BINAURALIZATION OF OMNIDIRECTIONAL ROOM IMPULSE RESPONSES - ALGORITHM AND TECHNICAL EVALUATION

Christoph Pörschmann [*], *Philipp Stade* [*] [†], *Johannes M. Arend* [*] [†]

[*]TH Köln, Institute of Communication Engineering, Cologne, Germany
[†]TU Berlin, Audio Communication Group, Berlin, Germany
christoph.poerschmann@th-koeln.de

ABSTRACT

The auralization of acoustic environments over headphones is often realized with data-based dynamic binaural synthesis. The required binaural room impulse responses (BRIRs) for the convolution process can be acquired by performing measurements with an artificial head for different head orientations and positions. This procedure is rather costly and therefore not always feasible in practice. Because a plausible representation is sufficient for many practical applications, a simpler approach is of interest.

In this paper we present the *BinRIR* (Binauralization of omnidirectional room impulse responses) algorithm, which synthesizes BRIR datasets for dynamic auralization based on a single measured omnidirectional room impulse response (RIR). Direct sound, early reflections, and diffuse reverberation are extracted from the omnidirectional RIR and are separately spatialized. Spatial information is added according to assumptions about the room geometry and on typical properties of diffuse reverberation. The early part of the RIR is described by a parametric model and can easily be modified and adapted. Thus the approach can even be enhanced by considering modifications of the listener position. The late reverberation part is synthesized using binaural noise, which is adapted to the energy decay curve of the measured RIR.

In order to examine differences between measured and synthesized BRIRs, we performed a technical evaluation for two rooms. Measured BRIRs are compared to synthesized BRIRs and thus we analyzed the inaccuracies of the proposed algorithm.

1. INTRODUCTION

Binaural synthesis is a powerful tool for headphone-based presentation of virtual acoustic environments (VAEs). It can be applied for auralization purposes in various areas like audio engineering, telecommunication, or architectural acoustics. For many of these applications, a plausible presentation is sufficient; an authenticity reproduction of the sound field is often not pursued. In this context plausibility refers to the illusion that the scenario being depicted is actually occurring [1] while authentic refers to a perception that the scenario cannot be distinguished from a real reference.

Binaural room impulse responses (BRIRs) can be applied, which are either simulated or measured with an artificial head for different orientations (and positions). Finally the BRIRs are convolved with anechoic signals in a binaural renderer. By considering the listener's head movements in the auralization, localization accuracy increases [2], front-back confusion can be decreased [2] and externalization of virtual sound sources improves [3][4]. Several commercial or scientific rendering engines are available, which adapt the sound field presented through headphones according to the orientation of the listener in real time (e.g. [5][6][7][8]). Depending on the head movements, which shall be considered in the auralization, these measurements need to be done for circular orientations in the horizontal plane or even for spherical head orientations considering horizontal and vertical rotations. However, measuring such BRIR datasets requires a large amount of time and the use of complex devices (e.g. a rotatable artificial head). Furthermore, for each listening position, another set of BRIRs needs to be captured. Thus, for many applications in the field of spatial audio and virtual environments, the effort is so high that circular sets of BRIRs are not used. To approach this issue, we developed the *BinRIR* (Binauralization of omnidirectional room impulse responses) algorithm, which aims for an auralization based on a simple measurement procedure. Only one single measured omnidirectional room impulse response (RIR) is required to obtain a plausible auralization when using dynamic binaural synthesis. The algorithm even allows to shift the listener position. Thus, one single measured RIR is sufficient to synthesize a BRIR dataset for a freely chosen head orientation and position in the room.

In literature, several approaches to obtain BRIRs from measured RIRs have been described. In [9][10] a synthesis of BRIRs from B-format measurements has been proposed. The spatial impulse response rendering (SIRR) method applies a decomposition of the sound field into direct and diffuse parts. While the diffuse part is decorrelated, vector-based amplitude panning is used to distribute the direct sound on different loudspeakers. In [11] a directional audio coding (DirAC) method is proposed which can capture, code, and resynthesize spatial sound fields. DirAC analyzes the audio signal in short time frames and determines the spectrum together with direction and diffuseness in the frequency bands of human hearing. As this method does not work impulse response-based it is quite different to the one presented in this paper. Another simple approach to synthesize BRIRs has been presented by Menzer. In [12][13] RIRs measured in the B-format are used to synthesize BRIRs. Direct sound, spectral shape and temporal structure are extracted from the RIR. Additionally, the incidence direction of the direct sound is estimated from the measured data. No specific treatment of the early reflections is proposed. All reflections and the diffuse reverberation are synthesized by performing an adequate reconstruction of the interaural coherence.

In this paper, we present research results on the binauralization of omnidirectionally measured RIRs. Parts of the studies including the basic idea and a basic description of the approach as well as results of a perceptual evaluation have already been published [14][15][16]. This paper is organized as follows: In section 2 we introduce and explain the *BinRIR* algorithm performing the spatialization of omnidirectional RIRs in detail. In section 3 we describe the results of a technical evaluation. We compare measured BRIRs of two different rooms to the synthesized counterparts and elaborate differences caused by the simplifications of the algorithm. Finally, section 4 concludes the paper and provides an outlook.

2. ALGORITHM DESIGN

2.1. General Structure

The basic idea of the *BinRIR* algorithm is to use only one measured omnidirectional RIR for the synthesis of BRIR datasets which can be used for dynamic auralization. The algorithm was implemented in Matlab and applies predictable information from sound propagation in enclosed spaces as well as knowledge regarding the human perception of diffuse sound fields. For processing, the RIR is split into different parts. The early part contains the direct sound and strong early reflections. For this part, the directions of incidence are modeled reaching the listener from arbitrarily chosen directions. The late part of the RIR is considered being diffuse and is synthesized by convolving binaural noise with small sections of the omnidirectional RIR. By this, the properties of diffuse reverberation are approximated. The algorithm includes an additional enhancement: The synthesized BRIRs can be adapted to shifts of the listener and thus freely chosen positions in the virtual room can be auralized.

The *BinRIR* algorithm incorporates several inaccuracies and deviates significantly from a measured BRIR. The directions of incidence of the synthesized early reflections are not in line with the real ones. Hence, differences in the perception of spatial properties (e.g. envelopment) between the original room and the synthesized room may occur. Furthermore, a point source is assumed for all synthetic BRIR datasets. Thus, it is not possible to rebuild source width and other properties of the source correctly. Finally, the diffusely reflected part of the early reflections cannot be precisely reconstructed.

The basic structure of the *BinRIR* algorithm is shown in Figure 1. As input data the algorithm only requires the omnidirectional RIR and the position of the sound source. Furthermore, the algorithm accesses an appropriate set of HRIRs and a preprocessed sequence of binaural noise. Both were obtained from measurements with a Neumann KU100 artificial head [17].

The algorithm is only applied to frequencies above 200 Hz. For lower frequencies the interaural coherence of a typical BRIR is nearly one and the omnidirectional RIR can be maintained. 7th order Chebyshev Type II filters are used to separate the low frequency part from the rest of the signal.

2.2. Direct sound and early reflections

Onset detection is used to identify the direct sound in the omnidirectional RIR. The direct sound frame starts with the onset and ends after 10 ms (5 ms followed by 5 ms raised cosine offset). The following time section is assigned to the early reflections and the transition towards the diffuse reverberation. For small and non-reverberant rooms (V $< 200\,\mathrm{m}^3$ and $\mathrm{RT}_{60} < 0.5\,\mathrm{s}$) a section length of 50 ms is chosen, otherwise the section is extended to 150 ms. In order to determine sections with strong early reflections in the omnidirectional RIR, the energy is calculated in a sliding window of 8 ms length and time sections which contain high energy are marked. Peaks which are 6 dB above the RMS level of the sliding window are determined and assigned to geometric reflections. A windowed section (raised cosine, 5 ms ramp) around each of the peaks is considered as one reflection. If several very dense reflections occur in adjacent sections, these sections are merged. Following this procedure, small windowed sections of the omnidirectional RIR are extracted describing the early reflections. The incidence directions of the synthesized reflections base on a spatial reflection pattern adapted from a shoebox room with non-symmetric positioned source and receiver. Thus a fixed lookup-table containing the incidence directions is used. By this a simple parametric model of the direct sound and the early reflections is created. Amplitude, incidence direction, delay and the envelope of each of the reflections are stored. The design of the algorithm identifying the reflections and its parameterization (e.g. window length, peak detection) was done based on empiric tests. Informal listening experiments during the development have shown that the exact way in which the reflections are determined is not substantial.

By convolving each windowed section of the RIR with the HRIR(φ) of each of the directions, a binaural representation of the early geometric reflective part is obtained. To synthesize interim directions between the given HRIRs, interpolation in the spherical domain is performed [18].

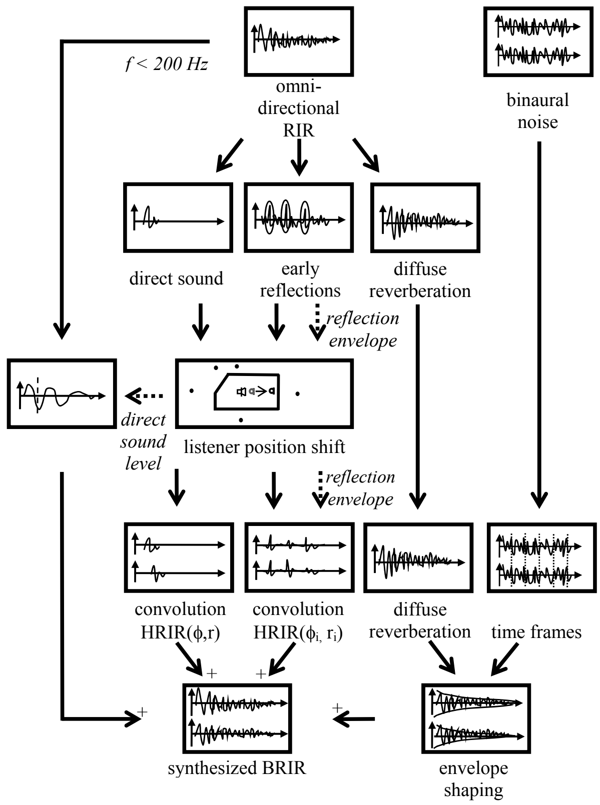

Figure 1: Block diagram of the *BinRIR* algorithm for synthesizing a BRIR based on one single omnidirectional RIR.

2.3. Diffuse Reverberation

The diffuse reverberation is considered reaching the listener temporarily and spatially equally distributed. Several studies have shown that after a so-called perceptual mixing time, no specific differences to completely diffuse reverberation can be perceived and thus, an exact reconstruction of the temporal and spatial structure is not required (e.g. [19]). It was found that the perceptual mixing time is room dependent and can be chosen according to predictors which are calculated based on geometric room prop-

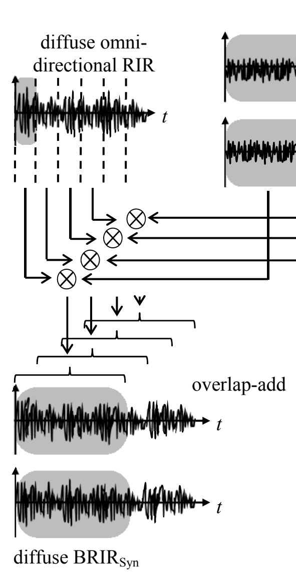

Figure 2: Synthesis of the binaural diffuse reverberation: Sections of the diffuse omnidirectional RIR (0.67 ms; 32 taps at 48 kHz sampling rate) and the binaural noise (2.67 ms; 128 taps at 48 kHz sampling rate) are convolved. Both sections are raised-cosine windowed. The diffuse BRIR is synthesized by summing up the results of the convolutions applying overlap-add.

erties. However, in [20] it has been shown that small perceptual differences still remain. Thus recent studies (e.g. [21][22]) proposed models applying an overlap between early reflections and the diffuse part instead of using a fixed mixing time. By this some diffuse energy is embedded in the early part of the BRIR. A similar approach is used in the *BinRIR* algorithm: All parts of the RIR excluding the sections of the direct sound and the detected early reflections are assigned to the diffuse part.
To synthesize the binaural diffuse reverberation we developed two different methods. In [14][15] the RIR was split up into 1/6 octave bands by applying a near perfect reconstruction filter bank [23] and the binaural diffuse part was synthesized for each frequency band. In [16] we proposed another method for the synthesis of the diffuse part which is used in this publication. The diffuse reverberation is synthesized by convolving small time sections (0.67 ms) of the omnidirectional RIR with sections of 2.67 ms binaural noise (both sections raised-cosine windowed). The results of the convolutions of all time sections are summed up with overlap-and-add. Figure 2 explains the synthesis of the diffuse reverberation in greater detail. By this, both the binaural features (e.g. interaural coherence) of the binaural noise and the frequency-dependent envelope of the omnidirectional RIR are maintained. The lengths of the time sections were determined by informal listening tests during the development of the algorithm. This method requires less computational power than the one proposed in [14][15]. Informal listening tests showed that both methods are perceptually comparable.

2.4. Listener position shifts

The algorithm includes a further enhancement: The synthesized BRIR can be adapted to listener position shifts (LPS) and thus freely chosen positions of the listener in the virtual room can be

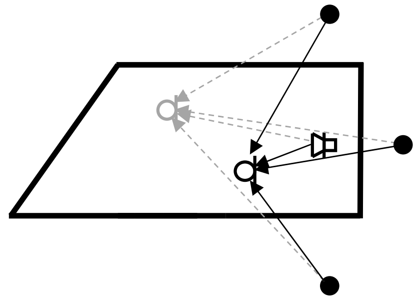

Figure 3: Basic principle of the listener position shifts (LPS): A mirror-image model is applied to modify the amplitude and the temporal structure of the direct sound and the early reflections. The receiver is moved from an initial position (grey) to a modified position (black). By this the paths of the direct sound and the reflections are changed.

auralized. For this, a simple geometric model based on mirror-image sound sources is used. The distance between the listener and each of the mirror-image sound sources is determined from the delay of the corresponding reflection peak to the direct sound peak. In a next step, a shifted position of the listener is considered and amplitudes (based on the 1/r law), distances, and directions of incidence are recalculated for each reflection (Fig. 3). Optimizing an earlier version of the *BinRIR* algorithm (e.g [16]) we modified the low-frequency component below 200 Hz when applying LPS. If, for example, the listener approaches the sound source, the amplitude of the direct sound increases and the low-frequency energy for the direct sound needs to be adapted accordingly. For this the low-frequency part of the direct sound (first 10 ms followed by 10 ms raised cosine set) is adjusted according to the 1/r law.

2.5. Synthesis of Circular BRIR sets

The synthesis of the BRIRs is repeated for constant shifts in the azimuthal angle (e.g. 1°) for the direct and the reflected sound. Thus, a circular set of BRIRs is obtained, which can be applied by different rendering engines for dynamic binaural synthesis. The synthesized sets of BRIRs can be stored in various formats, e.g the miro-Format [24], a multi-channel-wave format to be used by the SoundScape Renderer [7] and can be converted to the SOFA-format [25].

3. TECHNICAL EVALUATION

To analyze the performance of the algorithm, a series of listening experiments has already been conducted [14][16]. These experiments mainly aimed at a quantification of the perceptual differences between measured and synthesized BRIRs. In this paper we focus on a technical evaluation of the algorithm and compare different properties of the synthesized BRIRs to the properties of measured BRIRs. Therefore we analyzed the detected reflections regarding their direction, their time of incidence, and their amplitude to measured data. Furthermore, we compared the reverberation tails and the reverberation times of the synthesized BRIRs to the ones of measured BRIRs. We investigated to what extent the clarity (C_{50}) of the synthesized room matches the measured room's

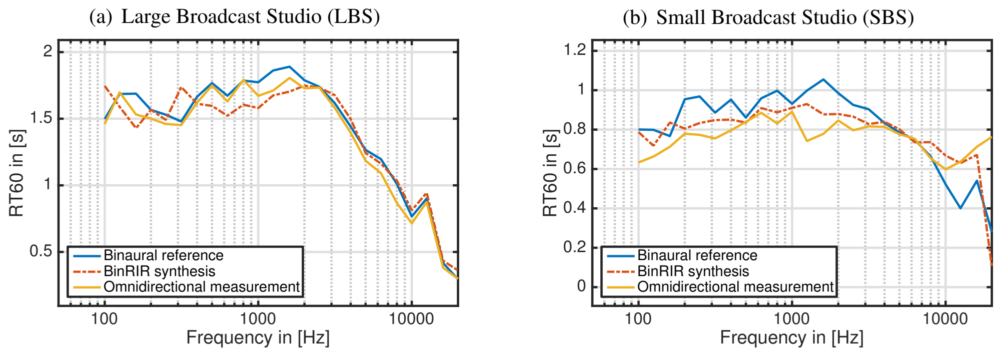

Figure 4: Reverberation Time (RT_{60}) of the Large Broadcast Studio (a) and the the Small Broadcast Studio (b). In each plot the RT_{60} for the binaurally measured reference, for the synthesis with the ***BinRIR*** algorithm and for the omnidirectional measurement are shown. The RT_{60} was calculated in 1/3 octave bands in the time domain

clarity. Finally we looked briefly in which way the use of the LPS influences the early part of the BRIR.

3.1. Measured rooms

The performance of the algorithm was analyzed for two different rooms. Both rooms are located at the WDR radiobroadcast studio in Cologne and are used for various recordings of concerts and performances. The "*KVB-Saal*" (Large Broadcast Studio - LBS) has a volume of 6100 m^3, a base area of 579 m^2 and can seat up to 637 persons. We measured the impulse responses in the 6th row ($Distance_{SrcRec} = 13.0$ m). The "*kleiner Sendesaal*" (Small Broadcast Studio - SBS) has a volume of 1247 m^3, a base area of 220 m^2 and 160 seats. The $Distance_{SrcRec}$ in this room was 7.0 m. In order to evaluate the algorithm, measured impulse responses from these rooms were used. In addition to the omnidirectional RIRs, which are required to feed the ***BinRIR*** algorithm, we measured circular BRIR datasets at the same position as a reference. This dataset was measured in steps of 1° on the horizontal plane with a Neumann KU100 artificial head. Finally we used data from spherical microphone array measurements, conducted with the VariSphear measurement system [26]. For this, we applied a rigid sphere array configuration with 1202 sample points on a Lebedev grid at a diameter of 17.5 cm. The omnidirectional RIRs and the array data were measured with an Earthworks M30 microphone. As sound source, a PA stack involving an AD Systems Stium Mid/High unit combined with 3 AD Systems Flex 15 subwoofers was used. The complete series of measurements is described in detail in [27].
Based on the microphone array measurements, we identified reflections in the room using sound field analysis techniques [28]. For this the array impulse responses were temporally segmented (time resolution = 0.5 ms) and transformed into the frequency domain. Applying a spatial Fourier transformation, the impulse responses were transformed into the spherical wave spectrum domain [29]. Then the sound field was decomposed into multiple plane waves using the respective spatial Fourier coefficients. Data was extracted for a decomposition order of $N = 5$, a spherical composite grid with 3074 Lebedev points at a frequency $f = 4500$ Hz. For this frequency quite robust results for the detection of the reflections were found. By this, a spatio-temporal intensity matrix of the sound field at the listener position was calculated. Each time slice of the matrix was analyzed with a specific algorithm in order to classify reflections which are represented as local maxima in the matrix [30]. The detected reflections were stored with their attributes "time", "direction" and "level" in a reflection list and can be used for further comparison with the ***BinRIR*** algorithm.

3.2. Direct Sound and early reflections

In a first step we looked at the early part of the synthetic BRIRs and compared the direct sound and the early reflections determined by the ***BinRIR*** algorithm to the reflections which are identified using sound field analysis techniques based on the microphone array measurements. To describe the directions of the reflections we used a spherical coordinate system. The azimuthal angle denotes the orientation on the horizontal plane with $\varphi = 0^\circ$ corresponding to the front direction. The elevation angle is $\delta = 0^\circ$ for frontal orientation and $\delta = 90^\circ$ for sound incidence from above.
Table 1 and 2 show the reflections detected by ***BinRIR*** as well as the 14 strongest reflections which were identified based on the array data. For each reflection, the time of arrival relative to the direct sound, the incidence direction, and the level are shown. The temporal structure and the energy of many of the reflections show similarities. For example in the LBS for reflection # 7, # 10 and # 13 and in the SBS for reflection # 13 the level and the time of arrival match quite well. Furthermore, some of the reflections were detected in the array measurements as impinging from different directions nearly simultaneously. These reflections are typically merged to one reflection with an increased level by the ***BinRIR*** algorithm (e.g. LBS: # 4 and # 5; SBS: # 6 and # 7; # 9 - # 11).
As the incidence directions of the synthetic reflections are chosen by ***BinRIR*** based on a lookup-table, it is not surprising that there are significant differences compared to the directions of the measured reflections. However, if the proportions as well as source and listener position of the synthesized room to some extent match the modelled shoebox room some of the incidence directions are appropriate. Thus, at least for the first reflection and as well for some other reflections detected by ***BinRIR***, an acceptable congruence exists both for the LBS and the SBS. The azimuthal deviation of the first reflection is less than 40° and in elevation no relevant deviation exists.
As already explained in 2.2 the ***BinRIR*** algorithm starts detecting reflections 10 ms after the direct sound. Thus no reflections reaching the listener within the first 10 ms can be found. The time frame

	Array measurement				**BinRIR algorithm**			
#	**Delay [ms]**	**Azimuth [°]**	**Elevation [°]**	**Level [dB]**	**Delay [ms]**	**Azimuth [°]**	**Elevation [°]**	**Level [dB]**
0	0.0	0	3	0.0	0.0	0	0	0.0
1	5.0	8	-58	-27.6				
2	7.5	142	-18	-27.5				
3	17.5	0	-8	-27.3				
4	24.0	309	0	-25.6	24.9	319	-2	-21.0
5	26.5	58	5	-27.9				
6	28.0	278	0	-28.1				
7	31.0	0	58	-20.3	31.1	10	-2	-19.4
8	38.0	253	-2	-28.0	38.4	85	-3	-22.9
9	50.5	180	20	-27.2				
10	79.5	178	17	-12.7	79.6	343	4	-14.9
11	92.5	353	73	-26.6	92.5	319	-54	-40.5
12					108.2	0	67	-34.4
13	117.5	183	35	-26.3	117.6	85	-50	-31.4
14	174.0	356	8	-25.0				
15	202.0	3	3	-24.3				

Table 1: Properties of the direct sound and the early reflections for the Large Broadcast Studio (LBS). Left side: Reflections determined from the analysis of the array data. Right side: Reflections detected by the *BinRIR* algorithm

	Array measurement				**BinRIR algorithm**			
#	**Delay [ms]**	**Azimuth [°]**	**Elevation [°]**	**Level [dB]**	**Delay [ms]**	**Azimuth [°]**	**Elevation [°]**	**Level [dB]**
0	0.0	5	5	0.0	0.0	0	0	0.0
1	1.5	7	-22	-6.9				
2	3.5	16	-76	-21.9				
3	7.5	210	-22	-27.5				
4	12.0	204	-17	-27.7				
5	15.5	27	66	-23.9				
6	20.0	284	0	-21.1	22.2	319	-2	-19.2
7	23.0	58	75	-24.5				
8	35.0	203	2	-27.7	35.4	10	-2	-21.6
9	50.0	152	1	-21.5				
10	50.5	151	-2	-21.5	51.3	85	-3	-18.5
11	51.5	147	29	-25.5				
12	57.0	129	0	-21.4				
13	59.0	173	7	-19.7	59.6	343	4	-15.9
14	99.5	331	-2	-19.5	100.3	319	-54	-29.3

Table 2: Properties of the direct sound and the early reflections for the Small Broadcast Studio (SBS). Left side: Reflections determined from the analysis of the array data. Right side: Reflections detected by the *BinRIR* algorithm

determining the geometric reflections ends after 150 ms and no reflections are detected by *BinRIR* after this period.
Furthermore, several reflections which can be extracted from the array data measurements are not detected by *BinRIR* (e.g. # 3 and # 9 in the LBS and # 5 and # 12 in the SBS). However, in total more than 2/3 of the reflections in the section from 10 ms to 150 ms determined from the array data correspond to a reflection determined by *BinRIR*.

3.3. Energy Decay and Reverberation Time

In a next step we compared the energy decay curves of the synthesis and of the binaurally measured reference BRIRs. As already explained, the *BinRIR* algorithm synthesizes diffuse reverberation by applying a frame-based convolution of the omnidirectional RIR with binaural noise. Thus in addition to the synthesized and the measured BRIR (reference) we analyzed the energy decay of the measured omnidirectional RIR as well. The following analysis is based on the impulse responses for the frontal viewing direction ($\varphi = 0°$, $\delta = 0°$). Analyzing the reverberation time RT_{60} (Figure 4) we observed that the general structure of the curves is similar, but variations between the three curves exist. The average unsigned deviation between the synthesis and the reference is 0.10 s for the LBS and 0.09 s for the SBS, the maxima are 0.26 s (LBS) and 0.23 s (SBS).

3.4. Interaural Coherence

Next, we compared the interaural coherence (IC) of the synthesized BRIRs and of the reference BRIRs (Figure 5). We calculated the IC according to [31] applying hamming-windowed blocks with a length of 256 taps (5.33 ms) and an overlap of 128 taps (2.67 ms). In each plot the IC calculated with three different starting points is shown. For reference and synthesis in both rooms the IC is significantly different when direct sound is included in the calculation ($t > 0$ ms). For the medium condition (LBS: $t > 150$ ms; SBS: $t > 50$ ms) significant differences between synthesis and reference can be observed. This is not surprising as no shaping of the IC is performed in the *BinRIR* algorithm. However, this difference is smaller for the SBS, because the impulse response is probably nearly diffuse at 50 ms. For the condition with the maximal starting point (LBS: $t > 300$ ms; SBS: $t > 150$ ms) which mainly comprises the diffuse reverberation the IC of the synthesized BRIR matches the reference quite well.

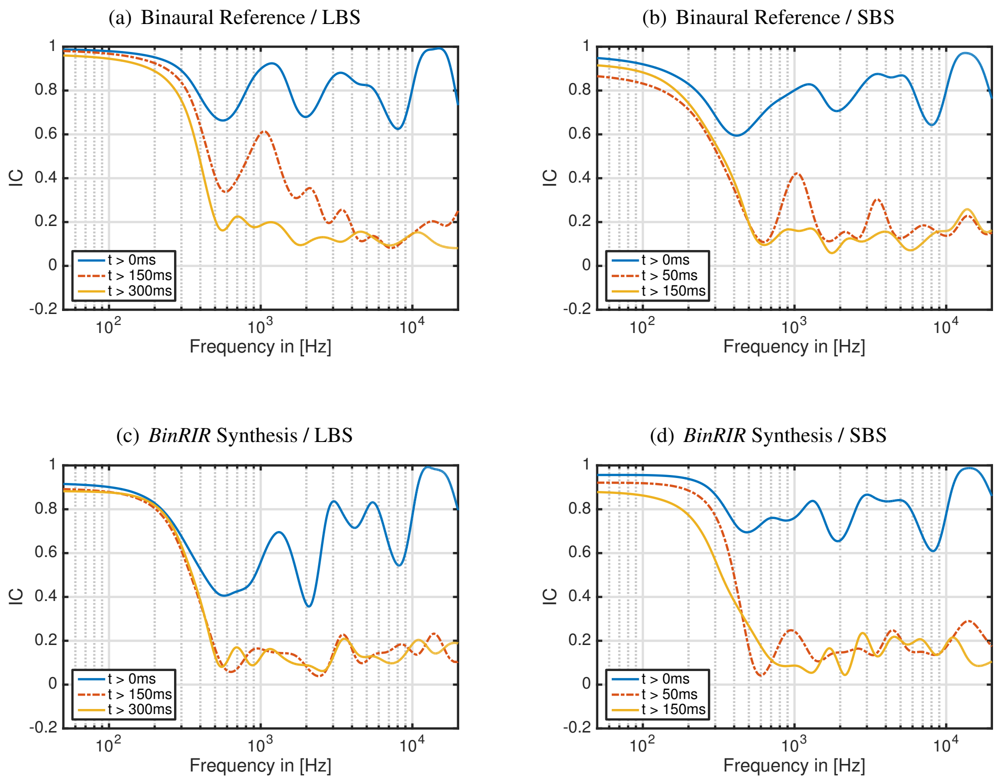

Figure 5: Interaural coherence (IC) of the Large Broadcast Studio (LBS) and the the Small Broadcast Studio (SBS). In plot (a) and (b) the data for the binaural reference is shown, in (c) and (d) the data for the *BinRIR* synthesis. For the LBS the IC is plotted for t > 0 ms, t > 150 ms and t >300 ms, for the SBS for t > 0 ms, t > 50 ms and t >150 ms.

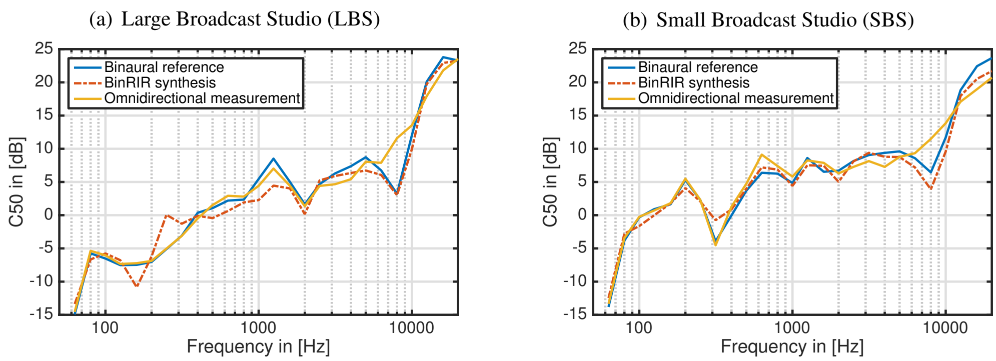

Figure 6: Clarity (C_{50}) of the Large Broadcast Studio (LBS) and the the Small Broadcast Studio (SBS). In each plot the C_{50} for the binaurally measured reference, for the synthesis with the *BinRIR* algorithm and for the omnidirectional measurement are shown.

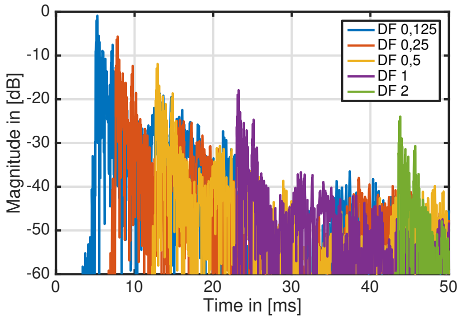

Figure 7: Influence of the Listener Position Shift (LPS) on the early part of the time response for the Small Broadcast Studio (SBS). The time responses for Distance Factors (DFs) from 0.125 - 2 are shown in different colors

3.5. Clarity

Next, we examined the clarity C_{50} over frequency for each of the conditions (Figure 6). The differences between the omnidirectional RIR, the synthesis and the reference are minor. The average unsigned deviation between the synthesis and the reference is 1.4 dB for the LBS and 1.1 dB for the SBS. The maxima are 5.2 dB (LBS) and 3.2 dB (SBS). Thus the ratio of the energy of the early part of the BRIR and the late diffuse part of the BRIR can be regarded as appropriate.

3.6. Listener position shifts

Finally we analyzed for the SBS in which way the early part of the synthesized BRIR is changed when listener position shifts (LPS) are performed. The results are shown in Figure 7 for synthesized $\text{Distance}_{\text{SrcRec}}$ between 0.875 m and 14 m (distance factor 1/8 - 2 of the original distance). It can be observed that the level and the time of arrival of the direct sound are changed significantly (according to the 1/r distance law) when performing LPS. The influence of the LPS on reflections is hard to observe from the plot, but changes in amplitude and time of arrival according to the geometric room model can be found here as well. The diffuse part and thus the complete late reverberation remain unchanged when performing LPS (not shown in Figure 7).

4. CONCLUSION

In this paper, the *BinRIR* algorithm was presented, which aims for a plausible dynamic binaural synthesis based on one measured omnidirectional RIR. In two different rooms, RIRs were measured and binauralized applying the presented *BinRIR* algorithm, so that synthetic BRIR datasets were generated. The presented method separately treats direct sound, reflections and diffuse reverberation. The early parts of the impulse responses are convolved with HRIRs of arbitrary chosen directions while the reverberation tail is rebuilt from an appropriately shaped binaural noise sequence. In an extension, the algorithm allows to modify the sound source distance of a measured RIR by changing several parameters of an underlying simple room acoustic model.

The synthetic BRIRs were compared to reference BRIRs measured with an artificial head. Due to missing information on spatial aspects, a perfect reconstruction of the sound field is generally not possible. An analysis of the early reflections showed that neither all reflections are detected by the *BinRIR* algorithm nor their directions match to the physical ones of the room. However, the reflections which were identified by the *BinRIR* algorithm correlate with the times of incidence and partly with the direction of incidence of the physical reflections in the room quite well. For the diffuse part, small differences in the reverberation time and the interaural coherence were observed. However, in general, the synthesis can be regarded as appropriate. An evaluation of the reverberation time RT_{60} and of the clarity C_{50} only showed minor differences between reference and synthesis. Analyzing the perceptual influences of the determined differences is not covered in the study presented here. Please refer to [14][16] for an analysis of these topics.

The approach presented in this paper can be combined with other modifications of measured RIRs. In [32][33], we discussed a predictive auralization of room modifications by an appropriate adaptation of the BRIRs. Thus the measurement of one single RIR is sufficient to obtain a plausible representation of the modified room. Furthermore, the opportunity to shift the listener position freely in the room can be employed when perceptual aspects of listener movements shall be investigated as e.g. proposed in [34].

5. ACKNOWLEDGEMENTS

The research presented here has been carried out in the Research Project MoNRa which is funded by the Federal Ministry of Education and Research in Germany. Support Code: 03FH005I3-MoNRa. We appreciate the support.

The code of the Matlab-based implementation and a GUI-based version of the *BinRIR* algorithm which is available under the GNU GPL License 4.1 can be accessed via the following webpage: http://www.audiogroup.web.th-koeln.de/DAFX2017.html

6. REFERENCES

[1] Mel Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, 2009.

[2] Jens Blauert, *Spatial Hearing - Revised Edition: The Psychoacoustics of Human Sound Source Localisation*, MIT Press, Cambridge, MA, 1997.

[3] Etienne Hendrickx, Peter Stitt, Jean-Christophe Messonnier, Jean-Marc Lyzwa, Brian FG Katz, and Catherine de Boishéraud, "Influence of head tracking on the externalization of speech stimuli for non-individualized binaural synthesis," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 2011–2023, 2017.

[4] W. Owen Brimijoin, Alan W. Boyd, and Michael A. Akeroyd, "The contribution of head movement to the externalization and internalization of sounds," *PLoS ONE*, vol. 8, no. 12, pp. 1–12, 2013.

[5] Ulrich Horbach, Attila Karamustafaoglu, Renato S. Pellegrini, and Philip Mackensen, "Design and Applications of a Data-based Auralization System for Surround Sound," in *Proceedings of 106th AES Convention, Convention Paper 4976*, 1999.

[6] Jens Blauert, Hilmar Lehnert, Jörg Sahrhage, and Holger Strauss, "An Interactive Virtual-Environment Generator for Psychoacoustic Research I: Architecture and Implementation," *Acta Acustica united with Acustica*, pp. 94–102, 2000.

[7] Matthias Geier, Jens Ahrens, and Sascha Spors, "The soundscape renderer: A unified spatial audio reproduction framework for arbitrary rendering methods," in *Proceedings of 124th Audio Engineering Society Convention 2008*, 2008, pp. 179–184.

[8] Dirk Schröder and Michael Vorländer, "RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environments," *Forum Acusticum*, pp. 1541–1546, 2011.

[9] Juha Merimaa and Ville Pulkki, "Spatial impulse response rendering I: Analysis and synthesis," *Journal of the Audio Engineering Society*, vol. 53, no. 12, pp. 1115–1127, 2005.

[10] Ville Pulkki and Juha Merimaa, "Spatial impulse response rendering II: Reproduction of diffuse sound and listening tests," *Journal of the Audio Engineering Society*, vol. 54, no. 1-2, pp. 3–20, 2006.

[11] Ville Pulkki, Mikko-Ville Laitinen, Juha Vilkamo, Jukka Ahonen, Tapio Lokki, and Tapani Pihlajamäki, "Directional audio coding-perception-based reproduction of spatial sound," *International Workshop On The Principles And Applications of Spatial Hearing (IWPASH 2009)*, 2009.

[12] Fritz Menzer, *Binaural Audio Signal Processing Using Interaural Coherence Matching*, Dissertation, École polytechnique fédérale de Lausanne, 2010.

[13] Fritz Menzer, Christof Faller, and Hervé Lissek, "Obtaining binaural room impulse responses from b-format impulse responses using frequency-dependent coherence matching," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 19, no. 2, pp. 396–405, 2011.

[14] Christoph Pörschmann and Stephan Wiefling, "Perceptual Aspects of Dynamic Binaural Synthesis based on Measured Omnidirectional Room Impulse Responses," in *International Conference on Spatial Audio*, 2015.

[15] Christoph Pörschmann and Stephan Wiefling, "Dynamische Binauralsynthese auf Basis gemessener einkanaliger Raumimpulsantworten," in *Proceedings of the DAGA 2015*, 2015, pp. 1595–1598.

[16] Christoph Pörschmann and Philipp Stade, "Auralizing Listener Position Shifts of Measured Room Impulse Responses," *Proceedings of the DAGA 2016*, pp. 1308–1311, 2016.

[17] Benjamin Bernschütz, "A Spherical Far Field HRIR / HRTF Compilation of the Neumann KU 100," *Proceedings of the DAGA 2013*, pp. 592–595, 2013.

[18] Benjamin Bernschütz, *Microphone Arrays and Sound Field Decomposition for Dynamic Binaural Recording*, Dissertation, TU Berlin, 2016.

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

[20] Philipp Stade, "Perzeptive Untersuchung zur Mixing Time und deren Einfluss auf die Auralisation," *Proceedings of the DAGA 2015*, pp. 1103–1106, 2015.

[21] Philipp Stade and Johannes M. Arend, "Perceptual Evaluation of Synthetic Late Binaural Reverberation Based on a Parametric Model," in *AES Conference on Headphone Technology*, 2016.

[22] Philip Coleman, Andreas Franck, Philip J.B. Jackson, Luca Remaggi, and Frank Melchior, "Object-Based Reverberation for Spatial Audio," *Journal of the Audio Engineering Society*, vol. 65, no. 1/2, pp. 66–76, 2017.

[23] Wessel Lubberhuizen, "Near perfect reconstruction polyphase filterbank, Matlab Central, www.mathworks.com/matlabcentral/fileexchange/15813 assessed 04/04/2017," 2007.

[24] Benjamin Bernschütz, "MIRO - measured impulse response object: data type description," 2013.

[25] Piotr Majdak, Yukio Iwaya, Thibaut Carpentier, Rozenn Nicol, Matthieu Parmentier, Agnieszka Roginska, Yôiti Suzuki, Kanji Watanabe, Hagen Wierstorf, Harald Ziegelwanger, and Markus Noisternig, "Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions," in *Proceedings of the 134th Audio Engineering Society Convention 2013*, 2013, number May, pp. 262–272.

[26] Benjamin Bernschütz, Christoph Pörschmann, Sascha Spors, and Stefan Weinzierl, "SOFiA - Sound Field Analysis Toolbox," in *Proceedings of the International Conference on Spatial Audio - ICSA*, 2011.

[27] Philipp Stade, Benjamin Bernschütz, and Maximilian Rühl, "A Spatial Audio Impulse Response Compilation Captured at the WDR Broadcast Studios," in *27th Tonmeistertagung - VDT International Convention*, 2012, pp. 551–567.

[28] Benjamin Bernschütz, Philipp Stade, and Maximilian Rühl, "Sound Field Analysis in Room Acoustics," *27thth Tonmeistertagung - VDT International Convention*, pp. 568–589, 2012.

[29] Earl G. Williams, *Fourier Acoustics - Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, UK, 1999.

[30] Philipp Stade, Johannes M Arend, and Christoph Pörschmann, "Perceptual Evaluation of Synthetic Early Binaural Room Impulse Responses Based on a Parametric Model," in *Proceedings of 142nd AES Convention, Convention Paper 9688*, Berlin, Germany, 2017, pp. 1–10.

[31] Fritz Menzer and Christof Faller, "Stereo-to-Binaural Conversion Using Interaural Coherence Matching," in *Proceedings of the 128th AES Convention, London UK*, 2010.

[32] Christoph Pörschmann, Sebastian Schmitter, and Aline Jaritz, "Predictive Auralization of Room Modifications," *Proceedings of the DAGA 2013*, pp. 1653–1656, 2013.

[33] Christoph Pörschmann, Philipp Stade, and Johannes M. Arend, "Binaural auralization of proposed room modifications based on measured omnidirectional room impulse responses," in *Proceedings of the 173rd Meeting of the Acoustical Society of America*, 2017.

[34] Annika Neidhardt and Niklas Knoop, "Binaural walkthrough scenarios with actual self-walking using an HTC Vive Real-time rendering of binaural audio Interactive binaural audio scenes Creating scenes allowing self-translation," in *Proceedings of the DAGA 2017*, 2017, pp. 283–286.