3-D Audio Reproduction in non-ideal Environments

Michael Kohnen[1], Sönke Pelzer[1], Michael Vorländer[1]
[1] *Institute of Technical Acoustics, Kopernikusstraße 5, 52074 Aachen, Germany, E-Mail: mko@akustik.rwth-aachen.de*

Introduction

3-D audio reproduction aims at the authentic rendition of an acoustic composition. This can comprise realistic or artificial material as well as different depth of conveyed information, such as position, width, distance, etc. The applications of 3-D rendering are crossing manifold domains in research, education, development and entertainment. Different techniques for 3-D reproduction have been researched since decades and can be grouped into loudspeaker-based and headphone systems. Therefore the signal processing for feeding 3-D rendering systems is well researched.

While it is well known how the signals must be conditioned for each single loudspeaker or the user's headphone, and while it is also well known how the signals must arrive at the user's ears, it is not well investigated how the path in between alters the final perception of the original acoustic composition. For headphone reproduction the (negative) effect of the transfer path can be controlled by equalization filters [1]. But in case of loudspeaker-based reproduction the equalization is not trivial to achieve, since loudspeakers are typically placed in rooms and because reflections inside this room result in very complex transfer paths. An exception would be a fully anechoic environment, but since these are not typical settings for most loudspeaker setups, a solution must be found for conventional environments.

The conclusion could be drawn, that headphones are to be favored for accurate 3-D reproduction, but since these two approaches are very different in many facets, the pros and cons are discussed controversially with many good arguments in favor of loudspeaker setups. As an example the aixCAVE virtual reality systems shall be mentioned here, as shown in Figure 1 and presented in [2], because of the immense effort that was put into the idea to free the user of any hardware that would touch the user's body or restrict his or her freedom of movement. A headphone would contradict the whole principle of this apparatus, while at the same time the best possible reproduction quality is desired to build a perfect multi-modal illusion. Unfortunately this specific reproduction room is a worst-case scenario for loudspeaker-based format, since the user is located inside a cube of glass walls, with nearly no sound absorbing materials except from the ceiling.

This contribution introduces a basic approach to analyze and handle the impact of the listening room on the reproduction of an acoustic composition. In detail it aims at the alteration of auralizations of virtual rooms that have been prepared using a room acoustical simulation, and therefore the room acoustical properties are of importance. To describe the acoustics the traditional set of room acoustical parameters will be used. The effect of the listening room on the auralization of parameters, such as reverberation time, clarity and strength will be investigated. Furthermore considerations are made to identify guidelines for rooms that are suitable for room auralizations.

As a result basic criteria on the room acoustical parameters of the reproduction room are stated in order to keep the alteration of the original auralization below the just noticeable difference limen (JND). Vice versa for a given reproduction room it can deduced which properties the auralized room should have so that it can be rendered without suffering from negative effects of the listening space.

Figure 1: The aixCave in Aachen. Five sides are glass walls (with closed door) and the ceiling is acoustically treated. A loudspeaker system is mounted under the ceiling.

Simulation Set-Up

To investigate the influence of the reproduction room on the target parameters, i.e. the acoustical parameters in the virtual room, a series of simulations have been performed. The series changes the absorption individually in each, virtual and reproduction, room. Then the deviation between acoustical target parameters in the virtual room and those of the overall reproduction is determined. For the virtual room a room size of 6.1m x 4.7m x 3m was chosen. The material parameters on each wall were the same. Source and receiver were positioned off-center to avoid symmetry effects and to simulate a realistic application. To gain results for more than one reproduction technique a single loudspeaker set-up was chosen obtaining fundamental data. The geometric set-up for the reproduction room was 8m x 5m x 3m with every wall having equal material characteristics. Again, to avoid effects of geometric symmetry source and receiver were positioned off-centered. In many listening environments the receiver is positioned somewhere in the middle of the room, therefore the offset from the center position of the room for the receiver is 10cm in all three directions. The source is positioned in a corner of the room at a distance of about 2.6m away from the listener. Simulations have been performed for all combinations of absorption coefficients in both rooms and the room impulse responses (RIR) were then convolved to represent the overall impulse response of the reproduction system.

Investigation of the Room Impulse Response

To understand the influence on the acoustic parameters a closer look at the behavior of the RIRs is given. Figure 2 shows the RIR of a virtual scene with a reverberation time (RT) of 2.5s on the left side. The RIR of the reproduction room also results in a RT of 2.5s and is not shown. The overall impulse response (IR) of the reproduction chain is shown on the right side. The reproduction unveils two obvious differences. The first is that the early part has a steady state part and does not rapidly fall down to the beginning of an exponential energy decay. The second is a different shape and slope of the energy decay curve for the late part of the IR.

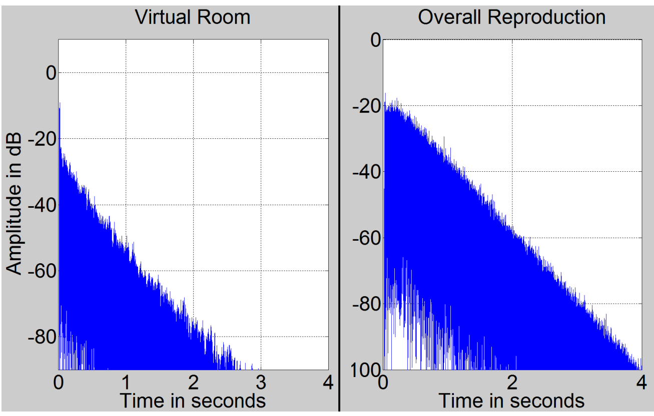

Figure 2: On the left side an IR with a RT of 2.5s at 1kHz for the virtual scene. On the right side the overall IR of an reproduction of the virtual scene in a reproduction room with a RT of 2.5s.

Investigation of the Acoustic Parameters

The simulation results of the room acoustical parameter study can be seen in figure 3. The x-axis shows the RT in the reproduction room with the loudspeaker array, the y-axis the target reverberation time, i.e. the initially intended in the virtual scene, both at 1 kHz. The color indicates the deviation of the RT of the reproduction from the target RT. The deviation is given in percent. The step size of the measurements is 0.1 seconds for the RTs. Values in-between are interpolated. It can easily be found that for an anechoic reproduction room all types of simulation rooms can be reproduced without introducing error regarding the RT, as expected. For reproduction rooms with longer RTs a minimal necessary RT in the virtual room can be found. Assuming a JND (Just Noticeable Difference) of 10% for the perception of RTs the transition from dark to light blue indicates exceeding this limit. For example a virtual scene with a RT of 1.5s would demand a listening room with less than 1.2s to result a deviation less than 10%. At last the linear dependency is to be pointed out.

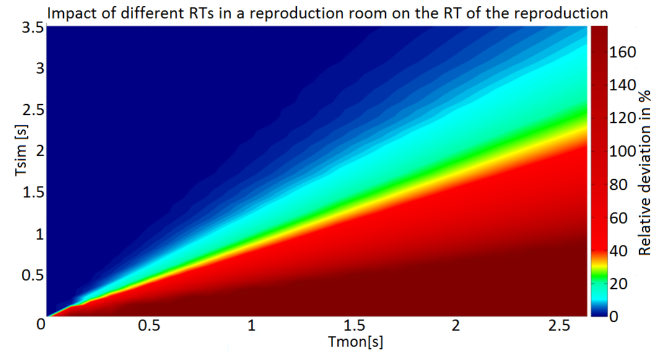

Figure 3: Results of a simulation series. The absorption coefficient in the virtual scene and in the reproduction room was changed and the over-whole IR was calculated. The color indicates the deviation of the RT of the whole reproduction from the target RT (i.e. the RT in the virtual scene).

The results for the early decay time (EDT) can be found in figure 4. The EDT is known for representing the perceived reverberation [3]. The results show that for the EDT stricter requirements have to be met. Regarding to a 10% JND a virtual scene with a EDT of 1.5s would demand a reproduction room with an EDT less than 0.6s. This is plausible as the overall IR increases especially for the early part for longer RTs in the reproduction room as can be seen in figure 2.

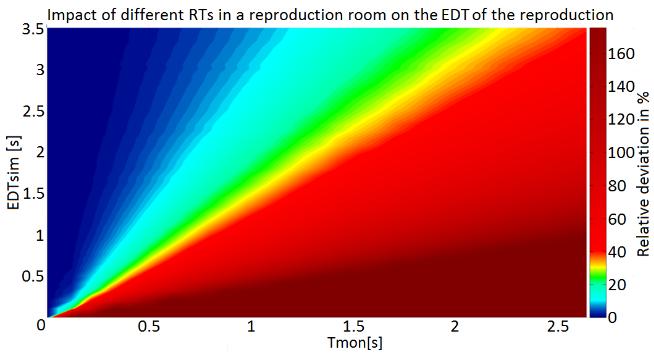

Figure 4: Results of a simulation series. The absorption coefficient in the virtual scene and in the reproduction room was changed and the overall IR was calculated. The color indicates the deviation of the EDT of the whole reproduction from the target EDT (i.e. the EDT in the virtual scene).

Figure 5 shows the results for an investigation on the clarity parameter C50 which relates the energy arriving during the first 50ms to the energy arriving after 50ms. It has to be noted that the deviation is now given in dB. A JND value of 1 dB is found in literature [4]. Clarity is very sensitive to the interaction between reproduction room and virtual scene as it strongly depends on the early energy.

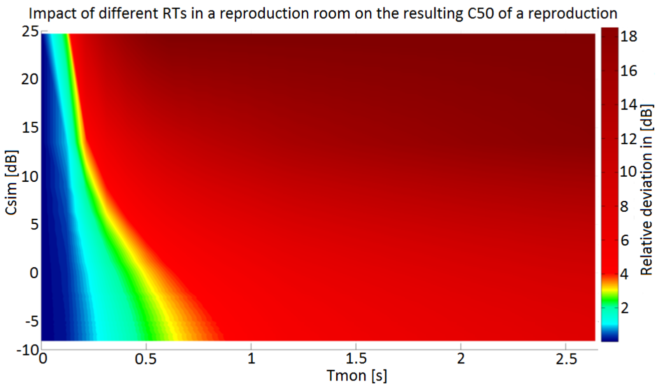

Figure 5: Results of a simulation series. The absorption coefficient in the virtual scene and in the reproduction room was changed and the overall IR was calculated. The color indicates the deviation of the C50 of the whole reproduction from the target C50 (i.e. C50 in the virtual scene).

Compensating the listening Room

As shown in the previous section, the selection of appropriate listening rooms for a generally valid reproduction of also smaller or dry virtual scenes is very limited. Most rooms will not fulfill these criteria, so that a way of compensating the errors introduced by the reproduction rooms are discussed in this section. The two main differences between the IR in the virtual scene and in the overall reproduction are early reflections behavior and the slope of the energy decay curve. Early reflections of a source exciting the reproduction room can be canceled out by using active cancellation. Successful performance of this technique can be found in [7] for crosstalk cancellation systems and in [6] for wave field synthesis, the latter one only in a 2-D set-up. A novel method for compensating the energy decay curve of the overall reproduction can be found in [5]. The absorption coefficients are adapted by matching the resulting energy decay curve to a target energy decay curve. The presented adaption allows a natural sounding modification of the virtual RIR response.

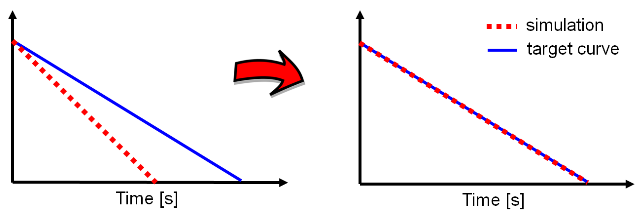

Figure 6: Schematics of the iterative adaption of the material characteristics in a virtual room to match the target energy decay curve.

Conclusion

Basic investigations and considerations regarding loudspeaker-based reproduction of room auralizations in real, non-anechoic environments were made with a focus on representative acoustic parameters. Design criteria were given to find minimum requirements for a reproduction room depending on the virtual room that is to be reproduced. The other way around minimum requirements for a virtual scene can be found depending on a given reproduction room. A method for optimizing reproductions by using active cancellation methods for early reflections and an adaption of the energy decay curve of the reproduction by adapting the energy decay in the virtual room was proposed.

References

[1] Masiero, B.: Individualized Binaural Technology. Dissertation, RWTH Aachen University (2012)

[2] Wefers, F.: Audiotechnik des aixCAVE Virtual Reality-Systems. 41. Jahrestagung für Akustik (DAGA 2015), Nürnberg (2015)

[3] Kuttruff, H.: Room Acoustics. Fourth Edition (1999). Taylor & Francis Group

[4] Bradley, J.S.; Reich, R.; Norcross, S.G.: A just noticeable difference in C50 for speech. Applied Acoustics, Vol. 58, Issue 2, p. 99-108 (1999)

[5] Pelzer, S.; Vorländer, M. : Inversion of a room acoustics model for the determination of acoustical surface properties in enclosed spaces. In Proceedings of Meetings on Acoustics Vol. 19, No. 1, p. 015115 (2013)

[6] Spors, S.; Buchner, H.; Rabenstein, R.; Herbordt, W.: Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering. Journal of the Acoustical Society of America 122, 354 (2007)

[7] Jungmann, J.O.; Mazur, R.; Kallinger, M.; Mertins, A.: Robust Combined Crosstalk Cancellation and Listening-Room Compensation. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (2011)