Perceptual Comparison of Measured and Simulated Sound Fields in Small Rooms

Marc Aretz, Lucas Jauer

Institut für Technische Akustik, RWTH Aachen University, 52056 Aachen, Deutschland, Email: mar@akustik.rwth-aachen.de

Introduction

Sound field simulations in small rooms, which are conducted by means of energy based geometrical acoustics (GA) tools, generally suffer from a physically inaccurate representation of the low to mid frequency range where the sound field is dominated by just a few distinct eigenmodes. In order to alleviate this problem we developed a combined simulation approach, where the frequency range below the Schroeder frequency f_s is simulated by a finite element (FE) approach (to realistically capture the modal structure of the sound field) and the frequency range above is simulated using a hybrid GA approach.
By means of subjective listening tests the present study investigates if the presented combined (FE-GA) approach yields "more realistic" simulation results than are obtained with purely GA based simulation tools. We therefore carried out measurements and simulations in three differently sized small rooms with largely varied acoustic characteristics. In the listening tests a perceptual comparison of measured and simulated sound fields was carried out, where the listeners were asked to assess the plausibility of the sound stimuli as well as the similarity between measured and simulated results.

Simulation Methods and Test Rooms

In order to include a wide variety of room acoustic conditions (both regarding reverberation as well as complexity of geometry and boundary conditions) in our study, sound fields have been measured and simulated in a car passenger compartment, a reverberation chamber and a control room of a recording studio. Table 1 gives an overview of the volume, surface area and measured and simulated reverberation times (RT). All simulations have been carried out using the FE fluid solver *SoundSolve* and the hybrid GA software *RAVEN*, which combines an image source method for the computation of early-order reflections with a stochastic raytracing algorithm for determining the reverberant sound field. The simulation tools are currently developed at ITA of RWTH Aachen University. Both FE and GA simulations account for the source characteristics (i.e. source directivity and frequency characteristics) as well as the acoustic, frequency dependant reflection characteristics of the room boundaries by using either measured data or calculated data based on idealized theoretical models. Especially for the recording studio and the car passenger compartment the acquisition of all relevant material and source data was tedious and despite using different measurement and calculation techniques not all surfaces could be characterized satisfactorily. Details on the data acquisition for these rooms can be found in [1] and [2] respectively.

Table 1: Characteristics of studied rooms

	Volume m^3	Surface m^2	RT_{mid} sec, meas.	RT_{mid} sec, sim.
passenger compartment	2.6	16.7	0.059	0.075
reverberation chamber	123.4	178.0	5.782	5.971
recording studio	58.2	127.0	0.140	0.309

Listening Tests

In order to better quantify differences between measured and simulated auditory events a suitable terminology and corresponding measures have to be defined that can be used in subjective listening tests. With respect to the terminology we suggest that the quality of a simulated auditory event shall be rated with regard to its

- **plausibility**
- **realism with respect to the listeners expectations**
- **similarity to the corresponding measured event**

The Rating in these categories can be expanded to different aspects of the auralization like localization, timbre and sense of space. In this study three different types of listening tests are proposed to evaluate the sound stimuli on the basis of this classification. The design of the listening test on plausibility and the according results are presented in the following paragraph. Realism and Similarity tests are only briefly discussed. As test stimuli we used excerpts of anechoic speech, pop music and classical music that were convolved with the measured and simulated room impulse responses for the three rooms. The stimuli were auralized via headphones. 15 probands (9 without and 6 with feedback) participated in the tests.

Plausibility Test

Our plausibility test is adapted from Nilsson [3]. In this test we played matching pairs of measured and simulated sound stimuli to the probands. The probands were informed about the type of room they were currently listening to. After listening to both stimuli in a pair, probands were asked to decide which stimulus of the two was the measurement. We conducted two runs of the listening test. In the first run probands were not given any feedback on their answers, thus there decision could only be based either on artifacts in the simulation or on a discrepancy between the expected and perceived acoustic characteristics in the presented room. In the second run probands were given feedback after each pair whether their answer was right, thus giving them the chance to learn about characteristic differences between measured and simulated stimuli, even if both stimuli by themselves sounded plausible to them. The percentage of right answers was analyzed for both tests and different combinations of stimuli and rooms. The results in fig-

ure 1 show that without feedback the probands were on average not able to recognize the simulation in the pair (overall average recognition rate of $\bar{r} = 0.53$ indicates a random guess). Only in the case of the recording studio with speech as a stimulus the recognition rate is significantly higher than 0.5, which is believed to be due to the overestimation of the RT in the simulated recording studio. On the contrary, with feedback the overall average recognition rate of $\bar{r} = 0.78$ is significantly higher than 0.5. A closer analysis of the results with feedback indicates that the distinction between simulation and measurements was easiest for the speech stimuli, which is attributed to the fact that the room characteristics are more apparent in the short speech pauses than for example in a continuous music signal.
Summing up, the results of this test imply that simulations (both FE-GA and GAonly) and measurements appear equally plausible to the probands as long as the reverberation time characteristics of the simulated room sufficiently matched those of the measured room. On the other hand both simulations appear to include subtle persistent characteristics, that could be correctly assigned to the simulation when the probands were given feedback. Only in the case of the reverberation chamber, where the quality of the FE simulation is exceptionally good thanks to the rather simple boundary conditions, the average recognition rate with feedback turned out to be closer to 0.5 for the FE-GA ($\bar{r} = 0.67$) than for the GA-only simulations ($\bar{r} = 0.89$). This might be interpreted as an indicator that a physically more accurate low frequency response in the FE-GA simulations is also subjectively appreciated, when the quality of the FE simulation is sufficiently high.

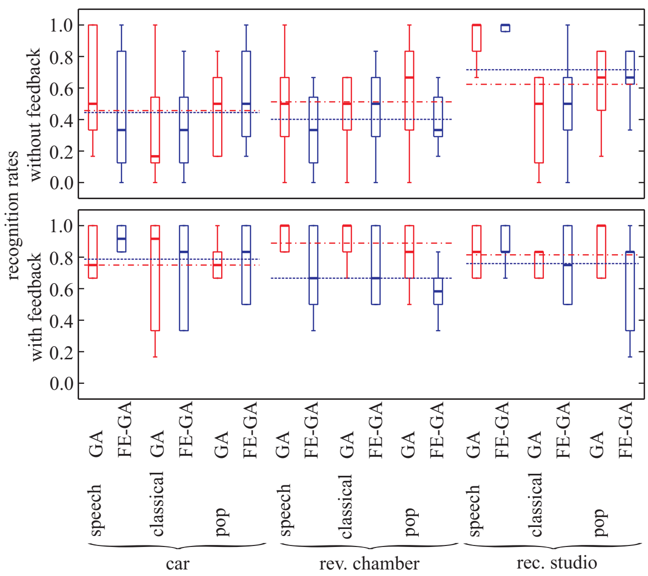

Figure 1: Results of plausibility listening test with and without feedback for different rooms, stimuli and simulation types.
The boxplots show median and 25% and 75% percentils. The whiskers extend to the most extreme data points not considered outliers. Dotted horizontal bars indicate room means for different sim. types.

Similarity and Realism Test

In the "realism test" we go one step further than in the plausibility test. Listeners are asked to assess the perceived realism of binaural simulations presented via headphones while actually being in the environment for which the sound field is simulated. By appealing to all senses and including the visual clues of the room size, source positions and material types, the listeners can assess the "realism" of the stimuli based on their expectation on the sound in this room.
The probably most demanding test simulation-wise is the similarity test. Here, the listeners are asked to evaluate the "similarity" of pairs of measured and simulated sound stimuli on a continuous scale from "identical" to "very different". Similar test procedures have been suggested by Nilsson [3] and Lokki [4]. However, our preliminary results of this test (not presented in this paper) unveiled that the lower bound ("very different") of the scale appeared very fuzzy to the probands and somewhat missing an anchor, thus resulting in a considerable scatter in the results. The test therefore has to be refined. A possibly useful test procedure which compares more than two stimuli at a time (e.g. measured vs. simulated stimuli of different quality) and thus might profit from a ranking effect is presented in Rec. ITU-R BS.1534-1. A thorough analysis of the similarity and realism test will be presented in a subsequent study.

Conclusions

The present paper proposes a terminology and suitable listening tests for the comparison of measured and simulated auditory events. We suggest that the quality of room acoustic simulations shall be evaluated on the basis of its plausibility, its realism with respect to the listeners expectations and its similarity to the according measured auditory events. In particular we investigate if the quality of room acoustic simulations in small rooms can be enhanced by using a combined wave- and ray-based simulation approach to better capture the modal characteristics of the sound field below f_s. Despite a considerable scatter in our listening test results, our analysis provides some indication that an enhanced simulation of the low frequency response using the combined wave- and ray-based approach is also subjectively appreciated by the listeners (see results of rev. chamber with feedback). However, this enhancement can only be achieved provided that realistic boundary conditions for all wall surfaces are known.

References

[1] Aretz, M.: Specification of realistic boundary conditions for the FE Simulation of low frequency sound fields in recording studios. Acta Acustica, 95, 874-882, 2009

[2] Aretz, M., Knutzen, J.: Sound field simulations in a car passenger compartment using combined wave- and ray-based simulation methods. Euronoise (2009), Scotland

[3] Nilsson, M.E. et al.: Perceptual Evaluation of real-time auralization tool. Euronoise (2009), Scotland

[4] Lokki, T., Järveläinen, H.: Subjective Evaluation of Auralization of physics-based room acoustics modeling. Int. Conf. on Auditory Display (2001), Finland