Optimization of Input Parameters for the Real-time Simulation of Room Acoustics – Revisited

Dirk Schröder[1], Björn Starke[1], Hans-Joachim Maempel[2]

[1] *Institute of Technical Acoustics, RWTH Aachen University, 52066 Aachen, E-Mail: dsc@akustik.rwth-aachen.de*
[2] *Audio Communication Group, TU Berlin, 10623 Berlin, E-Mail: hans-joachim.maempel@tu-berlin.de*

Introduction

Immersive environments mostly aim to simulate opto-acoustical scenes in a plausible way. For real-time auralization, methods of geometrical acoustics [1] provide quite accurate results within a reasonable computation time. The best results are achieved by combining deterministic methods for the computation of early specular reflections with stochastic approaches, such as ray tracing, for the computation of the reverberant sound field. The computations must be performed at interactive rates, thus their costs have to be minimized. This is typically done by reducing the order of image sources and/or the number of traced rays, as these parameters strongly influence both, the computational costs and the perceptual accuracy of the simulation. Since the degradation caused by the reduction does not have to be audible, a parameter threshold (PT) is a measure for an optimum parameterization since no perceptible improvement can be achieved above this threshold. The PTs of the number of traced rays have been experimentally determined for three orders of image sources as well as for an acoustical and an opto-acoustical stimulus generated in a CAVE-like environment. Therefore, a listening test applying a criterion-free 3AFC-paradigm with two different assessment methods and with the participation of expert listeners has been performed. The 3×2-design reveals the interaction of relevant technical and perceptual conditions. Results show that a lowered accuracy, hence faster computation time of the simulation, is not noticeable when a convergent opto-acoustical stimulus is presented.

Listening Tests

Due to the promising results of the preliminary studies presented at NAG/DAGA 2009 [2], several refined listening test series were carried out for further investigations [3]. The results presented here include additional listening tests that were accomplished in collaboration with the Audio Communication Group (ACG) at TU Berlin. In this joint series, the same geometrical room model of a concert hall

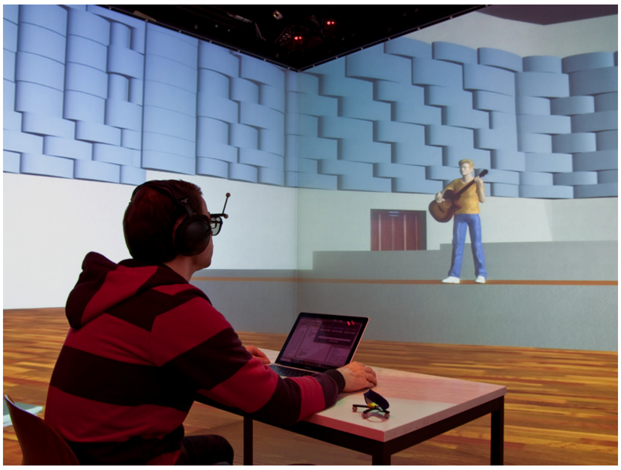
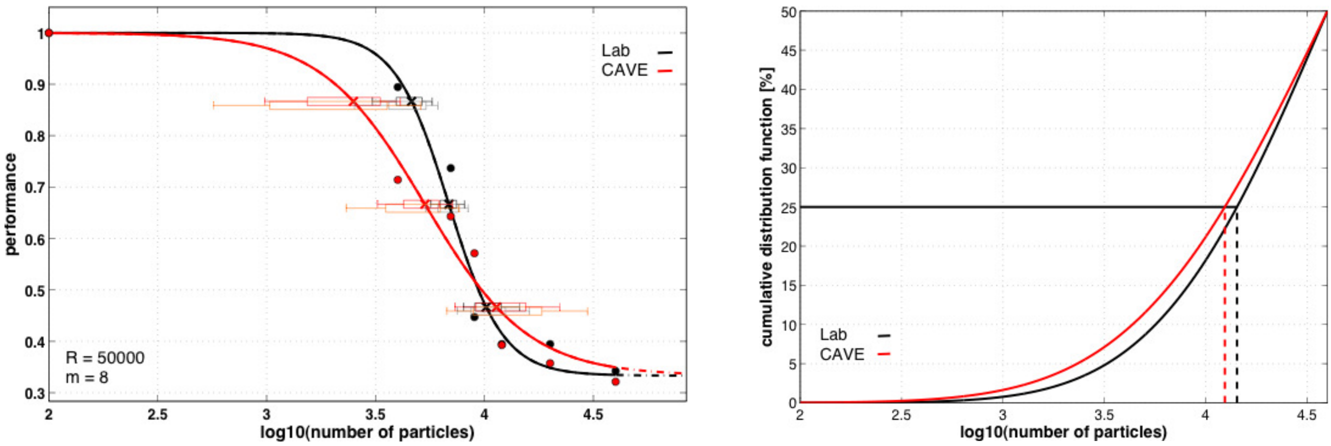


Figure 1: Test subject doing listening tests inside the CAVE-like environment at RWTH Aachen University

(volume 14372m^3, T(Sabine) 1.3 sec) was used as in the preliminary study, but with different auditory stimuli and a different receiver position. 21 Stimuli were created from convolving differently simulated impulse respones with a dry recording. With regard to the facilitation of the detectability of potential artefacts, a slowly picked acoustic guitar was chosen as audio content. For the simulation parameterization, the settings of the preliminary study were refined with focus on the PTs that were obtained at that time. Thus, models based on 3 orders of image sources and, at a time, 7 numbers of particles per simulated frequency band (100, 4000, 7000, 9000, 12000, 20000, 40000) were applied for the production of the stimuli. The numbers range from a level causing an easily detectable distorted sound (100 rays) to a state of the art simulation, which was defined as reference (40000 rays). The listening tests were performed according to the method of constant stimuli following a 3AFC-paradigm. The task was to discriminate between the test stimulus and the reference by determining the position of the differing stimulus. In order to improve its reliability, the test (see Table 1) was repeated once. A GUI guided the subjects through the test.

Measure: PT to the reference [No. of particles]		**Factor I (repeated measures)** Fixed image source order		
		1	2	3
Factor II (grouping variable) Stimulus modality	acoustical (Lab)	N_a=19		
	opto-acoustical (CAVE)	N_{oa}=14		

Table 1: Overview of the test design. N denotes the number of subjects.



(a) Psychometric functions fitting the results of the test series. A performance of 1 means that the subjects have always found the differing stimulus while a performance of 0.33 shows that the subjects have always guessed. Here, the PT relates to a performance of 0.66.

(b) Cumulative distribution function that is derived from the fitted Gaussian-distributed probability density functions. Here, the parameter threshold relates –per definition- to the 25% point of the collective psychometric function of all subjects.

Figure 2: Two methods of PT determination were applied.

All stimuli were played back using a Sennheiser (ITA)/ Stax (ACG) head-phone. The mono-modal listening tests (denoted as ‚Lab' in the following) were performed in a darkened room and subjects were asked to close their eyes in order to avoid distraction from looking at the laptop screen. The second test series were carried out by additionally presenting stereoscopic images of the simulated scene using the CAVE-like environment at RWTH Aachen University [4]. Here, subjects were asked to look at the avatar of the guitar player while listening to the stimuli (see Figure 1).

Evaluation

To find out the respective PTs, the test results were analyzed in two ways, one by using the `psignifit` Matlab toolbox by Wichmann and Hill [5] and the other by analyzing the cumulative distribution function and applying a 25% paradigm [6]. Figure 2 shows exemplarily the computed functions for both methods where the image source order was fixed (order 3) and the number auf rays was varied. Here, the number of rays is always plotted logarithmically in order to stay perceptionally adequate.

Figure 2(a) shows the psychometric functions for both test series, Lab and Cave, obtained by the maximum-likelihood estimation method after Wichmann and Hill. The red/black dots show the average performance of the subjects in respect to the stimulus. The 68.2% and 95.4% confidence intervals are given in horizontal bars at the 0.2, 0.5 and 0.8 level of the respective psychometric functions. The PT relates to a performance of 0.66 which results for the given example in 7900 rays (Lab) and 6300 rays (CAVE), respectively. The cumulative distribution functions shown in Figure 2(b) are derived from the fitted Gaussian-distributed probability density functions and represent the collective psychometric function of all subjects. Applying the 25% rule leads to a PT of 14300 (Lab) and 12400 (CAVE) rays in the given example. The higher number of rays is not surprising as the 25% rule yields a much stronger constraint for determining PTs in contrast to the approach by Wichmann and Hill.

Most noticeable, all functions showed similar behavior as in the preliminary tests. All PTs were significantly shifted to a lower number of rays if the visual stimuli were additionally presented (again an average of 2000 rays was observed). The total number of rays were revised according to the new assessment method, resulting in 12500 rays per frequency band. This relates to a workload of only 2 s for the ray tracing (10 frequency bands with 12500 particles each) on a standard personal computer (Intel Core 2 Quad Q6600/2.4 GHz). However, these results are only generalizable for the presented type of room geometry.

References

[1] Vorländer, M., *Auralization*, Springer Berlin, 2007

[2] Schröder, D., *Optimization of input parameters for real-time room acoustics simulation*, Proceedings of the NAG/DAGA, Rotterdam, The Netherlands, 2009

[3] Starke, B., *Untersuchung des Einflusses von Parametern akustischer Simulationen auf die Wahrnehmung in virtuellen Umgebungen*, Master Thesis, RWTH Aachen University, Germany, 2009

[4] Schröder, D., Wefers, F., Pelzer, S., Rausch, D., Vorländer, M., Kuhlen, T., *Virtual Reality System at RWTH Aachen University*, Proceedings of the ISRA, Melbourne, Australia, 2010

[5] Wichmann, F. A. & Hill, N. J., *The psychometric function: I. Fitting, sampling and goodness-of-fit*, Perception and Psychophysics 63(8), 1293-1313, 2001

[6] Gelfand, S. A., Hearing. An Introduction to psychological and physiological acoustics, 4th edition, Dekker New York, 2004