BNAM 2018
Baltic-Nordic Acoustics Meeting
15-18 April 2018
Harpa, Reykjavík, Iceland

Acoustic Virtual Reality – Methods and challenges

Finnur Pind[1,2], Cheol-Ho Jeong[1], Hermes Sampedro Llopis[1], Kacper Kosikowski[1], Jakob Strømann-Andersen[2]

1. Acoustic Technology, Department of Electrical Engineering, The Technical University of Denmark

2. Henning Larsen, Vesterbrogade 76, 1620, Copenhagen, Denmark

First author's contact info:fpin@henninglarsen.com

Virtual reality is a technology that has seen increasing usage in architecture and building design in recent years. It can add value to the design process by, for example, making it easier to communicate design considerations with relevant stakeholders, such as clients, developers, engineers and architects. It also helps the designers themselves by providing a more immersive and realistic view of the modelled building and a better sense of scale. VR is also used in several other fields, such as entertainment (video games in particular), training, education and healthcare. Incorporating sound and acoustics into the virtual reality sphere adds another dimension to the experience. It both makes the immersion more believable, and in the context of building design, makes it easy and intuitive to try out different acoustic designs and soundscapes. In traditional auralization, although a very powerful tool in itself, the receiver location is usually fixed. In VR, the receiver can move around in the modeled space and switch between different designs with a click of a button, and this way get a better feeling for the acoustics of the space. In this paper, a brief overview of some of the current technologies used in acoustic virtual reality will be outlined, where the pros and cons of different approaches will be discussed. Furthermore, some examples of how the technology has been used at Henning Larsen on chosen projects will be given.

1 Introduction

Room acoustic simulations have been a topic of active research since the 1960's [1, 2]. In the early 1990's commercial room acoustics software started to be widely available [3] and today most acousticians, along with many architects and other building designers, rely on room acoustic simulations for the optimization of room design with respect to acoustics. Although the software available today is very useful in many cases, there is still a plethora of ongoing research in the field of room acoustic simulations, where the objective is to make the simulations more accurate, more flexible (e.g. in terms of geometry handling) and faster. See e.g. [4, 5, 6] for relatively recent reviews of the current status of room acoustic simulations, and ongoing challenges that are being researched.

Initially, the simulation results in room acoustic modelling where first and foremost presented in terms of objective room acoustic parameters, primarily the reverberation time but also parameters such as clarity and strength. Later, auralizations where introduced [7, 8], which involve convolving a simulated room impulse response, calculated for a given source-receiver pair, with an anechoic ("dry") recording. This way, one could hear "how the simulated room sounds" or hear how a certain sound source, e.g. a particular instrument, sounds in the room.

Recently, there is a new trend emerging, where auralizations are taken one step further. This is the usage of virtual reality (VR) to "experience" the results of the room acoustic simulation, what is referred to as acoustic virtual reality (AVR) in this paper. Here, the user is presented with both visual and auditory stimuli at the same time, e.g. by means of a head mounted display (HMD) and headphones. This approach has a number of additional benefits, compared to traditional auralizations.

The purpose of this paper is to give an overview of the technologies used in acoustic virtual reality. A number of approaches can be used in this context, and each approach has certain pros and cons. The possible benefits of using this technology within the context of building design are discussed. Finally, some initial experiments of using the AVR in practice by the architectural firm Henning Larsen are described.

2 Room acoustic simulations

Simulating the acoustical behaviour of rooms is an inherently challenging task, mainly because of the wide range of frequencies / wavelengths which are of interest, spanning three orders of magnitude. There are primarily two distinct approaches which are used when simulating room acoustics, namely, the *geometrical methods* and the *wave-based methods*. Within each of these two categories are a number of different methods. A very simplified summary of the differences between the two approaches is that the geometrical methods have been around longer, they tend to be faster and are more commonly used in practice. But their accuracy is compromising, especially in some cases. The wave based methods on the other hand are more accurate but they are significantly more computationally intensive. These two approaches are summarized here below in sections 2.1 and 2.2. The output of a room acoustic simulation is typically a room impulse response (RIR) and derived acoustic parameters such as reverberation time, clarity and strength.

The concept of auralizations is well known within the acoustics community and is widely used in acoustics consulting and various types of acoustics research. Here, the simulated impulse response of the room, for a given, fixed source-receiver pair, is convolved with an anechoic signal recording and "spatialized" for stereo or multi-channel reproduction. The purpose of the spatialization step is to emulate the differences in sound heard at each ear, which is crucial for localization in 3D space [9]. The process involves convolving the monaural signal with a head-related transfer function (HRTF). The HRTF is specific for each person and is dependent on the incidence angle of the sound reaching the receiver (i.e. dependent on the orientation of the receiver).

With auralizations, it is possible to "hear how the recording will sound in the room", commonly described as an "acoustical rendering". The anechoic recording can for example be a speech signal, in order to assess speech intelligibility in the modelled space or it can be a recording of an instrument, when assessing acoustic qualities of performance spaces. Using auralizations as a part of the building design process is advantageous for many reasons. It allows for subjective evaluation of the quality of the acoustic design, and it makes it easier to communicate different acoustical designs and acoustical considerations to non-experts, e.g. by allowing clients and architects to hear the difference between "design A" and "design B".

The quality of the auralization is of course dependent on how accurate the simulated spatial room impulse response is. This will depend on the simulation method used (discussed here below), the accuracy of boundary conditions (which can be a large source of errors [10]), the quality of the architectural computer model used, how realistically the source directivity pattern is represented and finally the quality of the spatialization process, e.g. whether personalized or general HRTF functions are used and what type of HRTF interpolation is used.

2.1 Geometrical methods

In geometrical methods, a number of simplifying approximations regarding sound propagation and reflection are made, in order to make the computational task more manageable. The main approximation is that the sound wave is approximated as a ray or particle. These rays/particles are then propagated in the room, where the propagation is dictated by the geometrical laws of optics. This approximation is only appropriate in cases where the sizes of reflecting planes and dimensions of the room are very large compared to the wavelength of the acoustic wave, i.e. usually only at high frequencies and in large rooms. Some examples of methods which fall under this category are the ray tracing method [11], the image source method [12] and the beam tracing method [13].

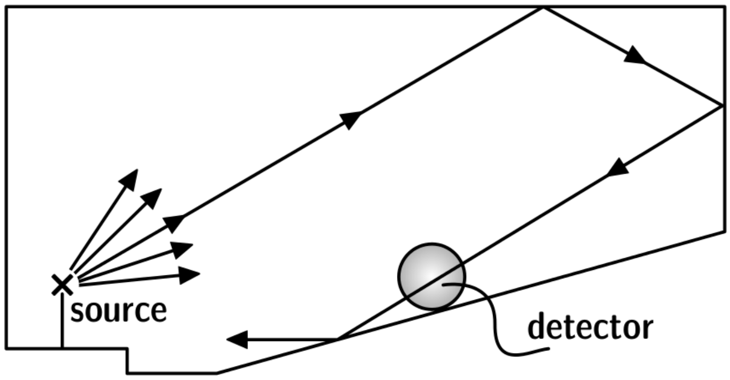

Figure 1: The principle of the classical ray tracing method (taken from [6])

By approximating a sound wave as ray which only contains energy, several wave phenomena such as diffraction, interference, phase and scattering is lost. In some rooms, the impulse response of the room will be significantly influenced by these phenomena, which renders (traditional) geometrical methods unsuitable for simulating these cases. This typically occurs in rooms where the room dimensions and sizes of objects in the room are small compared to wavelength. Therefore, it is particularly small and medium sized rooms (in the acoustical sense) and/or low frequencies which are not accurately modelled by geometrical methods. Typical classrooms, open plan offices, hospital wards, restaurants and music studios are all examples of rooms where a significant part of the frequency spectrum will be dominated by phenomena which geometrical methods do not capture accurately. In addition to this, there is evidence that even in large rooms, certain important acoustic phenomena is not captured by geometrical methods. Examples of this include the famous seat-dip effect [14], which occurs in large concert halls, and focusing due to dome-shaped or curved surfaces [15].

State of the art geometrical methods apply various types of correcting measures to try to account for diffraction, interference, phase and scattering, see examples in [16, 17, 18]. Actually, almost all contemporary commercial acoustic software includes scattering, although the assignment of scattering coefficients is usually based on crude visual inspections and guesswork. These measures can improve the accuracy to some degree, but they also come with some additional computational cost.

2.2 Wave based methods

In wave based methods, a wholly different approach is taken for simulating the acoustics. Here, the governing physics equations (typically the wave equation or the equivalent linearized Euler equations) are solved numerically. This means that no approximations regarding wave propagation and wave reflections are made, other than the numerical approximations involved in the discretization process. And since all acoustic phenomena, such as diffraction, interference, scattering and phase, is accounted for in the governing equations, these methods offer high accuracy. Methods in this category include the finite-difference time-domain method (FDTD) [19], the boundary element method (BEM) [20], the finite element method (FEM) [21], the finite volume method (FVM) [22] and the pseudospectral time-domain method (PSTD) [23]. The typical procedure usually involves subdividing the domain (the room or the room boundary) into a mesh of cells or a grid of points (see an example in figure 2), and then numerical approximations to the governing equations are solved on this mesh.

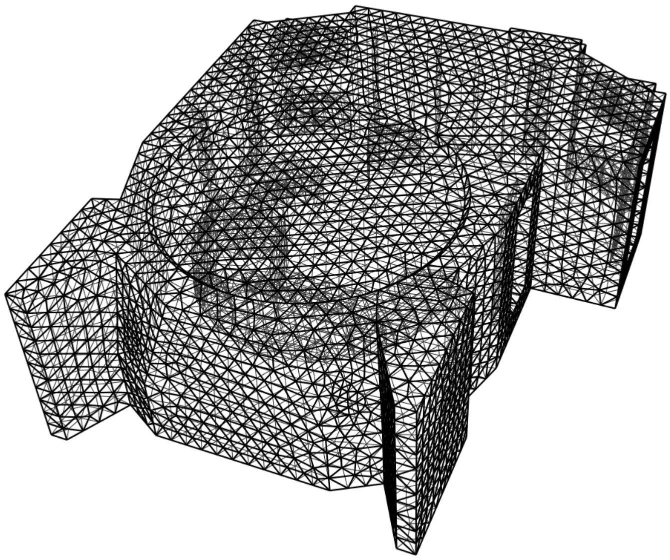

Figure 2: An example of a meshed room, suitable for e.g. FEM simulations (taken from [6])

The drawback of wave based methods, however, is that they are computationally vastly more demanding than their geometrical counterparts. Despite continuous advances in computational resources, these methods are currently still unpractical for broadband simulations over large domains.

A decent portion of current research of wave based methods is focused on bringing the computation time down. This includes implementation of algorithms on state of the art multi-core hardware, e.g. GPU's [24, 25], which can reduce computation time significantly. Another approach has been to combine wave-based and geometrical methods into "hybrid" algorithms. Here the wave based methods are usually restricted to low and medium frequencies, whereas the geometrical methods are made to cover the high frequencies [26, 27]. Recently, research into the application of so-called "high order" wave based methods, which can be significantly more efficient than typical low order wave based methods, has indicated promising results for improving the efficiency of these methods and making them more practical to use [28, 29].

3 Acoustic virtual reality

Acoustic virtual reality (AVR) builds on the foundation of auralizations, but here the main difference is that the receiver, and in some cases also the source(s), are not stationary, but can move around the modelled space [30]. This is then typically coupled with an immersive visual representation of the architectural model, e.g. by means of a head-mounted display (HMD) (see figure 3) or by using a CAVE VR environment. In some cases additional sensory stimuli such as smell or wind can even be added into the mix as well.

Figure 3: Typical AVR gear: a head-mounted display, headphones and a hand controller.

The use of AVR, compared to traditional auralizations, can add considerable additional value to the building design process. It allows for easy assessment of how the acoustics vary with position within the modelled space. Examples of how this could be used in practice include investigating how speech intelligibility changes as one moves away from the speaker or how disturbance and annoyance change within an open plan office environment. Another important feature is that it promotes holistic design practices, because it allows for assessing various design parameters concurrently. A practical example of this could be assessing how lighting, aesthetics and acoustics change when walls in a modelled space are covered with sound absorbing wall panels. The AVR setup can be programmed such that different design setups can be compared by clicking a button on the hand controller – this allows for easy A/B comparison and makes it easier to detect subtle changes in the modelled acoustic environment. Finally, a well-known fact is that the addition of sound and acoustics into the VR sphere makes the immersion more believable [31].

Since the room impulse response is dependent on the source and receiver position, in acoustic virtual reality it is necessary to continuously update the simulated room impulse response as the receiver (and potentially source as well) moves around the space. Furthermore, the spatialization process must be recalculated as the head orientation changes. These facts result in some computational challenges which are not present, or at least not nearly as time-critical, in traditional room acoustic simulations and auralizations.

There are mainly two approaches for simulating acoustic virtual reality, a *real-time calculation* approach and a *pre-calculated* approach. These are described briefly here below and their respective pros and cons are discussed in the discussion section (section 5).

3.1 Real-time calculation approach

As the name indicates, in this approach the entire room acoustic simulation and the spatialization are calculated in real-time. Whenever receiver and source position and/or orientation changes, the calculation is re-run. In order for this approach to be viable, considerable simplifications must be made, such that the computation is manageable within the strict time constraint of real-time performance. A common rule of thumb for the maximum allowed latency in real-time audio is 100 ms [32], which means that the entire process of simulation, spatialization, convolution of anechoic signal and playback must be performed within this timeframe. Making these simplifications to the simulation will naturally decrease the accuracy, and it is therefore important to make these approximations in an optimal manner (e.g. based on human perception of sound), such that one gets the most accuracy possible for minimal computational effort.

Wave-based methods are ruled out in this approach, since their computational overhead is too large for real-time performance, at least as of now. Instead, geometrical methods are applied. Some of the simplifications that are commonly made are the following. The geometry is typically simplified considerably, by making use of only a small number of faces (polygons) [33]. One of the most time-consuming tasks of a geometrical acoustics algorithm is to perform intersection tests, to determine which polygon the ray/particle hits, and by only having a few faces, then this aspect of the computation is reduced significantly. Another simplification, rooted in human perception of sound, is to update the direct sound and the early reflection portion of the impulse response frequently (25-100 Hz update rate), whereas the late reverberation tail is updated infrequently (1-5 Hz update rate) [33]. This is because the direct sound and early reflections are more important for localization and because for small changes in position/orientation, the early part of the room impulse response changes in a perceptually noticeable way. Finally, diffraction modelling, scattering and other augmentations to basic geometrical methods, done in order to improve accuracy, are typically omitted in real-time calculations.

In geometrical methods the angle of incidence of the "incoming" sound field at the receiver location is inherently known, due to the ray nature of the simulation. This makes it easy to implement the spatialization process, where the energy and phase of each incoming ray is adjusted according to the HRTF (which is angle dependent), effectively boiling down to a time domain convolution between the HRTF and the room impulse response. This brute-force *per ray* approach is fine for offline auralizations, but in real-time calculation, this approach can be too computationally intensive [34]. Again setting the focus on direct sound and early reflections, one can apply spatialization in this brute-force manner only to this early portion of the impulse response and then use amplitude panning for the late trail [34]. This will reduce computation time. However this comes at a price of reduced spatial (localization) accuracy. Another approach commonly used for decreasing the HRTF convolution computation time is to encode the HRTF in finite-order spherical harmonics. This has been shown to improve efficiency without introducing perceptual inaccuracies in localization [35]. The required order of the spherical harmonics should be as low as possible for maximum efficiency, while still maintaining the desired spatial accuracy. Typically higher order spherical harmonics are required for higher frequencies. The order can even be made to vary over the timespan of the impulse response, where higher orders are typically used in the early part of the response and lower orders in the late part. This can improve efficiency of the convolution significantly [36].

3.2 Pre-calculated approach

In this approach, the bulk of the computation is done in a "pre-calculation" stage, and only a small part of the computation is done during run-time. This way, virtually any type of simulation method can be used, although some methods might be more suitable than others. Time domain wave-based methods are particularly suitable for this approach, because of their high accuracy and because the calculation time is independent of how many receivers are used.

One way of implementing this is to calculate the room impulse response in a grid of receiver locations across the modelled space prior to run-time, typically for fixed source locations. This series of impulse responses could be calculated with a highly accurate simulation method, and therefore account for phenomena such as diffraction and phase. The creation of the monaural auralized signal for each of the simulated impulse responses on the grid would also be done prior to run-time. Then during run-time, as the user walks through the modelled space, the only calculation needed would be interpolation between the impulse responses on the grid nearest to the current receiver location at any given time and the convolution with the HRTF, in order to spatialize the sound.

Here it is necessary to store the simulated impulse responses in some way which contains spatial information, e.g. by means of spherical harmonics (ambisonics). When using wave based methods it is necessary to use plane wave decomposition to extract the spatial information of the simulated impulse response [37]. Spherical harmonics can then also be used to perform the HRTF convolution efficiently.

However, a drawback with this implementation is that it becomes somewhat impractical for dynamic (moving) sources. If these types of sources are allowed, then the amount of stored impulse responses becomes too large in terms of memory requirement. Another issue with this implementation which requires more research, is the question of how fine the grid needs to be for sufficient accuracy.

Another pre-calculated approach based on the equivalent source method (which is yet another wave based method) was developed by Mehra et al. [38]. Their approach can account for dynamic sources and receivers and has the high accuracy associated with wave based methods.

Finally, the image source method is also very suitable for pre-calculated AVR, when using a static source and dynamic receivers. The determination of image sources can be done in the pre-calculation step and then during run-time the only calculation necessary is to perform a visibility check of the image sources.

4 Examples of usage of acoustic VR

Some initial experiments of using AVR have taken place at Henning Larsen. They are described briefly here below.

4.1 Carl H. Lindner College of Business – classroom acoustics

Henning Larsen designed the Carl H. Lindner College of Business, a new building which is part of the University of Cincinnati campus. Henning Larsen's responsibilities included architectural design, lighting design and acoustical design. The building is 22.500 m^2 and construction is currently ongoing.

As a part of the interior design of the classrooms in the building, a virtual reality mock-up of a typical classroom was set up. There were two acoustical designs which were being discussed with the client, a cheaper but less optimal solution which involved the use of ceiling absorbers only (resulting in a reverberation time of around 1 sec) and a more expensive but optimal design which involved ceiling and wall absorbers (resulting in a reverberation time of around 0.6 sec).

In the acoustic virtual reality mock-up there were four versions of the same classroom set up side by side, the only difference between the four versions being the acoustical treatment. The user could walk between the different classrooms in the AVR mock-up and experience the different acoustics. In the first classroom there was no acoustic treatment whatsoever (resulting in a reverberation time of roughly 2.5 sec), mainly to give a frame of reference. In the second classroom was the cheaper-but-less-optimal design. The third classroom was the optimal design and finally the fourth classroom had *too much* acoustic treatment (reverberation time of roughly 0.4 sec), again mainly to give some frame of reference. Figure 4 shows the building in question and a screen capture from the AVR mock-up.

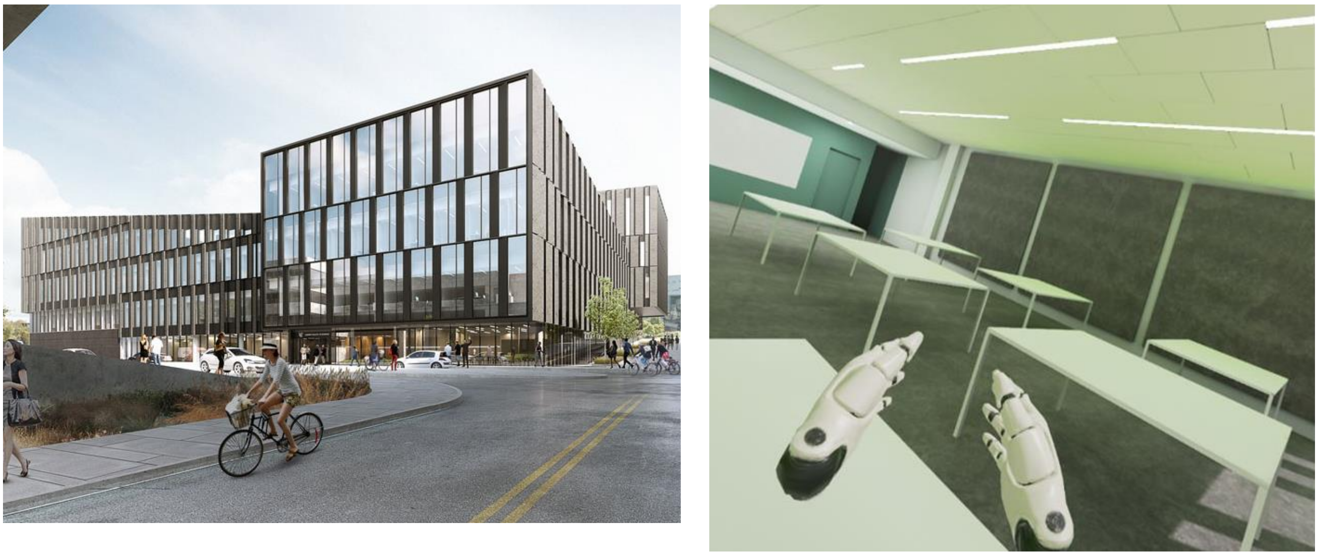

Figure 4: Architectural rendering of the building (left) and screen capture from the AVR mock-up (right).

The acoustic VR was created in a relatively naïve, pre-calculated manner. A room impulse response for a fixed source-receiver pair was simulated in Odeon, where the source was located in the front of the room, where the teacher would typically stand, and the receiver was in the middle of the seating area. The monaural convolved signal was also created in Odeon, using an anechoic speech signal. The virtual reality model was set up using the Unreal game engine. This game engine has some audio features, including simple free-field spatialization based on generic HRTFs and free-field level adjustment based on source-receiver distance. This was used to spatialize the sound.

This AVR mock-up proved to be valuable in the discussions with the client, and within the architectural design team. Previous discussions, which centered around objective room acoustic parameters such as reverberation time weren't very fruitful, whereas once everyone involved had tried the AVR, there was a good understanding of the importance of good acoustics in the classrooms and what it actually means to go from roughly 1 sec reverberation time to 0.6 sec reverberation time. Furthermore, it was clear from the mock-up what aesthetic influences the added room acoustic treatment would have.

It should of course be mentioned that this approach of setting up the AVR is not very accurate, since only one impulse response is made to represent the entire room, albeit adjusted in level based on source-receiver distance and spatialized based on free field spatialization. This means for example that the direct-to-reverberant ratio will be the same wherever the user stands in the room, which is of course highly inaccurate. Nevertheless, this naïve approach proved useful in this context. In a more complex room and/or with more complex sound sources this approach would probably have been too unrealistic to be useful.

4.2 AVR using STEAM Audio engine

Recently some experiments using the STEAM audio engine in conjunction with the Unity game engine have also been undertaken at Henning Larsen. The STEAM audio engine takes a real-time approach, based on the ray tracing method. It is designed for creating realistic video game audio. The user can specify several parameters, such as the order of the spherical harmonics of the spatial impulse response (higher orders improve localization but can lead to audible latency and glitches), the number of rays used and reflection order, sound absorption coefficients of surfaces at low, mid and high frequencies (800 Hz, 4000 Hz and 15000 Hz respectively). Further information about the STEAM audio engine can be found here [39].

Although the engine has not been used on actual projects at Henning Larsen yet, some informal tests have been carried out. They indicate that the engine can be useful for creating decently accurate acoustic virtual reality, although for more complex spaces the realism of the simulation becomes questionable, e.g. due to the lack of diffraction modelling. Also when the accuracy is turned up, e.g. by increasing the amount of rays used, reflection order, spherical harmonics order etc. the simulation becomes laggy and glitchy, at least on the "normal-but-powerful" desktop computer used in these informal tests. It should also be mentioned that the STEAM engine is in rapid development, and will without a doubt improve in accuracy with future releases. Furthermore, not all features of the audio engine have been thoroughly investigated at this stage.

4.3 Current work

Currently there is ongoing work in using a grid based pre-calculated approach, where the impulse responses on the grid are encoded using spherical harmonics (to store spatial information of the impulse response). This work takes inspiration from [40] – which uses a similar approach but only for a static receiver location (but allowing for head movement).

The model under development uses Odeon, Unity and Oculus Rift VR gear. The physical space is discretized in a grid of points. The RIR of each point is computed in Odeon and exported in ambisonics (B-Format). Ambisonics is used here because it encodes the directional information of a specific three-dimensional sound field in either four channels (first order) or nine channels (second order). This approach allows for keeping some information of the directionality of the rays, producing a more realistic reproduction of the sound field.

For all points on the grid, each channel is convolved with an anechoic sound. Then, the convolved channels are decoded to feed a specific virtual loudspeaker array inside Unity which reproduces the given sound field. The way in which both visuals and acoustics are linked is as follows: once the model is imported into Unity, the player is always followed by an array of virtual loudspeakers which reproduces the decoded ambisonics convolved sounds of the nearest grid point. When the player moves his head, the sound field will not be changed, but the subject will perceive the change of directionality due to the HRTF that Unity and Oculus Rift applies. Interpolation between grid points is used to make the transition between grid points smoother.

The required number of files in this approach could be extremely high, if the modelled space is large and the grid is fine. In order to optimize the computational load, a C# script was developed which manages the instantiation and destruction of the audio files sources. Further work could consist of sending the position and head tracking information from Unity to an external processor in order to achieve real-time processing. This way, the pre-processing step (not to be confused with the pre-calculation of the RIR) and the storage of all the audio files can be avoided.

5 Discussion

The use of acoustic virtual reality will likely replace traditional auralizations in the coming years, due to the added benefits described in section 3.

The different approaches described above for simulating AVR have certain pros and cons. The real-time approach has lower accuracy, due to the strict latency threshold. The accuracy is nevertheless most likely acceptable in some cases, e.g. in large rooms and in early design stages. And with ever increasing computational power, more and more calculations can be done within the constraints of real-time, which will result in increased accuracy. The main benefit of the real-time approach however, is the flexibility that comes with it. Moving sources are handled with ease and geometry and materials can be changed on the fly. This can be very useful during the design process, especially in early design stages where various vastly different designs are usually being considered.

The pre-calculated approach has more or less the opposite pros and cons. The accuracy is much greater, which is probably necessary in most building design cases, such as in small and medium sized spaces and when realistic immersion is important. The drawback is that every time geometry and materials are changed, the pre-calculation needs to be run again. If a highly accurate simulation method, e.g. a wave based method, is used for this step, then the computation time of the pre-calculation can be expected to be long.

Acoustic virtual reality for building design purposes is tightly linked with video game audio. The objective is roughly the same, to have realistic virtual acoustics and to improve the immersion into the VR sphere / game. Recently some computer games have started to use simulation based approaches akin to those described in this paper to create realistic audio [41]. However, in video game audio there is significantly more “artistic freedom”, because here the objective is more focused on creating the most impressive audio experience, whereas in building design the objective is simply to come as close to reality as possible (although the two goals are often related!).

In addition to room acoustics, it would be beneficial to include building acoustic simulations into the AVR as well, mainly sound isolation simulations. See a discussion on this topic in [33].

A final thought worth mentioning is that of the concept of soundscapes. In order to create realistic acoustic virtual reality or realistic virtual soundscapes of complex acoustical scenarios, it is important to not only have accurate simulation techniques, but also to have the appropriate anechoic dry recordings, along with information on sound source power and directivities. Consider e.g. a complex acoustic scene such as an open plan office. If one wants to set up an AVR model of an open plan office, it is necessary to insert into the model all the ambient sounds that make up the soundscape of the open plan office. This includes keyboard clicking, printers, phones ringing, noise from technical installations, drum noise from walking, people talking, ambient noise from outside and so on and so forth. It becomes quite a task to collect, calibrate and model all these different sounds at once – much more complicated than e.g. simulating a single talker in an auditorium or arguably even a group of instruments in a performance space.

6 Conclusion

This paper summarizes the main trends and methods for creating acoustic virtual reality, i.e. augmenting virtual reality representations of building models with sound and acoustics. There are many approaches that can be used, and they can roughly be divided into two main categories, namely the real-time calculation approach and the pre-calculated approach. There are different approaches within these categories, but perhaps the main tradeoff between the two approaches is that of accuracy versus flexibility. In the pre-calculated approaches high accuracy can be obtained, e.g. by means of using wave-based methods. However these methods are restricted to static scenes and in some cases to fixed sources as well. The real-time approach has a rather severe latency threshold, which limits the accuracy considerably, but instead it can easily handle moving sources, and changing geometries and materials, which can be valuable in some cases. Some experiments with AVR are described and they indicate that AVR can indeed add value to the building design process.

References

[1] M. R. Schroeder and K. H. Kuttruff, On frequency response curves in rooms. Comparison of experimental, theoretical, and Monte Carlo results for the average frequency spacing between maxima, *The Journal of the Acoustical Society of America*, 34(1):76-80, 1962.

[2] A. Krokstad, S. Strom, and S. Soersdal, Calculating the acoustical room response by the use of a ray tracing technique, *Journal of Sound and Vibration*, 8(1):118-125, 1968.

[3] G.M. Naylor, ODEON - Another hybrid room acoustical model, *Applied Acoustics*, 38(2-4):131-143, 1993.

[4] M. Vorlander, Computer simulations in room acoustics: Concepts and uncertainties, *The Journal of the Acoustical Society of America*, 133(3):1203-1213, 2013.

[5] L. Savioja and U. P. Svensson, Overview of geometrical room acoustic modeling techniques, *The Journal of the Acoustical Society of America*, 138(2):708-730, 2015.

[6] M. Vorlander, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, Springer, 2008.

[7] M. Kleiner, B.-I. Dalenbäck, P. Svensson, Auralization – An Overview, *Journal of the Audio Engineering Society*, 41(11), 861-875, 1993.

[8] J. H. Rindel and C. L. Christensen, Room acoustic simulation and auralization - how close can we get to the real room?, *Proceedings of Eighth Western Pacific Acoustics Conference*, 2003.

[9] H. Fastl and E. Zwicker. *Psycho-Acoustics*. 3rd ed., Springer, 2007.

[10] C.-H. Jeong, G. Marbjerg, J. Brunskog, Uncertainty of input data for room acoustic simulations, *Proceedings of Bi-annual Baltic-nordic Acoustic Meeting*, 2016.

[11] A. Kulowski, Algorithmic representation of the ray tracing technique, *Applied Acoustics*, 18(6):449-469, 1985.

[12] H. Lee and B.-H. Lee, An efficient algorithm for the image model technique, *Applied Acoustics*, 24(2):87-115, 1988.

[13] S. Laine, S. Siltanen, T. Lokki, and L. Savioja, Accelerated beam tracing algorithm, *Applied Acoustics*, 70(1):172-181, 2009.

[14] J. LoVetri, D. Mardare, and G. Soulodre, Modeling of the seat dip effect using the finite-difference time-domain method, *The Journal of the Acoustical Society of America*, 100(4):2204-2212, 1996.

[15] M. L. S. Vercammen, Sound concentration caused by curved surfaces, PhD thesis, Eindhoven University of Technology, The Netherlands, 2011.

[16] G. Marbjerg, J. Brunskog, C.-H. Jeong, and E. Nilsson, Development and validation of a combined phased acoustical radiosity and image source model for predicting sound fields in rooms, *The Journal of the Acoustical Society of America*, 138(3):1457-1468, 2015.

[17] C.-H. Jeong, J.-G. Ih, and J. H. Rindel, An approximate treatment of reection coefficient in the phased beam tracing method for the simulation of enclosed sound fields at medium frequencies, *Applied Acoustics*, 69(7):601-613, 2008.

[18] J.H. Rindel, G.B. Nielsen, and C.L. Christensen, Diffraction around corners and over wide barriers in room acoustic simulations, *Proceedings of the 16th International Congress on Sound Vibration*, 2012.

[19] D. Botteldooren, Finite-difference time-domain simulation of low-frequency room acoustic problems, *The Journal of the Acoustical Society of America*, 98(6):3302-3308, 1995.

[20] J. A. Hargreaves and T. J. Cox, A transient boundary element method model of Schroeder diffuser scattering using well mouth impedance, *The Journal of the Acoustical Society of America*, 124(5):2942-2951, 2008.

[21] W. Ahnert, M. Bansal, and S. Feistel, Large scale FEM analysis of a studio room, *Audio Engineering Society Convention 120*, 2006.

[22] S. Bilbao. Modeling of complex geometries and boundary conditions in finite difference/finite volume time domain room acoustics simulation, *IEEE Transactions on Audio, Speech, and Language Processing*, 21(7):1524-1533, 2013.

[23] M. Hornikx, T. Krijnen, and L. van Harten, OpenPSTD: The open source pseudospectral time-domain method for acoustic propagation, *Computer Physics Communications*, 203:298-308, 2016.

[24] L. Savioja, Real-time 3D finite-difference time-domain simulation of low-and mid-frequency room acoustics, *13th International Conference on Digital Audio Effects*, 2010.

[25] J. J. Lopez, J. M. Navarro, D. Carnicero, and J. Escolano, Some comments about graphic processing unit architectures applied to finite-difference time-domain room acoustics simulation: Present and future trends, *The Journal of the Acoustical Society of America*, 133(5):3614-3614, 2013.

[26] M. Aretz and M. Vorlander, Combined wave and ray based room acoustic simulations of audio systems in car passenger compartments, part I: Boundary and source data, *Applied Acoustics*, 76:82- 99, 2014.

[27] A. Southern, S. Siltanen, D. T. Murphy, and L. Savioja, Room impulse response synthesis and validation using a hybrid acoustic model. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(9):1940-1952, 2013.

[28] J. van Mourik and D. Murphy, Explicit higher-order FDTD schemes for 3D room acoustic simulation, *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(12):2003-2011, Dec 2014.

[29] B. Hamilton and S. Bilbao, FDTD methods for 3D room acoustics simulation with high-order accuracy in space and time. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(11):2112-173, 2017.

[30] L. Savioja, J. Huopaniemi, T. Lokki, R. Väänänen, Creating interactive virtual acoustic environments, *Journal of the Audio Engineering Society*, 47(9), 675-705, 1999.

[31] M.S. Gazzaniga, R. B. Ivry, and G. R. Mangun. *Cognitive Neuroscience: The Biology of the Mind*. W. W. Norton, 1998.

[32] J. Sandvad, Dynamic aspects of auditory virtual environments, *Audio Engineering Society Convention 100*, 1996.

[33] M. Vorlander, D. Schroder, S. Pelzer and F. Wefers, Virtual reality for architectural acoustics, *Journal of Building Performance Simulation*, 8:1, 15-25, 2015.

[34] C. Schissler, P. Stirling and R. Mehra, Efficient construction of the spatial room impulse response, *Proceedings of the IEEE Virtual Reality Annual International Symposium*, 122-130, 2017.

[35] G. D. Romigh, D. S. Brungart, R. M. Stern and B. D. Simpson, Efficient Real Spherical Harmonic Representation of Head-Related Transfer Functions, *IEEE Journal of Selected Topics in Signal Processing*, 9(5), 921-930, 2015.

[36] S.E. Favrot and J. Buchholz, LoRA: A Loudspeaker-Based Room Auralization System, *Acustica United with Acta Acustica*, 96(2), 364-375, 2010.

[37] R. Mehra, L. Antani, S. Kim and D. Manocha, Source and Listener Directivity for Interactive Wave-Based Sound Propagation, *IEEE Transactions on Visualization and Computer Graphics*, 20(4), 495-503, 2014.

[38] R. Mehra, A. Rungta, A. Golas, M. Lin and D. Manocha, WAVE: Interactive Wave-based Sound Propagation for Virtual Environments, *IEEE Transactions on Visualization and Computer Graphics*, 21(4), 434-442, 2015.

[39] https://valvesoftware.github.io/steam-audio/

[40] E. Ballestero, P. Robinson, S. Dance, Head-tracked auralisations for a dynamic audio experience in virtual reality sceneries, *24th International Congress on Sound and Vibration*, 2017.

[41] N. Raghuvanshi, A. Allen, and J. Snyder, Numerical wave simulation for interactive audio-visual applications, *The Journal of the Acoustical Society of America*, 139(4):2008-2009, 2016.

[42] M. Kleiner, B.-I. Dalenbäck, P. Svensson, Auralization – An overview, *Journal of the Audio Engineering Society*, 41(11), 861-875, 1993.