DOI: 10.61782/fa.2023.1263

forum acusticum 2023

WALKABLE AURALIZATIONS WITH CONGRUENT INTERACTIVE VISUALS VIA GAME ENGINES IN AN IMMERSIVE CLASSROOM

Samuel Chabot[1*] **Mincong Huang**[1] **Jonas Braasch**[1]
[1] The Experimental Media and Performing Arts Center, Rensselaer Polytechnic Institute,
Troy, NY, USA

ABSTRACT

Students experience and learn consequences of room design through audio-visual-congruent walkable auralizations in a human-scale immersive environment. A 128-channel wave-field synthesis system presents auralizations created with an efficient two-dimensional ray-tracing algorithm and publicly available floor plans. Additional volumetric data inform late reverberation parameters. Congruent visuals are produced using several methods: student-generated imagery is automatically formatted and transmitted to the display using a web-based interface accessible via personal devices. This interface also provides near-instant access to innumerable locations available in massive online user-contributed databases via their application programming interfaces, such as those offered by the Google Maps Platform. Finally, more dynamic visuals are created through the use of game engines, such as Unreal Engine, which render sites interactive. Multiple generated receiver points allow students to actively move the listening position and explore changes across a venue. In Rensselaer's course Aural Architecture, students of both architecture and acoustics combine their skillsets to produce these dynamic and interactive audiovisual (re)creations of spaces and historical concert venues. Listeners can walk through these in Rensselaer's CRAIVE-Lab over an extended user area (12 m × 10 m). However, the method is extensible to other immersive environments with adequate audiovisual fidelity.

Keywords: *auralization, immersive environment, ray-tracing, game engine*

1. INTRODUCTION

"Situated learning" is the concept of locating a learning experience in a specific setting, environment, or activity [1], often, the context the learner would be using their acquired knowledge within. This "learning through doing" approach is growing within the science, technology, engineering, and math (STEM) education community. Hands-on demonstrations and activities apply classroom knowledge to real-world examples. The Education Resource Information Center (ERIC) reports: "Applied to the classroom, situated learning is not only reflecting upon and drawing implications from previous experiences but is immersion in and with the experience" [2].

One method for achieving high levels of immersion in a variety of activities and environments is the harnessing of virtual reality (VR) technologies. As early concepts of VR began appearing, William Bricken drew connections between the technology and its potentials for education [3]. The development of CAVE systems [4] renders it possible for a classroom of learners to enter these virtual environments collectively. Previous work to locate learners within appropriate visual [5]–and aural [6]–contexts has been done. Students utilize a CAVE-like environment to reinforce architectural acoustics knowledge acquired in a traditional classroom setting.

Foundational to architectural acoustics education is the acoustic impulse response and its relationship with the built environment. The impulse response is used to characterize and auralize interior spaces, both existing and in development [7]. These can be measured between a source (loudspeaker) and receiver (microphone) location in-situ, and also estimated using a computer-generated virtual space. The convolution with sufficiently dry au-

---

*__Corresponding author__: chabos3@rpi.edu.

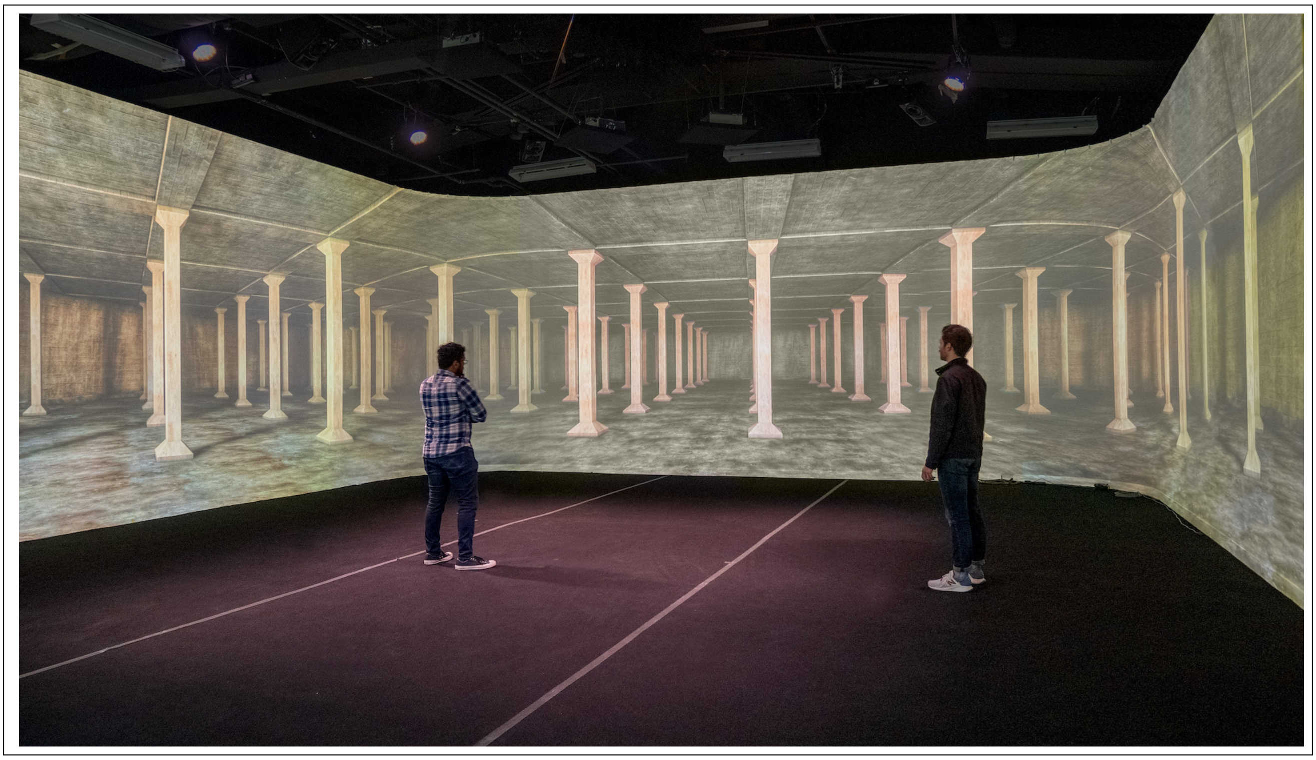

Figure 1. Two users stand within a simulated environment.

dio files produces an auralization simulating a recording captured with the source-receiver pair [8]. For a deeper understanding see [9]. In educational scenarios, these auralizations create opportunities for students to draw connections between perceptual experiences and the complex built environment and apply classroom learning to real-world examples [10].

These auralizations can be produced for various spatial audio playback systems, including wave-field synthesis (WFS). This technique utilizes Huygen's concept of elementary waves [11] and Fresnel's principle of interference [12] to simulate the propagation of sound waves using a dense array of loudspeakers [13]. An advantage of this spatial audio technique is the ability to reproduce sound accurately across larger listening areas.

The environment in which much of this research occurs is Rensselaer's *Collaborative Research Augmented Immersive Virtual Environment*, or CRAIVE-Lab. shown in Fig. 1. The space hosts up to 49 participants (per fire code) in its floor area of 10 × 12 m. The lab seeks equal emphasis on both audio and visual content for a congruent experience. Visuals are produced using eight short-throw projectors blended together on a 4.3 m-tall, nearly 360° micro-perforated screen. The perforation allows the screen to remain acoustically transparent for the dense horizontal array of 128 loudspeakers located behind it. Six more hung from the ceiling provide elevation cues. The system achieves a low-frequency cutoff of 37 Hz and a spatial aliasing frequency of about 900 Hz.

The method discussed in Sec. 2 takes a 2D approach over more traditional 3D methods, conveying a few advantages in the educational setting. The production and simulation of 3D models requires much greater time and complexity. This 2D method creates auralizations from accessible floor plans in some cases in less than an hour. The volume of the simulated space is used to calculate an appropriate late reverberation tail.

In the past, paired visuals have been shown in the form of static panoramic images. These are produced either in-situ by the user or sourced from online crowd-sourced repositories such as Google Street View. The latter has streamlined the process tremendously and opened up the possibility of visualizing sites that would likely otherwise be geographically inaccessible to the user. The in-

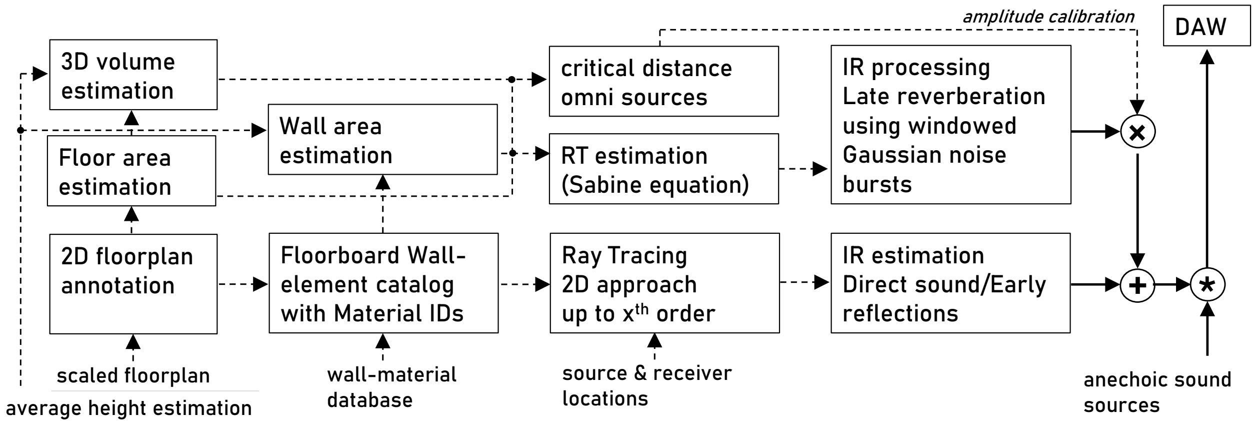

Figure 2. System architecture for the auralization algorithm.

terface for exploring available locations is browser-based and therefore accessible via users' personal devices, reducing necessary training or configuration. However, these images are still static and do not allow users to move about the virtual environment.

To this end, the inclusion of dynamic, explorable visuals is enabled through the use of game engines, such as Unity or Unreal Engine. Environments are modeled in software such as Blender to create visual recreations. These are then imported into a game engine environment in which users have control to move the viewpoint (i.e. the ability to explore). Auralizations for multiple source-receiver pairs can be rendered so that exploration of the visual environment corresponds to changes in the acoustical environment. This allows students to experience perceptual differences across an environment, deepening the understanding of the complex relationship between built environment and aural perception.

2. RAY-TRACING AURALIZATION

In the auralization method, depicted in the flow diagram of Fig. 2, a ray-tracing algorithm is used to estimate the early reflections at a receiver position. A floor plan with known dimensions or scale provides the basis for the boundary conditions. The floor plan is annotated using an image editing software such as Photoshop to enumerate the walls, corners, and scale of the space. Source and receiver positions are also annotated on the floor plans. The receiver positions simulate virtual microphones for each loudspeaker in the WFS setup. Wall and floor absorption coefficients are assigned using materials listed in the Deutsches Institut für Normung (DIN) database [14].

The program emits rays at equidistant azimuth angles covering the full horizontal plane. At each boundary, the reflection angle of an intersecting ray is calculated using Sahl/Snell's law [15,16], which estimates the reflected angle equals the incoming angle, $\cos(\alpha_r) = \cos(\alpha_i)$. The next order ray continues in the new direction until reaching a maximum order specified by the user.

These rays are collected by the enumerated receivers (virtual microphones). A virtual array of receivers are positioned at the corresponding real-world loudspeaker positions, rendering a separate impulse response for each. The algorithm determines which rays intersect a defined radius about each receiver. The intensity of each ray is reduced by distance traveled according to the inverse-square law, as well as by the corresponding absorption coefficients applied to the boundary. Dissipation effects of air are also considered for high frequencies.

A frequency-specific late diffuse reverberation is also generated using estimated volume and wall, floor, and ceiling dimensions. With the dimensions and applied absorption coefficients, reverberation times can be calculated using the Sabine equation:

$$T_{60} = 0.161 \cdot V/(A + 4m \cdot V)\mathrm{s}. \tag{1}$$

where total absorption, A, is calculated defined as the sum of all surface elements, S_n, multiplied with their applied absorption coefficient, α_n:

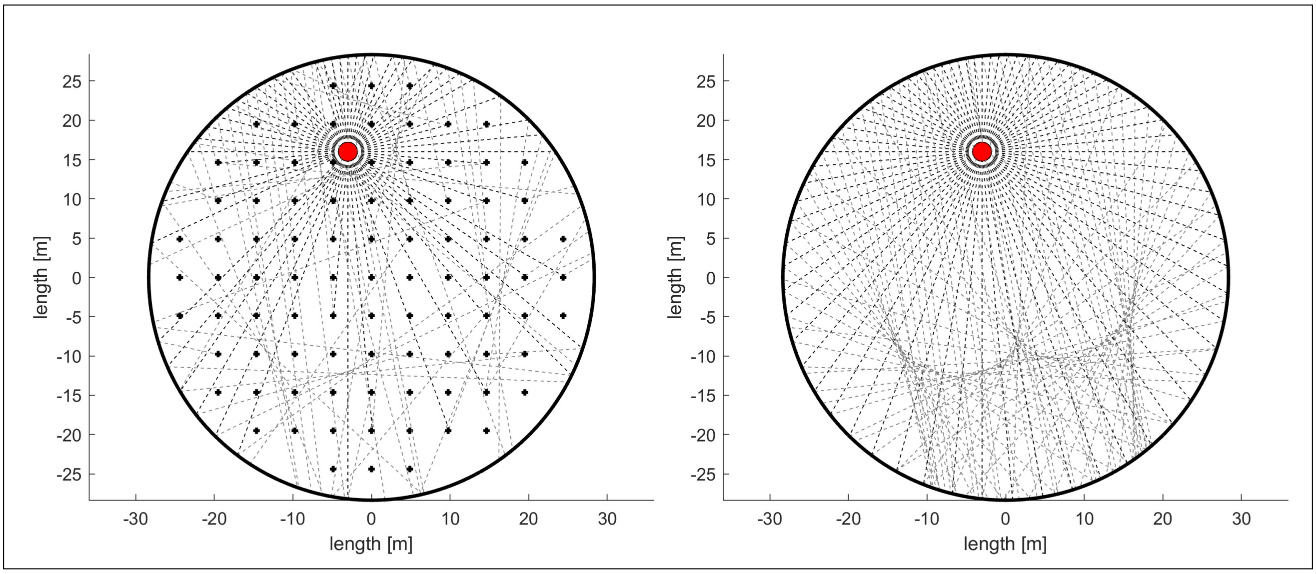

Figure 3. Ray-tracing algorithm performed on the Dan Harpole cistern, with columns (left) and without (right).

$$A = \left(\sum_{k=1}^{K} \alpha_k \cdot S_k \right). \tag{2}$$

A unique Gaussian noise sample is processed for each receiver using an exponentially-decaying window adjusted to the frequency-specific reverberation time. This method has been shown to benefit from a large number of independent reverberation channels [17]. This subsequent late reverberation is adjusted to balance the total energy of the impulse response and the sum of calculated direct and reverberant energy. A cross-fade method blends the decay of early reflections into the late reverberation.

3. GAME ENGINE VISUALIZATION

In order to facilitate rapid content creation, an interface has been developed which allows non-expert users to contribute imagery acquired in-situ or from crowd-sourced repositories such as Google Street View [6]. This interface is browser-based and can therefore be accessed from users' personal devices. While this method has been successful at rapidly producing congruent visuals, these images are static.

To offer more dynamic visuals that allow for exploration, a more in-depth method is undertaken. Virtual recreations of the auralized environments are produced using 3D modeling software such as Blender. These programs allow for the fabrication of the environment and texturing of its surfaces. These 3D models are then imported into game engines, such as Unity or Unreal Engine. For the CRAIVE-Lab, a Unity package has been developed by coauthor Huang which utilizes a 360° camera rig and first-person controls. Content such as the modeled environment is then brought into the game and placed with respect to the rig. This determines its placement on the screen. The built "game" makes use of the entire panoramic screen, offering a complete horizontal field-of-view of the modeled environment. Users navigate with familiar or custom input devices: keyboard and mouse, traditional game controller, custom iPad interface, etc.

Currently, multiple source-receiver pairs are rendered for the auralization so that the user can experience the simulated acoustics at various positions in the environment. During visual exploration, they have the opportunity to toggle between these available rendered source-receiver pairs. These measured positions can be annotated in the model or the game engine to convey where users can experience congruent auralizations.

4. EXAMPLE AND FUTURE WORK

To demonstrate this work, an example is created for the Dan Harpole Cistern in Fort Worden, WA. Built in 1907 and originally an underwater tank, the space is now a massive, empty concrete structure featuring a reported 45 sec-

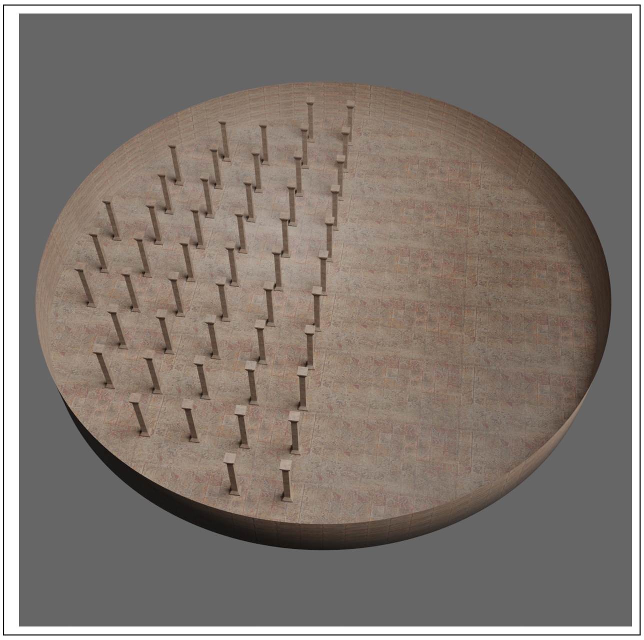

Figure 4. Model of the Dan Harpole Cistern created using Blender. Half of the cistern is shown with the existing columns, while half is shown without. Various iterations of these environments can be created for auralization. Modeled content is then imported into a game engine for real-time exploration.

ond reverberation time.

The space is first modeled in Blender. Figure 4 depicts half of the space with its numerous columns and half without. For this environment in particular, the 89 columns diffuse what would otherwise be audible focusing effects due to the cylindrical walls. In his book *Acoustics and the Performance of Music*, Jürgen Meyer points out the importance of columns in churches, noting that their effectiveness depends on their thickness, with a lower limit as a reflector around 1000 Hz to 1500 Hz and complete bending around columns expected only below 200 Hz to 300 Hz [18]. This phenomenon is demonstrated to students using this approach. The results of the ray-tracing algorithm can be seen in Fig. 3, with and without columns at left and right, respectively. A strength of these dynamic visualizations is the ability to simply toggle between different iterations of an environment.

This methodology is utilized regularly by students of acoustics to experience the complex relationship between the built environment and perceived room acoustics. Two students within a rendering of the Dan Harpole cistern can be seen in Fig. 1. Students engage in and become familiar with an acoustic simulation workflow similar to that they will use after graduating. Whereas previously students only sourced static visuals, the production of dynamic visuals provides exposure to a greater variety of iterations, and familiarity with game engines is seen as an advantage in a professional setting where their usage is only increasing.

While these examples are rich learning opportunities, they are also more complicated to produce. In order to assuage some of this added overhead, coauthor Huang is developing a real-time ray-tracing algorithm for game engines in order to produce real-time auralizations of acoustic environments. This would circumvent the current steps required to auralize individual source-receiver pairs, creating an increasingly dynamic application.

Finally, a binaural manikin has validated the walkable nature of the auralizations, and subsequent assessments should measure the benefit to students, perhaps by comparing the knowledge of those with and without exposure to the facility.

5. CONCLUSIONS

These dynamic visualizations congruent with walkable auralizations are utilized in courses of architectural acoustics such as *Aural Architecture*. Students choose a site of interest and create their own walkable auralizations. They then explore an aural and visual environment to better understand how the built environment shapes acoustic perceptions. They report a richer grasp of the acoustical knowledge and terminology acquired in a traditional classroom setting through the direct application in these experiential examples. Students collaborate to produce the complex visual and aural environments to practice the workflow of a professional setting. The list of available sites continues to grow, and current auralized venues with static or dynamic congruent visuals include the Cologne Cathedral, St. Patrick's Cathedral in New York City, and the Pantheon in Rome.

6. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant IIS-1909229.

7. REFERENCES

[1] J. Lave and E. Wenger, *Situated Learning*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[2] D. Stein, "Situated learning in adult education," Tech. Rep. ED418250, ERIC Clearinghouse on Adult, Career, and Vocational Educ., Columbus, OH, USA, 1998.

[3] W. Bricken, "Learning in virtual reality," Tech. Rep. ED359950, Washington Univ., Seattle, WA, USA, 1990.

[4] C. Cruz-Neira, D. J. Sandin, T. A. Defanti, R. V. Kenyon, and J. C. Hart, "The CAVE: Audio visual experience automatic virtual environment," *Commun. ACM*, vol. 35, pp. 64–72, Jan. 1992.

[5] S. Chabot, J. Drozdal, M. Peveler, Y. Zhou, H. Su, and J. Braasch, "A collaborative, immersive language learning environment using augmented panoramic imagery," in *Proc. Immersive Learn. Res. Netw. Conf. 2020*, (San Luis Obispo, CA, USA), pp. 225–229, 2020.

[6] S. Chabot and J. Braasch, "Walkable auralizations for experiential learning in an immersive classroom," *The Journal of the Acoustical Society of America*, vol. 152, pp. 899–910, 8 2022.

[7] D. T. Murphy, "Archaeological acoustic space measurement for convolution reverberation and auralization applications," *Proc. 9th Int. Conf. Digit. Audio Effects, DAFx 2006*, pp. 221–226, 2006.

[8] M. Vorländer, *Auralization*. Springer, 2020.

[9] J. Braasch, "Convolution, Fourier analysis, cross-correlation and their interrelationship," in *Springer Handbook of Systematic Musicology* (R. Bader, ed.), pp. 273–284, Cham, Switzerland: Springer, 2018.

[10] S. Pelzer, L. Aspöck, D. Schröder, and M. Vorländer, "Integrating real-time room acoustics simulation into a CAD modeling software to enhance the architectural design process," *Buildings*, vol. 4, no. 2, pp. 113–138, 2014.

[11] C. Huygens, *Traité de la lumière où sont expliquées les causes de ce qui luy arrive dans la reflexion, & dans la refraction, et particulierement dans l'etrange refraction du cristal d'Islande [Treatise on light in which are explained the causes of that which occurs in reflection, in refraction, and particularly in the strange refraction of Iceland spar]*. chez Pierre Vander Aa marchand libraire, 1690.

[12] A. Fresnel, "Mémoire sur la diffraction de la lumière [memorandum on light diffraction]," *Mémoires de l'Académie des Sciences, da p. 339 a p. 475: 1 tav. ft; AQ 21*, p. 339, 1819.

[13] A. J. Berkhout, "A holographic approach to acoustic control," *Journal of the audio engineering society*, vol. 36, no. 12, pp. 977–995, 1988.

[14] Deutscher Normenausschuß, ed., *Schallabsorptionsgrad-Tabelle*. Beuth-Vertrieb GmbH, Berlin, 1968.

[15] R. Rashed, "A pioneer in anaclastics: Ibn Sahl on burning mirrors and lenses," *Isis*, vol. 81, no. 3, pp. 464–491, 1990.

[16] M. Zghal, H.-E. Bouali, Z. B. Lakhdar, and H. Hamam, "The first steps for learning optics: Ibn Sahl's, Al-Haytham's and Young's works on refraction as typical examples," in *Education and Training in Optics and Photonics*, p. ESB2, Optical Society of America, 2007.

[17] S. Agrawal and J. Braasch, "Impression of spatially distributed reverberation in multichannel audio reproduction," in *Audio Engineering Society Convention 145*, p. 10076, Audio Engineering Society, 2018.

[18] J. Meyer, *Acoustics and the Performance of Music*. New York, NY: Springer, 5th ed., 2009.