PERCEPTUAL AND ROOM ACOUSTICAL EVALUATION OF A COMPUTATIONAL EFFICIENT BINAURAL ROOM IMPULSE RESPONSE SIMULATION METHOD

Torben Wendt,
Acoustics Group, Medical Physics
and Cluster of Excellence "Hearing4all",
University Oldenburg
Oldenburg, Germany
`t.wendt@uni-oldenburg.de`

Steven van de Par,
Acoustics Group and Cluster of Excellence "Hearing4all",
University Oldenburg
Oldenburg, Germany
`steven.van.de.par@uni-oldenburg.de`

Stephan D. Ewert,
Medical Physics and Cluster of Excellence "Hearing4all",
University Oldenburg
Oldenburg, Germany
`stephan.ewert@uni-oldenburg.de`

ABSTRACT

A fast and perceptively plausible method for synthesizing binaural room impulse responses (BRIR) is presented. The method is principally suited for application in dynamic and interactive evaluation environments (e. g., for hearing aid development), psychophysics with adaptively changing room reverberation, or simulation and computer games. In order to achieve a low computational cost, the proposed method is based on a hybrid approach. Using the image source model (ISM; Allen and Berkley [J. Acoust. Soc. Am. Vol. 66(4), 1979]), early reflections are computed in a geometrically exact way, taking into account source and listener positions as well as wall absorption and room geometry approximated by a "shoebox". The ISM is restricted to a low order and the reverberant tail is generated by a feedback delay network (FDN; Jot and Chaigne [Proc. 90th AES Conv., 1991]), which offers the advantages of a low computational complexity on the one hand and an explicit control of the frequency dependent decay characteristics on the other hand. The FDN approach was extended, taking spatial room properties into account such as room dimensions and different absorption characteristics of the walls. Moreover, the listener orientation and position in the room is considered to achieve a realistic spatial reverberant field.

Technical and subjective evaluations were performed by comparing measured and synthesized BRIRs for various rooms. Mostly, a high accuracy both for some common room acoustical parameters and subjective sound properties was found. In addition, an analysis will be presented of several methods to include room geometry in the FDN.

1. INTRODUCTION

Room acoustical simulations are desirable for many purposes, such as developing or testing signal processing algorithms, or to e. g. test the effect of reverberation on speech intelligibility. Furthermore, they are of interest for audio-visual simulation environments (e. g, for training and rehabilitation) and in entertainment, e. g. in computer games, all requiring a real-time adaptation of the virtual environment, depending on the movement of the listener and/or the sound sources.

One traditional way to emulate the acoustics of a certain room is to measure binaural room impulse responses (BRIRs) and to convolve dry source signals with the BRIRs. However, such measurements are time consuming and their usage is restricted to static scenarios. Furthermore, one is restricted to actually existing rooms. Alternatively, room acoustics can be simulated, enabling different degrees of realism, ranging from simple artificial reverb generation to complex room acoustical simulation (image source model [1], CATT [2], ODEON [3]), even for dynamic scenarios (e. g. [4], [5], [6]).

Depending on the application, physically correct rendering of a soundfield is required or a perceptually convincing auralization, implying plausibility and authenticity, is sufficient. For room simulations used in psychoacoustic research, rehabilitation or in computer games perceptual aspects are most important, implying accordance of room acoustical parameters, e. g. reverberation time, definition, and measures like speech intelligibility. In this case simplifications can be made to reach computational efficiency allowing for real-time rendering of dynamic acoustic scenes, in which the positions of sources and receivers can be changed interactively.

Several approaches exist to synthesize room impulse responses. If the wavelength of a sound is small compared to the characteristic dimensions of reflecting objects, concepts of geometric acoustics (GA), such as the image source model [1] or the ray tracing method [7] can be applied. Both methods have been used and further developed in various room acoustics simulation algorithms, mostly as hybrids together with other algorithms (e. g. [3], [8]). However, these methods still have high computational complexities.

If the exact room geometry is neglected, artificial reverberation can be synthesized very efficiently and with predefined reverberation time. Here, a common approach are feedback delay networks (FDNs), based on Schroeder's pioneering work on parallel delay lines with feedback [9] and further developed (amongst others) by Stautner and Puckette [10] and Jot and Chaigne [11].

One way to achieve real-time performance while maintaining the advantages of the more “accurate” GA-based BRIR synthesis and reverberation algorithms is their combination in a hybrid approach: The initial part of the impulse responses is computed based on a GA method. The reverberant tail is generated by a more effective reverberation algorithm. Perception motivates such an approach as the early sound reflections create the impression of a certain spatial source width on the one hand and support speech intelligibility on the other hand. The following reverberant tail contains diffuse reflections and its frequency dependent decay characteristics conveys information about the wall absorption and room size.

Here, a hybrid approach was evaluated which combines the image source model (ISM) for a shoebox geometry to simulate early reflections up to a low order, and an FDN for creating a diffuse reverberation tail. The FDN was extended to be directly linked to the room geometry used in the ISM, and to be able to spatially render the reverberation tail in order to generate BRIRs. For low ISM orders, BRIRs can be simulated very efficiently with this approach. In technical and subjective evaluations, the ability of the algorithm to create plausible and authentic simulations was assessed for single and connected (coupled) shoebox rooms. Two different approaches for spatial reverb distribution rendering were compared, taking room dimensions and receiver position into accout.

2. SIMULATION METHOD

A hybrid approach [12] was used to synthesize BRIRs. Early sound reflections are computed by an image source model up to a low order. The late reverberation was generated by a feedback delay network.

The auralization steps are described explicitly for the case of headphone presentation, reflected by the application of head-related impulse responses (HRIRs). The adaptation to arbitrary loudspeaker-based playback systems, such as higher order ambisonics or wave field synthesis, can be easily achieved by replacing them by respective loudspeaker-controlling functions.

2.1. Image source model

The ISM regards a sound reflection as the direct sound of a mirrored version of the original source. This so-called image source differs from the original source by its time delay and its attenuation due to the distance to the receiver, as well as the respective wall reflection coefficient. The sound of an image source is reflected again at other walls, creating higher order image sources. In this way, arbitrarily complex reflection paths can be modeled.

The ISM implementation in the proposed simulation method is restricted to empty shoebox-shaped rooms, where the six wall surfaces are represented each by frequency dependent absorption coefficients. These shoebox-shaped rooms enable a very efficient calculation of image source positions in comparison to arbitrary room geometries [13]. Nevertheless, for a shoebox room the number of image sources up to reflection order N is of order $\mathcal{O}(N^3)$, which considerably affects computational efficiency for higher reflection orders. Another limitation of the ISM is that it inherently assumes only specular instead of diffuse reflections, although they are of importance to describe room acoustics.

In the ISM implementation, the following signal processing steps are performed for each image source: A “$1/\text{distance}$” attenuation factor and a time delay due to distance to receiver; an “effective reflection filter”, being the (frequency domain) product of all wall reflection coefficients that are involved to “create” the current image source; an HRIR, according to the azimuth and elevation position of the image source relative to the receiver’s head orientation. Finally, the binaural signals for all image sources are added up to one two-channel output.

2.2. Extended feedback delay network

The extended FDN used here is based on the general multichannel network as suggested by Jot and Chaigne [11] and consists basically of a set of parallel delay lines whose outputs are fed back via a feedback matrix A.

The number of parallel channels (delay lines) was set to 12, with four channels associated to each (showbox) room dimension (two channels per wall) reflected in several parameter choices. Firstly, the delay units $\tau_j, j \in \mathbb{N}_{\leq 12}$ were directly related to the room dimensions via sound propagation speed (plus a random jitter per channel). Secondly, the absorption filters with transfer functions H_j^{abs} simulate the frequency dependent sound attenuation due to the wall reflections and air absorption. After Jot and Chaigne [11] the frequency dependent reverberation time $T_{60}(f)$ conveyed by the resulting RIR is controlled explicitly by the following frequency responses, if all other processing steps are energy preserving:

$$20\lg|H_j^{\text{abs}}(f)| = -60\tau_j/T_{60}(f)\,. \tag{1}$$

In the simulation method, the reverberation time is predicted from the wall absorption coefficients via Sabine’s formula. Thirdly, the feedback matrix A redistributes the outputs back to the input channels. This process is energy preserving if A is an orthogonal matrix. Here, a randomly created unitary matrix was chosen, providing a high variety of pulse amplitudes.

Two last processing steps per channel, referred to as “binauralization steps”, extend the FDN to introduce spatiality distributed and externalized reverberation. (1) Via HRIR filtering the FDN channels are mapped to 12 points (directions) around the head, with two points positioned on each wall. (2) Reflection filters—identical to those applied to the first order image sources in the ISM—simulate a direction dependent sound intensity of reverberation, due to the different acoustical wall properties.

Two possible principles are suggested to map the 12 directions around the receiver’s head, which are sketched in Fig. 1 (microphone symbol: recveiver, big “⊗”: direct sound source, small “⊗”: reverb source). The first one (lhs of Fig. 1) is called “cube” condition. Here, the 12 directions are mapped to points on a cube around the receiver’s head. The cube always moves with the receiver (receiver is always in its centre) and is axis aligned with the room.

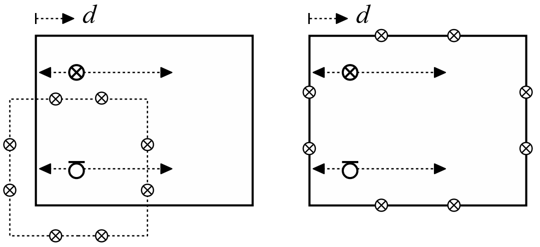

Figure 1: *Illustration of two possible techniques of spatial reverb distribution: “cube” (lhs) and “box” (rhs) method. See text for explanation. (Arrows will be explained in sec. 3.2.*

In this way all 12 incidence directions are more or less equally distributed around the head. In the second "box" condition the 12 directions are mapped to points on the actual six wall surfaces like depicted in the rhs part of Fig. 1. Here, the sound incidences are warped according to the room dimensions and the actual receiver position. Differences between both methods will clearly be audible for rooms with large differences in dimensions. If not specified, in the following the cube condition will be used as standard rendering method.

2.3. Combination of ISM and FDN

For a smooth transition from the early-reflections part (ISM) to the late-reverberation part (FDN), i. e. a straight decay of the BRIR on dB scale, the energy and initial delay of the FDN input signal have to be suitable. For this purpose the FDN input signal consists of the Nth-order ISM pulses before HRIR filtering. In order to avoid comb-filter coloration effects, which occur if a fixed temporal pattern of pulses is fed into the FDN, the ISM output is distributed to the FDN channels. Because the number of image sources of order N does in general not equal the number M of FDN channels, the ith ISM pulse is fed into the FDN channel $[(i-1)\mathrm{mod} M]+1$.

2.4. Simplified model for coupled rooms

In addition to the single shoebox-shaped room a strongly simplified method to simulate the acoustics of two connected shoebox rooms that are acoustically coupled, e. g. by an open door is suggested.

It is assumed that a source S is located in room 1 and a receiver R in room 2 as depicted in Fig. 2. The sound transmission from room 1 to room 2 is then simulated by a single virtual source S′ located in the door which is exciting room 2 as in the case of the single shoebox simulation described above. The virtual source radiates the monaural impulse response of room 1 for a source position specified by the coupled-room arrangement and a "monaural" receiver R′ inside the open door. Thus, the effective BRIR is obtained as the convolution product of the monaural RIR of room 1 with the BRIR of room 2.

Depending on the source position in room 1, it is either visible or invisible for a receiver in room 2. If it is not visible, no direct sound will arrive at the receiver but only reflections and diffractions. In this case, the direct sound pulse of the RIR of room 1 is discarded in the current approach.

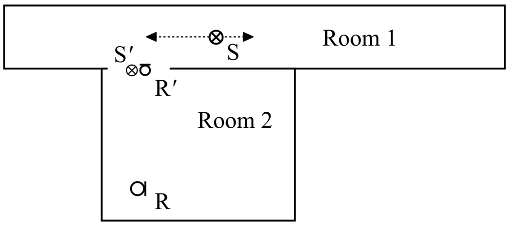

Figure 2: *Sketch of two adjacent rooms, that are acousticly coupled by an open door. See text for explanation. (The arrow will be explained in sec. 3.2.)*

3. EVALUATIONS

Three main aspects of the proposed simulation method were evaluated. Firstly, for a set of real-existing rooms, subjectively rated sound properties of measured and respectively synthesized BRIRs were compared. Secondly, the two approaches to realize the binauralization steps of the extended FDN (see sec. 2.2) were evaluated with respect to binaural parameters. Thirdly, the approach to simulate the acoustics of two coupled rooms was evaluated.

To perform these evaluations, a test-database containing measured and synthesized BRIRs was created. BRIRs were measured for various rooms of different size and reverberation time, as well as a for a few source-receiver configurations in two connected rooms. Additionally, some measured BRIRs were taken from the AIR database [14].

The BRIR measurements were performed using an omnidirectional loudspeaker based on a ring-radiator and an artificial head MK2 by Cortex. Rooms were excited with a logarithmic sweep [15] (50 Hz to 18 kHz) offering removal of nonlinear harmonic distortions from the recorded and inverse filtered signal. BRIRs were each calculated as the mean of BRIRs from 10 single recordings and equalized by the inverse loudspeaker transfer function.

For the BRIR synthesis a single mean wall absorption coefficient was used. It was determined for each room from its reverberation time via the inverse form of Sabine's formula, which ensures that reverberation times of measured and synthesized BRIRs are in good accordance.

The HRIRs used in the simulation were from the same artificial head as used for the BRIR recordings. The database [16] offers HRIRs with azimuth angles in 2° steps at elevations near the equatorial level. Towards the poles, the azimuth angle sampling decreases. Elevation angles are sampled in 2° steps.

A varying synthesis parameter was the maximum image source order. The goal was to find a trade-off between accuracy and computational efficiency.

In the following, for all different rooms and source-receiver configurations, the term "room condition" will be used. In contrast, different types of BRIR synthesis, differing in the choice of simulation parameters, will be referred to as "synthesis conditions". All room- and synthesis conditions will be introduced in the following.

3.1. Subjective sound properties

3.1.1. Room- and synthesis conditions

BRIRs were chosen from four rooms of different size and reverberation time, specified in Tab. 1. For the synthesized BRIRs, the maximum image source order N was varied in $\{1, 3\}$. For one room the BRIR was synthesized only by the ISM with $N = 20$. In the extended FDN, the cube condition was chosen (see 2.2). Two dry source signals, female spoken speech and a guitar play (steel strings) were convolved with the recorded and simulated BRIRs. Presentation sound pressure levels ranged from 60 to 65 dB SPL, depending on the source-receiver distance and the room reverberation.

3.1.2. Subjects and procedure

15 normal-hearing subjects (7 female, 8 male) aged 24 to 32 years participated in the experiment. Sounds were presented via headphones in a sound attenuating booth. Since the synthesis method was implemented as an offline simulation, no head tracking and adaptively changing soundfield was employed.

The sound properties which were to be rated on a seven-point scale were "naturalness" and "room size". A test and a retest were

Table 1: *Rooms, whose BRIRs were used in the subjective evaluation. Reverberation times T_{60} were obtained from measured BRIR (broadband).*

Room	Dimensions (m)	T_{60} (s)
Aula	(12.0, 30.0, 10.0)	4.8
Empty chamber	(1.88, 2.74, 2.82)	2.5
Lecture room	(10.90, 10.80, 3.15)	0.8
Laboratory	(4.97, 4.12, 3.00)	0.3

performed in two sessions, each with a randomized order of presented sounds. Before the actual experiment was performed, sound examples illustrating extremal distinctions of the sound properties had been presented.

3.1.3. Results

Fig. 3 shows the results from the subjective sound property ratings as mean values over all subjects and source signals. Each panel shows results for one sound property. The avarage ratings are plotted for all synthesis conditions against rooms. Error bars indicate inter-subject standard errors.

For naturalness (left panel) ratings differ strongly between rooms. Whereas the BRIRs of the laboratory were rated to sound most natural, lowest naturalness was perceived for the empty chamber. BRIRs of the aula and lecture room were rated to have a medium to high naturalness. Between synthesis conditions, almost no differences are visible, and for most rooms, differences between synthesized and measured BRIRs are very low. Moreover, for some conditions even the synthesized BRIRs were rated to sound slightly more natural than the measured one. This shows, that the proposed simulation method is able to synthesize BRIRs that sound as natural as measured ones. Remaining differences in perceived naturalness between rooms might be due to familiarities of subjects with these acoustic environments in daily life, since the Laboratory sounds as dry as an ordinary living room, whereas the empty chamber sounds rather unusual, even by the measured BRIR. This might also be due to the unusual relation of its very small room size and its high reverberation time (see Tab. 1).

For room size (right panel), again clear differences between rooms are perceived. The order is well in accordance with reverberation times and, except for the empty chamber, with the actual room sizes (see Tab. 1). Differences within synthesis conditions and between synthesises and measurements are practically not existent. This shows that firstly the simulation method is able to represent different room sizes and secondly to achieve this independently from maximum image source order as far as tested.

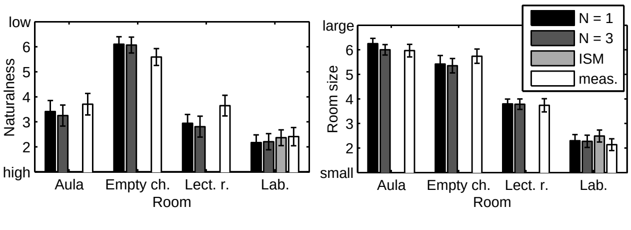

Figure 3: *Subjective sound property ratings of measured and synthesized BRIRs for four rooms, averaged over all subjects and source signals (speech and music). Error bars indicate inter-subject standard errors.)*

As a consequence of the experimental design, no direct mapping of synthesized BRIRs to actual rooms was performed. Given that no head tracking was employed, a potential effect of head rotations on the subjective ratings could so far not be assesed. Future research will apply the system in a real-time environment and will address this issue.

3.2. Evaluation of spatial properties of the extended FDN

3.2.1. Room- and synthesis conditions

Two rooms each with different configurations of source- and receiver positions as well as wall properties, were used to evaluate the spatial reverb rendering. Fig. 1 depicts schematically the conditions for one room. Tab. 2 specifies the room dimensions and the absorption coefficients for (250, 500, 1k, 2k, 4k) Hz. While room 1 has an almost square base area, room 2 represents a long corridor. Side wall absorption coefficients were specified in two different ways: In the “closed” condition, all side wall absorption coefficients were equal as given in line 3 in Tab. 2. By this way, spatial sound properties in azimuth can be investigated in dependence of room geometry in connection to the positions of the receiver and all virtual “reverb” sources. In the “open” condition, the left side was completely open, meaning that no wall was existent. This was technically represented by a broadband absorption coefficient of 0.99, whereas the absorption of all other side walls did not differ from those of the closed condition. By this, the spatial rendering was to be evaluated in a challenging condition for the model.

Also in both rooms and for the closed- and open condition, the distance d of source and receiver to the left wall was chosen to be 0.3 m or 5 m (see Fig. 1). The source was always in the front direction of the receiver, yielding a direct sound with no interaural differences. All differences are thus due to reflections and reverberation.

Since no suitable real rooms were found for BRIR measurements, the ISM with reflection order 20 was used as reference. For all room conditions, the cube- and box condition were compared against each other and the reference. ILDs and IACCs were determined from the BRIR up to the time $\min\{T_{60}(f)\}$, where $T_{60}(f)$ is the frequency dependent reverberation time. Positive ILDs indicate higher signal energy on the right.

Table 2: *Specification of two virtual rooms used for evaluation of the binauralization steps in the extended FDN. See text for further explanation.*

Dimensions room 1	(10.9, 10.8, 3.15) m
Dimensions room 2	(3.9, 30.0, 3.15) m
Absorption side walls	(0.05, 0.10, 0.13, 0.16, 0,22)
Absorption open side	0.99
Absorption floor	(0.03, 0.03, 0.03, 0.03, 0.02)
Absorption ceiling	(0.70, 0.60, 0.70, 0.70, 0.50)

Table 3: *Results of the evaluation of the FDN binauralization steps: Comparison of cube and box method with purely ISM-created BRIRs in terms of ILDs and IACCs. See text for explanation of conditions.*

	ISM	cube	box	ISM	cube	box
ILD (dB)				IACC		
R 1, closed						
$d = 0.3\,\mathrm{m}$	-0.3	$\mathbf{-0.7}$	1.3	0.5	0.5	0.5
$d = 5.0\,\mathrm{m}$	0.3	**0.7**	0.8	0.7	0.7	0.7
R 1, open						
$d = 0.3\,\mathrm{m}$	0.9	**0.9**	1.8	0.9	0.8	0.8
$d = 5.0\,\mathrm{m}$	1.2	1.1	**1.2**	0.8	0.8	0.8
R 2, closed						
$d = 0.3\,\mathrm{m}$	-0.6	$\mathbf{-0.2}$	3.1	0.6	**0.5**	0.4
$d = 5.0\,\mathrm{m}$	-0.6	**0.7**	3.7	0.7	**0.6**	0.5
R 2, open						
$d = 0.3\,\mathrm{m}$	0.5	**1.0**	3.2	0.9	**0.8**	0.6
$d = 5.0\,\mathrm{m}$	0.5	**1.0**	3.0	0.9	**0.8**	0.6

3.2.2. Results

The results for all room- and synthesis conditions are shown in Tab. 3. Comparing results for the cube and box condition, values that are closer to those of the ISM reference, with a difference of at least 0.1 dB (ILD) or 0.1 (IACC) to the less matching condition, are printed in bold face for clarity. To interpret the results just-noticeable differences (JNDs) of ILD have to be considered as, e. g. determined in [17] for musical instruments in several reverberant conditions: 1.0–1.4 dB for $T_{60} = 1.3$ s; 0.8–1.2 dB for $T_{60} = 0.8$ s; 0.4–0.8 dB for the anechoic condition.

For room 1 (“R 1”) (closed) overall small absolute ILD values are observed in the range of the JNDs. It has to be kept in mind, that these ILDs originate from reflections only, given that the direct sound was always located in the front direction. For the box condition ILDs are larger and differ considerably more from the reference. Clear mismatches to the reference ILD are obtained in the close-to-wall position ($d = 0.3$ m). Overall small or vanishing ILDs are plausible for the closed conditions since all side walls are equal in absorption coefficient. For the closed room 2 (“R 2”), both ISM and cube condition show again small absolute ILD values. However, ILDs for the box condition differ clearly from those of the cube and reference conditions. This is not surprising because the majority of virtual reverb sources lie clearly to the right hand side of the receiver (see also scheme in Fig. 1). For the open versions of the rooms, ILDs obtained from ISM-created BRIRs and the cube condition have very similar values. Largest differences are again obtained for the box condition in room 2.

The IACC results reveal overall no distinct differences between the synthesis conditions for room 1. For room 2, where maximum differences are 0.3, the cube condition yields IACCs that are closer to those created by the ISM.

In conclusion, it can be said that the cube condition mostly creates spatially more realistic BRIRs in terms of ILD and IACC than the box condition, especially when the room geometry and receiver position are challenging. In addition, also an informal subjective listening test yielded highest perceptive similarity between the cube and the reference condition.

3.3. Evaluation of simulation of coupled rooms

3.3.1. Room- and synthesis conditions

The two adjacent rooms, an office and a corridor, acoustically coupled by an open door, are specified in Tab. 4 in terms of dimensions and absorption coefficients for (250, 500, 1k, 2k, 4k) Hz. The arrangement of both rooms and positions of source S and receiver R are depicted in Fig. 2. Two source positions were investigated. In the “visible” condition the source is placed at the left end of the double arrow, and in the “invisible” condition it is placed at the right end. Measured BRIRs for two real rooms from which data in Tab. 4 were obtained served as reference. The ISM condition and the proposed hybrid method with $N = 3$ and $N = 1$ were evaluated. In both hybrid conditions, the “cube” synthesis was used.

Besides a comparison of the BRIRs in the time domain, ILDs and IACCs were determined as described in sec. 3.2.1 and compared with the reference.

Table 4: *Specification of rooms used in the evaluation of the simulation of coupled rooms.*

Room 1 (corridor)	
Dimensions	(30.0, 1.94, 2.50) m
Absorption coeff.	(0.16, 0.16, 0.13, 0.15, 0.17)
Room 2 (office)	
Dimensions	(4.43, 4.50, 3.00) m
Absorption coeff.	(0.25, 0.30, 0.35, 0.32, 0.28)

3.3.2. Results

Fig. 4 shows normalized BRIR time signals for the measured (upper panels) and synthesized ($N = 3$, lower panels) case on an arbitrarily scaled ordinate. As expected for coupled rooms, the measured BRIR in the invisible condition (rhs) shows a rising amplitude in the beginning. This effect can hardly be observed in the simulated BRIR. In this simple approach here, only one convolution of two single RIRs was used which cannot mimick real coupling of the rooms.

Tab. 5 shows ILDs and IACCs obtained from BRIRs of all conditions. For all of them a clear dominance of sound energy on the left is obtained (negative ILDs), which is primarily due to the direction of the (virtual) direct sound (source S' in Fig. 2). The ISM-created BRIRs, which can be assumed to simulate the real rooms best, have indeed ILDs that are closest to those of the measured BRIRs. The ILDs of the hybrid method BRIRs differ maximally 3.4 dB from measurement condition, which is clearly above the JND in reverberant conditions, at least for frontal source positions [17].

For the IACC, all room- and synthesis conditions yield very small values. A slightly higher accordance with the measurement is obtained for the ISM synthesis, but it is questionable, whether these differences were audible.

Concluding, the evaluation showed that this simple approach has limitations if the acoustics of coupled rooms should be simulated in a convincing way. Improvements should consider removal of the direct path between the virtual source and the receiver in the invisible condition. In a second step diffraction of the direct sound

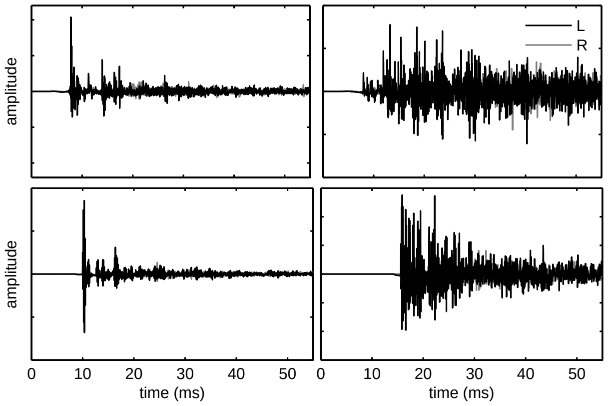

Figure 4: *BRIR time signals on arbitrary amplitude scale. Lhs: visible condition, rhs: invisible condition. Upper panels: measured BRIRs, lower panels: synthesis with $N = 3$.*

Table 5: *Results of the evaluation of the coupled-rooms simulation. Comparison of ILDs and IACCs. See text for explanation of conditions.*

		meas.	ISM	$N = 3$	$N = 1$
ILD (dB)	visible	-6.0	-6.6	-9.4	-8.3
	invisible	-4.5	-6.3	-7.8	-7.8
IACC	visible	0.1	0.1	0.2	0.2
	invisible	0.1	0.1	0.2	0.2

can be taken into account by inclusion of lowpass-filtered versions of the direct path, depending on the geometric relation of the door opening and the source and receiver positions.

4. SUMMARY AND CONCLUSIONS

A hybrid approach for synthesizing binaural room impulse responses of shoebox-shaped rooms was presented. It computes geometrically exact early reflections using the image source model up to a low reflection order, and approximates the reverberant tail by a high efficient feedback delay network. The FDN was extended to enable a spatial reverb rendering, taking into account room geometry as well as wall absorption and source- and receiver positions.

The proposed simulation method was evaluated with respect to different properties using subjective and technical measures. In a subjective evaluation subjects rated the naturalness and room size of measured and respectively synthesized BRIRs. The ratings show that the simulation method is able to represent perceived naturalness and room size very well and independently from maximum image source order, whereas differences in these properties between rooms are clearly conveyed.

For the extended FDN, two spatial reverb rendering techniques (sec. 2.2) were compared in a technical evaluation assessing interaural level differences and interaural cross correlation coefficients. It was shown that synthesized spatial reverberation has better accordance with purely ISM-created reference BRIRs if the reverberation emitting virtual sound sources are equally distributed around the listener's head. In comparison, positioning these sources on the actual room wall surfaces yielded worse results (sec. 3.2.2).

A first, simple approach to simulate the acoustics of two adjacent coupled rooms was evaluated by comparing time signal representations, ILDs and IACCs for measured and synthesized BRIRs. While the results for this approach were not fully convincing future improvement with refined approximations can be expected.

In conclusion, the evaluation showed that the suggested computationally efficient approach for synthesizing binaural room impulse responses is suited for applications where perceptual plausibility and authenticity is acceptable.

5. ACKNOWLEDGEMENTS

This work was supported by the DFG FOR 1732 and the Cluster of Excellence EXC 1077/1 “Hearing4all”.

6. REFERENCES

[1] J. B. Allen and D. A. Berkley, “Image method for efficiently simulating small-room acoustics,” *J. Acoust. Soc. Am.*, vol. 66, no. 4, pp. 943–950, 1979.

[2] B.-I. Dalenbäck, “Engineering principles and techniques in room acoustics prediction,” in *BNAM, Bergen, Norway, May 2010*, 2010.

[3] G. M. Naylor and J. H. Rindel, “Predicting Room Acoustical Behaviour with the ODEON Computer Model,” in *124th ASA meeting New Orleans, November 1992*, 1992.

[4] B.-I. Dalenbäck and M. Strömberg, “Real Time Walkthrough Auralization – The First Year,” Tech. Rep., CATT (Dalenbäck), Valeo Graphics (Strömberg), 2010.

[5] D. Schröder, F. Wefers, S. Pelzer, D. S. Rausch, M. Vorländer, and T. Kuhlen, “Virtual Reality System at RWTH Aachen University,” in *Proceedings ICA 2010, 20th International Congress on Acoustics: 23–27 August 2010, Sydney, New South Wales, Australia*, 2010.

[6] A. Silzle, P. Novo, and H. Strauss, “IKA-SIM: A system to generate auditory virtual environments,” in *Audio Engineering Society Convention 116*, 2004.

[7] A. Krokstad, S Strøm, and S. Sørsdal, “Calculating the acoustical room impulse response by the use of a ray tracing technique,” *J. Sound Vib.*, vol. 8, no. 1, pp. 118–125, 1968.

[8] Steven M. Schimmel, Martin F. Müller, and Norbert Dillier, “A fast and accurate »shoebox« room acoustics simulator,” Tech. Rep., 2009.

[9] M. R. Schroeder, “Natural Sounding Artificial Reverberation,” *Journal of the Audio Engineering Society*, vol. 10, no. 3, pp. 219–223, 1962.

[10] J. Stautner and M. Puckette, “Designing Multi-Channel Reverberators,” *Computer Music Journal*, vol. 6, no. 1, pp. 52–65, 1982.

[11] J.-M. Jot and A. Chaigne, “Digital delay networks for designing artificial reverberators,” in *90th AES Convention*, 1991.

[12] T. Wendt, S. van de Par, and S. D. Ewert, “A computational efficient and perceptually plausible algorithm for binaural room impulse response simulation,” *subm. to Journal of the Audio Engineering Society*.

[13] J. Borish, "Extension of the image model to arbitrary polyhedra," *J. Acoust. Soc. Am.*, vol. 75, no. 6, pp. 1827–1836, 1984.

[14] M. Jeub, M. Schäfer, and P. Vary, "A binaural room impulse response database for the evaluation of dereverberation algorithms," Tech. Rep., 2009.

[15] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *Audio Engineering Society Convention 108*, 2 2000.

[16] G. Geißler and S. van de Par, "Messung von HRTF am Kunstkopf MK 2 von Cortex," AG Akustik, Carl-von-Ossietzky-Universität Oldenburg, 2012.

[17] S. Klockgether and S. van de Par, "Just Noticable Differences of Spatial Perception in Directly Manipulated Binaural Room Impulse Responses," in *AIA/DAGA 2013, Merano, Italy*, 2013.