MICHAEL VORLÄNDER
Institute for Hearing Technology and Acoustics, R WTH Aachen University,
Kopernikusstraße 5, 52074 Aachen
mvo@akustik.rwth-aachen.de

AURALIZATION OF ENVIRONMENTAL NOISE

KEYWORDS: Auralization, computer simulation, real-time audio signal processing

INTRODUCTION

Imagine if time travel were possible and one could experience the future, i.e. to see and hear an urban or rural environment, which does not exist yet. In fact, this is possible with the technology of Virtual Reality. In virtual reality, almost any environment can be experienced auditory-visually, even if it is in the future. It goes without saying that a halfway correct description of the environment is a very big challenge. But once all characteristics of the environment and the sound sources have been specific and implemented, the task of technical realization still remains. This requires a complex Virtual Reality (VR) technology, namely a display device for visual presentation (“3D glasses”) and a 3D audio playback technology (“surround sound technology”). The point is that with technology, both 3D viewing and 3D listening are available in virtually every home. Smartphones already provide rudimentary approaches. Head-Mounted Displays (HMD) have arrived on the computer games market and at correspondingly affordable prices. Binaural playback via headphones is both old-fashioned and once again the focus of current research when it comes to perfected individualized solutions.

The technology is basically available, what do we do with it now? First of all, acoustics in research and practice and also interdisciplinary research with other sciences!

Before the introduction to the acoustic-technical basics of virtual acoustics, here is a look at examples: VR technology can be used to plan runways or flight routes near airports. The major challenge here is to characterize the generation of aircraft noise with sufficient accuracy and to generate sound source signals from this, which then reach the receiver via models of atmospheric sound propagation. This application will be discussed in more depth below.

FROM SIMULATION TO AURALIZATION TO ACOUSTIC VIRTUAL REALITY

The progress of virtual acoustics over the last decades can be divided into the phases of the development of simulation techniques, audio signal processing, 3D audio technology and integration into VR system technology. In the fields of acoustics, the first steps were taken in room acoustics, then in building acoustics and vehicle acoustics, from which modular approaches were created through the generalization of components and work steps, which can now be used in practically all applications of audible sound.

Early as 1962, Manfred Schroeder formulated a vision of computer simulations in room acoustics, which were later developed by Krokstad et al. (1968) were actually presented. In the 1990s, these methods were finally developed to such an extent that results in great detail could be achieved after just a few minutes or hours on standard PCs (which only then became widespread), if necessary “over a weekend”. These programs represent a very useful addition to the model measurement technology, which also has its advantages, but is not exactly easy to use for the rapid prediction of room-acoustic impulse responses. From around 1990, first auralizations were presented, which enable to create audio files from the computed impulse responses (Kleiner et al. 1993).

Reproducing a previously made recording of a sound source in that exact environment should ideally provide exactly the same listening experience. While this is not an auralization in particular, it is a good reference for auralization validations. The key feature of auralization is that it considers the source and transmission environment separately, see Fig. 1 (Vorländer 2020). As a result, different sound sources can be listened to, for example, with existing audio filters, without all of these having been previously recorded in the relevant transmission environment. Or a source can be listened to in different environments.

Accordingly, auralization consists of a modular approach in the separation of sound source, sound propagation components and receiver. In principle, all parts are interchangeable, so that when the components are varied, the resulting effect on the auditory event can be perceived immediately. It is now only a small step to integrate it into the system technology of virtual reality. However, it took 30 years to go there, rather there were many small steps from the first multimodal interactive VR implementations such as “SCATIS” 1992–1995 (Blauert et al. 2000) and “DIVA” (Savioja et al. 1999) up to today's established VR systems

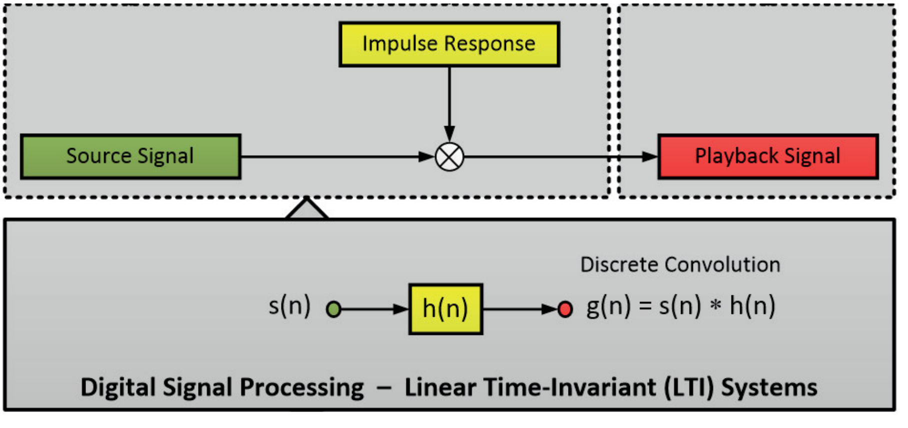

Figure 1. Key features of auralization: sound signal, sound propagation filter (impulse response), and sound playback.

such as "TASCAR" (Grimm 2015 [1]) or "VA" (Wefers and Vorländer 2018 [2]). Interestingly, in both of the early VR systems, multimodal investigations of haptics and acoustics were in the foreground, while the visual component was of only rudimentary importance. This can be explained by the fact that 3D computer graphics were still in the early stages of development.

Today, next to haptics, computer graphics is the most important component of VR technology and is evident in the majority of VR applications and publications in the field. However, the acoustics are playing an increasingly important role, especially since the feeling of presence in the virtual world ("immersion") is increased enormously with a plausible 3D acoustic simulation.

VR tools are now an integral part of research and development. An overview of the system components is shown in Fig. 2. The environment model contains all the input data needed for the simulation, and just like with auralization, these are sound source position, sound power, orientation, directivity, etc. as well as all the input data for the environments, be it indoors or outdoors. The receiver must also be placed, also with regard to the

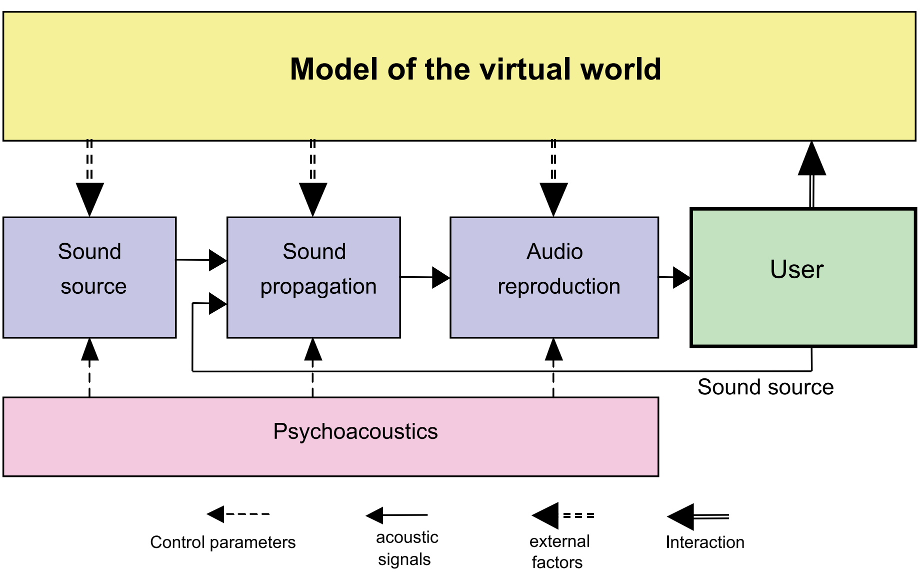

Figure 2. Implementation of an acoustic VR system

[1] https://github.com/gisogrimm/tascar
[2] http://virtualacoustics.org/

viewing/listening direction. The acoustic virtual reality must then be produced and presented in “real time” in such a way that it is perceived as quasi-real. And this in synchronization with the simulations for the other sensory perceptions in virtual reality (visual, tactile, tactile, olfactory), of which not all, but as many as possible should interact in order to achieve complete immersion. At this point, psychoacoustics provides valuable information on the perception thresholds of details in spectral, temporal and spatial sound, loosely based on the motto “simulate only as complex and precise as humans can resolve differences”.

STEPS TO CREATE VIRTUAL ACOUSTIC SCENES

It should be pointed out once again that not only “sound effects” but physically-based virtual acoustic scenes are to be created. Depending on the application, it must be decided which psychoacoustic criteria are particularly important and in which components the auralization must be applied particularly precisely. It is indeed the case that every auralization only creates an illusion, but never a real image of reality, even if an existing environment is supposed to be reproduced exactly. This consideration brings us to the question of validation against comparisons between reality and virtual reality, but more on that later.

SOUND SOURCE CHARACTERIZATION

The work steps in virtual acoustics begin with the recording or synthesis of sound signals. However, one important condition must be taken into account here: Examples of pieces of music recorded in 3D are multi-channel recordings and simulations (e.g. Rindel et al. 2004), which can also be used with sound-dependent radiation characteristics. The source characterization of musical instruments and the human voice is therefore much more than a simple "recording" like in a music studio. Many other factors have to be considered, e.g. suitable data formats with sufficient resolutions in spatial and spectral dimensions.

Now it would be nice if the same concept could be more generally applicable with an enveloping microphone array. Unfortunately, this is not possible with moving sources whose sound emissions are to be recorded during normal operation. We think of sources of traffic noise such as vehicles, trains or airplanes, which are the relevant sources in environmental noise. Theoretical source models, computer simulations or experimental methods must be used here, also in combination, in order to at least approximately estimate the spectral envelopes of the partial sound power of stochastic noise components (jet noise, tire noise, wind noise, etc.). In a synthesis step, previously neutral noise signals (white, pink) can be parametrically filtered from this data in such a way that the spectral and directional sound intensities correspond to those of the traffic sources. The signal components from periodic or rotational processes in machines or engines, in the case of rolling noise, a synthesis can be added, which is controlled using revolutions per minute data. This was, for example, by Pieren et al. (2017) for a rail noise simulation and, as described in the overview by Rizzi (2016) for the aircraft noise simulation, most recently also by Dreier and Vorländer (2020) for the auralization of commercial aircraft.

Now that the source signals are at least available in an approximate form, one can now ask how the sound propagates into the environment. Here, too, there are simply cases and situations with extreme challenges. In many cases, the radiated power is independent of the environment. This is not generally the case with primary sources of structure-borne noise, but it is definitely the case with the vast majority of airborne noise sources.

SOUND PROPAGATION MODELS

In general, the simulation of the impulse responses must include all relevant effects on the sound wave propagation, such as reflection, scattering, transmission, diffraction, refraction, attenuation, wind speeds and temperature profiles in the atmosphere, etc.

A more elegant solution is the direct solution with simulation methods that deliver results in real time, i.e. within around 50 milliseconds. In an indoor scene, for example, the VR user could look around or “walk around in the room”, sound sources can also move, which creates the need to constantly calculate new impulse responses. This is especially true for a virtual scene with aircraft noise, where the source is moving very quickly while looking around the virtual environment. However, real-time impulse response calculations are only possible with the help of approximations, which entail substantial acceleration of the calculation time. This applies in particular to “geometric acoustics” (Savioja and Svensson 2015). This is known, among other things, in the form of image source methods and ray tracing. In geometrical acoustics, sound paths are constructed by connecting the source point and receiving point with a straight line, which “rays” are considered to be the counterpart to waves (ray/particle-wave dualism), just as in ray optics. It is not so well known that curved paths can also be calculated, namely in the case of refraction in stratified media such as the atmosphere (Wilson 2015). For the propagation of sound in addition to the direct line of sight between the source and receiver, i.e. via reflection, scattering and diffraction, the computer has to find the relevant beam paths in no time at all. Special algorithms will not be discussed here, but it should be said that parallel to the historical development of virtual acoustics there has been a rapid development of numerical methods, some of which are ideally suited for real-time applications (Savioja and Svensson 2015).

Source data and sound propagation models are now available, which can be used to simulate the sound field at the reception point. But the receiver is no longer a mathematical point. There is someone standing there, and this person hears the virtual sound event in a three-dimensional space.

3D HEARING AND SEEING IN VIRTUAL WORLDS

Playback technology for “3D audio” is an essential part of VR systems, which must be able to meet high quality standards with regard to the psycho-acoustically relevant aspects of perception. Details may vary from one VR application to the next. Some applications require accurate localization, while for others monaural spectral features such as reproduction with exact loudness and timbre are more important.

In a visual analogy, modern "shutter glasses" based on polarization filters or green-red filters in conjunction with high-resolution video displays deliver very good stereoscopic images, as do the head- mounted displays (HMD) or "VR goggles". HMDs have two small video displays built into goggles or a helmet. Two slightly shifted images for the left and right eyes are created for the illusion of depth in the video, similar to old stereoscopic devices. Binocular vision allows distances to be estimated when looking at nearby objects. With the right eye, we see a nearby object projected onto a different part of the retina than the left, and this difference becomes more significant when the object is close. In this respect, technologies for three-dimensional vision are already available in good quality solutions. This is now also the case with loudspeaker-based 3D audio technology such as Ambisoncs (Gerzon 1985).

A 3D audio playback system for VR applications in research should not be confused with surround sound systems in consumer electronics. The main difference is that VR applications are based on physical models and the highest possible degree of realism in the components of sound and vibration generation, transmission and reproduction. This is a different goal than that pursued by an audio engineer for a music production. Even when recording live, he or she uses recording techniques and strategies for aesthetic optimization and instrument placement to achieve the best result for a home speaker setup, which then reproduces stereo or 5.1 signals.

APPLICATION IN ENVIRONMENTAL NOISE

The acoustics of outdoor environments are subject of research in environmental noise and in soundscape studies. This may apply to nature and in particular to the built environment in the context of human well-being and health. The sound sources are natural sources such as animals and waterfalls, and man-made objects such as vehicles, etc. Traffic noise play a very large role in these studies, and the sound propagation from the road, rail or air vehicles in the outdoor environment must be computed.

LINKING ATMOSPHERIC AND URBAN SOUND PROPAGATION

Real-time computation of atmospheric sound propagation can be implemented by using approximations of geometrical acoustics. A common method is ray tracing, which provides estimates of the acoustic paths. However, the atmosphere is a dynamic inhomogeneous medium. More specifically, speed of sound variances as well as wind have to be taken into account. Therefore, refraction and translation effects, which arise from interaction of the sound wave with the medium are included in the ray tracing algorithm. The method delivers so-called eigenrays, which connect the (moving) source with the receiver. Since the propagation through inhomogeneous media causes curved rays, the identification of eigenrays is not trivial. In an efficient implementation (ray zooming), the simulation is carried out with a low angular resolution. Based on the results, the resolution is stepwise increased around certain emission angles. This procedure is repeated multiple times until the receiver is reached within a certain accuracy. With this method, eigenrays between an aircraft and a receiver on the ground (reflecting plane) can be efficiently determined (Schäfer and Vorländer 2021). This does not include, however, any realistic urban environment with buildings, etc.

In order to extend the applicability to full scenarios of the built environment, Schäfer and Vorländer (2022) introduced the so-called virtual source method (VSM). It combines the properties of a curved free-field path calculated with the atmospheric model to sound paths of the urban model. For this purpose, a virtual source position is determined from the atmospheric path, which is then used for the urban simulation as shown in Fig. 3. This position is further serving as apparent source position, so that the incident direction at the receiver as well as the atmospheric propagation delay are maintained for the direct urban sound path. Thus, the distance between apparent source and the receiver is calculated using speed of sound of the urban domain. With this, the atmospheric propagation path can be smoothly integrated into the urban propagation simulation algorithm.

Real-time sound path computation in urban areas is based on geometrical acoustics as well. The main propagation effects are

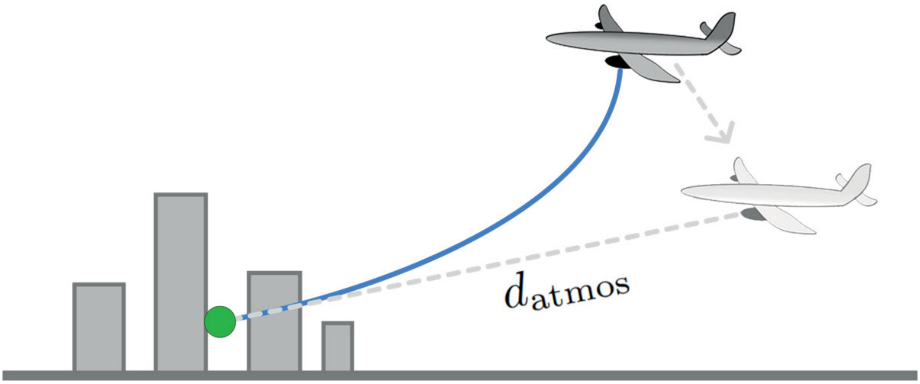

Figure 3. Definition of virtual source based on an atmospheric free field path (after Schäfer and Vorländer 2021).

reflection, scattering, and diffraction. While reflection and scattering components can be determined with classical ray tracing similar as applies to room acoustics, higher-order diffraction is crucial due to rather sparse impulse responses. Pathfinder algorithms and strategies for culling, i.e. extraction of irrelevant (inaudible) paths, is essential in order to achieve real-time performance. First steps are published by Erradj et al. (2021) and will be continued in the context of traffic noise auralization (Dreier et al. 2022) and, as mentioned above, in the combination of atmospheric and urban sound propagation. With this overarching approach, complex scenarios consisting of aircrafts and road or rail vehicles can be auralized.

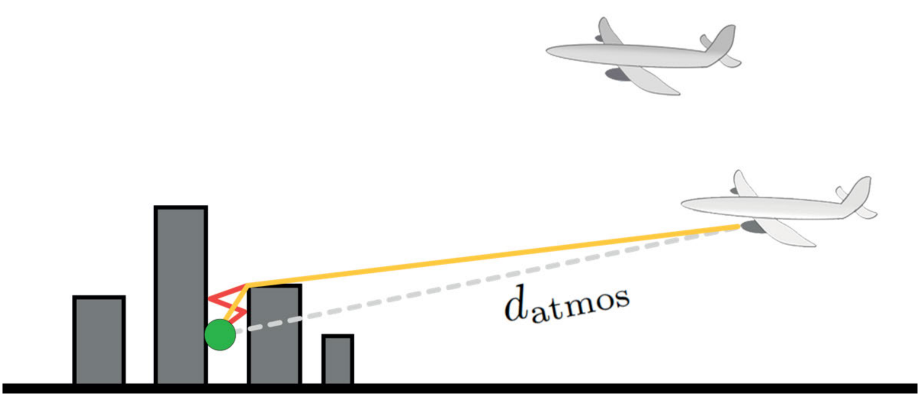

Figure 4. Atmospheric free field path (dashed, grey) and determination of urban paths (yellow and red) based on this virtual source position and an urban propagation model.

As shown in Fig. 4, after the relevant sound paths are known each path with its features (spectrum and delay) can be further processed by means of digital filters.

REAL-TIME SIGNAL PROCESSING

A real-time audio processing framework must be able to acoustically render virtual outdoor environments with dynamic, moving sources, partly fast-moving sources and interactive receivers in a built environment. To comply with aspects of real-time auralization constraints, it is required to implement dynamic scene handling, acoustic simulation and audio processing. The approach introduces an auralization timeline to record and up-sample scene updates and sort simulation results from a scalable scheduler instance in a chronological, layered history (Stienen 2022). Based on the history data, a smooth treatment of propagation delays is achieved by interpolation and extrapolation methods and chronologically correct event scheduling. A Digital Signal Processing (DSP) network processes propagation effects such as spectral effects from reflection and diffraction. Time-variant Doppler shift for many individual geometrical sound paths in real time using low-order Infinite Impulse Response (IIR) filter units. Key feature of the design is the implicit capacity to adapt to fast motion by introducing Single-Input Multiple-Output Variable Delay Lines (SIMO-VDLs). For spatial audio reproduction, binaural technology with a directional clustering routine at the receiver operates at a quasi-constant computational load.

The method uses Head-Related Impulse Response (FIR_1, FIR_2, ...) convolutions and adjusts the Inter-aural Time Difference (ITD) mismatch using Fractional Delays or each individual propagation path, while maintaining efficiency by merging Doppler interpolation and ITD correction into a single routine. The complete layout of the DSP network that represents the core of the audio rendering module is visualised schematically in Fig. 5.

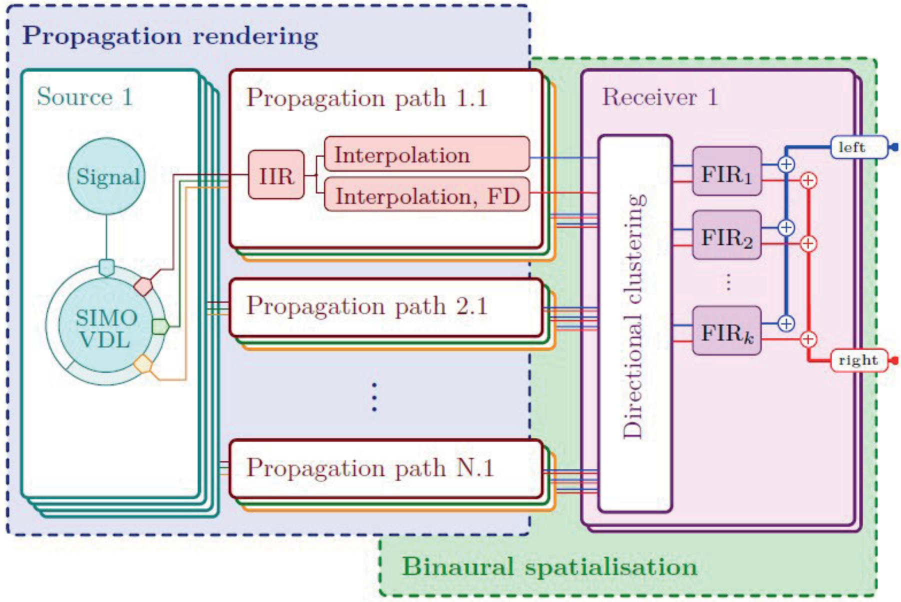

Figure 5. DSP network for real-time auralization of dynamic, outdoor environments. Each of *N* sources provides a signal that feeds a SIMO-VDL. Read cursors per source receiver pair transfer a delayed propagation path signal including spectral propagation effects.

CONCLUSION

Using a combined scenario, an aircraft flyover over an urban area was auralized[3]. The resulting audio demonstration underlines how considering reflections and diffraction at urban structures, increases the stereophonic realism significantly. This holds particularly regarding the localization of the source when the direct path is occluded, so that secondary sound paths become dominant.

Next steps towards improvements and extensions are in progress. This concerns listening tests with variants of diffraction filters in order to optimize computational effort with reference to perception. Also, inclusion of surface scattering is relevant when it comes to complex building facades with roughness tex-

[3] Audio-visual demonstration of an aircraft flyover. ### Youtube or IHTA-website, link will be added in the author proofs ###

ture on various scales. Regarding atmospheric propagation, turbulence is an important aspect, which can be included in the auralization process using time-variant filters. While this was successfully done for auralizations based on the Atmospheric Ray Tracing framework alone where only two sound paths are considered, this is ongoing work for the presented approach. Finally, the combination of the aircraft sound component with that of other sound sources in urban environments will lead to the ultimate goal of establishing an open-source Virtual Reality toolbox for soundscape research.

It can be seen that each application requires careful consideration of the entire simulation chain from source to receiver and beyond to experimental design. This is exactly what distinguishes virtual acoustics in acoustic research and practice from 3D sound in computer games, in which impressive 3D scenarios are presented, also highly interactive, but without any reference to the physically real conditions in the respective scene.

ACKNOWLEDGEMENTS

The author would like to thank the former and current members of the VR working group at the Institute for Hearing Technology and Acoustics at RWTH Aachen University: Lukas Aspöck, Christian Dreier, Anne Heimes, Simon Kersten, Michael Kohnen, Tobias Lentz, Josep Llorca-Bofí, Imran Muhammad, Sönke Pelzer , Philipp Schäfer, Oliver Schmitz, Dirk Schröder, Jonas Stienen and Frank Wefers as well as the VR group in the IT department headed by Prof. Torsten Kuhlen. The work of Michael Vorländer and his group was mainly supported by the Deutsche Forschungsgemeinschaft.

FURTHER MULTIMEDIA EXAMPLES

Virtual acoustics at the IHTA of the RWTH Aachen University: https://bit.ly/36ZIqUw

Park and Convention Center: https://bit.ly/33L4foW

Aircraft noise: https://bit.ly/33KzgcB
Interactive scene in a park: https://bit.ly/2peKdnE
A look into a sports arena: https://bit.ly/2X8CxQs
Interacting with Virtual Humans: https://bit.ly/373l0hd

REFERENCES

Blauert J, Lehnert H, Sahrhage J and Strauss H (2000). An Interactive Virtual-Environment Generator for Psychoacoustic Research. I: Architecture and Implementation. Acustica united with Acta Acustica 86, 94

Dreier C and Vorländer M, Psychoacoustic optimization of aircraft noise - challenges and limits. Proc. Internoise 2020, Seoul, Korea

Dreier C, Hahn J, Heck J, Llorca-Bofi J and Vorländer M (2022) Real-time vehicle pass-by noise synthesis for urban sound auralization. Proc. ICA 2022 Gyeongju, Korea.

Erradj A, Stienen J and Vorländer M (2021) The Image Edge Model. Acta Acustica 5, 17.

Gerzon MA (1985) Ambisonics in multichannel broadcasting and video. J. AES 33, 859

Grimm G, Luberadzka J, Herzke T and Hohmann V (2015) Toolbox for acoustic scene creation and rendering (TASCAR): Render methods and research applications. Proc. Linux Audio Conference, 9

Kleiner M, Dalenbäck BI and Svensson P (1993) Auralization – an overview. J. AES 41, 861

Krokstad A, Strøm S and Sørsdal S (1968) Calculating the acoustical room response by the use of a ray-tracing technique. J. Sound Vib. 8, 118

Pieren R, Heutschi K, Wunderli JM, Snellen M and Simons DG (2017) Auralization of railway noise: Emission synthesis of rolling and impact noise. Applied Acoustics 127, 34

Rindel, JH, Otondo, F and Christensen CL (2004). Sound source representation for auralization. Proc. International Symposium on Room Acoustics: Design and Science, Awaji, Japan, April 11-13, 2004.

Rizzi SA (2016) Toward reduced aircraft community noise impact via a perception-influenced design approach. Proc. Internoise, Hamburg, Germany, 220

Savioja L, Huopaniemi J, Lokki T and Väänänen R (1999) Creating Interactive Virtual Acoustic Environments. J. AES 47, 675

Savioja L and Svensson UP (2015) Overview of geometrical room acoustic modeling techniques. J. Acoust. Soc. Am 138(2), 708

Schäfer P and Vorländer M (2021) Atmospheric Ray Tracing: An efficient, open-source framework finding eigenrays in a stratified, moving medium. Acta Acustica 5, 26.

Vorländer M (2020) Auralization - Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality. 2nd edition. Springer Nature Switzerland AG 2020. (383 pages)

Wefers F and Vorländer M (2018). Flexible data structures for dynamic virtual auditory scenes. Virtual Reality 22(4), 281

Wilson DK (2015) Outdoor Sound Propagation Calculator. In VE Ostashev and DK Wilson, Acoustics in Moving Inhomogeneous Media (Second Edition). Taylor & Francis, 2015. https://www.routledge.com/downloads/Y105698/Y105698_Web_Download.zip