular instrumentation, e.g., higher-order microphones. Among the datasets available in the literature, we can mention the Aachen Impulse Response (AIR) dataset [8], which contains binaural room impulse responses of Aula Carolina Aachen, i.e., a former church known for its marked reverberation. The OpenAir Library[1] contains, instead, the impulse response of several different churches, such as Heslington Church and York Minster. To the best of our knowledge, REVERBDATA [9], i.e., an acoustic cultural heritage dataset of Portuguese churches, is the only publicly available dataset to contain multichannel IRs acquired with a higher-order microphone (in particular, with a 3rd-order microphone) in different locations. Hence, the literature still lacks datasets recorded with higher-order microphones able to satisfy the requirements of novel deep learning-based methodologies.

In this article, we present a dataset of church impulse responses acquired in the church of Saints Marcellino and Pietro with the aim of providing researchers with novel data to test and develop their spatial audio algorithms. Placed in the heart of the famous city of Cremona (Italy), the church of Saints Marcellino and Pietro stands as a testament to centuries of architectural evolution and acoustic richness [10].

This work has been funded by Regione Lombardia and Fondazione Cariplo under the "Casa Grasselli, Il palazzo della musica: poli e proposte culturali per una Città Creativa" project (ref. 2020-1324).

[1] https://www.openair.hosted.york.ac.uk

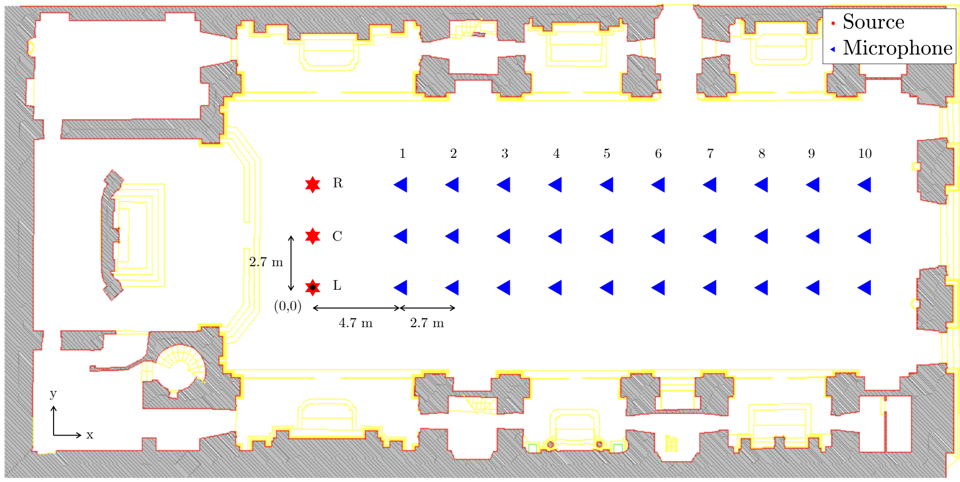

Fig. 2: Floor plan of the Saints Marcellino and Pietro church. The blue triangles show the measurement points, while the red hexagrams show the location of the three sources.

Constructed during the time span 1606-1620, its design reflects the stylistic and functional needs of its time, blending traditional architecture with unique structural choices. The use of materials like wood, plaster, and marble contributes to both the visual grandeur and the acoustic characteristics of the space. With its aisleless rectangular shape and high ceiling, the church was built to accommodate large congregations while enhancing the clarity and resonance of liturgical music [10].

Some Italian institutions, like Regione Lombardia and Fondazione Cariplo, have embraced the mission of preserving and restoring this space with the ultimate purpose of returning it as a concert hall for the local community. Indeed, since 2024, the church of Saints Marcellino and Pietro is among the locations of the famous Monteverdi Festival, an event that gathers people passionate about baroque music and early Italian operas from all over the world. The study of the acoustics of the church, therefore, became of utmost importance [11]. In particular, we acquired the impulse responses using both an omnidirectional microphone and a 4th-order microphone in thirty different points of the stall area and for three different locations of the source, which was placed in the stage area as shown in Fig. 1. Finally, it is important to highlight that the presented dataset [12] not only serves as a precious collection of IRs for spatial audio applications, but also for cultural heritage and for driving the design and architectural projects concerning acoustic renovation [11].

II. Dataset Description

The dataset presented here consists of thirty IRs acquired for three different source locations in the space where the seats are supposed to be placed. We acquire the data using an omnidirectional microphone and a higher-order microphone, which are positioned in correspondence with the seated listeners' heads to replicate their listening conditions. The humidity in the church is measured to be 57% and the temperature 26.9 °C. Said IRs are recorded using as stimulus the logarithmic sine sweep

$$s[k] = A_{\text{in}} \sin\left(2\pi f_1 L \exp\left(\frac{k}{f_{\text{s}} L}\right)\right) , \quad (1)$$

where k is the sample index, $A_{\text{in}} = 0.5$ is the sweep amplitude, $f_{\text{s}} = 48$ kHz is the sampling frequency, and $L = T/\log(f_2/f_1)$. The initial and final frequencies are, instead, set to $f_1 = 22$ Hz and $f_2 = 22$ kHz, while the duration of the sweep as $T = 7$ s. The omnidirectional microphone is connected to the Presonus FireStudio Project audio interface, which is, in turn, daisy-chained to the spatial microphone interface via FireWire; we use Reaper as Digital Audio Workstation (DAW) to acquire the audio signals. Then, starting from the raw recordings, the final IRs are obtained employing the deconvolution method presented in [13], followed by a normalization step involving the IR maximum absolute value.

A. Church Environment

The floor plan of the Saints Marcellino and Pietro's church is shown in Fig. 2 [10], where we highlight the positions of the microphones (blue triangles) and the sources (red hexagrams). With the aim of characterizing the environment as is, we acquired the impulse responses with bare floor and no benches, chairs or other installments. The church is 51.08 m long, 24.89 m wide, and has a maximum height of about 23 m. The environment is irregular, not equivalent to a shoe-box room, and presents several niches, altars, as well as two wooden confessionals. Moreover, a large wooden panel is present opposite to the main altar, as visible in Fig. 1. The walls are covered with different materials, such as frescos, marble, wood, and plaster, all characterized by different absorption coefficients, and present irregular textures and volumes. Altogether, materials and geometry contribute to creating a complex acoustic environment, making the presented dataset interesting from various standpoints, e.g., sound propagation analysis and spatial audio rendering.

B. Source Setup

We acquire the church impulse responses considering three source positions as to simulate different musicians spread in the area where the stage is supposed to be realized. In particular, we employ the Ntek OMNI 4” HP omnidirectional sound source in order to achieve isotropic sound radiation. Such a dodecahedral source is powered by the Brüel & Kjær 2716 power amplifier and connected to the Presonus audio interface. We measure the mean noise floor via sound level meter, obtaining a value of 32 dB, and we set the amplifier so as to obtain in the 10th row a level of about 40 dB above said noise floor. This ensures the recording of the impulse responses with a sufficient signal-to-noise ratio, avoiding, at the same time, background interference and allowing for an accurate analysis of the church acoustics. It follows that the use of T_{30} turns out to be more reliable than T_{60} for computing the reverberation time [14]. Finally, with respect to the origin reported in Fig. 2 (i.e., source S_{L}), the three sources are located at positions $S_{\mathrm{L}} = [0, 0, 1.87]$, $S_{\mathrm{C}} = [0, 2.7, 1.87]$, and $S_{\mathrm{R}} = [0, 5.4, 1.87]$, where the subscripts “L,” “C,” and “R” stand for “left,” “center,” and “right,” respectively. The positions are referred to the acoustic center reported in the operating manuals of the instrumentation.

C. Microphone Setup

We employ the Beyerdynamic MM 1 as omnidirectional microphone and the MH acoustics Eigenmike® em32 as higher-order microphone (4th-order, 32 channels). We position the microphone capsule (or the center of the spherical array) as to point in the direction parallel to the walls at a height $z = 1.269$ m. The first microphone row is placed at a distance of 4.7 m from the source row. Then, the microphones are positioned on a 10×3 grid with 2.7-meter spacing as shown in Fig. 2. The spherical microphone is connected to its interface by means of a CAT6 cable, which provides both power and synchronization, whereas the omnidirectional microphone to the audio interface by a studio-quality XLR cable. Due to the availability of a single spatial and omnidirectional microphone, we repeat the acquisition thirty times, moving only the microphones whilst maintaining the same environment configuration.

D. Data Format

The IRs are acquired considering a sampling frequency of 48 kHz and are truncated at 7 seconds. The recordings are provided either as single (for the case of the omnidirectional microphone) or multichannel (for the case of the spherical microphone) `.wav` files, saved at 24 bit per sample. The naming convention is chosen as follows:

- we name the files according to the label `S`**`pos`**`_M`**`pos_mic_row`**`.wav`, where `S` stands for “source,” and `M` stands for “microphone;”
- the label **`pos`** indicates the particular column with respect to Fig. 2 and takes value in the set {`L`, `C`, `R`};
- the label **`mic`** indicates the microphone used to acquire the IR and takes value in the set {`OMNI`, `EIG`};

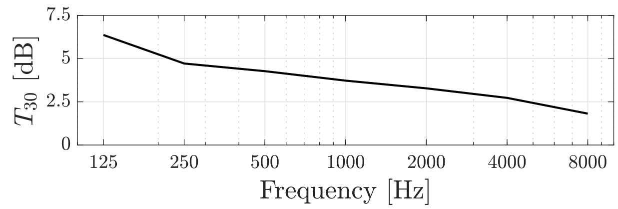

Fig. 3: Octave-band reverberation time T_{30} averaged over all positions and sources.

- the label **`row`** indicates the row with respect to Fig. 2 and takes value in the set {`1, 2, 3, 4, 5, 6, 7, 8, 9, 10`};

For example, the label of the IR acquired by the left spherical microphone in the first row when the right source is active is formed as `SR_ML_EIG_1.wav`. Finally, the position of each of the measurement points shown in Fig. 2 is reported in the `pos.csv` file.

III. Evaluation of Measurements

In this section, we introduce both acoustic and spatial descriptors as to evaluate and present the ChurchIR dataset.

A. Reverberation Time

With the purpose of characterizing the acoustics of the church of Saints Marcellino and Pietro, we compute the reverberation time employing T_{30} [14] since, as mentioned in Sec. II-B, said metric is deemed reliable given the noise floor conditions and the source level. In particular, we employ the backward integration method proposed by Schroeder [15], and we consider the IRs acquired through the omnidirectional microphone. The T_{30} averaged over all positions and sources is 4.12 s. Fig. 3 shows, instead, the reverberation time computed for the octave bands centered in $\{125, 250, 500, 1000, 2000, 4000, 8000\}$ Hz. The church presents higher reverberation at low frequency, with $T_{30} = 6.37$ s at 125 Hz, and increasingly fast decay when we consider higher frequencies, reaching a value of 1.82 s at 8000 Hz.

B. Acoustic Metrics

We here present additional metrics to further characterize the church environment, focusing on those designed for music applications. According to the ISO-3382-1 standard [14], we decide to take into account:

- **Early Decay Time (EDT)** - it measures the initial sound decay in the environment. It is defined as the time (in seconds) needed for the energy to drop by 10 dB below the peak (which once normalized is at 0 dB) multiplied by a factor 6. Hence,

$$EDT = 6 \cdot (t_0 - t_{-10}), \qquad (2)$$

  where t_x is the time for which the normalized EDC is x dB. Since the EDT is strongly affected by the initial part of the EDC, which varies significantly from seat to seat, the EDT varies more than T_{30} for a given space.

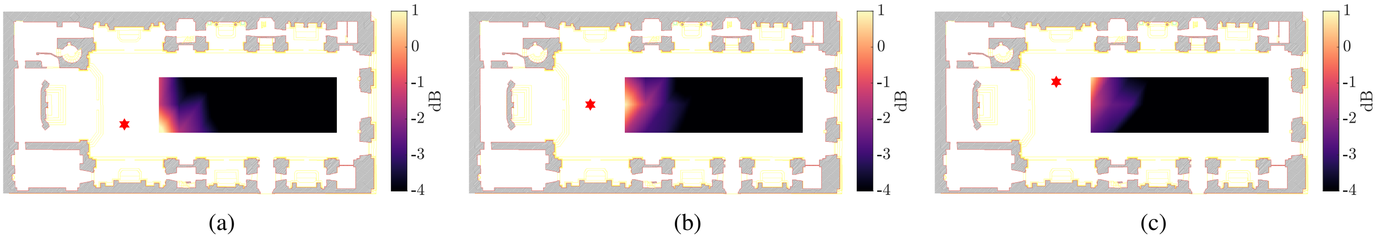

Fig. 4: Clarity index in space considering the three source positions. (a) Source S_{L}; (b) source S_{C}; (c) source S_{R}.

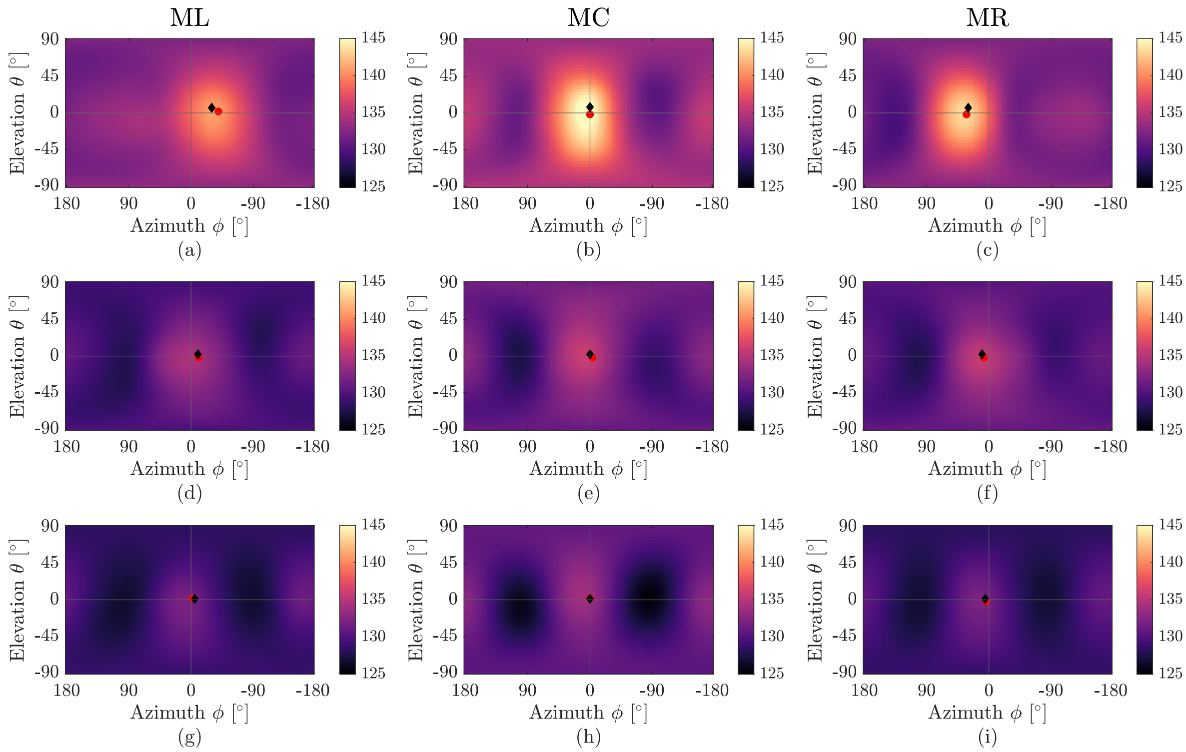

Fig. 5: Pseudo-spectra computed considering source S_{C}. The first column is referred to microphones `ML`, the second to `MC`s, and the third to `MR`s, whereas the rows are referred from top to bottom to rows `1`, `5`, and `10` of Fig. 2, respectively.

- **Center Time (CT)** - it describes the position in time of the center of gravity of the squared IR. It is computed as
$$CT = \frac{\sum_{n=0}^{L-1} n \cdot h^2[n]}{f_{\mathrm{s}} \cdot \sum_{n=0}^{L-1} h^2[n]}\,, \tag{3}$$
where h is the IR, n is the sample index, L is the IR's length in samples, and f_{s} is the sampling frequency. For music, a $CT \in [0.07, 0.15]$ s is desirable.
- **Clarity (C80)** - it measures how clear the sound quality is. For musical performances, a desired Clarity is obtained when the ratio of early (within 80 ms) to later sound energy of the direct sound is high. In particular, we compute it as
$$C_{80} = 10\log_{10}\frac{\sum_{n=0}^{L_{80}-1} h^2[n]}{\sum_{n=L_{80}}^{L-1} h^2[n]}\,, \tag{4}$$
where L_{80} is the length in samples of 80 ms. A range deemed acceptable for a concert hall is $[-4, 1]$ dB.
- **Bass Ratio (BR)** - it is a metric designed for musical performances and identifies whether reverberation is present in the low or higher regions of the spectrum. Indeed, a higher reverberation at low frequency is desired as this creates a feeling of "warmth." The Bass Ratio is computed using
$$BR = \frac{T_{125} + T_{250}}{T_{500} + T_{1000}}\,, \tag{5}$$
where T_f is the reverberation time at frequency f. For reverberant halls, a $BR \in [1.1, 1.25]$ is desirable.

We average all the metrics over the thirty positions and the three sources. The results but BR's are reported in Table I, which presents said descriptors also considering the octave bands already introduced in Sec. III-A. The EDT starts with a high value at low frequencies and progressively decreases toward higher frequencies, thus showing that low-frequency energy decays more slowly, while higher frequencies decay

TABLE I: Octave-band acoustic metrics computed in the Saints Marcellino and Pietro's church.

	125	**250**	**500**	**1000**	**2000**	**4000**	**8000**	**Avg.**
EDT	4.89	4.68	4.26	3.75	3.30	2.69	1.67	**3.6**
CT	0.34	0.31	0.27	0.26	0.27	0.186	0.10	**0.25**
C80	−4.89	−5.55	−5.20	−4.69	−3.91	−2.83	−0.64	**−3.7**

more quickly, probably due to the absorption of walls and air. The C80 follows a pattern that is expected in reverberant environments, with negative values at all frequencies, indicating a prevalence of late reflections over direct sound. However, it improves at higher frequencies reaching -0.64 dB at 8000 Hz, revealing that musical details become more intelligible in the upper region of the spectrum. The CT remains relatively stable in the midrange but shows a noticeable drop at higher frequencies, reinforcing the idea of a space where sound energy disperses more efficiently as frequency increases. As far as BR is concerned, we obtain, instead, a value of 1.21, giving further insights into the church characteristics for musical performances. Finally, Fig. 4 shows the evolution of C80 (for the three sources) in space by interpolating measures at the thirty points. As expected, we can clearly see that the clarity drops when we move the microphone far away with respect to the sources, but still, the values remain within the limits imposed by the ISO-3382-1 standard [14]. We refer the reader interested in a more thorough analysis of the church acoustics, also in the presence of an audience, to [11].

C. Pseudo-Spectra and Source Localization

As a further test, in order to show applications of the proposed dataset for spatial audio, we consider the scenario of source localization, defined as the task of detecting the Direction of Arrival (DOA) of a sound source [16]. To do so, we take into account the multichannel IRs acquired through the spherical microphone, and we apply a simple Maximum Directivity (MD) beamformer [17] to verify that the DOA is compliant with the acquisition data. Fig. 5 shows the pseudo-spectra obtained considering source S_{C}. In particular, the subfigures on the left column are related to microphones `ML 1`, `5`, and `10`; similarly, the center and right columns correspond to microphones `MC` and `MR`, respectively. The actual and estimated source locations of the source are denoted by black diamonds and red circles, respectively. By looking at the pseudo-spectra, we can clearly recognize that the location of the central source is matched, since the microphones in the first row, from left to right, see the sound coming from right, center, and left, respectively. The slight mismatch between the two markers can indeed be related to some errors in the placement of the microphones or geometry annotations, but especially to the high-reverberant environment in which the measurements took place.

IV. Conclusions

We presented ChurchIR, a dataset of multichannel IRs acquired in the church of Saints Marcellino and Pietro, in Cremona, Italy. The church dates back to the sixteenth century, and it is now used as a hall for chamber, choral, or baroque concerts. The dataset is recorded employing both an omnidirectional microphone and a 4th-order spherical microphone, and contains the IRs of thirty points for three different source locations. We presented the measurements providing spatial and acoustic descriptors, such as reverberation time and clarity index, as well as pseudo-spectra. Possible applications of the proposed dataset include, for example, interactive auralization of church acoustics, spatial upsampling, and virtual acoustics.

References

[1] S. Girón, L. Álvarez Morales, and T. Zamarreño, "Church Acoustics: A State-of-The-Art Review After Several Decades of Research," *J. Sound Vibration*, vol. 411, pp. 378–408, Dec. 2017.

[2] H. Rosseel and T. Van Waterschoot, "A State-of-the-art Review on Acoustic Preservation of Historical Worship Spaces Through Auralization," *Signal Processing*, vol. 234, pp. 1–25, Sep. 2025, 109992.

[3] H. Rosseel, M. Jälmby, and T. Van Waterschoot, "Low-Rank Multi-Channel Fast Convolution for Real-Time Auralization of Church Acoustics," in *Proc. of the 32nd Eur. Sign. Proc. Conf. (EUSIPCO)*, Lyon, France, Aug. 2024, pp. 1017–1021.

[4] A. I. Mezza, A. Bernardini, and F. Antonacci, "Large-Scale Room Impulse Response Dataset Compression with Neural Audio Codecs," in *Proc. of the 5th Int. Symp. on the Internet of Sounds (IS2)*, Erlangen, Germany, Sep. 2024, pp. 1–8.

[5] U. Berardi, "Simulation of Acoustical Parameters in Rectangular Churches," *J. Build. Perform. Simul.*, vol. 7, no. 1, pp. 1–16, Jan. 2014.

[6] L. Alvarez-Morales and F. Martellotta, "A Geometrical Acoustic Simulation of the Effect of Occupancy and Source Position in Historical Churches," *Applied Acoustics*, vol. 91, pp. 47–58, Apr. 2015.

[7] F. Miotello, F. Terminiello, M. Pezzoli, A. Bernardini, F. Antonacci, and A. Sarti, "A Physics-Informed Neural Network-Based Approach for the Spatial Upsampling of Spherical Microphone Arrays," in *Proc. of the 18th Int. Workshop Acoust. Enhancement (IWAENC)*, Aalborg, Denmark, Sep. 2024, pp. 215–219.

[8] M. Jeub, M. Schafer, and P. Vary, "A Binaural Room Impulse Response Database for the Evaluation of Dereverberation Algorithms," in *Proc. of the 16th Int. Conf. on Digital Signal Process.*, Santorini, Greece, 2009, pp. 1–5.

[9] D. Alarcão and A. C. Pinto, "The REVERBDATA Room Impulse Response Database," in *Proc. of the 156th Audio Eng. Conv. (AES)*, Madrid, Spain, Apr. 2024.

[10] A. Giuseppe Landi and E. Zamperini, "Technique at the Service of a New Liturgical Model: the Timber Roof of the Church of Saints Marcellino and Pietro in Cremona," *TEMA*, vol. 08, no. SI, Oct. 2022.

[11] S. Parrinelli, R. Giampiccolo, L. Angelo Giuseppe, and F. Antonacci, "Improving the Acoustics of the Church of Saints Marcellino and Pietro in Cremona (Italy) for Musical Performances," *Acoustics*, vol. 7, no. 2, Jun. 2025.

[12] R. Giampiccolo, S. Parrinelli, and F. Antonacci, "ChurchIR: A Dataset of Multichannel Church Impulse Responses for Spatial Audio Applications," Feb. 2025. [Online]. Available: https://doi.org/10.5281/zenodo.14901751

[13] F. Angelo, "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique," in *Proc. of the 108th Audio Eng. Conv. (AES)*, Paris, France, 2000.

[14] ISO, "Acoustics - Measurement of Room Acoustic Parameters - Part 1: Performance Spaces," Geneva, Switzerland, 2009, iSO 3382-1.

[15] N. Xiang, "Schroeder Integration: Foundation for Advanced Sound Energy Decay Analysis," *J. Acoustic. Soc. America*, vol. 129, no. 4, pp. 2429–2429, Apr. 2011.

[16] S. Gombots, J. Nowak, and M. Kaltenbacher, "Sound Source Localization – State of the Art and New Inverse Scheme," *Elektrotech. Inftech.*, vol. 138, no. 3, pp. 229–243, Jun. 2021.

[17] B. Rafaely and D. Khaykin, "Optimal Model-Based Beamforming and Independent Steering for Spherical Loudspeaker Arrays," *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 7, pp. 2234–2238, Sep. 2011.