Buenos Aires – 5 to 9 September, 2016
Acoustics for the 21st Century…

PROCEEDINGS of the 22nd International Congress on Acoustics

Plenary Lecture: Paper ICA2016-481

From acoustic simulation to virtual auditory displays

Michael Vorländer

Institute of Technical Acoustics, RWTH Aachen University, mvo@akustik.rwth-aachen.de

Abstract

Simulation and auralization techniques are used in engineering, architecture, sound design and in applications in hearing research. The components of this technique are acoustic simulations and signal processing tools and the data interfaces in between, for which well-established solutions exist. The main bottlenecks are lack of data of 3D characterization of sound sources and material parameters, and interfaces to spatial audio technology. These problems are subject to research. Whether the virtual environment is considered sufficiently accurate or not, depends on many perceptual factors, and on the pre-conditioning and the degree of immersion of the user in the virtual environment. In this presentation the processing steps for creation of Virtual Acoustic Environments are briefly presented and illustrated in examples including room acoustics, music, history, and hearing research.

Keywords: Acoustic simulation, auralization, Virtual Reality

From acoustic simulation to virtual auditory displays

1 Introduction

Acoustical simulation methods are applied in a large variety of tasks in acoustic engineering, in research on human perception and in arts. When it comes to real-time processing, the auralization, the so-called rendering of acoustic spaces offers a direct feedback between the user and the environment. As concerns the direct feedback, interaction may take place, so that the user changes the scene conditions, the positions of sources and receivers, or the whole environment.

Virtual auditory displays are basically the auditory component of Virtual Reality (VR) [1]. VR is an emerging technology. It is the representation and simultaneous perception of reality and its physical attributes in an interactive computer-generated virtual environment. And it should be multimodal: visual and auditory cues, possibly more senses such as force feedback (haptics) or tactile sense should be included in the man-machine interface. It is obvious that in architectural applications like a virtual walk through a complex of buildings, auditory information can significantly assign meaning to visual information and the overall impression, thus giving meaning to the whole complex problem.

Today, acoustic phenomena are mostly described by using so-called “single-number metrics”. The discussion with urban planners and architects, for example, is usually based on the sound pressure level, or the sound reduction index, in decibels. This creates ambiguity in the communication with the responsible authorities and in the discussion about investment in noise control measures. In fact this is a drastic simplification. Human auditory perception is a complex physiological and psychological process. A significant part of acoustic engineering is related to acoustic comfort, well-being, and to the beauty of sound in a concert hall or to the efficiency of a public-address system in a train station. A general analysis of acoustic match and comfort, however, is not completely represented by a single metric such as the decibel, since the actual situation of the sound or noise problem, the activities and expectations of humans affected and the context of the situation must be taken into account. Therefore the area of noise effects, annoyance research and related fields (e.g. communication) can be considered to be given more importance in the future.

In auralization, 3D “sound images” are created by combining computer aided design (CAD) models of outdoor or indoor spaces with acoustic sources models, computational sound propagation models, audio signal processing and sound reproduction by loudspeakers or headphones [2]. Auralization can be applied in acoustic design and optimization of urban squares, airport halls, or classrooms and concert halls, for example. Auralization and Acoustic Virtual Reality have been developed with significant contributions of the applicant during the last two decades [3]. Other key contributions include Kleiner et al from 1993, who published an overview on the emerging technique of auralization in concert hall acoustics [4]. The state of the art, the methods, the challenges and implementation techniques required for auralization and virtual acoustics have become a rather complex field of research, particularly with the rapid progress in real-time signal processing and integration in virtual reality systems.

In this contribution the real-time simulation tools for acoustics are briefly reviewed and discussed concerning applications and evaluations of such systems.

2 Fundamentals of auralization

The process of auralization contains the separation into the processes of sound generation, sound propagation and sound reproduction into system blocks, and the corresponding representation of these blocks with tools from system theory (see Fig. 1, from [5]). In the figure, the discrete source signal, s(n), is called a “dry” sound. The resulting signal after sound propagation in (between) rooms, g(n), contains the features of both the sound source and the sound propagation or transmission system. The performance of a sound propagation system is represented by the system’s impulse response, h(n). The sound signal at the receiver position is then achieved by convolving the original dry sound signal with the impulse response (the impulse response is usually represented by a digital filter).

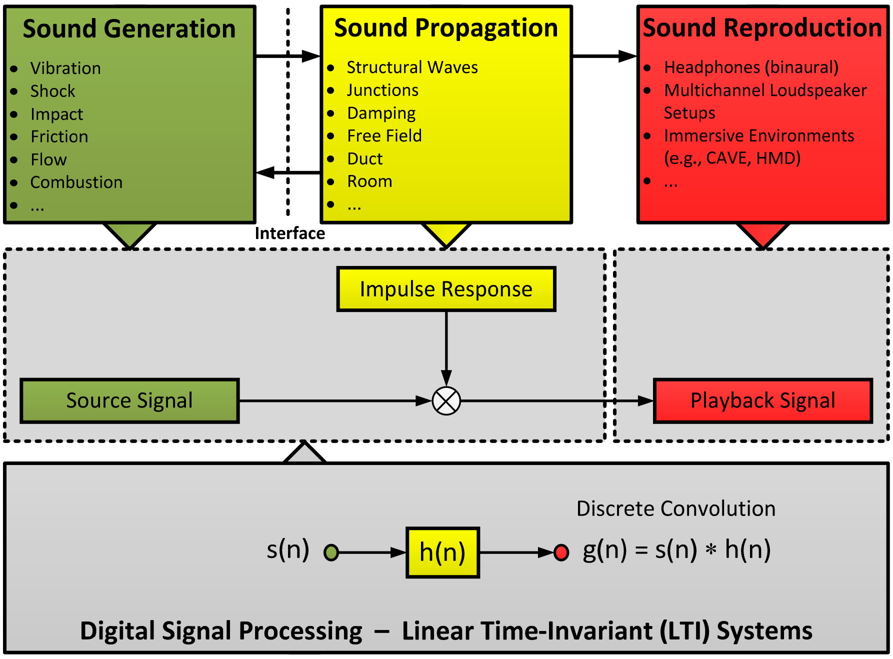

Figure 1: Generation and propagation of sound and its representation in the physics domain (top) and in the domain of acoustic signal processing (bottom), from [5]

3 Acoustic simulation techniques

Acoustic computer simulations are already applied in various design processes with good success. Sophisticated simulation algorithms help to create information about room acoustics, building acoustics, vehicle acoustics, noise control already during the early design and planning.

From a psycho-acoustical point of view, the impulse response of the propagation path (in the following referred to as Impulse Response (IR)) can be divided into three parts - the direct sound, early distinct reflections and the late (diffuse) reverberation. These parts require not the same attention or precision. Following the human's perception of sound, each part of the IR features individual requirements. For instance, small deviations of timing and spectral information for the direct sound and distinct reflections affect the subjective sound source localization. In contrast, our hearing evaluates the late part of the IR (e.g. late reverberation) with a much lower temporal resolution, where only the overall intensity by diffraction and specular and scattered reflections in a certain time slot has to be energetically correct [6, 7].

In outdoor spaces the conditions are very different because the reverberation, if any, has a very low energy compared with direct and reflected energy. The challenge is to calculate the energies and the delays of the few direct paths and reflections accurately, including diffraction.

4 Audio reproduction technology

Spatial sound fields can be created by using one of two general concepts. One can try to reproduce ear-related signals, taking advantage of the fact that the hearing sensation only depends on the two input signals to the ears [8]. Also, loudspeakers arranged around a listening area ("sweet spot") may serve for a spatially distributed incident sound field. The potential to involve more than one listener in the second approach illustrates the conceptual difference between the two methods.

In binaural technology [9, 10], a mono source signal, properly characterized and calibrated according to well-defined specifications can be processed in such a way that its perceptual cues are amended by a spatial component. A binaural mixing console can be used for processing headphone or loudspeaker signals by using head-related transfer functions, HRTF [10]. With a database of HRTF or the corresponding head-related impulse responses, HRIR, any direction of sound incidence can be simulated, when a mono source signal is convolved with a pair of head-related impulse responses by using a FIR filter.

5 Real-time auralization

Now, the method of auralization can be integrated into the technology of "Virtual Reality". As new challenge, the latency in the input-output auralization chain from tracking, audio hardware, signal convolution, and audio reproduction further reduce the maximum permissible computation time for both acoustics simulations and reproduction (rendering) [11]. Real-time processing is only possible with significant reductions of complexity. Here, physical and psychoacoustic evaluations usually help to find the space between simplifications and the period. In the following, data management and convolution problems are briefly discussed with respect to real-time processing.

The real-time convolution engine is typically based on partitioned block convolution and processes the monaural audio signals of the virtual sources with these filters [12]. For each listener, the signals of adjacent sound paths are summed up. As the sound propagation changes (e.g. movement or rotation of the listener), the room acoustic simulation is re-run and the filters, or parts of them, are exchanged. The non-uniform filter partitioning is chosen to support the required filter update rates for the application [12].

6 Applications of Virtual Auditory Displays

6.1 Hearing research, diagnosis and training

The potential of visual and acoustic VR systems will also contribute to create real-world situations for advanced psychoacoustic experiments on localization, attention, and higher-level cognition and for audiological diagnostics. Up to now, these tests are performed in rather simple environments such as an approximated free or diffuse sound field.

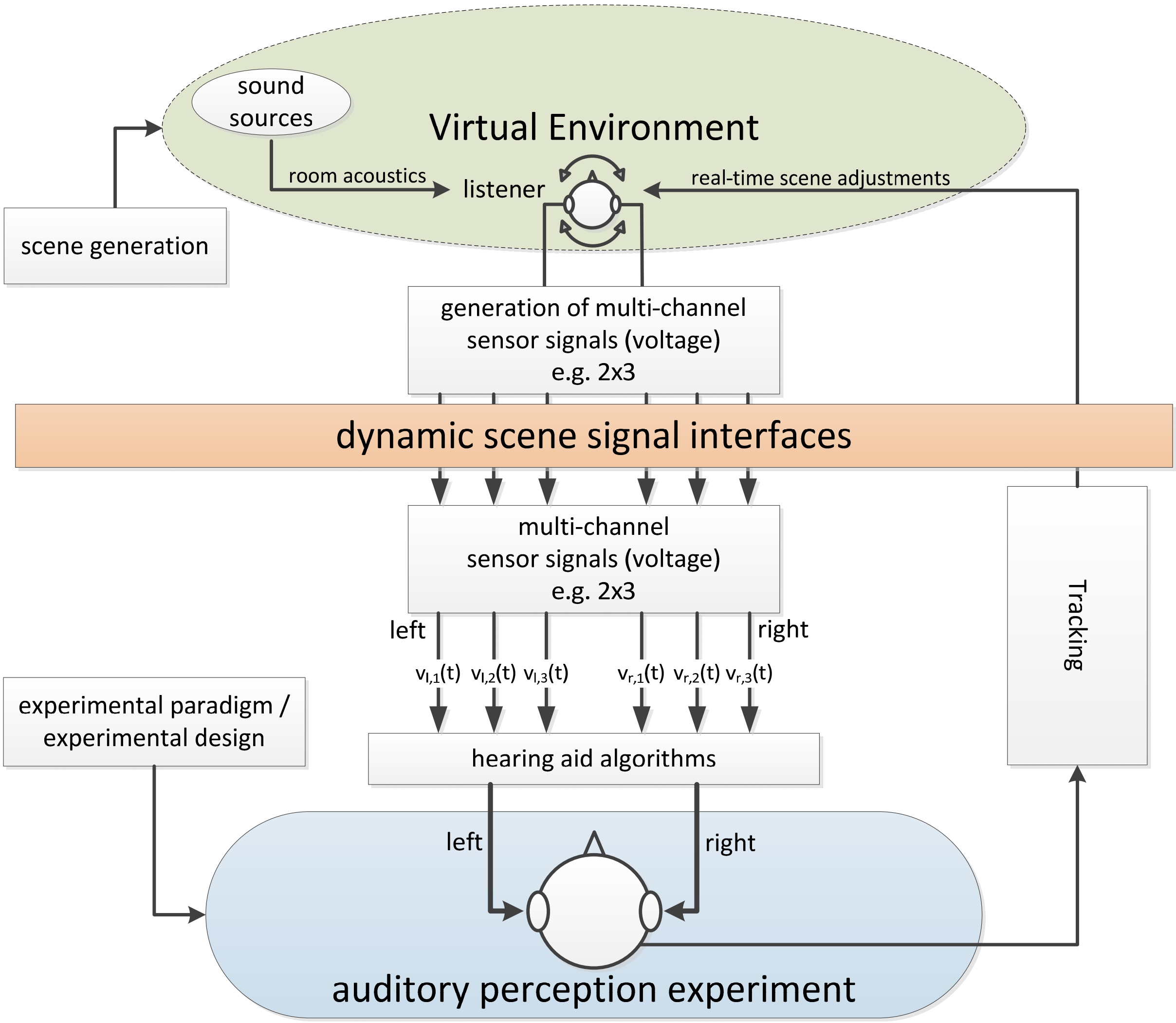

Figure 2: Concept of a dynamic acoustic room simulation for hearing aids with binaural technology

The corresponding sound propagation paths in terms of impulse responses or transfer functions, however, can also be precisely simulated for a human listener with or without hearing aids or cochlea implants, and in this respect the binaural synthesis technique will be extended to an innovative audio reproduction system that allows an evaluation of different auditory scenes through either the audio-input of hearing aids and/or by simulation of the correct sound field at the ear of the hearing-impaired person using an expanded binaural technology. For hearing aids with more than two microphones, this approach can be extended towards a multi-channel HRTF database and integrated into real-time dynamic room auralization software. Then, the dynamic virtual test environment can be freely chosen as an outdoor rural or urban or indoor environment with multiple dynamic sources and a head-tracked listener. This approach is not only relevant for better diagnostic assessments of hearing impaired subjects but also for other patient groups suffering from marked deficits to concentrate and communicate in noisy environments [13, 14].

The approach of the binaural synthesis for the direct sound and of the early reflections is plausible. Through individual reflections and with high precision in the multichannel HRTF data, all parts of the impulse response are well pictured. For indoor and outdoor scenes with a reverberant sound component, however, the exact illustration of sound pressure, pressure differences, and coherence between the microphones by means of those algorithms that simulate the reverberant sound has to be assured. Any stochastic, artificial, or recorded reverberant sound fails in this interrelationship. Instead the reverberant sound field has to arrive as a sum of plane waves with a specific spatial distribution, which can then be processed with a coherent binaural synthesis.

In binaural technology, static and dynamic acoustic scenes have to be examined separately. In dynamic scenes, the user benefits from head movements. To obtain ideal results in static scenes, the individual characteristics of the user should be reproduced as exactly as possible. Therefore, individual or at least individualized HRTFs as well as individual rectification techniques should be used.

6.2 A reconstruction of the historic Montreux Jazz Festival

In the city of Montreux, Switzerland, one of the most famous Jazz festivals is organized since 1967. The venue was the Casino building, which was completely destroyed in a fire disaster in 1971. Today, a virtual acoustics reconstruction was created by Schröder et al. [15] by implementing an audio-visual simulation of the historic festival venue and by playing concerts of historic recordings in the simulated environment in a room with video screen and surround sound system. This way it is possible to recreate the impression and the special atmosphere of the historic place, which is by far more than just a documentation of “old times” but an inspiration for further studies in history and musicology.

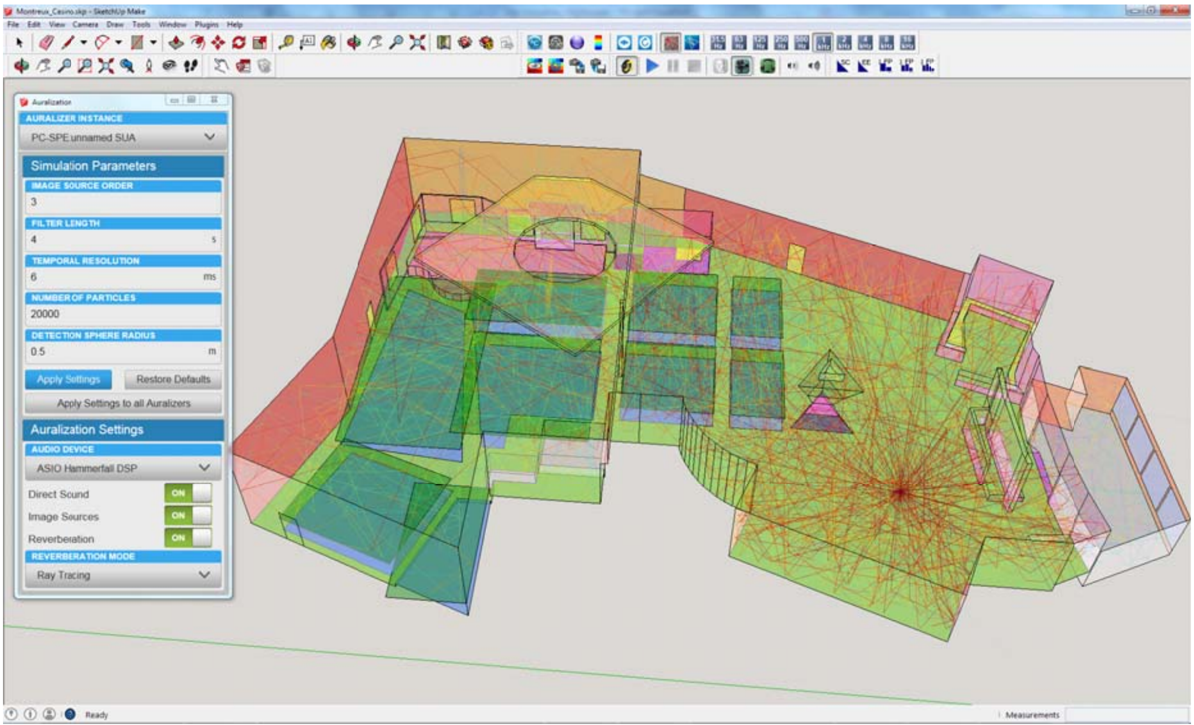

Figure 3: CAD model of the Montreux Casino before 1971 (Schröder et al. [15]).

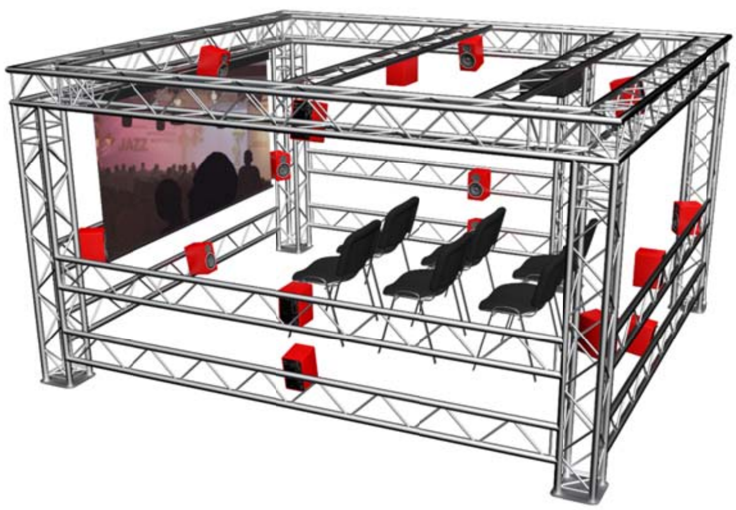

Figure 4: AVR Installation of the Montreux Casino reconstruction (Schröder et al.[15]).

6.3 Virtual Reality: What is Immersion?

As concerns the discussion of plausibility or even authenticity of a virtual environment, the components of real-time processing and interaction must be considered and finally the sensation of "immersion" [16, 17]. Being immersed makes the user feel present in the virtual environment, thus allowing studies of quite complex behaviour under realistic conditions, or high-quality productions of virtual spaces with quasi realistic impressions for any kind of communication and interaction between humans or humans and machines.

Today, to the author's knowledge there is no metric available for the sensation of immersion. In the final example, it is therefore presented a step towards immersion studies related to auditory stimuli in a complex virtual environment, and the required degree of complexity and the required number of sound events at which a reasonable sensation of immersion is achieved.

In a recent project, it was investigated the effect of the specific spatial audio technology on the perceived immersion. This was done in two steps, namely by introducing a questionnaire and a well-designed psychophysical test in a VR environment. The results are promising, as significant items of the questionnaire (related to room and source perception, attention and attribution) could be identified which correlate with immersion, and, furthermore, significant differences were observed between spatial audio technologies used [18, 19].

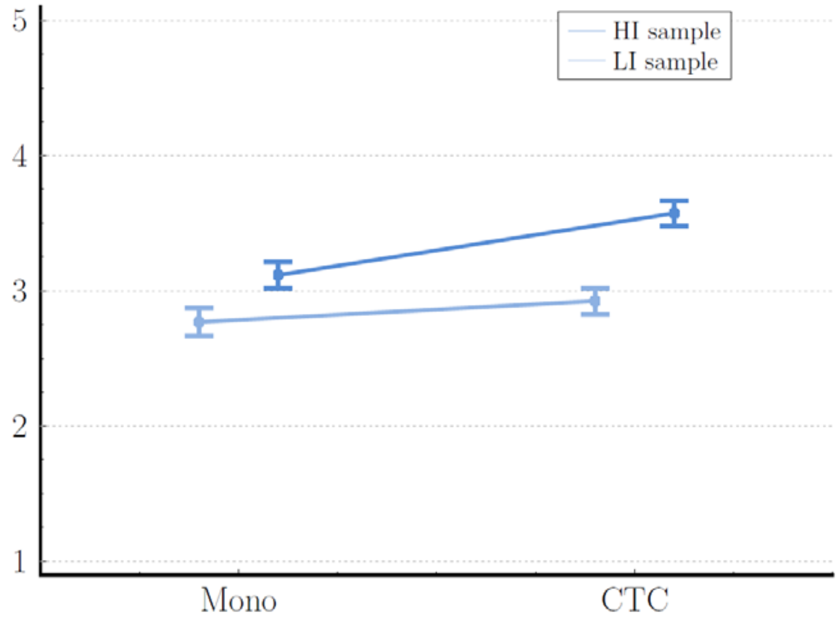

Figure 5: Results of between-group test trial: Y axis shows the average rating of all 40 items for mono group (N = 17) and CTC group (N = 18). Depicted are the mean values and the error bars for the "low immersion" sample and the "high immersion" sample of both groups.

7 Summary and Conclusions

Simulation and auralization is state of the art since several decades. In real-time implementation, Virtual Auditory Displays can be created, and new research directions can be established, and these can go beyond the focus of room acoustics and music played in rooms. For hearing aid and cochlear implant technology, such technology can be applied to use quasi-realistic listening conditions of environments of daily life, such as busy streets or rooms, communication situations with spatially distributed noise, attention experiments, etc. For studying perception of music in rooms, historic spaces can be reconstructed virtually, giving chances to enjoy lost venues. Perceptual studies of VR technology are always required in order to check the quality of the system. It will be most challenging to transport VR technology to the consumer market but as smartphones with little added devices can be transformed into head-mounted displays rather easily, it can be expected that VR will penetrate the marker quickly, and not just for games but for serious applications in research, development and engineering.

Acknowledgments

The author is grateful for contributions of the acoustic VR group at ITA (in particular Dr. Frank Wefers, Dr. Dirk Schröder, Sönke Pelzer and Lukas Aspöck) and of the Virtual Reality Group led by Prof. Torsten Kuhlen at RWTH's IT Center.

References

[1] M. Vorländer, B. Shinn-Cunningham, Virtual Auditory Displays. In: K. Stanney, Handbook of Virtual Environments, chapter 4, 2nd edition. CRC press, New Jersey 2013

[2] M. Vorländer, Auralization - Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality. Springer, 2008.

[3] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen, Creating interactive virtual acoustic environments. J. Audio Eng. Soc. 47 (1999) 675

[4] M. Kleiner, B. I. Dalenbäck, and U. P. Svensson. Auralization - An Overview. J. Audio Eng. Soc. 41 (1993) 861

[5] M. Vorländer, D. Schröder, S. Pelzer, F. Wefers, Virtual Reality for Architectural Acoustics. Journal of Building Performance Simulation (2014) doi:10.1080/19401493.2014.888594

[6] S. Pelzer, L. Aspöck, D. Schröder, M. Vorländer, Integrating Real-Time Room Acoustics Simulation into a CAD Modeling Software to Enhance the Architectural Design Process. Buildings 2 (2014) 113

[7] D. Schröder, F. Wefers, S. Pelzer, D. Rausch, M. Vorländer, T. Kuhlen, Virtual Reality System at RWTH Aachen University. In Proceedings of the International Symposium on Room Acoustics (ISRA), Melbourne, Australia, 2010

[8] J. Blauert, Spatial Hearing – The Psychophysics of Human Sound Localization, revised edition. MIT Press, 1996

[9] R. Nicol. Binaural Technology. Audio Engineering Society 2010

[10] B. Masiero, Individualized Binaural Technology - Measurement, Equalization and Perceptual Evaluation. PhD thesis, RWTH Aachen University, 2012.

[11] D. Schröder. Physically Based Real-time Auralization of Interactive Virtual Environments. PhD thesis, RWTH Aachen University, 2011.

[12] F. Wefers, Partitioned convolution algorithms for real-time auralization. PhD thesis, RWTH Aachen University, 2014.

[13] F. Pausch, Z. Peng, L. Aspöck, J. Fels, Speech perception by children in a real-time virtual acoustic environment with simulated hearing aids and room acoustics. Proc. 22nd International Congress on Acoustics: ICA 2016 Buenos Aires.

[14] L. Aspöck, F. Pausch, M. Vorländer, J. Fels, Dynamic real-time auralization for experiments on the perception of hearing impaired subjects. Proc. 41st German Conference on Acoustics, DAGA 2015 Nuremberg, Germany

[15] D. Schröder, D., S. Pelzer, M. Vetterli, M. Vorländer, Through the Hourglass: A Faithful Audiovisual Reconstruction of the Old Montreux Casino. Acoustics Australia 1 [2015], 49

[16] L. Aspöck, A. Colsman, M. Kohnen and M. Vorländer, Investigating the immersion of reproduction techniques for room auralizations. DAGA Aachen, 2016

[17] A. Colsman, L. Aspöck, M. Kohnen and M. Vorländer, Development of a questionnaire to investigate immersion of virtual acoustic environments. DAGA Aachen, 2016

[18] A. Lindau, S. Weinzierl, Assessing the Plausibility of Virtual Acoustic Environments. Acta Acustica united with Acustica, 98 (2012), 804

[19] A. Lindau, V. Erbes, S. Lepa, H.-J. Maempel, F. Brinkmann, S. Weinzierl, A Spatial Audio Quality Inventory for Virtual Acoustic Environments (SAQI)", in: Acta Acustica united with Acustica, 100(5): 984