

Article

Designing a Virtual Reality Platform to Facilitate Augmented Theatrical Experiences Based on Auralization

Eftychia Lakka ^{1,*}, Athanasios G. Malamos ², Konstantinos G. Pavlakis ² and J. Andrew Ware ¹ ¹ Faculty of Computing, Engineering and Science, University of South Wales, Treforest CF37 1DL, UK² Media, Networks and Communications Lab, Department of Electrical and Computer Engineering, Hellenic Mediterranean University, GR 71004 Heraklion, Greece

* Correspondence: efilakka@gmail.com

Received: 21 April 2019; Accepted: 13 June 2019; Published: 1 July 2019



Abstract: In the last few years, the immersive theater has become a new trend for modern performances. Venues increasingly utilize widely available computer technologies, such as virtual/augmented reality and spatial sound, to help facilitate the realization of different ideas. Motivated by this current trend, a prototype platform has been developed that enables the design and implementation of an augmented theatrical experience based on spatial sound immersion. This paper describes the implementation of the platform and, through several use case scenarios, its evaluation. The paper concludes with a discussion of the results and offers some thoughts on future developments.

Keywords: spatial sound; augmented reality; virtual reality; 3D; intelligent systems

1. Introduction

Recent technological developments are having a fundamental impact on the field of art design. Technologies such as holography, head-mounted displays, fulldome immersive video projection, kinesthetic communication (haptic technology), transparent monitors, three-dimensional (3D) sound, and electronic sensors facilitate development of sophisticated and interactive environments using augmented and virtual reality. These environments allow individuals to engage more fully with the art environment, make contact with the artist, and even participate in the creation of the art. Moreover, modern digital technologies play a key role in many theatrical performances, enabling immersive and kinetic theater scenography.

Augmented reality (AR) differs from virtual reality (VR) in that AR “supplements, rather than supplants, the real world” [1]. In other words, VR replaces the real world with a simulated experience, whereas AR allows a virtual world to be experienced while simultaneously experiencing the real one [2]. AR involves the interactive experience of a real-world environment in which the objects that reside within it are, through computer-generated information, augmented in such a way as to enhance the engagement experience. In a theater, the main purpose is not absolute virtualization that might reduce the artistic and historical value of the environment but rather an enhanced experience through augmentation of the actual environment. Hence, the goal of this augmentation is participant immersion, which is the ultimate goal of an effective augmented experience. Immersive installations satisfy the human desire to escape beyond the borders of reality and feel the experience as art [3].

Recent years have witnessed significant growth in the domain of the immersive environments. In the literature, “immersive” tends to be used to refer to the viewer’s sense of virtual incorporation in an environment that is either a real-world simulation or completely artificial. Vanguard technologies and innovative practices are now routinely combined with more traditional theatrical elements, including

not only image but also sound. It is generally accepted that the aurality constitutes an integral part of the production theatrical performance and offers additional details and a visceral sense to the immersive theater. Within a theater, aurality can be considered the germane sounds of the space, the sound effects, recorded music, live music, performers' voices, audience voices, and so forth.

Particularly, the immersive theater aims to enhance the experience of "traditional" scenes in which users experience the integration of virtual elements into the real world in order to enable more dynamic and interactive systems. One of the most effective immersion tools in art is sound [4]. The sound design and experience are usually referred to as aurality. Aurality encloses the synthesis, directionalization, spatialization, and reception of the sound in a virtual world [5].

The focus of this paper relates to one aspect of current developments, that is, the design of an augmented theatrical experience based on spatial sound auralization created by the efficient mixing of mature and innovative VR techniques. Section 2 provides an overview of the background to the work, Section 3 highlights the motivation of this work, Section 4 analyzes the implementation and the use case scenarios, and Section 5 draws conclusions from the work.

2. Background

2.1. Immersion in Theater

The evolution of VR and AR technology has provided new powerful tools and changed the interaction between the actors and the audience in the theater. Thus, traditional art could be affected by VR and/or AR technology and can take advantage of the pioneering immersive experience in order to disrupt the boundaries between the observers and observed objects.

Generally speaking, in the theater, the main purpose is to engage the audience, through as many senses as possible, in the action that takes place on the stage. Unlike the cinema, which has focused on improving the visual connectivity with the audience through such developments as 3D technology, in the theater, the dominant sense targeted to facilitate the viewer's immersion is hearing.

Immersive theater usually relates to aural experience and there are many examples of such theaters as illustrated later in this section. Immersive theater requires complex and expansive scenography and high-end technological means [5] in order to generate emotions.

In 2015, David Rosenberg [6], cofounder of the Shunt theater company in London, supported by Britain's Royal National Theatre, directed an extremely interesting thriller show that the audience watched in complete darkness while wearing headphones, so that the sound played a key role in the theatrical narrative. The sound of this performance was recorded using binaural technology (two microphones placed apart by the same distance as that between a human's ears) giving the viewer a significant sense of intimacy and immediacy.

In the same period, another category of performances, the podplays (internet downloadable plays designed using specific sound effects) appeared. Some of them were best listened to in a certain location, while for others, the listener watched the play while traversing a particular route. Moreover, many theatrical groups have used even more accessible technological media, such as mobile phones, to offer a different theatrical experience to the audience. For example, the theater group "This Is Not a Theatre Company", based in New York [7], presented a series of theatrical performances where the "audience" was required to download a recording of the show and then listen to it through headphones, thus having the ability to hear its "sound" in a space different from that of a theater. For instance, in the "Ferry Play", the viewer watched the show on a ferry on route to Staten Island in New York. The scene of the theater was the ferry; the viewers listened to the recording while watching nearby attractions, and through the sense of smell, became cognizant of the odors of the surrounding area.

2.2. Sound in Theater

Taking into account the heritage of the ancient Greek period, theaters were traditionally designed, with an emphasis on acoustics, with the aim of providing the best sound performance, as it was directed from the stage to the audience.

As Vitruvius stated, theater is for the “ear of the spectator”: people are required to listen to the actions. From the formal Ancient Greek amphitheaters to the Renaissance London Theater, the acting was advanced through theater sounds, as it was voices that made visible the performance [4].

One of the principal characteristics of the traditional theater was the fact that the performances relied on the natural sound of the voices and the movement of the actors. This method of sound propagation is ideal when there are a small number of actors that can be distinguished by their clear and strong speech, even if not speaking loudly. However, theatrical performances that are based on natural sound usually suffer from unrealistic rendering of emotional shades. For instance, two different scenes from a play maybe based in, for example, a church and an open area; however, the actors’ voices do not change, and as a result, they do not convey the locality of the scene to the hearer.

One of the first insertions into the sound projection mechanisms of the theater was the use of loudspeakers and microphones mounted in series on the stage; their purpose was to reinforce the acoustics, especially in areas without conducive natural sound. This linear location of sound equipment solves the problem of volume but does not facilitate personalization to a specific actor. This changed when personal wireless microphones were introduced.

A novel requirement in theaters is to augment an actor’s speech with pertinent spatial sound effects. For example, when a particular scene is based in a room, the addition of reverberation to the actor’s voice can help create authenticity in the sound. Such techniques can help imbue the atmosphere of the scenography and not only the emotions of the script in the overall sound experience. With this in mind, the last decade has witnessed significant growth in the introduction of spatial sound effects in theater stage performances. For example, at the Théâtre National de Chaillot [8], which has one of the largest concert performance halls in the city of Paris with a capacity of 1270 seats, they recently announced (December 2016) the installation of a spatial sound system using wave field synthesis (WFS) technology. In the same year, the New National Theater of Tokyo installed a sound system based on a sound analyzer and multichannel sound management (TiMax2 SoundHub-S32) [9]. This method offered automated vocal localization to the audience through live position tracking of the performers in the scene when listening to the voices of the actors. Additionally, N. Steinberg and J. Crystal [10] created the “Hamilton” performance in 2016 at the Richard Rodgers Theater in New York where they performed a spatial sound reproduction based on sound analyzer capabilities and a multichannel console using 172 independent speakers.

This paper introduces a groundbreaking alternative method to represent the sound in a theatrical experience based on spatial sound immersion, providing the spatial sound with physical attributes of sound to the audience.

2.3. Spatial Sound Technology

The spatial sound provides the user of virtual environments (VEs) the ability to recognize the location of sound source(s) [11], deduce information concerning the environment around the sound source(s), and, in general, to perceive the immersive environment in the same way as a listener recognizes the sound in the real world. Therefore, it is necessary that the sound field in a 3D space be computed, and thus from the physical/algorithmic point of view, sound propagation techniques must be used to simulate the sound waves as they travel from each source to the listener by taking into account the interactions with various objects in the scene [12]. In other words, spatial sound rendering in a VE goes far beyond traditional stereo and surround sound techniques through the estimation of physical attributes which are involved in sound propagation [13]. Phenomena such as surface reflection, diffusion, reverberation, as well as wave phenomena (interference, diffraction) can be included in the formation of spatial impressions within a virtual 3D scene. A brief discussion of these phenomena and the importance of their consideration for giving a listening impression of presence in a closed room follows.

Reflection: During the propagation of a sound wave in an enclosed space, the wave hits objects or room boundaries and its free propagation is disturbed. Moreover, during this process, at least a portion

of the incident wave will be thrown back, a phenomenon known as reflection. If the wavelength of the sound wave is small enough with respect to the dimensions of the reflecting object and large compared with possible irregularities of the reflecting surface, a specular reflection occurs. This phenomenon is illustrated in the Figure 1a, in which the angle of reflection is equal to the angle of incidence. In contrast, if the sound wavelength is comparable to the corrugation dimensions of an irregular reflection surface, the incident sound wave will be scattered in many directions. In this case, the phenomenon is called diffuse reflection (Figure 1b).

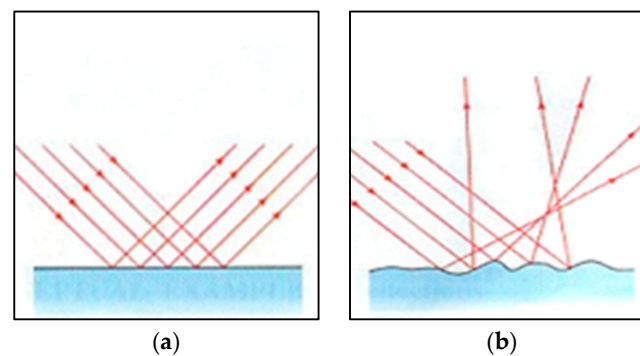


Figure 1. Reflection—sound physical phenomenon. (a) Specular reflection phenomenon. (b) Diffuse reflection phenomenon.

The brain of the listener works out subconsciously the early reflections (arrival time less than 40 ms after the direct sound from sound source) from objects in a closed room and takes information about the room, the surfaces, and their distances from the listener and increases the loudness of the initial sound field appropriately. Sound reflections that is received later, at times greater than 40 ms after the direct sound, contribute to what is known as reverberation.

Reverberation: Reverberation is the extended duration of a sound for a certain period after the termination of the sound source. It is caused by the successive late reflections of sound on the boundaries (walls, floor, ceiling, and other objects) inside a closed room. Again, this is a very important phenomenon because it gives the listener the feeling of its presence in an enclosed space instead of an open area as well as a perception of the size and the characteristics of the room.

Diffraction: Another easily experienced phenomenon of a sound wave is the diffraction perceivable when, for example, hearing but not seeing another person from behind a door. Diffraction is the spread of waves around corners (Figure 2b), behind obstacles, or around the edges of an opening (Figure 2a). The amount of diffraction increases with wavelength, meaning that sound waves with lower frequencies, and thus with greater wavelengths than obstacles or openings dimensions, will be spread over larger regions behind the openings or around the obstacles [14].

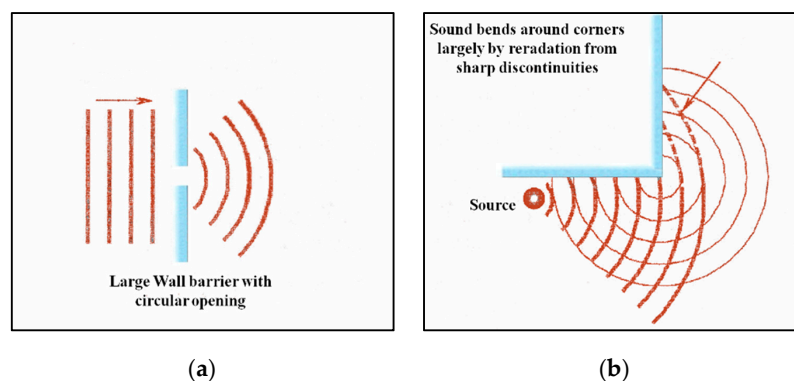


Figure 2. Diffraction—sound physical phenomenon. (a) Behind obstacles or around the edges of an opening. (b) Around corners.

Refraction: Refraction is the change in the propagation direction of waves when they obliquely cross the boundary between two mediums where their speed is different (Figure 3). This phenomenon should be considered for a realistic sound simulation [15].

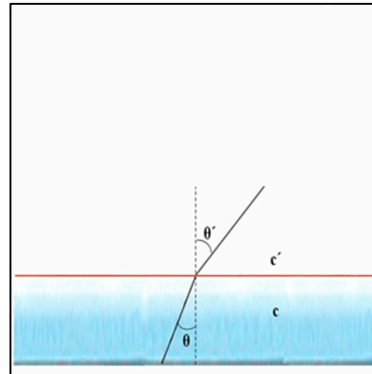


Figure 3. Refraction—sound physical phenomenon.

3. Motivation

This paper underlines the importance of real-time spatial sound immersion in live theater, enabling more dynamic projected content and audience interactivity. Specifically, the creators of a performance would have all the information about the sound field for every position in the scene during the performance. Hence, they would be able to locate the listener-viewer of the performance at different positions on the stage depending on the plot.

The prototype platform, which is presented in this paper, offers the opportunity for the viewer to intervene and modify the listening point by moving it to different parts of the stage, for example, listening to an ancient Greek tragedy from the location of a member of the chorus, the location of a specific actor, and so forth. Furthermore, the recorded or synthesized sounds are integrated into the scene and are included into the virtual and augmented environment in order to immerse the viewer into the world of the play.

Thus, the principal advantage of the prototype platform is the innovative combination of traditional theatrical elements, including VR and AR using 3D sound. In particular, Table 1 summarizes the effects that have been implemented in comparison with previous approaches. Taken together, it will enrich the theatrical experience of a play through the simulation of its sounds' physical phenomena.

Table 1. Comparison of sound effects between the paper implementation and the previous approaches.

Sound Effect	Previous Approaches	New Approach
Doppler	✓	✓
Filters	✓	✓
Gain	✓	✓
Panning	✗	✓
Reverberation	✗	✓
Absorption Coefficient	✗	✓

4. Methodology

4.1. Prototype Platform

The prototype platform is based on:

- Extensible 3D (X3D) [16], which is an ISO standard and the successor to VRML.
- The X3DOM framework [17], as reported by Behr et al. [14]. It is an open source JavaScript framework, which incorporates a set of technologies including X3D, WebGL, HTML, CSS, and JavaScript.

- The introduction of spatial sound components in the X3DOM framework, based on X3D specification and Web Audio API [18].

Web development tools, rather than the more traditional graphics engines (gaming machines and graphics libraries), were chosen due to the fact that there has been significant growth in the number of web applications. The global adoption of JavaScript and its derivative architectures (React, Angular) and the empowerment of hybrid technologies has created a landscape where the internet is a platform rather than a communication mechanism. Moreover, advancements in cloud computing, capable of supporting a wide variety of demanding services, has made it possible to shift technology to the internet almost completely.

From a practical point of view, the use of online technology offers the possibility of distributed processing and load balancing between multiple systems. This feature is valuable when addressing the crucial issue of real-time sound that has a high sensitivity to delay time (<50 ms). The feasibility of using internet technology to resolve corresponding problems has already been positively explored and a series of interactive real-time examples in that area have already been presented [19].

4.1.1. Spatial Sound System

This section is devoted to the implementation of the propose design. The implementation takes advantage of recent improvements in web browser technology (HTML5 and the embedding of a number of new libraries that implement new objects and services). Previous work [18] introduced an extended version of X3DOM platform that incorporates spatial sound capabilities within a virtual scene using the widely supported Web Audio API proposed by the World Wide Web Consortium (W3C) and implemented in all browser platforms.

Web Audio API [20] is a high-level JavaScript API which can be used to synthesize audio in web applications. The approach of Web Audio API is based on the concept of audio context, which presents the direction of audio stream flows, between sound nodes. In every node, the properties of the sound can be adapted and changed depending on the application requirements. Additionally, it is open source and supported by most browsers. Furthermore, multichannel audio is available and is integrated with Web Real-Time Communications (WebRTC). High-level sound facilities such as filters, delay lines, amplifiers, and spatial effects (such as panning) are offered. At the same time, audio channels can have 3D distribution determined by the position, speed, or direction of the viewer and the sound source. In addition, it is characterized by compositionality using audio node structure, which can be linked together in order to form an audio routing graph.

In the first instance, the AudioListener object represents the position and orientation of the unique person listening to the audio 3D scene using a panning algorithm. Its properties have the positionX, positionY, and positionZ parameters, which represent the location of the listener in 3D Cartesian coordinate space. Additionally, the forwardX, forwardY, and forwardZ parameters represent a direction vector in 3D space. Both a forward vector and position vector are used to determine the orientation of the listener. Another fundamental module of the approach is the PannerNode object (based on panning algorithm below) that has an orientation vector representing in which direction the sound is projecting. For the reason that both sound source and listener can be moved, they have a velocity vector associated with them that is used to record the speed and direction of movement. Together, these two velocities can be used to simulate the Doppler effect.

Panning Algorithm: Both mono-to-stereo and stereo-to-stereo panning should be supported. The first type of processing is used when all connections to the input are mono. Otherwise, stereo-to-stereo processing is used. Hence, the panning algorithm identifies how the sound spatialization is calculated. It includes two main models: head-related transfer function (HRTF) panning and equal-power panning.

1. HRTF Panning

This model requires a set of HRTF impulse responses recorded at a variety of azimuths and elevations. As a result, the development needs a highly optimized convolution function, which is

more computationally intensive than the equal-power panning technique but offers more perceptually spatialized sound.

2. Equal-Power Panning

This model is a simple and low-complexity algorithm with satisfactory results. It uses equal-power panning in which the elevation is ignored.

Distance Effects: Closer sounds appear louder, while sounds further away appear quieter. Exactly how a sound's volume changes according to distance between the listener and the sound source depends on the distance model (Figure 4).

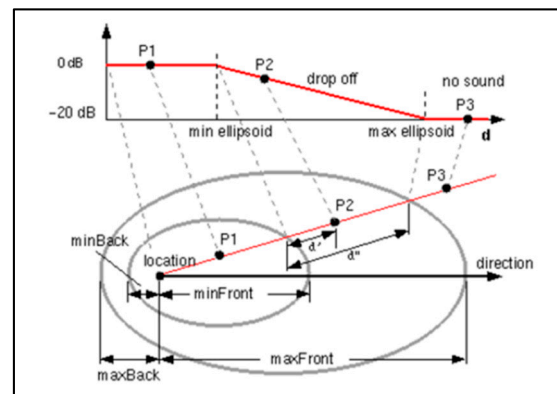


Figure 4. Distance model. From Extensible 3D (X3D), Part 1: Architecture and base components, sound component [21].

Development of Sound Physical Phenomena

The W3C Web Audio API [20] does not support physical phenomena in sound except for the Doppler effect. Within this paper, the standard was expanded and amplified with additional parameters including echo, reflection, absorption, and diffraction of sound. In addition, on the objects of the X3DOM 3D platform, the absorption coefficient was added as a new property in order to describe the material of the objects in the 3D environment.

For the implementation of these sound physical phenomena (echo, reflection, absorption, and diffraction), the sound is separated into individual frequency parts. This is facilitated by filters that are endogenously supported by the Web Audio API. For further development of the sound physical phenomena per frequency band, it is first necessary to identify the point at which the sound impacts the surface of the virtual space. Traditionally, geometric methods are the most commonly utilized techniques to address complicated issues, such as the analysis of sound in bands of frequencies. However, a growing body of literature has proposed hybrid methods as effective techniques for the development of sound physical phenomena for the reason that these methods combine different algorithms and incorporate the benefits of all of them. Based on that, the combined model of ray tracing and image sources (Figure 5) is the most suitable for analyzing sound in bands of frequencies, as it can increase the calculation cost in real-time applications [15].

Reverberation Algorithm Feedback Delay Networks

The propagation of the physical phenomenon of reverberation is achieved by simulating the feedback delay networks that create the reflected sound wave, taking into account the geometry of the scene [22].

Correspondingly, the sound physical phenomena of reflection and diffraction are analyzed using the above hybrid model of ray tracing and image-sources method for each frequency band. Finally, absorption is implemented using the absorption coefficient, which is a characteristic of the materials of the objects in the 3D environment.

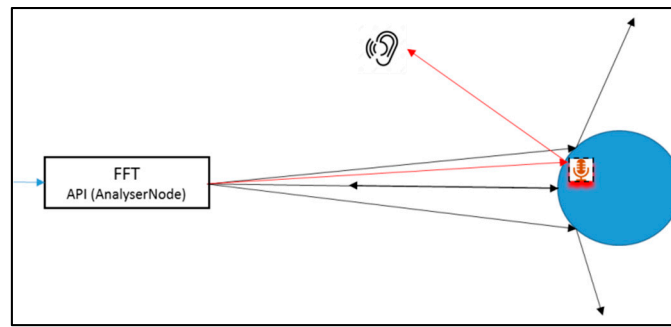


Figure 5. A hybrid model combining ray tracing and image sources method for analyzing sound in bands of frequencies.

4.1.2. Tracking System of Actors

The method to enable the tracking of actors is based on several studies [23–25]. This part of the implementation includes either the visual tracking of actors' movement on the stage or tracking, using Bluetooth technology, the location of wearable sensors attached to the actors' clothes. The most appropriate method depends on the requirements and circumstances of the theater stage.

4.2. Use Case Scenarios

In order to confirm the validity of the method, the use case scenarios presented below were considered:

Use Case 1:

In the first use case scenario, the prototype platform was connected to the DOLBY Atmos (available online: <https://www.dolby.com/us/en/brands/dolby-atmos.html>) and Auro 3D (available online: <https://www.auro-3d.com/>) systems (Figure 6). The purpose of this scenario was to investigate the addition of the interactivity to the 3D platforms that are based on preforms and are primarily aimed at use in cinemas. Additionally, based on objective acoustic measurements, using binaural microphones, a system for evaluating and correcting spatial sound was developed. In addition, based on the evaluation system, a procedure was developed to optimize the number and position of speakers in the space.

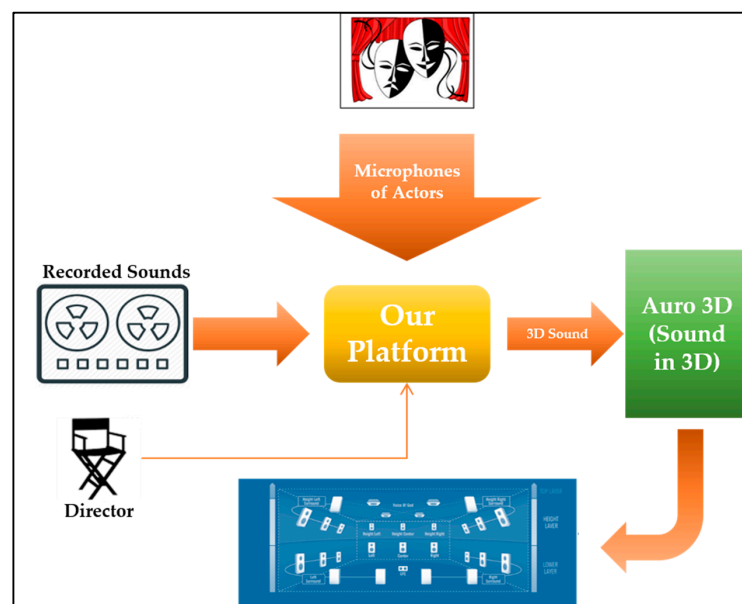


Figure 6. Use case scenario 1—spatial sound using speakers based on the Auro3D protocol.

Use Case 2:

In the second use case scenario, the audience used individual headphones through the wireless transmission process (broadcasting). In this case, the sound was not diffused in the scene but it provided spatial sound from the platform to the listeners (audience), as Figure 7 illustrates. The requirements were increased and the Production Director was able to allocate different listening points through the real and fictitious environment of the play to different audiences (for example, an ancient tragedy, some of the users watching as if they were in the chorus and some from the center of the stage). Also, the listening point could change depending on the phase of the theatrical play.

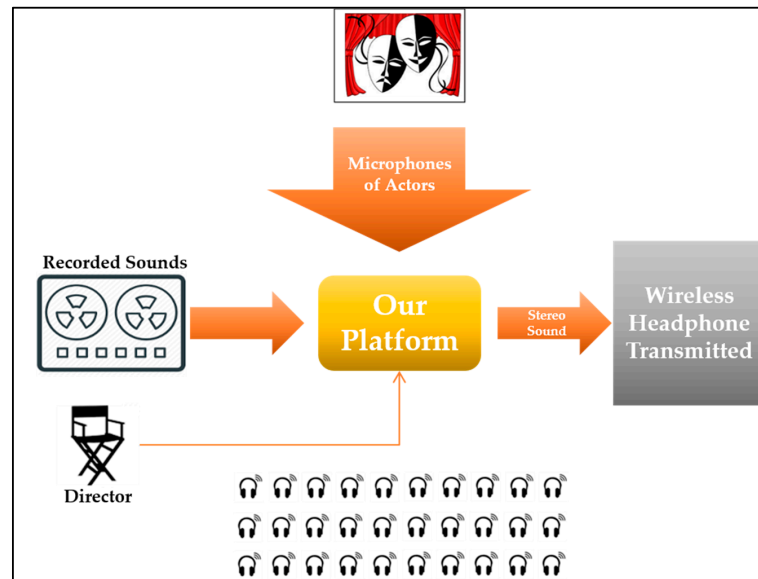


Figure 7. Use case scenario 2—spatial/personalized sound with cohearing of the audience through individual headphones.

Use Case 3:

The third use case scenario was a generalization of use case 2, in which users received a personalized and customizable sound from their headphones. In this, the transmission of the spatial sound was achieved via the internet (Figure 8). This scenario required the distributed execution of part of the platform on the personal device (tablet, smartphone, etc.) of the audience member.

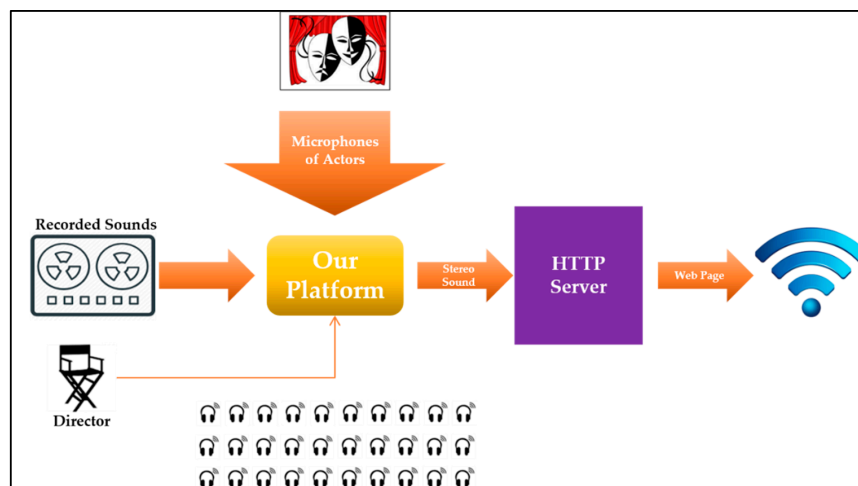


Figure 8. Use case scenario 3—spatial/personalized sound with cohearing of the audience through personal devices.

5. Conclusions

This paper describes a prototype of an innovative platform for the design and implementation of an augmented theatrical experience based on spatial sound immersion. The purpose of a fully developed platform will be to offer audiences more immersive participation in a theatrical performance. In particular, 3D scenes will be adaptable depending on the plot of the storyline portrayed. As a result, the Production Director will be able to add a wide variant of spatial sound effects in order to better express the idea of the plot and change traditional artistic processes, creating a new perception for the audience.

To support the evaluation of the prototype platform, three use case scenarios were considered. Namely, the objectives of the evaluation of the prototype platform were successfully achieved. No technical issues showed up in the three use cases scenarios and user interaction was not disturbed by system crashes. The evaluation subjects concentrated on assessing the functionality as well as the usability of the system.

While these scenarios demonstrated the potency of the prototype, the need to reduce the time delay when sound rendering in a 3D scene was identified. As with any other AR immersive system, the realism of the experience can be compromised by such delay; hence, the focus on reducing the delay represents a relevant entry point for future work which is centered on the process enhancement aspect both of the real-time sound spatialization platforms [26] and the immersive theater.

Author Contributions: E.L. and A.G.M. conceived of the presented idea. E.L., A.G.M. and K.G.P. developed the model and performed the computations. E.L. and A.G.M. wrote the manuscript with support from K.G.P. and J.A.W., A.G.M. and J.A.W. helped supervise the project. All authors provided critical feedback and helped shape the research, analysis and manuscript.

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

References

1. Azuma, R.T.; Bishop, G. Improving Static and Dynamic Registration in an Optical See-Through HMD. In Proceedings of the SIGGRAPH '94—the 21st Annual Conference on Computer Graphics and Interactive Techniques, Orlando, FL, USA, 24–29 July 1994; ACM: New York, NY, USA, 1994; pp. 197–204.
2. Mann, S.; Furness, T.; Yuan, Y.; Iorio, J.; Wang, Z. All Reality: Virtual, Augmented, Mixed (X), Mediated (X,Y), and Multimeditated Reality. *arXiv* **2018**, arXiv:1804.08386.
3. Mitchell, B. The immersive artistic experience and the exploitation of space. In *Proceedings of the 1st International Conference on Ideas before Their Time: Connecting the Past and Present in Computer Art*; BCS Learning & Development Ltd.: Swindon, UK, 2010; pp. 98–107.
4. Patrick, F. *Sound for the Stage*; Entertainment Technology Press Ltd.: Cambridge, UK, 2002.
5. Kendrick, L. *Theatre Aurality*; Palgrave Macmillan: London, UK, 2017.
6. Drama in the Dark. Available online: <https://www.cam.ac.uk/news/drama-in-the-dark> (accessed on 17 February 2019).
7. This is not A Theatre Company. Available online: <https://www.thisisnotatheatrecompany.com/about-the-company> (accessed on 17 February 2019).
8. Chaillot National Theatre Adopts 3D Sound. December 2016. Available online: <http://www.mondodr.com/chaillot-national-theater-adopts-3d-sound/> (accessed on 17 February 2019).
9. TiMax SoundHub and Tracker Selected for Tokyo's New National Theatre. Available online: https://www.prosoundweb.com/channels/av/timax_soundhub_and_tracker_selected_for_tokyos_new_national_theatre/ (accessed on 17 February 2019).
10. 'Hamilton' Is Revolutionizing the Art and Science of Broadway Sound Design. Available online: https://motherboard.vice.com/en_us/article/mg73pq/hamilton-is-revolutionizing-the-art-and-science-of-broadway-sound-design (accessed on 17 February 2019).
11. Hamid, F.; Kapralos, B. A Review of Spatial Sound for Virtual Environments and Games with Graphics Processing Units. *Open Virtual Real. J.* **2009**, *1*, 8–17. [CrossRef]

12. Yeh, H.; Mehra, R.; Ren, Z.; Antani, L.; Manocha, D.; Lin, M. Wave-ray coupling for interactive sound propagation in large complex scenes. *ACM Trans. Graph. (TOG)* **2013**, *32*, 165. [CrossRef]
13. Lakka, E.; Malamos, A.G.; Pavlakis, K.G.; Ware, J.A. Spatial Sound Rendering—A Survey. *Int. J. Interact. Multimed. Artif. Intell.* **2018**, *5*, 33–45. [CrossRef]
14. Labschütz, M. Realistic real-time audio rendering in virtual environments. *cg.tuwien.ac.at*. 2011. Available online: <https://www.cg.tuwien.ac.at/research/publications/2011/labschuetz-2011-rsr/labschuetz-2011-rsr-paper.pdf> (accessed on 17 February 2019).
15. Vorländer, M. *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms, and Acoustic Virtual Reality*; Springer: Berlin/Heidelberg, Germany, 2008.
16. Daly, L.; Brutzman, D. X3D: Extensible 3D Graphics Standard. *IEEE Signal Process. Mag.* **2007**, 130–135. [CrossRef]
17. Behr, J.; Jung, Y.; Keil, J.; Drevensek, T.; Zoellner, M.; Eschler, P.; Fellner, D. A Scalable Architecture for the HTML5/ X3D Integration Model X3DOM. In Proceedings of the Web3D '10 Proceedings of the 15th International Conference on Web 3D Technology, Los Angeles, CA, USA, 24–25 July 2010; ACM: New York, NY, USA, 2010. [CrossRef]
18. Stamoulas, A.; Lakka, E.; Malamos, A.G. Wrapping X3DOM around Web Audio API. *Int. J. Interact. Multimed. Artif. Intell.* **2015**, *3*, 36–46. [CrossRef]
19. Spatial Sound. Available online: <http://www.medialab.teicrete.gr/minipages/x3domAudio/index.html> (accessed on 17 February 2019).
20. Web Audio API. Available online: <https://www.w3.org/TR/webaudio/> (accessed on 17 February 2019).
21. Extensible 3D (X3D). Part 1: Architecture and Base Components. Sound Component. Available online: <http://www.web3d.org/documents/specifications/19775-1/V3.2/Part01/components/sound.html> (accessed on 17 February 2019).
22. Valimaki, V.J.; Parker, D.; Savioja, L.; Smith, J.O.; Abel, J.S. Fifty Years of Artificial Reverberation. *IEEE Trans. Audio Speech Lang. Process.* **2012**, *20*, 1421–1448. [CrossRef]
23. Kosmopoulos, D.; Voulodimos, A.; Doulamis, A. A System for Multicamera Task Recognition and Summarization for Structured Environments. *IEEE Trans. Indust. Inform.* **2013**, *9*, 161–171. [CrossRef]
24. Protopapadakis, E.; Doulamis, D.; Doulamis, N. Tapped delay multiclass support vector machines for industrial workflow recognition. In Proceedings of the 2013 14th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Paris, France, 3–5 July 2013; pp. 1–4.
25. Lalos, C.; Voulodimos, A.; Doulamis, A.; Varvarigou, T. Efficient tracking using a robust motion estimation technique. *Multimed. Tools Appl.* **2014**, *69*, 277–292. [CrossRef]
26. Cuevas-Rodríguez, M.; Picinali, L.; González-Toledo, D.; Garre, C.; de la Rubia-Cuestas, E.; Molina-Tanco, L.; Reyes-Lecuona, A. 3D Tune-In Toolkit: An open-source library for real-time binaural spatialization. *PLoS ONE* **2019**, *14*, e0211899. [CrossRef] [PubMed]

