A REAL-TIME AURALIZATION PLUGIN FOR ARCHITECTURAL DESIGN AND EDUCATION

Lukas Aspöck, Sönke Pelzer, Frank Wefers, Michael Vorländer

Institute of Technical Acoustics,
RWTH Aachen University,
Aachen, Germany
{las,spe,fwe,mvo}@akustik.rwth-aachen.de

ABSTRACT

The role of acoustics in architectural planning processes is often neglected if the designer lacks necessary experience in acoustics. Even if an acoustic consultant is involved he might be presented with limited options after the initial planning process. Some disadvantageous decisions might be hard to reverse then. To improve and facilitate the construction process permanent immediate feedback should be given to the designer. Planning cannot be imaged today without live 3D visual rendering. But also acoustics should be rendered in real-time to provide the same type of intuitive feedback. Therefore a real-time room acoustics auralization was implemented into a popular CAD-Modeling tool. Binaural room impulse responses are continuously updated using image sources and ray tracing algorithms and convolved in real-time with audio feed from recorded sounds or the user's microphone. The CAD model can be freely modified during the simulations including geometry, surface materials and source and receiver positions. Using streaming low-latency convolution, an immediate feedback is provided to the user.

1. INTRODUCTION

Nowadays the most important tool during the design process of an architect is the CAD-editor. It enables the architect not only to use it as a planning tool, but it also offers the possibility to experience his ideas in three dimensions without erecting the physical structure. Especially when it comes to buildings and rooms which are used for the presentation of acoustical signals, such as concert halls or conference rooms, it becomes important to also include the acoustical characteristics in the design process. For the visual appearance, it is possible to use rendering tools producing a photorealistic image of the model. To achieve the same level of quality for the audio feedback, an auralization based on room acoustics simulations is required. Room acoustic simulation algorithms are usually based on the assumption that a sound wave is interpreted as a ray which behaves in a similar way as a light ray (Geometrical Acoustics). Often a hybrid model is used to simulate a room impulse response, combining the image sources methods[1] and the ray tracing algorithm [2].

Based on the geometry as well as on surface properties, acoustic simulation software (e. g. CATT-acoustic [3]) is able to provide realistic auralizations of the room as well as a reliable prediction of the room acoustic parameters. The same simulation models are also integrated in immersive Virtual Reality systems[4], allowing the user to interact intuitively in a virtual environment while experiencing multimodal feedback. Similar room acoustics simulation techniques are expected to be applied in the entertainment industry soon, e. g. for real-time sound rendering in computer games [5]. These models however focus not on providing the realistic acoustical reproduction of the environment, but on rendering plausible audio feedback with reduced computational effort, without interfering the high computational demands of the visual rendering.

However, all of these listed applications of room acoustic simulations do not represent a suitable easy-to-use tool, which the architect can use to analyse and experience the room acoustics in his building. The available room acoustic tools are often expert tools, requiring extensive knowledge about room acoustics and the simulation techniques. Virtual Reality systems are very expensive and also do not represent a convenient tool due to the system's complexity. By integrating an auralization software into the popular 3D modeling software *SketchUp*, we are able to provide an easy to operate tool, which helps to understand the effects of room acoustics as well as enriches the architectural design process.

2. STATE OF THE ART

Extensive research in the areas of room acoustics, binaural hearing and spatial reproduction made it possible to develop algorithms to simulate [6] and auralize [7][8] virtual sound fields. Because of the steady increase of computational power, today's room acoustic simulation software (e. g. ODEON [6], CATT or EASE) is able to calculate accurate results by generating room impulse responses in a reasonable amount of time, including effects such as absorption, scattering and diffraction[9]. In various surveys, the accurarcy of different room acoustic simulation software has been validated and compared to measurements[10][11][12]. These tools help the user to predict and understand the room acoustics in the room of interest. A binaural synthesis based on measured Head-Related-Transfer-Functions[13] offers the possibility to listen to virtual sources in the room. These auralizations are however often only possible for a static receiver-source situation in the room or, in case of dynamic situations, depend on precalculated and interpolated simulation results [14] or on the usage of computer clusters[15]. Although editing options for CAD models are given in most simulation tools, the possiblities are usually very limited and an external dedicated 3D design tool is much more powerful and often preferred. To eliminate this lack of interactivity with the model, this work presents the implementation of a real-time auralization software, directly integrated in the *SketchUp* modeling software and providing binaural feedback through headphones to the user.

3. REAL-TIME AURALIZATION IN SKETCHUP

The real-time auralization tool can be seperated in three parts: the *SketchUp* plug-in, the simulation client(s) and the real-time convolution module. The concepts and aspects of their implementation of these software modules are presented in the next sections. Fig. 1 gives an overview of the system. The plug-in (Section 3.1) not only includes the graphical user interface, but also acts as a server for the scene data. It contains a network interface which provides the relevant data for the room acoustic simulation to one or more simulation clients (Section 3.2). The results of the simulation (e. g. a binaural room impulse response) are sent to the convolution module (Section 3.3) and played back to the user.

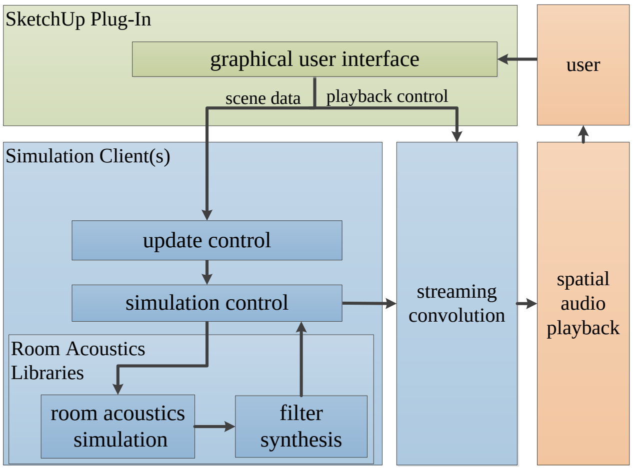

Figure 1: *System components of the auralization system*

3.1. *SketchUp* plug-in

SketchUp[1] is a popular modeling tool, which can easily be learned by beginners and efficiently be operated by expert users. Its functionality can be extended by adding plug-ins using the scripting language *Ruby*. In order to run a room acoustic simulation, the basic functionality of *SketchUp* (setting up a room and assigning wall materials) had to be extended. The implemented plug-in realizes the selection and positioning of sound sources and receivers and automatically sends all acoustically relevant data to the simulation client:

- Room polygons
- Assigned wall materials
- Receiver characteristics
- Source characteristics
- Auralization configuration

A data update is only sent if the situation was changed by the user. Different update rates are used, while source and receiver positions are updated at a high rate (~100 Hz), geometry changes are sent at lower rates (~10 Hz).

To properly modify the situation, the graphical user interface of *SketchUp* had to be extended with special functionality required for the auralization. These extensions are presented in the following subsections.

[1] www.sketchup.com

Sources and Receivers

Besides the room model, for an acoustic transfer path at least one source and one receiver must be defined. Therefore the plug-in introduces additional buttons on the program surface to place acoustic sources and receivers. The objects also contain an adequate visual model (see Fig. 2) and can freely be moved with the *move*-function of *SketchUp*. By right-clicking on the object it is possible to assign a source directivity, an anechoic sound file or soundcard input channel for the playback (sources), and HRTF data (receivers). If multiple receivers are added to the scene, the user has to pick one receiver as the *active* receiver. The sound sources of the scene can be muted and unmuted also by right-clicking on them. To run an auralization, the scene has to contain at least one source. However, adding one *dummy-head* receiver is not essential, because one receiver object is always represented by the position and orientation of the camera viewing the scene. By pressing the button in the toolbar, the user is able to switch between the *camera*- and *dummy-head-receiver* object.

Geometry and Materials

For the modification of the geometry and the wall materials, no extensions were required. Accurate results of the room acoustic simulations can however only be provided if the room does not contain any holes and all materials are assigned on the inside of the room. If the acoustical properties (absorption and scattering) should be considered by the room acoustic simulation, a database file with the same name of the assigned *SketchUp* material has to be included in the material database of the simulation. By pressing a button in the toolbar, the user can visualize the acoustical properties of the selected material (absorption and scattering coefficient for third-octave bands).

Control Panels

The control panel, displayed on the right side in Fig. 2, includes the most important settings of the room acoustic simulation (e. g. Image Source order, Ray Tracing particles) and of the auralization. The simulation components can be switched on and off. Buttons for playback and volume control are included in the toolbar. Another panel, which is currently being developed, enables the user to distribute the workload of the acoustical simulations to multiple simulation clients.

3.2. Simulation Client

Room Acoustic Simulation

The simulation is based on the software library *RAVEN* [16][17], developed at the Institute of Technical Acoustics, RWTH Aachen. The interfaces of the *RAVEN* module include functions to define the scene and run simulations in various configurations. Results of the *RAVEN* simulation models have been validated for various situations, e. g. in [18].

RAVEN defined state-of-the-art algorithms and includes hybrid acoustic simulation models to generate single components of a room impulse response. Fig. 3 shows an energetic room impulse response including the direct sound, the early reflections (calculated by the image sources method) and the reverberation (generated either by a ray tracing algortihm or a statistical artificial reverberation model). The artificial reverberation was integrated

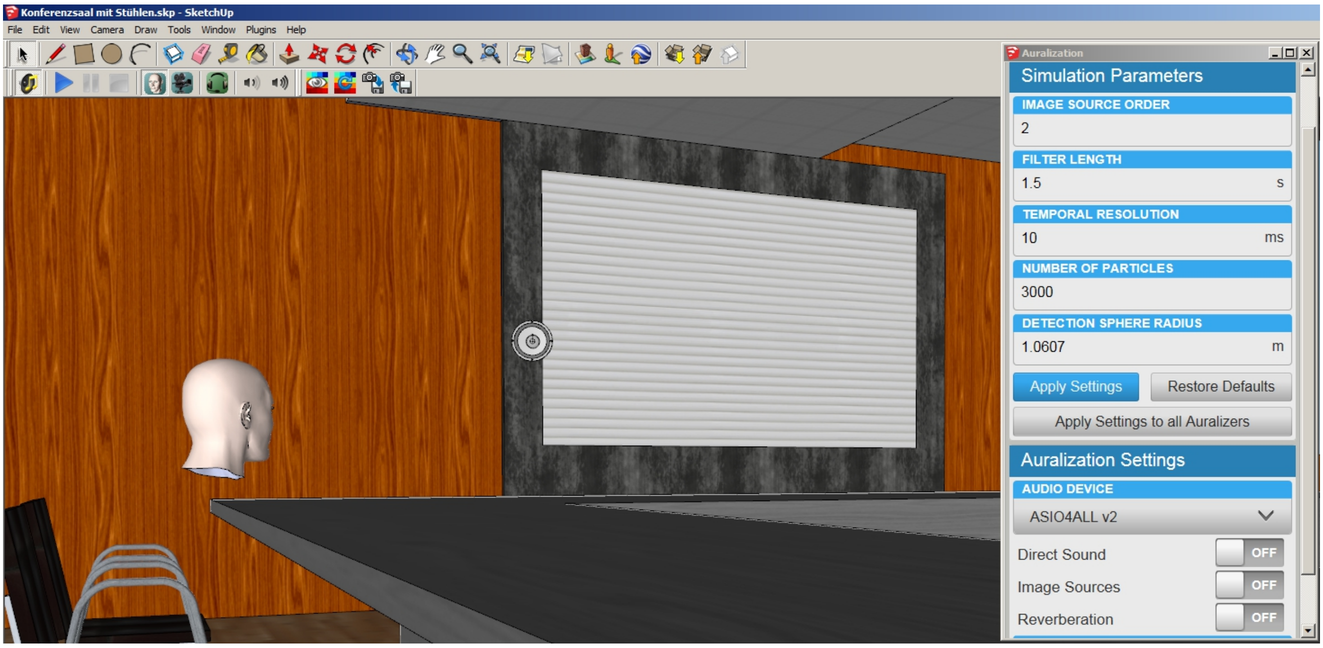

Figure 2: *Graphical user interface of the auralization tool integrated in the modeling software SketchUp*

to demonstrate the difference between and the two reverberation models. As the generation of an artificial reverberation requires significantly less computations in comparison to the ray tracing, it can also be used to provide plausible reverberation feedback without violating the real-time condition in case of only low computational capacities being available. The seperation of the simulation components is reasonable in terms of psychoacoustics, algorithms and data structures [17, 19]. It also offers the possibility to seperately auralize each part of the impulse response.

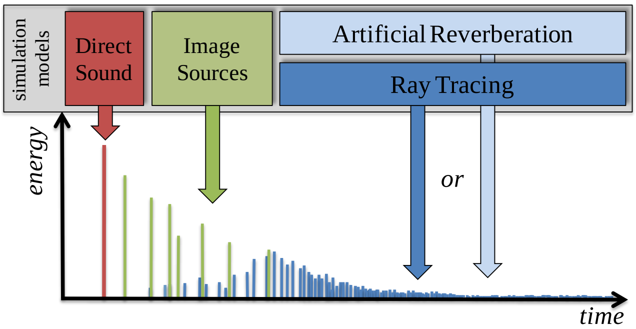

Figure 3: *Simulation models and their contribution to the energetic impulse response*

Simulation Objects

For each active source-receiver combination of the scene, one simulation object is created. For a simulation job, a multi-threading approach is applied, which generates parallel simulation threads according to the configuration of the object (see Table 1). Each simulation component has its own thread, running at different priority levels and update rates. Once one simulation task is finished, the contribution of the simulation result is exchanged in the total room impulse response. Adjusted to the psychoacoustical demands of a highly interative situation (e. g. quick receiver movement in the *camera* mode), the calculation of the direct sound audibility and filter response is executed at a higher priority level than the image sources and the reverberation threads.

Table 1: *Example of a simulation object, generating a binaural room impulse response containing direct sound and reverberation based on a ray tracing*

sourceID	1
receiverID	3
DirectSound	*enabled*
EarlyReflections	*disabled*
ReverberationMode	*Ray Tracing*
FilterMode	*Binaural*

The resulting filter parts of the simulations are cached. If the user enables one simulation component after it has been disabled, it is immediately available. This also applies for switching back and forth between different receiver positions - the room impulse response for the source to the reactivated receiver can be quickly loaded from the working memory without repeating a simulation. This keeps the latency for direct comparisons at a minimum level.

Update handling

A scene modification by the user can affect the auralization in different ways. While a small movement of the source might modify the situation of the direct sound and the early reflections significantly, the perceived reverberation of the room hardly changes.

To classify the modifications made by the user, a class was implemented to analyse the current modification. Fig. 4 shows all possible modification types. *HRTF* and *Directivity* occur if the users selects a different directional characteristic for the receiver or the source in the scene.

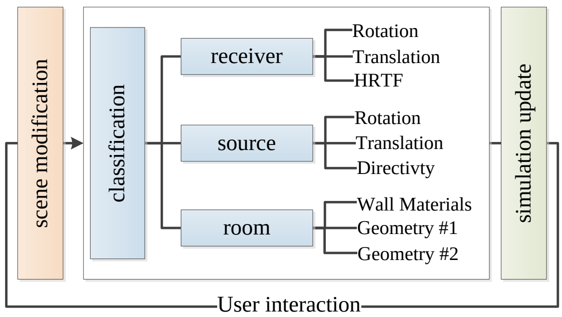

Figure 4: *Classification of a scene modification*

A room modification however represents the most defacing modification. While for an insignificant geometry modification (*Geometry #1*, e. g. dividing a plane of the room without any acoustical effect for the simulation), no update is required and a material exchange only leads to an update of the ray tracing simulation and image source filter, a normal room modification (*Geometry #2*, e. g. shifting a wall or adding objects to the room) necessitates an update of the spatial data structures as well. Spatial subdivision is usually applied to reduce the high number of intersection tests which occur during Geometrical Acoustic simulations [20][21]. For static room geometries with typical polygon counts, the usage of a Binary Space Partioning tree [22] leads to the lowest computation times. For modifiable room geometries however, it has been found that other spatial data structures, e. g. the Spatial Hashing (SH) technique, should be prefered because the update of a spatial hashmap is faster in comparison to the update of the BSP tree[23]. In the here presented auralization tool, both of these spatial data structures are used, according to the situation of the scene. In general, the BSP tree is used to reduce the number of necessary intersection tests within the direct sound, image source and ray tracing calculations. In case of a geometry modification, the calculations of the early parts of the room impulse responses are based on the SH structure, reducing the latency of these calculations significantly.

3.3. Convolution and Streaming

The real-time convolution and streaming module of the auralization tool can be used as a stand-alone application, which can be controlled via a network interface. During the standard application of the tool however, it runs on the same computer as the *SketchUp* software, providing the acoustic feedback through headphones directly to the user. The module supports the convolution and playback of multiple stereo channels, which allows to auralize multiple sound sources at once. The convolution is based on a Fast Fourier Transformation (FFT) convolution algorithm, performing an efficient block-based *Overlap-Save* convolution [24]. Subdividing of the room impulse responses in multiple parts of equivalent length realizes a quick convolution with low latency times. Cross fading is applied to avoid audible artifacts caused by the exchange of filters during the convolution. The headphone playback of the binaural signals is realized using low-latency *ASIO* sound card drivers. Crosstalk cancellation filters for the playback of binaural signals through loudspeakers [25] are currently being integrated.

3.4. Performance Analysis

The performance of the auralization software not only depends on the used hardware, but also on the configuration of the simulation as well as on the complexity of the auralized scene. Instead of an extensive performance analysis for various situations, the calculation times of only one typical example is discussed here. This situation is characterized by the parameters shown in Table 2.

Table 2: *Description of the example scenario*

Room model	*Concertgebouw* Amsterdam
Number of polygons	505
Room volume	20786.3 m^3
Image Source Order	1
Ray Tracing particles	1500 (per octave band)
Radius dection sphere	1 m
Length of Filter	2 s
Reverberation time	2.7 s

The analysed scenario contains one source and receiver, both located inside the concert hall. According to the described parameters and the equations given in [19], the standard deviation of the energy envelope of the late decay is less than $\sigma_L = 0.8\,dB$.

Fig. 5 visualizes the timeline of the most relevant steps in case of a geometry modification. The calculations were carried out on a common desktop computer (*Intel Core i7* @ 3.40 GHz processor, 8 GB RAM).

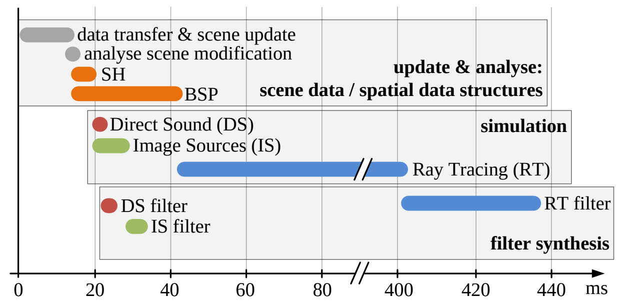

Figure 5: *Timeline of a simulation update after a geometry modification for the given example situation*

As a reaction on the user modifying the geometry, the scene data is updated and the modification is analysed by the simulation client. Both spatial data structures are recalculated in different threads by the geometry controller. The calculation of the Spatial Hashing update takes approximately 3 ms, while it takes 25 ms to update the BSP tree. As soon as the SH update is finished, the simulations of the direct sound and the image sources are executed. Including the binaural filter synthesis, the updated results are sent to the convolution engine after a total processing time of 24 ms (DS) and 34 ms (IS). The simulation thread of the ray tracing algorithm is started as soon as the BSP update has finished. The ray tracing calculation is *waiting* for the updated BSP-tree because the overall calculation time of a ray tracing based on SH would be significantly higher (700 ms). In a future version, the spatial data structure will be implemented transparently and the ray tracer will already start working on the SH table until the BSP is available. As the binaural filter synthesis of the reverberation is based on an highly accurate approach, which includes spatial information (HRTF data) for each reflection of the room impulse response, the calculation times are relatively high. Different methods for the synthesis of a perceptually equivalent late part of the impulse response are currently investigated [26] and will be implemented. The convolution and reproduction of the binaural signal adds another 5 to 10 ms of delay to the simulation latency, depending on the used buffer size of the sound card driver.

The range of calculation times (one source, one receiver) for similar conditions (configuration, model, hardware) are shown in Table 3.

Table 3: *Range of calculation times for the simulation jobs and the filter synthesis.*

Simulation Component	Simulation Time	Filter Synthesis
Direct Sound	0.5 .. 2 ms	1 .. 5 ms
Image Sources	1 .. 15 ms	1 .. 10 ms
Ray Tracing	250 .. 1500 ms	25 .. 75 ms

A total response time of almost 500 ms violates the real time requirement. However, as the user is not able to perceive minor changes in the reverberation and the most important changes (direct sound and early reflections) can be updated in less than 50 ms even in case of highly interative situations (e. g. flying through the virtual scene in the *camera mode*), the response times are regarded as sufficiently low. Additionally, because typically, a user of the *SketchUp* software is not able of doing more than one geometry modification within one second, a system response including the update of the room's reverberation within one second constitutes *immediate feedback* for the user.

4. APPLICATION

The auralization plug-in is currently successfully being used for demonstrations and exhibitions as well as in university courses including architecture, room acoustics and Acoustic Virtual Reality at different universities. Students are able to design and modify a room while receiving immediate acoustic feedback. This enables them to learn about room acoustics in a playful manner, e. g. exchanging wall materials helps to understand the effects of absorption and scattering. The tool provides a perceptual measure of the room acoustic parameters, which can be calculated and visualized with a tool named *SketchUp-Visualiser* [27], which is also developed at the Institute of Technical Acoustics, RWTH Aachen. A convenient feature of the tool is the investigation of different listening positions by placing multiple receivers at different seats in the audience. By using the *switch-receiver* function, it is possible to quickly switch between different listener positions in a concert hall (see Fig. 6). The dummy heads represent the listening positions, the orange dummy head indicates the currently active receiver.

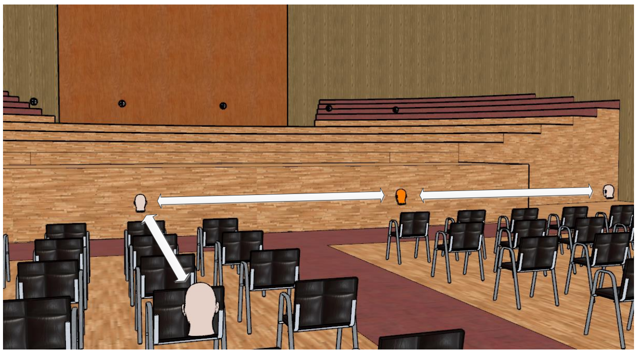

Figure 6: *Switching receiver positions in a concert hall*

5. CONCLUSION

This work describes the implementation and the possible applications of a real-time auralization tool which is embedded in the popular 3D modeling software *SketchUp*. The development of the graphical user interface was focused on a seamless integration into the software as well as on the intuitive control of the auralization also by non-expert users. The simulation client realizes a state-of-the-art room acoustic simulation, efficiently combining multiple simulation models to create a binaural room impulse responses which is directly processed by the low-latency convolution engine of the tool. The simultaneous use of two spatial data structures is applied to keep the latency of the auralization at a minimum level. Additionally, an elaborated update management reduces the simulation workload in various cases of user interaction.

Although not all simulation results can be calculated in real-time, the auralization tool is able to auralize interactive situations based on physically correct simulation results without significant delays. Because of the parallelized threading concept with a prioritization on quick updates of psychoacoustically dominant parts of the impulse response, the auralization of one sound source in a room can easily be calculated on a single standard desktop computer or laptop, providing an immersive listening experience with multimodal feedback. Currently it is successfully being used for teaching and demonstrations, e. g. in lab courses about room acoustics and spatial hearing. Combined with the calculation and visualization of room acoustic parameters in *SketchUp*, the presented auralization software represents a convenient and reliable tool also for room acoustic consultants.

In future, the usage of multiple simulation clients for the auralization of multiple sources at once will be tested and investigated. Models for priority management in case of larger scenes with multiple sources will be researched and applied to the software. The modular software concept also supports the extension of the software with different simulation and reproduction methods, with Higher-Order Ambisonics and VBAP recently being implemented [28].

6. REFERENCES

[1] J.B. Allen and D.A. Berkley, "Image method for efficiently computing small-room acoustics," *Journal of the Acoustical Society of America*, vol. 65(4), pp. 934–950, 1979.

[2] A. Krokstad, S. Strom, and S. Sörsdal, "Calculating the

acoustical room response by the use of a ray tracing technique," *Journal of Sound and Vibration*, vol. 8, no. 1, pp. 118 – 125, 1968.

[3] Bengt-Inge L. Dalenbäck, "Room acoustic prediction based on a unified treatment of diffuse and specular reflection," *The Journal of the Acoustical Society of America*, vol. 100, no. 2, pp. 899, 1996.

[4] Dirk Schröder, Frank Wefers, Sönke Pelzer, Dominik Stefan Rausch, Michael Vorländer, and Torsten Kuhlen, "Virtual Reality System at RWTH Aachen University," in *Proceedings ICA 2010, 20th International Congress on Acoustics*, 2010.

[5] Nikunj Raghuvanshi, Christian Lauterbach, Anish Chandak, Dinesh Manocha, and Ming C. Lin, "Real-time sound synthesis and propagation for games," *Commun. ACM*, vol. 50, no. 7, pp. 66–73, 2007.

[6] Graham Naylor, "Treatment of early and late reflections in a hybrid computer model for room acoustics," *Journal of the Acoustical Society of America*, vol. 92(4), pp. 2345–2345, 1992.

[7] Mendel Kleiner, Bengt-Inge Dalenbäck, and Peter Svensson, "Auralization-an overview," *Journal of the Audio Engineering Society*, vol. 41, no. 11, pp. 861–875, 1993.

[8] K. Heinrich Kuttruff, "Auralization of impulse responses modeled on the basis of ray-tracing results," *Journal of the Audio Engineering Society*, vol. 41, no. 11, pp. 876–880, 1993.

[9] R. R. Torres, U. P. Svensson, and M. Kleiner, "Computation of edge diffraction for more accurate room acoustics auralization," *J. Acoust. Soc. Am.*, vol. 109, pp. 600–610, 2001.

[10] Michael Vorländer, "International round robin on room acoustical computer simulations," in *International Congress on Acoustics*, 1995.

[11] Ingolf Bork, "A comparison of room simulation software - the 2nd round robin on room acoustical computer simulation," *Acta Acustica united with Acustica*, vol. 86, pp. 943–956, 2000.

[12] Ingolf Bork, "Report on the 3rd round robin on room acoustical computer simulation - Part II: Calculations," *Acta Acustica united with Acustica*, vol. 91, pp. 753–763, 2005.

[13] Jens Blauert, *Spatial Hearing - The Psychophysics of Human Sound Localization*, The MIT Press - Cambridge, Massachusetts, 1996.

[14] B.-I. Dalenbäck and M. Strömberg, "Real time walkthrough auralization - the first year," in *Proceedings IOA Copenhagen*, 2006.

[15] Tobias Lentz, Dirk Schröder, Michael Vorländer, and Ingo Assenmacher, "Virtual Reality system with integrated sound field simulation and reproduction," *EURASIP: Journal on Advances in Signal Processing*, vol. 2007, pp. 187, 2007.

[16] D. Schröder and M. Vorländer, "RAVEN: A real-time framework for the auralization of interactive virtual environments," in *Proc. of EAA Forum Acusticum*, Aalborg, 2011, pp. 1541–1546.

[17] Dirk Schröder, *Physically Based Real-Time Auralization of Interactive Virtual Environments*, Ph.D. thesis, Fakultät für Elektrotechnik und Informationstechnik der Rheinisch-Westfälischen Technischen Hochschule Aachen, 2012.

[18] Sönke Pelzer, Marc Aretz, and Michael Vorländer, "Quality assessment of room acoustic simulation tools by comparing binaural measurements and simulations in an optimized test scenario," *Acta acustica united with Acustica*, vol. 97, no. S1, pp. 102–103, 2011.

[19] Michael Vorländer, *Auralization - Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, Springer, 2010.

[20] Thomas Funkhouser, Ingrid Carlbom, Gary Elko, Gopal Pingali, Mohan Sondhi, and Jim West, "A beam tracing approach to acoustic modeling for interactive virtual environments," *Proceedings of SIGGRAPH 98*, pp. 21–32, July 1998.

[21] M. Jedrzejewski and K. Marasek, "Computation of room acoustics using programmable video hardware," in *Proc. Computer Vision and Graphics International Conference, ICCVG 2004, Warsaw, Poland*, 2004.

[22] Dirk Schröder and Tobias Lentz, "Real-time processing of image sources using binary space partitioning," *Journal of the Audio Engineering Society*, vol. 54, no. 7/8, pp. 604–619, 2006.

[23] Dirk Schröder, Alexander Ryba, and Michael Vorländer, "Spatial data structures for dynamic acoustic virtual reality," in *ICA2010: 20th International Congress on Acoustics*, 2010.

[24] Frank Wefers and Michael Vorländer, "Optimal filter partitions for real-time fir filtering using uniformly-partitioned fft-based convolution in the frequency-domain," in *Proceedings of the 14th international conference on digital audio effects : September 19-23 : IRCAM, Paris, France / Institut de recherche et coordination acoustique musique (Paris)*. 2011, pp. 155–161, IRCAM-Centre Pompidou.

[25] Tobias Lentz, "Dynamic crosstalk cancellation for binaural synthesis in virtual reality environments," *Journal of the Audio Engineering Society*, vol. 54, no. 4, pp. 283–294, 2006.

[26] Lukas Aspöck, Sönke Pelzer, and Michael Vorländer, "Using spatial information for the synthesis of the diffuse part of a binaural room impulse response," in *DAGA 2014: 40. Deutsche Jahrestagung für Akustik - 10. - 13. März 2014 in Oldenburg*. 2014, pp. 71–72, Deutsche Gesellschaft für Akustik.

[27] Sönke Pelzer, Lukas Aspöck, Michael Vorländer, and Dirk Schröder, "Interactive real-time simulation and auralization for modifiable rooms," *International Symposium on Room Acoustics*, 2013.

[28] Sönke Pelzer and Michael Vorländer, "3d reproduction of room auralizations by combining intensity panning, crosstalk cancellation and ambisonics," in *Proceedings of the EAA Joint Symposium on Auralization and Ambisonics, Berlin.*, 2014.