

6dof Audio Rendering Using Spatial Impulse Responses In An Urban Context

A. Bevilacqua¹, A. Farina², L. Tronchin³

¹ Department of Architecture and Engineering, CIDEA, University of Parma, 43100, Parma, Italy

² CIRI, University of Bologna, 40141, Bologna, Italy

³ Department of Architecture, University of Bologna, 40141, Bologna, Italy

Corresponding author: Antonella Bevilacqua (AB), antonella.bevilacqua@unipr.it

Abstract. Many acoustic studies on urban environments focus on soundscapes and subjective perception, analysing environmental parameters such as loudness, fluctuation strength and roughness. However, room acoustic parameters have rarely been applied in this context, even though public squares are frequently used for musical events. Considering classical concerts in historical or archaeological sites in Italy, or pop music performances in public squares amplified through audio systems, a key challenge is that the acoustic characteristics of these spaces remain largely unknown. This study addresses this gap by conducting acoustic measurements in Piazzale San Francesco in Parma, a public square already used for summer musical events, and by implementing its audio-rendering in a virtual reality (VR) environment. The acoustic survey was carried out following the standards for enclosed rooms, in accordance with ISO 3382, while the digital model of the space was created using photogrammetry. The piece *Ti tradi quell'alma ingrata* from *Il Trovatore* was used as the audio signal in Unreal Engine, enabling a navigation of the digital space calibrated with measured impulse responses (IRs). The resulting VR rendering allows users to be fully immersed in the virtual space, freely exploring it with Six Degrees-of-Freedom (6DoF), as they would in reality while listening to the same musical performance.

1. Introduction

For decades, research on impulse responses (IRs) has focused primarily on enclosed spaces, whether or not they were intended for musical performances [1]. Such rooms have been extensively analyzed in relation to reflections generated by architectural elements and by the geometry of sidewalls and ceilings [2]. In contrast, the AGORA project, within which the present study is developed, investigates the acoustics of outdoor spaces, with particular emphasis on public squares. The objective is to make the sonic experience accessible to a broad audience, rather than limiting the discussion to acoustics specialists. This perspective shifts the focus towards outdoor environments, which are frequently used for musical events, especially in Italy during the summer season, where favorable thermo-hygrometric conditions can be combined with satisfactory listening quality [3]. However, the acoustic characteristics of open-air public squares have never been systematically assessed, which constitutes the main novelty of this study. Previous investigations have considered semi-enclosed settings such as palace courtyards [4, 5] or ancient open-air theatres [6, 7], yet these differ fundamentally from urban squares: the former are physically smaller and more confined, while the latter are substantially larger and more open. Existing literature has already emphasized the cultural significance of Italian public squares as spaces that embody civic, commercial, and social identity [8]. A less explored but equally

important aspect is their acoustic identity, an intrinsic feature independent of their functional use. The overarching goal of the AGORA project is therefore to capture the sonic identity of Piazzale San Francesco in Parma and translate its static acoustic environment into an immersive virtual-reality (VR) experience.

Piazzale San Francesco is itself a site layered with centuries of historical transformations. The church facing the square, from which it takes its name, was converted into a prison and later repurposed as a university study facility [9]. Today, the square also hosts musical activities promoted by the Casa della Musica, located directly opposite the church. These performances contribute to a vibrant cultural atmosphere, enhanced further by the nearby Casa del Suono, a dedicated listening space equipped with a Wave Field Synthesis system comprising 189 loudspeakers arranged along its perimeter, enabling advanced audio-visual spatialization [10].

In this study the visual-audio rendering has been created by using the six-degree-of-freedom (6DoF) in Unreal Engine, one of the most common real-time 3D creation tools [11]. The 6DoF is an audio technique that consists of both rotation and translation in a VR environment where the users can explore and navigate immersive audio experience rendering. Unlike earlier 3DoF binaural systems, true 6DoF rendering must update interaural cues and the scene's spatial filtering (early reflections, distance cues and reverberation) as the listener moves, in order to preserve plausibility and spatial stability.

Contemporary approaches to 6DoF audio fall into four broad categories:

- The object-based representation of scene elements as tagged audio objects with geometry and metadata, and their rendering through distance models and binaural filters, a flexible and production-friendly approach that naturally supports translation when object geometry is known [12, 13].
- The High-Order-Ambisonics (HOA) based capture of a full sound-field representation that decodes to binaural output; HOA handles rotation elegantly but requires very high orders or additional processing to represent listener translations accurately [14, 15].
- The spatial sampling is based on the interpolation grid capture, consisting of distributing HOA arrays on a dense grid for recording the scene from many positions [16], useful to render an accurate 6DoF although it is very time consuming.
- The parametric / model-based method infers compact room/signal parameters (direct path, early reflections, late reverberation) and synthesize binaural signals for new listener positions, offering runtime efficiency with model-dependent limitations [17].

Each of these approaches has inherent limitations, either because they rely on specific modelling assumptions or because real-time performance is constrained by the increasing demands of commercial applications. For this reason, many current products adopt hybrid pipelines that combine object metadata with HOA or parametric components, seeking a balance between perceptual realism and computational cost [18]. While the video-game industry prioritizes real-time VR performance, scientific research tends to focus more on the accuracy and perceptual fidelity of 6DoF audio rendering.

In this study, Reverb Zones were used as the primary tool for rendering the audio of a musical excerpt from Scene 2, Act I of *Il Trovatore* by G. Verdi. The audio signal was incorporated into the virtual reconstruction of Piazzale San Francesco, created through a photogrammetric workflow realized in Unreal Engine. The virtual model was divided into three virtual rooms (acoustic zones), each acoustically calibrated using the IR measured at corresponding source-receiver positions in

the actual square. This approach resulted in a realistic and immersive audio-video rendering that can be freely explored and navigated by a variety of users.

2. Historical overview of Piazzale San Francesco in Parma, Italy

Piazzale San Francesco, in the heart of Parma's historic centre, is a historically significant public space which evolved over the centuries alongside the city's cultural and social changes [19]. This square is framed by four historical buildings: Casa della Musica to the south and housed in the elegant historic Palazzo Cusani, the Franciscan monastery (that was later transformed into a prison and now part of the University of Parma) on the west, the Church of San Francesco del Prato on the north, while the east side is open to a narrow alley (i.e. Borgo degli Studi) characterised by a restricted traffic, on the opposite side of which there is a building known as Casa del Suono, multimedia museum famous for the acoustic and sound exhibitions, founded by Prof. A. Farina. A view of its particular layout is shown in Figure 1. In particular, the paving of the square is made of small natural stone blocks arranged in a pattern known as "peacock tail."

In the Middle Ages, the area of Piazzale San Francesco was an open field, used for markets and tournaments. In the early 13th century, Franciscan friars settled here, gaining local support from the community [20]. Construction of the Gothic Church of San Francesco del Prato began in 1240, which continued through the 14th century, making it one of Parma's largest churches. A convent was built alongside, hosting prominent Franciscans and noble burials. The adjacent open space functioned as a courtyard for preaching and gatherings, laying the foundations of the future square [21].

The 1767 Atlante Sardi, Parma's first official land registry map, depicts in detail the block of the church of San Francesco. Commissioned by Minister Guillaume Du Tillot (1759–1771) within the framework of Enlightenment-inspired administrative and urban reforms, the atlas reflects a systematic approach to territorial and urban documentation. The map clearly illustrates the special configuration of the complex and the surrounding historical buildings described above. Du Tillot, appointed by Louis XV to serve at the court of the Duke of Parma, oversaw finance, culture and institutional reforms, reorganizing bodies such as the Academy of Fine Arts, as well as several convents and religious institutions [22].

The Napoleonic period marked the abrupt end of convent life of Piazzale San Francesco. Following the suppression of religious orders, convents and monastic properties were confiscated, and the church was subsequently converted into a prison.

Nowadays, part of the former Franciscan convent is converted into a University of Parma building, while an auditorium has been built within Casa della Musica for music organizations, becoming a vibrant cultural square.

Figure 1.

The square's layout, with its overall dimensions of 28×50 m, makes it ideal for 3D acoustic surveys and environmental sound measurements, with the surrounding architecture as main reflecting surfaces.

3. Methodology

This section describes the methodological framework adopted in the study. Section 3.1 presents the visual survey of Piazzale San Francesco, carried out through photogrammetry to obtain an accurate 3D model of the square. Section 3.2 details the acoustic measurement campaign, including the instrumentation and impulse response acquisition procedure. Section 3.3 illustrates

the development of the immersive audio-visual VR environment, where the measured spatial impulse responses are integrated with the digital model for convolution-based rendering in 6DoF framework. A schematic overview of the entire workflow is provided in Figure 2 to map the methodological steps.

Figure 2.

3.1. Visual survey using photogrammetry

Traditional methods to carry on architectural surveys rely on simple equipment, such as tape measure and laser meter. Nowadays, modern digital technologies (i.e. drones and high-resolution cameras) can be used, especially for photogrammetry [22]. These tools allow a very accurate and detailed capture of the space for a fast 3D modelling, making the geometric survey more precise and time effective.

Photogrammetry is a technique used to shape and position an object using at least two different photos taken from separate angles. For this purpose, overlapping photos of buildings have been taken to create a 3D model of the square [23].

Drones, also known as Unmanned Aerial Vehicles (UAVs), are flying machines that do not need a pilot on board to be operated [24]. A DJI Mini 4 Pro drone (lightweight quadcopter) has been used in this study to solve visibility issues for wider areas and overcome many of the limitations of ground-based surveys [25]. In addition to the drone, a Sony Alpha 7 IV camera has also been used to enhance the level of detail in the photogrammetric dataset, which is a high-resolution full-frame mirrorless camera.

Using inverse triangulation, the spatial position of each point on the object is then calculated [26]. Today, these steps are automated using Structure from Motion (SfM) algorithms, which come from the field of computer vision [27]. The software automatically detects thousands of key points in each photo and matches them across overlapping images. Then, through an iterative process, it calibrates the internal camera settings and determines the position and orientation of each photo, resulting in a sparse point-cloud 3D space made up of millions of points [28]. A scheme of the SfM algorithm workflow is explained in Figure 3.

Figure 3.

From this dense cloud, a polygonal mesh is generated by connecting the points with triangular faces, giving a detailed 3D model of the scene. The final step is texturing, which means mapping the original photos onto the 3D model to create a photorealistic representation.

To enhance the perceived visual fidelity of the capture while preserving high rendering performance in the reconstructed scene, the photogrammetric model was partitioned, enabling different spatial resolutions to be assigned to distinct portions of the geometry. Ground-level features were prioritised, as these are more prominently visible to the user. In particular, the statue of Giuseppe Verdi seated on a bench was processed as an independent asset and subsequently aligned with the main square model once both elements were imported in Unreal Engine.

In aerial surveys of squares and historic city centers, the same data can also be used to generate vertical and georeferenced orthophotos. For this case study, both vertical (nadir) shots of the open area and oblique shots of the building façades have been combined [29]. Figure 4 shows the scheme of the photogrammetry procedure by using drones.

Figure 4.

The visual survey of Piazzale San Francesco was carried out on March 11, 2025. The average temperature was 11°C, relative humidity around 60%, while visibility remained stable at around 11 kilometers with a still wind condition. These meteorological conditions are important for the quality of the UAV-based survey other than for acoustic parameters [30]. One of the most beneficial aspects of the weather was the presence of a uniformly cloudy sky. While completely sunny days often result in strong shadows cast by buildings, the diffuse lighting on a cloudy day significantly reduces shadow contrast and improves the uniformity of aerial images. The flight altitude was fixed at 25 m above ground level, allowing safe navigation in an urban environment. The absence of wind contributed to stable flight conditions, reducing motion blur and helping to maintain a regular flight pattern. Figure 5 illustrates the equipment in operation and a sample image taken with the drone.

Figures 5A and 5B.

For this photogrammetric survey, a total of 2910 photographs were captured, comprising both aerial and ground-level shots, to ensure a medium-high level of detail for modelling the square. Table 1 summarizes the characteristics of the visual-geometric survey.

Table 1.

3.2. Acoustic survey

A campaign of acoustic measurements was taken in Piazzale San Francesco using a wide range of equipment, listed as follows:

- Omnidirectional active loudspeaker (Look Line SL203AC), equipped with 12 transducers.
- Spherical multi-channel microphone array (Eigenmike® em64).
- 4-channel microphone (Sennheiser Ambeo VR).
- Omnidirectional microphone (Behringer ECM8000).
- Dummy Head (Neumann KU100).
- Personal computer.

A Shanqiu Mini UPS FX5-48 Pro PoE injector was connected to the em64 microphone. This device provides both power to the microphone and data transmission through a single Ethernet connection. The signal was routed to a dedicated laptop for real-time processing, where the Dante protocol (Digital Audio Network Through Ethernet) allowed the multichannel audio signal to be sent digitally over the network with extremely low latency.

The loudspeaker was placed at 1.6 m from ground in two specific positions and was used to play a, exponential sine sweep (ESS) signal with a duration of 15 s and an equalized sound pressure level between 40 Hz and 20 kHz. All microphones were moved across the square in six specific

positions, to capture the IR of the square. Images in Figure 6 illustrate two moments captured during the acoustic survey, while Figure 7 illustrates a scheme of the equipment positions taken in the square.

Figure 6.

Figure 7.

Table 2 summarizes the distance between sound source and receiver positions. Currently, there are no regulations specifying the required number of source and receiver positions for measuring IRs in public squares. However, a total of 6×2 source-receiver combinations is considered sufficient for the calibration of the virtual model, as the measurements points were distributed across the entire area, representing all three zones identified in this case study.

Table 2.

3.3. Immersive audio-visual virtual reality (VR)

Piazzale San Francesco has long served as a venue for musical activities, from historic liturgical singing by the nearby church to modern open-air concerts organized by Casa della Musica. To explore how effectively a musical performance sounds within the square, a digital model integrating architecture and sound was realized, which has been processed with Unreal Engine 5.4 for immersive simulation [31].

For this purpose, Convolution Reverb system in Unreal Engine has been used to convolve a spatial IR with a dry audio signal, resulting in a highly accurate simulation. Unlike standard algorithm-based reverb effects, Convolution Reverb relies on actual physical measurements, which makes it very realistic, accurately reproducing the reverberation characteristics of a real space within a virtual environment [32]. For this operation, the main features are three, which will be discussed singularly:

- The HOA impulse response.
- The audio signal.
- The spatialization rendering with 6DoF, which combines both.

3.3.1. HOA impulse response

HOA impulse responses have been used for the convolutions, obtained from the multi-channel spherical array microphone utilized during the survey. The use of HOA allows the sound field to react dynamically to the listener's position and orientation, in terms of directivity and intensity as it would be in a real physical environment [33]. An example of omnidirectional impulse response (single channel) is shown in Figure 8; in contrast, a HOA impulse response is characterized by multiple channels, each recorded on separate tracks and representing a specific directivity component, as determined by the spatial arrangement of the capsules over the spherical array microphone.

Figure 8.

3.3.2. Sound signal selected for the virtual navigation

For the sound signal, a piece from classical music repertoire has been selected in line with the historical identity of this site with respect to Giuseppe Verdi, the famous native composer in Parma. The chosen piece is the cabaletta “*Ti tradi quell’alma ingrata*” from Scene 2 of Act I of *Il Trovatore*, composed in 1852 [34]. Notably, a background soundscape recorded in Piazzale San Francesco was added to enhance the immersive experience.

The selected piece was recorded in an anechoic chamber, which is essential for the convolution with the measured IRs that carry all the sonic characteristics of the square. The arrangement selected for this demo features the following instruments:

- Violin, carrying the main melody.
- Viola, contributing to the cohesion of the ensemble.
- Cello, supporting harmony and defining rhythm.
- Lyrical soprano, soloist of the aria.

3.3.3. Audio-visual rendering in 6DoF

In terms of audio-visual rendering in 6DoF, the immersive experience has been created in different ways so that a comparison between methodologies can be applied. Notably, for all methodologies the VR allows the user to virtually move freely in three-dimensional space by translation along the X, Y, and Z axes as well as by rotation through pitch (tilting up and down), yaw (turning left and right), and roll (tilting side to side).

In particular, the audio has been processed with Audio Volumes, which is an audio system within Unreal Engine which allows to define invisible audio volumes in a virtual scene where specific reverb effects are applied to audio files [35]. In this case, Convolution Reverb was applied to each volume using the nearest measured Impulse Response, with priority settings adopted in overlapping regions. Based on the current position of the virtual sound sources, the IRs corresponding to source position 1 were employed. This approach is consistent with the established robustness of room acoustics parameters to moderate spatial displacements of source and receiver positions.

Nevertheless, a further study could compare the present auralization results with those obtained from a full 6dof geometric acoustics simulation. Such a model would be calibrated against the measurement data at the reference positions, enabling analysis of the divergence between simulated and measured results as the distance from the calibration points increases.

Although the Unreal Engine project supports 6dof movement of both source and receiver, the current analysis assumes exact correspondence between measured source-receiver positions and their simulated counterparts, thereby limiting the evaluation to discrete validated locations.

Alternatively to the measured IR, it is also possible to assign custom reverberation parameters such as decay time and diffusion. “Fade Time” settings allow to regulate how smoothly reverberation transitions can be between zones. For this study, setting a Fade Time of 1-2 s has been selected to blend the transition when entering or leaving a zone, which is similar to the average reverberation time measured within the square [36].

The number of reverb zones used for this study are six, organized as shown in Figure 9.

Figure 9.

4. Results

4.1. Digital modelling reconstruction

RealityCapture is a Windows-based software that has been used for the 3D modelling reconstruction through photogrammetry. It is available through the Epic Games platform well known to process high-quality results [37].

After the dense cloud, it is possible to generate mesh surfaces, turning the point cloud into a continuous polygonal surface, which represents the actual shape of the square. For texturing, RealityCapture projects the original photos onto the 3D surface to create a photorealistic rendering [38]. For this study, 8912×8912 pixels have been selected. A view of the reconstructed digital model is given in Figure 10.

Figure 10.

4.2. Measured acoustic values

Sound is a fundamental part of how architecture and outdoor natural environment are experienced, yet it is often neglected in conservation and design. This study extends the evaluation of standard acoustic parameters for buildings to public squares, confirming that open urban spaces also develop a recognizable sound identity with their reflections from surrounding constructions, especially when used for performances or community events.

The post-production process of the recorded acoustic data was carried out using Adobe Audition 3.0 on a Windows operating system, together with the Aurora plug-in.

The acoustic parameters under evaluations are early decay time (EDT), reverberation time (T30), clarity index (C80), definition (D50), and strength (G). The measured acoustic data have been assessed in the frequency range between 125 Hz and 4 kHz, as shown in Figure 11; very low octaves have been discarded due to the interference of background noise with the measurements [39], as well as frequencies > 4 kHz due to the air absorption phenomenon that can make the sound decay less reliable [40, 41]. However, the 125-4000 Hz range is the most important range for acoustic analyses. The values were averaged across the six receiver positions; however, they were calculated separately for source positions 1 and 2.

Figures 11A, 11B, 11C, 11D, 11E.

Figure 11(a) shows that EDT values related to S1 range from approximately 1.8 s at low frequencies to 1.2 s at high frequencies, and from 2.0 s to 0.9 s for S2. A pronounced peak at 1 kHz is observed for S1 curve, evident at 1 kHz for S2, mainly caused by reflections from the surrounding building façades, which are made of highly reflective brickwork.

Regarding reverberation time, T30 values vary between 1.8 s and 1.5 s, values comparable to those typically found in medium-sized churches, as shown in Figure 11(b). Specifically, the two curves have different trendlines flatter around 1.5 s for S2 and progressively decaying for S1.

Figure 11(c) presents the clarity index for music. C80 increases from 3 dB at low octaves to 5 dB for S1 and to 9 dB for S2, at 4 kHz. These results make the perception of music clearer than an enclosed space.

The definition values, shown in Figure 11(d), fluctuate around 55% for S1 across the entire frequency spectrum, with a little raise at 2-4 kHz reaching 65%. For S2, the D50 values fluctuate around 70%, with the exception of 125 Hz that is 35%. This behaviour indicates a pretty balanced distribution of early and late energy, making the sound environment suitable for both speech and music, as a definition of approximately 50% represents an optimal equilibrium between the two.

Figure 11(e) shows that the G values follow a double-slope trend, with one slope up to 20 m and another from 20 m to 40 m. The difference between S1 and S2 is very small, and both remain within a range from approximately 1 dB near the sound source to about -12 dB at the farthest positions.

4.3. Immersive audio-visual virtual reality (VR)

The immersive audio-visual contents allows users to explore the space in 360° using a controller, to freely rotate their viewpoint to observe the entire reconstructed square. Furthermore, if the sound source moves within the virtual environment, the listener should be able to perceive corresponding changes in the sound field in real time. The video is also fully compatible with VR headsets and Ray-Ban Meta glasses equipped with integrated binaural audio. Alternatively, the immersive VR can be experienced in a dedicated listening room. Figure 12 shows a moment of the VR audio-video.

Figure 12.

While the present study relies exclusively on objective acoustic parameters (such as EDT, C80, D50, and STI) to compare the behavior of the different simulated environments, it is important to recognize that these descriptors do not always mirror perceptual experience. As highlighted in previous publications on room acoustics and auralization [42], discrepancies often emerge between measured values and subjective impressions reported by listeners. Several authors have shown that, in many cases, the perceptual impact of variations in acoustic parameters is weaker or less pronounced than what the numerical values might suggest, especially when differences fall within thresholds of just noticeable differences (JNDs) or when listeners are not trained experts [43].

For this reason, future developments of the study will include a dedicated subjective listening test, which is considered an essential step to validate the objective findings. A potential extension of this research could involve a comparison between Reverb Zones in Unreal Engine and the acoustic response obtained with a GA model simulation, considering that each approach presents distinct limitations:

- Reverb Zones calibrate the IR at each point of the space using a measured IR and estimate the acoustic response at other points through interpolation, based on the volume of the space and the boundary settings of vertical and horizontal surfaces [44]. In this approach, the transition zone between two adjacent rooms can reflect an averaged reverberation characteristic of the two spaces, potentially smoothing localized acoustic differences.
- GA models offer a useful understanding of the acoustic behavior of the investigated space, particularly in relation to the distribution of early reflections and the overall reverberant character at mid and high frequencies. Nevertheless, the simplified nature of ray-based methods becomes evident when approaching those phenomena that depend on the full complexity of wave propagation, including diffraction, interference, and low-frequency modal activity. These factors are fundamental to the listener's perception of warmth, envelopment, and spatial definition.

5. Conclusions

By capturing and reproducing the lively and socially significant sounds of public squares using advanced multichannel audio technologies, this study seeks to transform listening into a fully

immersive experience. While the acoustic properties of enclosed spaces have been thoroughly investigated, often through virtual reconstructions and 3D auralization techniques, historic urban squares remain largely unstudied. Moreover, the influence of sound on social interaction and community cohesion has received only a specific attention in the existing literature, mainly focused on environmental parameters, as a result of soundscape research.

More often, in many public squares the audio amplification for pop music events often falls short of its full potential. This is not due to a lack of equipment or technical skill, but rather to the inherent challenge posed by the unfamiliar acoustics of open urban spaces. Unlike concert halls or theatres, where acoustic behaviour is well-documented and predictable, public squares present complex geometries, reflective surfaces, and unpredictable interactions with surrounding buildings. As a result, sound designers frequently rely on tentative sound checks, adjusting levels by ear during the event itself, rather than guided by precise measurements or simulations. This improvisational approach can lead to uneven coverage, excessive reverberation in some areas, and a lack of clarity or balance, limiting the audience's experience despite the performance being otherwise vibrant and engaging.

This study represents a scientific approach to estimating the acoustics of public squares, spaces that are increasingly used for live musical performances. Understanding the acoustic response of these environments, including the distribution of reflections that contribute positively to clarity and sound impression, as well as the overall propagation of sound across the area, is a key factor in optimizing the listening experience for the audience. Such analyses can be conducted using a range of modeling and rendering techniques that can be easily achieved using less elaborated methods.

At the same time, a significant portion of the population, including people with mobility impairments or other constraints, may be unable to attend public events in person and are consequently excluded from cultural life. In this context, immersive VR experiences provide an avenue for greater social inclusion, enabling individuals to engage with space and the performance remotely. The implementation of 6DoF navigation enhances this immersive experience by allowing users to move freely within the virtual environment, exploring the acoustic characteristics of the square as if they were physically present.

This approach not only broadens access to cultural events but also offers an innovative tool for studying and documenting the acoustic qualities of public spaces, potentially informing future design and sound reinforcement strategies.

Acknowledgment

The authors would like to thank Matteo Fadda for his support during the acoustic measurements and for his valuable contribution to the data analysis.

References

- [1] L. Tronchin, A. Bevilacqua. Acoustic study of different sceneries at the São Carlos national theatre of Lisbon. *Applied Acoustics*, 2021, 180, 108102.
- [2] L. Tronchin, A. Bevilacqua. Immersive acoustics within the National Theatre in Iasi, Romania. Conference on Immersive 3D Audio (I3DA): from Architecture to Automotive, Bologna, Italy, 10-12 September 2025.

- [3] A. Bevilacqua, G. Tentoni, A. Martinetti, A. Farina, R. Yan, M. Fadda, L. Battisti, C. Van Tonder, L. Tronchin. Soundscapes across two historical cities: Piazza del Popolo in Faenza and Leicester Square in London. Conference on Immersive 3D Audio (I3DA): from Architecture to Automotive, Bologna, Italy, 10-12 September 2025.
- [4] D. Zhang, Y. Feng, M. Zhang, J. Kang. Sound field of a traditional Chinese Palace courtyards theatre. *Building & Environment*, 2023, 230, 109741.
- [5] G. Iannace. The use of historical courtyards for musical performances. *Building Acoustics*, 2016, 23(3-4): 207-222, <https://doi.org/10.1177/1351010X16678219>.
- [6] L. Tronchin, A. Bevilacqua. Historically informed digital reconstruction of the Roman theatre of Verona. Unveiling the acoustics of the original shape. *Appl. Acoustics*, 2022, 185, 108409.
- [7] A. Bevilacqua, W. Fuchs. Study of the Roman theatre of Volterra: Acoustic effects based on the presence of different architectural construction elements. *Building Acoustics*, 2024, 31(3): 225-246.
- [8] A. Farina, A. Bevilacqua, M. Fadda, L. Battisti, M.C. Tommasino, L. Tronchin. Immersive 3D Soundscape: Analysis of Environmental Acoustic Parameters of Historical Squares in Parma (Italy). *Urban Sci.* 2025, 9, 259. <https://doi.org/10.3390/urbansci9070259>.
- [9] F. Ottoni. Dialoghi di restauro: riflessioni e analisi critiche sui progetti di restauro, a Parma e dintorni, Ed. Quasar: Rome, Italy, 1(6) in MADlab, 2023 (in Italian).
- [10] P. Martignon, F. Adriaensen, D. Torelli, A. Farina. A Digitally Controlled Two Dimensional Loudspeaker Array. 32nd AES International Conference, Hillerød, Denmark, 21–23 September 2007.
- [11] “Unreal Engine” Accessed on 03 December 2025, available online at <https://www.unrealengine.com/en-US>.
- [12] A. Palazzi, L. Begamini, S. Calderara, R. Cucchiara. End-to-end 6-DoF object pose estimation through differentiable rasterization. 15th European Conference Computer Vision ECCV, Munich, Germany 8-14 September 2018.
- [13] R. Bleidt, A. Borsum, H. Fuchs, S. Merrill Weiss. Object-Based Audio: Opportunities for Improved Listening Experience and Increased Listener Involvement. SMPTE Motion Imaging Journal, 2015, 124(5): 13.
- [14] T. Carpentier. Normalization schemes in Ambisonic: does it matter? 142nd AES Convention, Berlin, Germany, 20-23 May 2017.
- [15] M. Kronlachner, F. Zotter. Spatial transformations for the enhancement of Ambisonic recordings. 2nd Int. Conf. on Spatial Audio (ICSA), Erlangen, Germany, 21-23 February 2014.
- [16] B. Rafaely, V. Tourbabin, E. Habets, Z. Ben-Hur, H. Lee, H. Gamper, L. Arbel, L. Birnie, T. Abhayapala, P. Samarasinghe. Spatial audio signal processing for binaural reproduction of

recorded acoustic scenes – review and challenges. *Acta Acustica*, 2022, 6: 19. DOI: <https://doi.org/10.1051/aacus/2022040>.

[17] J.M. Arend, S.V.A. Garı, C. Schissler, F. Klein, P.W. Robinson. Six-Degrees-of-Freedom Parametric Spatial Audio Based on One Monaural Room Impulse Response. *J. Audio Eng. Soc.*, 2021, 69(7/8): 557–575. DOI: <https://doi.org/10.17743/jaes.2021.0009>.

[18] J. Herre, J. Hilpert, A. Kuntz, J. Plogsties. MPEG-H 3D Audio—The New Standard for Coding of Immersive Spatial Audio. In *IEEE J. Sel. Top. Signal Process.*, 2015, 9(5): 770–779.

[19] E. Canniffe. *The Politics of the Piazza: The History and Meaning of the Italian Square*. ED. Routledge, 2008.

[20] “Storia,” San Francesco del Prato. Accessed on May 25, 2025. [Online]. Available at <http://www.sanfrancescodeiprato.it/chiesa/storia/>

[21] P. Rodriguez-Navarro, T. Gil-Piqueras, G. Verdiani. Drones for architectural surveying. Their use in Documenting towers of the Valencian coast. *Congreso Internacional de Expresión Gráfica Arquitectónica*, Alicante, Spain, 30 May – 1 June 2018.

[22] G. Tentoni, A. Martinetti, L. Guardigli, B. Turillazzi, A. Bevilacqua, A. Farina, L. Battisti, L. Tronchin. Two soundscapes in comparison: Piazza Cavour in Rimini and Buckingham Palace Square in London. 2025 Immersive and 3D Audio: from Architecture to Automotive (I3DA), Bologna, Italy, 10-12 September 2025.

[23] M. Siewczynska, T. Ziolo. Analysis of the Applicability of Photogrammetry in Building Façade. *Civil and Environmental Engineering*, 2022, 32(3): 182-206. DOI: 10.2478/ceer-2022-0035.

[24] V. Lambey, A.D. Prasad. Unmanned aerial vehicle applications of 3D modelling, visualization and parameter calculations. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Dehradun, India, 20–23 November 2018.

[25] C. Achille et al. UAV-Based Photogrammetry and Integrated Technologies for Architectural Applications—Methodological Strategies for the After-Quake Survey of Vertical Structures in Mantua (Italy). *Sensors*, 2015, 15(7): 15520–15539, doi: 10.3390/s150715520.

[26] A. Gaong, G.E. Bin Idris, L. Chong Luh, A.A. Ab Rahman, W.S.M. Wan Mohamed Sabri, A.H. Abdul Jalil. Comparative Evaluation of 3D Building Model Using UAV Photogrammetry and Terrestrial Laser Scanner (TLS). *Built Environment Journal*, 2025, 22(1). <https://doi.org/10.24191/bej.v22i1.1066>.

[27] A. Eltner, G. Sofia. Structure from motion photogrammetric technique. *Developments in Earth Surface Processes*, In book: *Remote sensing of Geomorphology*, 2020, 23: 1-4. DOI: 10.1016/B978-0-444-64177-9.00001-1.

[28] D. Tian, M. Gong, J. Li, J. Shi. 3D point cloud classification network with hybrid sampling enhancement and point energy attention. *Pattern recognition*, 2025, 112791.

- [29] D. Minh Bui, H. Kim, J. Youn, C. Kim. Automated 3D building model reconstruction using orthophotos and point clouds. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLVIII-G-2025, Dubai, UAE, 6–11 April 2025.
- [30] N. Granzotto, R. Yan, L. Tronchin. Measurements of Room Acoustic and Thermo-Hygrometric Parameters - A Case Study, *Appl. Sci.*, 2023, 13(5), doi: 10.3390/app13052905.
- [31] “Unreal Engine” Accessed on 24 November 2025. Available at <https://www.unrealengine.com/en-US/blog/unreal-engine-5-4-is-now-available>.
- [32] H. Baran Firat, L. Maffei, M. Masullo. 3D sound spatialization with game engines: the virtual acoustics performance of a game engine and a middleware for interactive audio design. *Virtual Reality*, 2022, 26(1121/1), DOI: 10.1007/s10055-021-00589-0.
- [33] G. Götz; S.J. Schlecht; V. Pulkki. A dataset of higher-order Ambisonic room impulse responses and 3D models measured in a room with varying furniture. 2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA), Bologna, Italy, 8-10 September 2021.
- [34] D. D’Orazio. Anechoic recordings of Italian opera played by orchestra, choir, and soloista. *J. Acoust. Soc. Am.*, 2020, 147(2): EL157–EL163, doi: 10.1121/10.0000739.
- [35] “Using Reverb Zones” Accessed on 24 November 2025, available at https://www.audiokinetic.com/en/public-library/2025.1.3_9039/?source=SDK&id=spatial_audio_roomsportals_reverbzones.html.
- [36] J.W. Murray. *Building Virtual Reality with Unity and SteamVR* (2nd ed.). CRC Press. 2020, <https://doi.org/10.1201/9780429295850>.
- [37] C. Richardt, P. Hedman, R.S. Overbeck, B. Cabral, R. Konrad, S. Sullivan. Capture4VR: from VR photography to VR video. *ACM SIGGRAPH 2019*.
- [38] H. Sun, S. Zollmann. Towards Casually Captured 6DoF VR Videos. 2022 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR): 176-179, California, USA, 12-14 December 2022.
- [39] A. Bevilacqua, G. Tentoni, A. Martinetti, L. Guardigli, B. Turillazzi, M. Fadda, A. Farina, L. Battisti, L. Tronchin. Measured soundscapes in Piazzale San Francesco in Parma and Trafalgar Square in London. 2025 Immersive and 3D Audio: from Architecture to Automotive (I3DA), Bologna, Italy, 10-12 September 2025.
- [40] K. Attenborough. Sound propagation in the atmosphere. In book: *Handbook of Acoustics*: 117-155, Springer, 2014. DOI: 10.1007/978-1-4939-0755-7_4.
- [41] O. Cramer. The variation of the specific heat ratio and the speed of sound in air with temperature, pressure, humidity, and CO2 concentration, *J. Acoust. Soc. Am.* 1993, 93: 2510-2516.
- [42] A. Pastusiak, L. Blasinski, J. Kocinski. (2025). Listening Effort in Reverberant Rooms: A Comparative Study of Subjective Perception and Objective Acoustic Metrics. *Archives of Acoustics*, 50(3), 321–329. <https://doi.org/10.24425/aoa.2025.154824>.

[43] E. Arvidsson, E. Nilsson, D. Bard Hagberg, O.J.I Karlsson. (2021). The Difference in Subjective Experience Related to Acoustic Treatments in an Ordinary Public Room: A Case Study. *Acoustics*, 3, 442-461. <https://doi.org/10.3390/acoustics3020029>.

[44] N. Zheng, N. Peters, E. Bates. (2025). On the relationship between objective room acoustic parameters and perceived differences in spatially interpolated RIRs. Immersive and 3D Audio: from Architecture to Automotive (I3DA), Bologna, Italy, 10-12 September 2025. DOI: 10.1109/I3DA65421.2025.11202090.

FIGURES

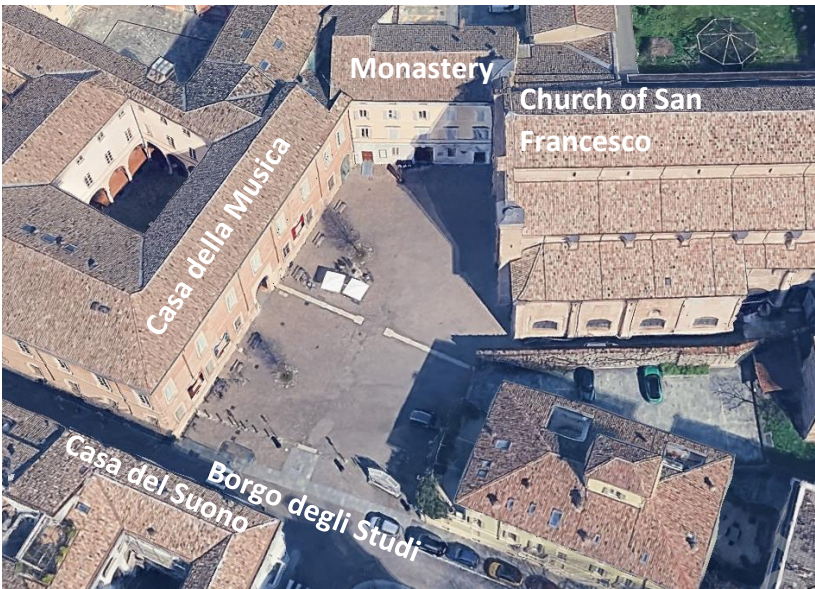


Figure 1. Please insert Figure 1 in the text body as indicated.

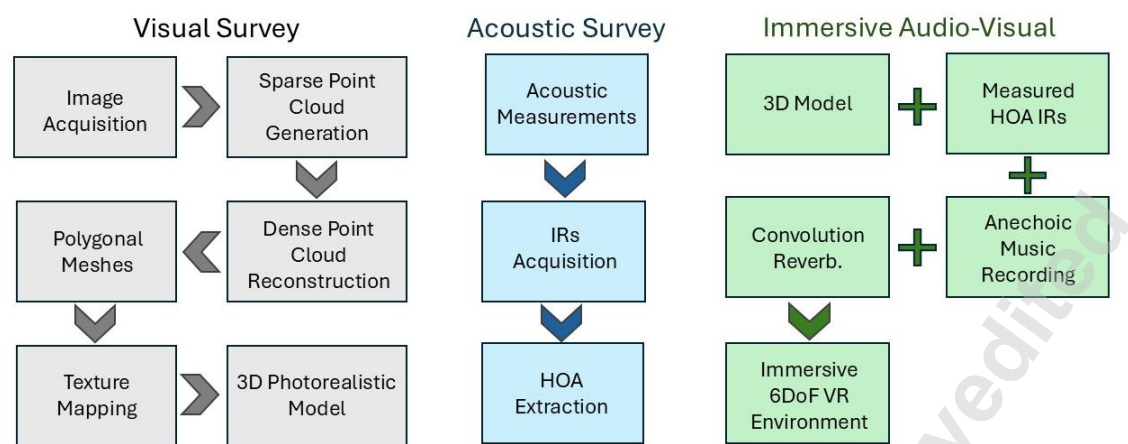


Figure 2. Please insert Figure 2 in the text body as indicated.

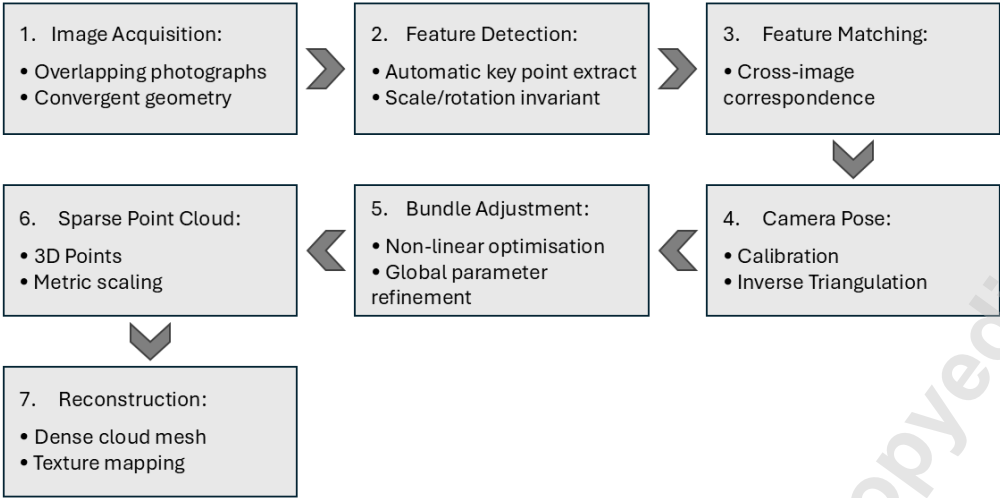


Figure 3. Please insert Figure 3 in the text body as indicated.

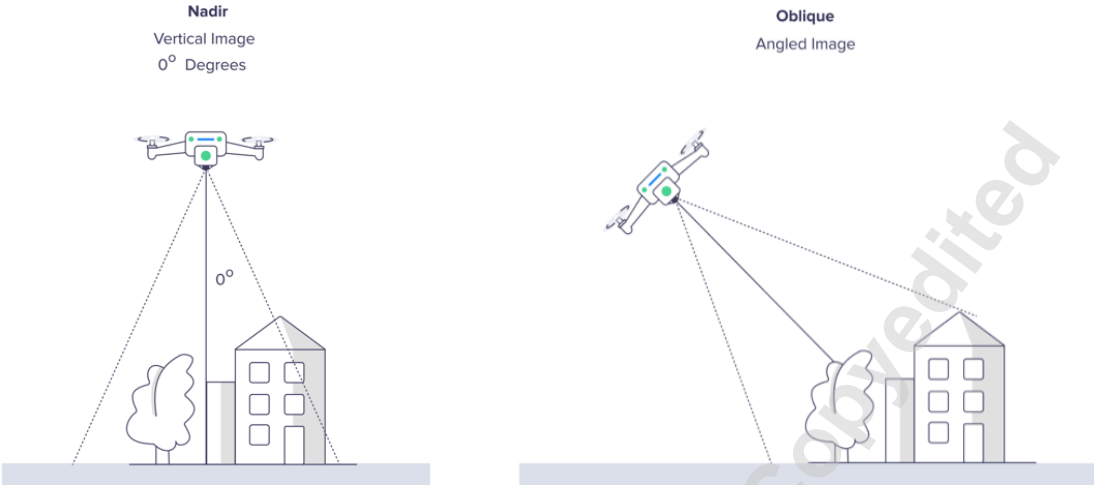


Figure 4. Please insert Figure 4 in the text body as indicated.



(a)



(b)

Figures 5A and 5B. Please insert Figures 5A and 5B in the text body as indicated.



Figure 6. Please insert Figure 6 in the text body as indicated.

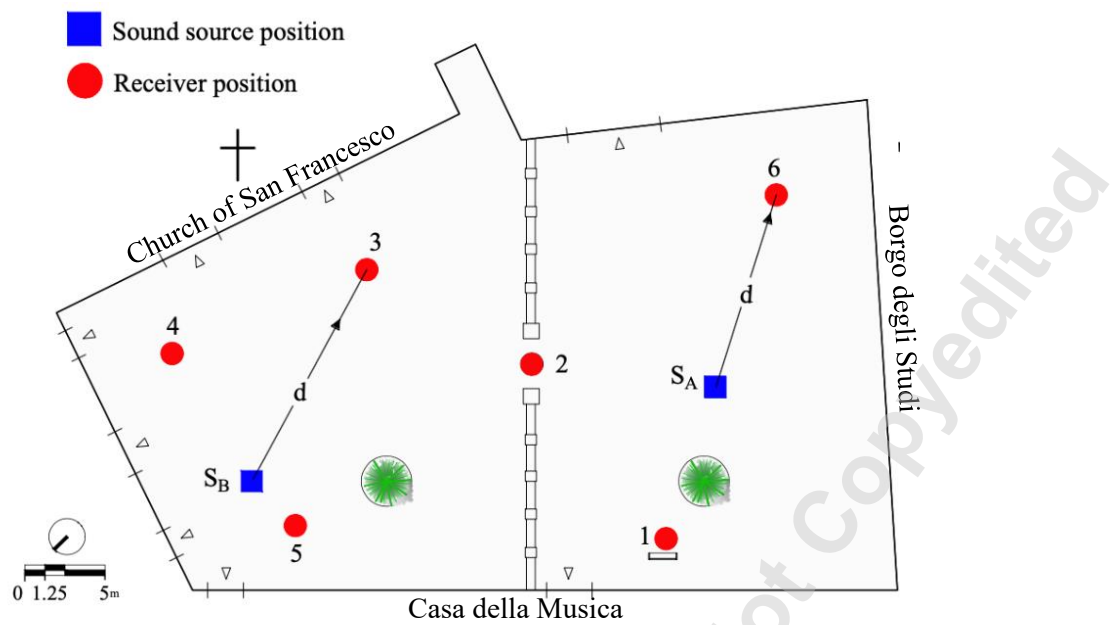


Figure 7. Please insert Figure 7 in the text body as indicated.

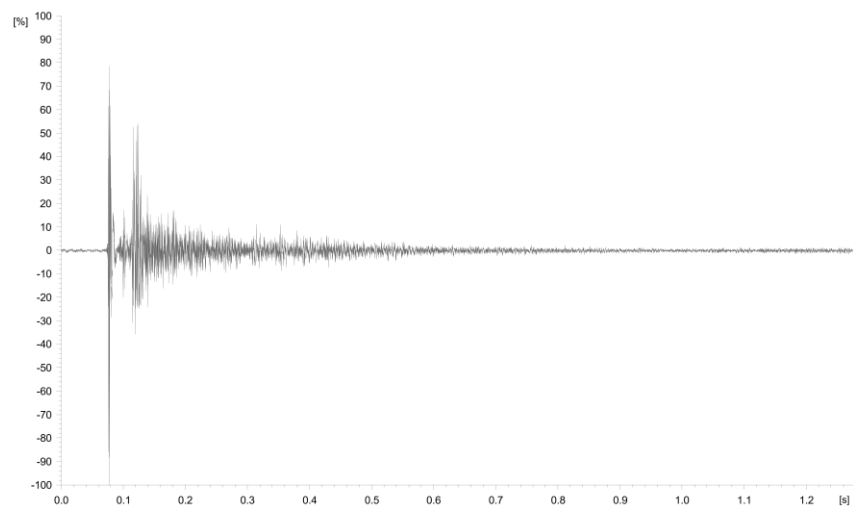


Figure 8. Please insert Figure 8 in the text body as indicated.

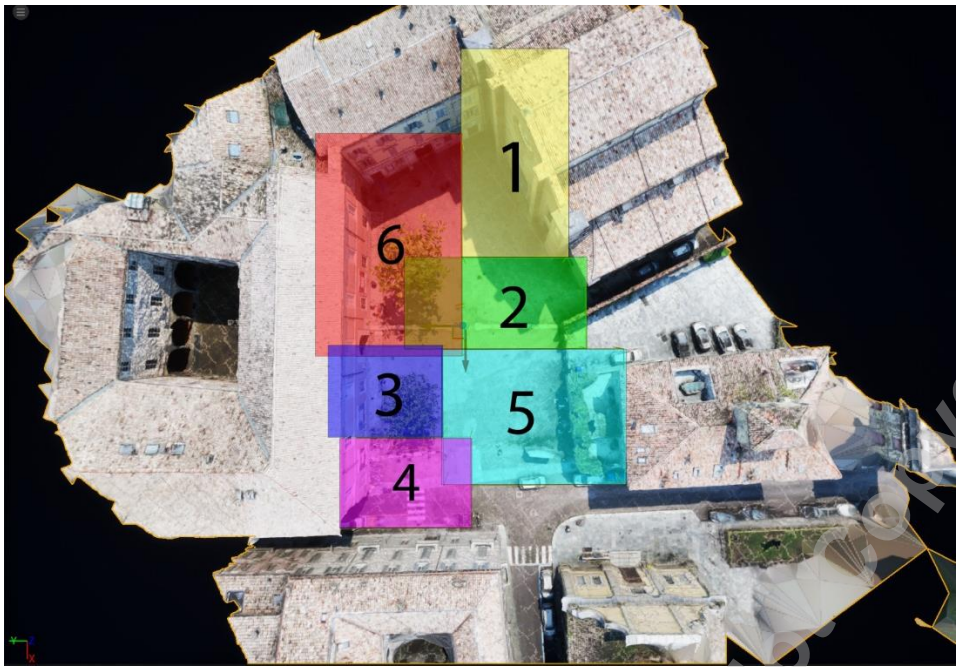
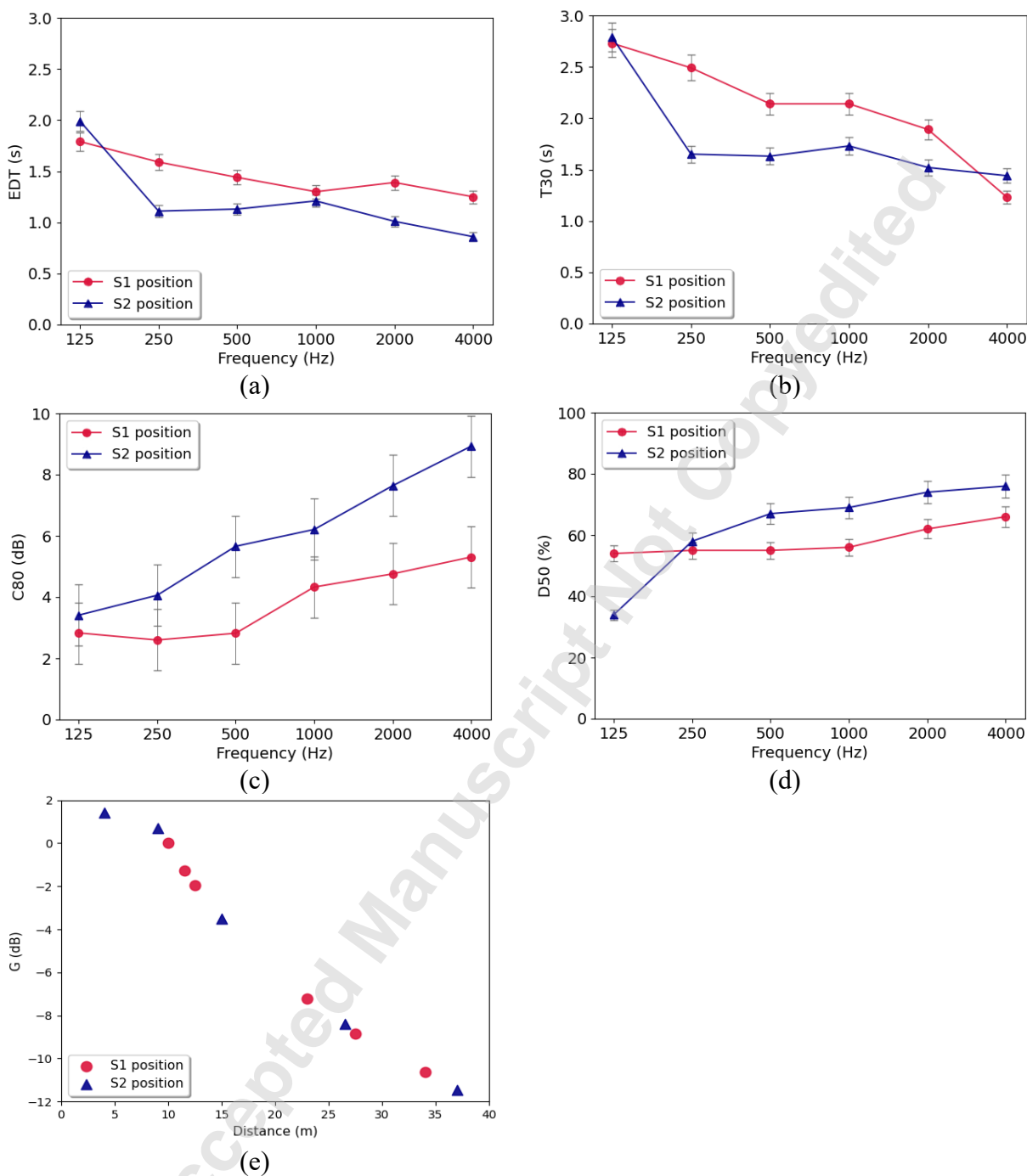


Figure 9. Please insert Figure 9 in the text body as indicated.



Figure 10. Please insert Figure 10 in the text body as indicated.



Figures 11A, 11B, 11C, 11D, 11E. Please insert Figures 11A, 11B, 11C, 11D, 11E in the text body as indicated.



Figure 12. Please insert Figure 12 in the text body as indicated.

TABLES

Type	No. Photos	Format	Size (GB)	Overlap
Drone	1479	JPEG + DNG	41,20	Approx. 80%
Camera	1431	RAW + JPEG	44,80	Manual alignment, High redundancy

Table 1. Please insert Table 1 in the text body as indicated.

Sound Source Position	Receiver Position	Distance (m)
A	1	9.98
A	2	11.61
A	3	23.03
A	4	34.09
A	5	27.70
A	6	12.54
B	1	26.28
B	2	19.04
B	3	15.00
B	4	9.26
B	5	4.03
B	6	37.40

Table 2. Please insert Table 2 in the text body as indicated.

List of Figures

- Figure 1. Piazzale San Francesco in Parma, Italy.
- Figure 2. Schematic overview of the methodological framework.
- Figure 3. Structure from Motion (SfM) computational process.
- Figure 4. Drone angles used for photogrammetry.
- Figure 5. In-situ visual data collection (a), view from the drone (b).
- Figure 6. Acoustic equipment deployed in Piazzale San Francesco for the acoustic survey.
- Figure 7. Schematic equipment positions taken in Piazzale San Francesco, Parma.
- Figure 8. Omnidirectional impulse response from measurements carried out within Piazzale San Francesco, Parma. Equipment position: S2 – R3.
- Figure 9. Reverb zones identification for the Audio-visual rendering of Piazzale San Francesco in 6DoF.
- Figure 10. Snapshot of the RealityCapture view for polygonal mesh resolution.
- Figure 11. Measured acoustic parameters within Piazzale San Francesco in Parma, Italy: Early Decay Time (a), T30 (b), C80 (c), definition (d), strength (e).
- Figure 12. View of virtual players in the digital environment of Piazzale San Francesco.

List of Tables

- Table 1. Characteristics of visual survey.
- Table 2. Distance between sound source and receiver positions.