INFLUENCE OF WALL IMPEDANCE MODELING ON PREFERENCE RATING OF A VIRTUAL ACOUSTIC ENVIRONMENT

Jan Joschka Wohlgemuth
jj.wohlgemuth@gmail.com

Christoph Hold
christoph.hold@telekom.de

Fabian Seipel
fabian.seipel@gmail.com

ABSTRACT

This paper presents the results of a listening test comparing three different approaches to model boundary layers in binaural acoustic simulations of cubic rooms. Simulations included an image source model and a stochastic decay. Results show enhanced preference ratings of emulations that used Butterworth lowpass filter or porous layer modeling over constant linear wall absorption behavior. The porous layer model is the most realistic and reaches the highest preference score.

1. INTRODUCTION

Acoustic room simulations have become an essential part of engineering or construction and especially play an important role in the upcoming context of virtual reality content production as an elegant and efficient way to predict acoustic environments. However, those simulations are neither perfect nor without drawbacks [1]. A simulation can only be perceived as realistic if room geometry, absorption and diffusion behaviour as well as other factors in the computer model reflect the properties of real enclosed spaces. Mostly, they remain flawed because as for now, only few computer based simulations take care of more complex wave effects like interference and diffraction [2]. In order to enhance listening experiences regarding realism and authenticity, simulation models need to take more acoustic effects of real-world enclosed spaces into account. Reflection effects evoke different results depending on the boundary element surfaces and their corresponding highly specific frequency responses. Consequently, we expect frequency dependent simulations of wall absorption and reflection to reach higher preference ratings among experienced as well as inexperienced listeners.

2. METHODS

The following steps were taken to binaurally simulate a three dimensional environment of a cubic room:

- Specifying anechoic audio material and room parameters (dimensions, source- and receiver position)
- Modeling early reflections with the image-source model (ISM) given by Allen[3] and Lehmann[4]
- Convolving every image source with the corresponding head related transfer function from the FABIAN library [5]
- Modeling late reverberation decay by stochastic decay with temporal envelopes
- Considering binaural coherence for late decay
- Modeling of the three different boundary layer impedances compared in the listening test
- Convolving ISM impulse responses M-times with frequency dependent boundary layer model impulse response. Hereby, M is the number of times a sound wave arriving at the receiver has been reflected by a boundary layer
- Convolving stochastic decay impulse responses N-times with frequency dependent boundary layer model impulse response. N is the mean number of times the sound waves arriving at the receiver been reflected by a boundary layer according to Lindau [6] at mixing time.
- Blending both models by cross fading the resulting impulse responses at the so-called mixing time, investigated by Lindau [6], to create a final room impulse response (RIR)
- Apply a simple generic headphone equalization filter for circum-aural or supra-aural headphones constructed from a parametric equalizer

2.1. Image Source Model

Binaural room impulse responses in a cubic room were obtained by superposition of image sources which can be represented by delta pulses with amplitude a and delay τ, convolved with their corresponding head related impulse response (HRIR) from the FABIAN library [5] depending on the angle of incidence (φ, ϑ):

$$\mathrm{BRIR}(t) = \sum_{u=0}^{1} \sum_{l=-N}^{N} a(u,l) \cdot \delta[t - \tau(u,l)] * \mathrm{HRIR}(\varphi(p_{SSQ,rel}), \vartheta(p_{SSQ,rel}), t) \qquad (1)$$

The image source model was implemented for room dimensions of L_x = 10m, L_y = 10m, L_z = 5m, source position x_s = 8m, x_s = 8m, x_s = 2,5m and receiver position x_r = 2m, x_r = 2m, x_r = 1m. Figure 1 shows the ISM for this setup and summing index N=1.

For reasons of further individual processing, the ISM produces a set of *M* individual impulse responses for each group of image sources with the same number *M* of simulated reflections. In this case frequency dependent absorption of boundary layers can be simulated by applying *M* filters to all image sources which would have been reflected *M*-times. A summing index of *N* = 1 already produces a maximum number of *M* = 9 reflections and therefore *M* = 9 groups of image sources as well as *M* = 9 impulse responses. Figure 1 displays the resulting BRIR's grouped by number of reflections as well as the BRIR for direct sound (IR dry).

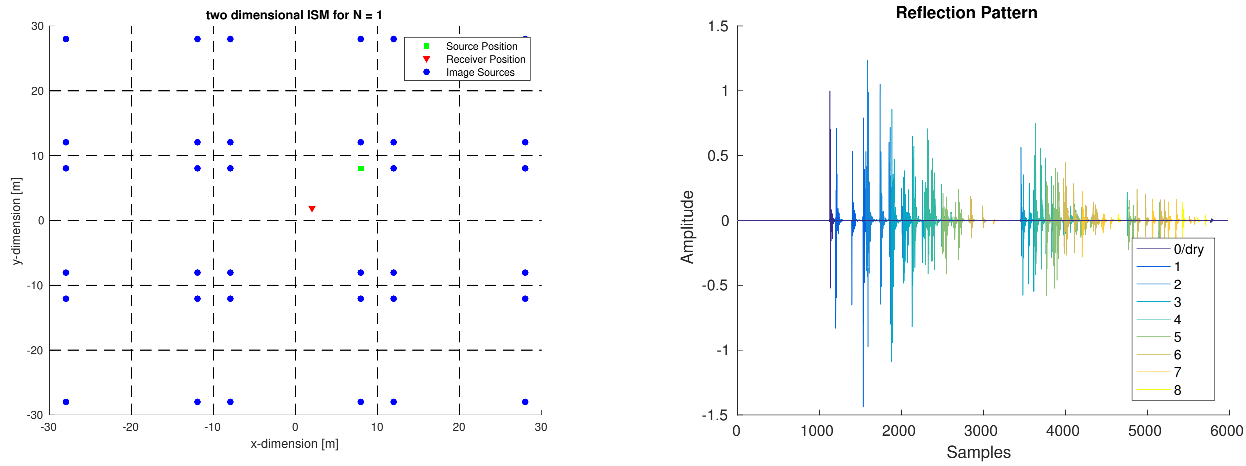

Figure 1: Image Source Model sketch (left) and its output (right). Output is colored according number of underlying reflections ($M = 0 \ldots 8$)

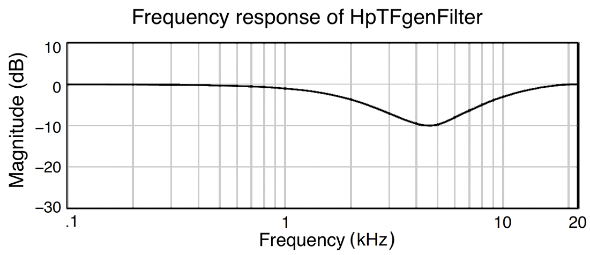

Figure 2: Magnitude response of the simple generic headphone compensation filter

2.2. Stochastic Decay

Under the assumption that the sound field becomes diffused, stochastic decay takes over after the sp-called mixing time, calculated according to Lindau [6]. The boundary layer absorption factor used here is the mean absorption coefficient of the employed model, averaged over nine octaves starting at 62.5 Hz. For the basic signal per channel we chose white Gaussian Noise because of its flat frequency response and maximum decorrelation. Sound energy in an enclosed space with Volume V decreases exponentially from a starting value w_0 [7, eq. 5.7 and 5.8]:

$$w(t) = w_0 \cdot exp(-\frac{c}{4}\frac{A}{V}), \tag{2}$$

where A denotes the *equivalent absorption area* of the room, in our case derived from mean alpha. The total length of the reverb tail is given by [7, eq. 5.9]:

$$T = 0.161 \cdot \frac{V}{A} \tag{3}$$

The resulting function $w(t)$ applies as a temporal envelope over the white noise. Since reverb does not reach both ears independently, the next step in the algorithm produces binaural coherence between both, so far statistically independent, ear signals. According to Borß and Martin [8], this can be modeled by a head size and frequency dependent signal crossfeed, whereas above a certain frequency head shadowing effects guarantee complete independence of both ear signals.

2.3. Boundary Layer Models

Three models for the boundary layer of the virtual room are implemented for this listening test. Model A describes a layer with constant absorption coefficient α of 0.25. One quarter of the energy of the incident wave is absorbed and three quarters are reflected regardless of the angle of incidence, frequency, amplitude and phase of the wave. Model B is a first order Butterworth lowpass filter with a 3 dB cutoff frequency at 4 kHz. Model C is based on an Vorlaender's impedance model [9] of a porous layer in front of a hard wall.

$$Z \quad = \quad \omega'' - jZ_0 cot(kd) \tag{4}$$

The modeled porous layer is 3 cm thick and has a flow resistivity of 50 Rayl/cm. According to Moeser [10] this value is situated in the middle of the technologically relevant range of 5 to 100 Rayl/cm. The 3dB cutoff frequency of the resulting transfer function is 4 kHz. As it turns out, the absorption factor curve of the modeled boundary layer is similar to the measured curve of 5 mm needle felt stuck to concrete, published in [9].

	Octave band frequency in Hz					
Material	250	500	1k	2k	4k	8k
Needle felt	0.02	0.05	0.15	0.3	0.4	0.4
Model C	0.01	0.03	0.09	0.3	0.5	0.6
Model B	0.02	0.03	0.08	0.2	0.5	0.8

Table 1: Absorption coefficients of Needle felt 5 mm stuck to concrete, of the modeled porous layer in front of a hard wall (both according to [9]) and of the lowpass filter (ModelB).

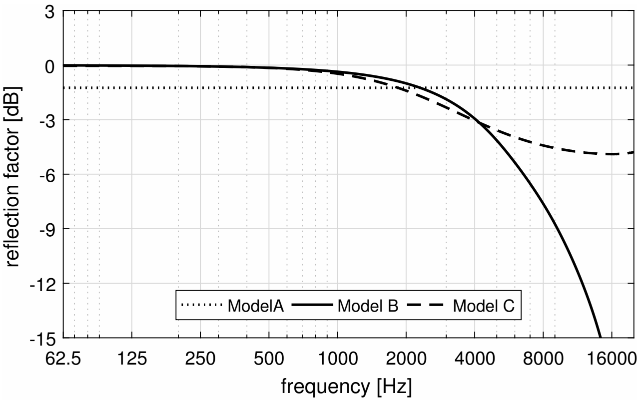

Figure 3: Reflection factors of the 3 boundary layer models. Model A has a constant absorbtion factor of 0.25, model B is a butterworth lowpass filter and model C is based on an impedance model of Vorlaender [9] of a porous layer in front of a hard wall.

2.4. Auralization

At this point we have calculated the M impulse responses of the ISM, the stochastic decay and the boundary layer impulse responses. The ISM impulse responses are first convolved M times with the boundary layer IR and then summed up. In order to predict the number of convolutions of boundary layer IR and stochastic decay, we calculated the mean number of reflections at the boundary layers for all incident sound waves before arriving at the receiver's position at the mixing time according to [6].

Consequently, the two resulting impulse responses are cross faded using a linear ramp of 10ms. Linear ramps were also employed in the evaluations of Lindau [6]. To compensate for the electro-acoustic behaviour of the average headphone, a simple generic compensation filter was utilized (Figure 2). The result of the above process is a stereo impulse response, that any anechoic audio signal can be convolved with in order to simulate the desired acoustic environment.

2.5. Listening Test

For the listening test, subjects were presented with a voice sample which was convolved with each of the different binaural room impulse responses from the three calculated models. The subjects used their individual headphone setups for the playback of the stimuli. Each subject was asked to rank the stimuli on a scale from 1 to 3 (1 meaning the least preferable and 3 the most preferable) by preference. The group of test subjects included five students, aged between 23 to 32 years.

3. RESULTS

The results of the test, summarized in table 2, indicate the lowest preference rating by all subjects for the constant linear absorption model. Its arithmetical mean $\mu_A = 1$ of the ranking points is by far the lowest without any deviation $\sigma_A = 0$. Two of the inexperienced subjects stated they could not clearly hear a difference between the Butterworth lowpass filter and the porous layer model, but they both preferred the first mentioned over the second. Model B achieved an arithmetical mean $\mu_B = 2.4$ with standard deviaion $\sigma_B = 0.49$, whereas Model C was ranked first with slightly higher mean score $\mu_C = 2.6$ but the same standard deviation $\sigma_C = 0.49$. Figure 4 shows the results of the listening test. Subjects reported model A to sound metallic, model B to sound muffled and smeared, and model C to sound clear and most authentic.

Table 2: simulated stimuli as used in listening test

Stimulus	Boundary Layer Models
A	Constant Linear Absorption
B	Butterworth Lowpass Filter
C	Porous Layer Model

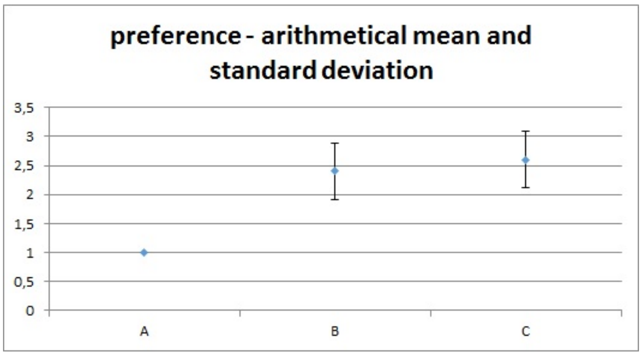

Figure 4: Results of the qualitative listening test. Mean score and standard deviation

4. DISCUSSION

The evaluation of the performance rating shows, that the influence of the boundary layer impedance simulation has a remarkable impact on the preference rating. Table 1 displays that the smallest difference of absorption coefficients occurs between the modeled porous layer and a measured surface absorption (needle felt). The bigger the difference between the behavior of a real absorber and the model, the less did the participants like the stimulus. Model B, a first order Butterworth filter, which is very simple to digitally implement, ranked only slightly worse than the porous layer model. Unexperienced participants were not able to distinguish between the filter and Vorländer's complex model. Experienced audiophile subjects however generally preferred Model C.

4.1. Limitations

The idealization of parameters used in simulations is mandatory. With regards to computational expense and model constraints there need to be certain limitations to the complexity of the implementation.
For comparison purpose, a reference HRTF from measured data of the FABIAN head and torso, was utilized. Any deviation of a subject's anthropometry from the FABIAN anthropometry may lead to decreased sensation. The generic headphone equalization is not optimally implemented. Since subjects used their own individual headphones, this may have led to deviations in the transfer paths amongst the subjects. Each wall in the binaurally simulated

room is identical. There is no distinction between more and less acoustically hard surfaces. In real world rooms however, the ceiling often differs from other walls in terms of absorption behaviour due to constructional reasons. In addition no objects or any type of furnishing inside the room were included in the simulation. Such objects alter the sound field by reflection, absorption and reverberation, which would presumably result in an increase of perceived authenticity. Due to the small number of participants, the listening test is also far away from being representative. It may indicate trends but cannot obtain solid results. All techniques used in the auralization are based on models. Every deviation from the model to a real environment has the potential to lead to a loss of perceived authenticity.

5. REFERENCES

[1] Pavel Zahorik, "Perceptually relevant parameters for virtual listening simulation of small room acoustics," *The Journal of the Acoustical Society of America*, pp. 779–779, 2009.

[2] A. K. Kloasak, "Computer modeling and simulation of sound field in 24 rectangular concert halls," *Acoustics'08 Paris*, p. 3099–3104, 2008.

[3] Jont B Allen and David A Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, pp. 943–950, 1979.

[4] Eric A Lehmann and Anders M Johansson, "Prediction of energy decay in room impulse responses simulated with an image-source model," *The Journal of the Acoustical Society of America*, vol. 124, no. 1, pp. 269–277, 2008.

[5] Fabian Brinkmann, Alexander Lindau, Markus Müller-Trapet, Michael Vorländer, and Stefan Weinzierl, "Cross-validation of measured and modeled head-related transfer functions," *Fortschritte der Akustik–DAGA*, pp. 1118–1121, 2015.

[6] Alexander Lindau, Linda Kosanke, and Stefan Weinzierl, "Perceptual evaluation of model-and signal-based predictors of the mixing time in binaural room impulse responses," *Journal of the Audio Engineering Society*, vol. 60, no. 11, pp. 887–898, 2012.

[7] Heinrich Kuttruff, *Room acoustics*, Spon Press, London, New York, 2009.

[8] Christian Borß and Rainer Martin, "An improved parametric model for perception-based design of virtual acoustics," in *Audio Engineering Society Conference: 35th International Conference: Audio for Games*, Feb 2009.

[9] Michael Vorländer, *Auralization: fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*, Springer, Berlin, 1st ed edition, 2008, OCLC: ocn171111969.

[10] Lothar Cremer and Michael Moser, *Technische Akustik*, VDI-Buch. Springer, Berlin, 5. aufl edition, 2003.