Proceedings of Meetings on Acoustics

Volume 30 http://acousticalsociety.org/

Acoustics `17 Boston

173rd Meeting of Acoustical Society of America and 8th Forum Acusticum
Boston, Massachusetts
25-29 June 2017

Architectural Acoustics: Paper 4aAAd5

A parametric model for the synthesis of binaural room impulse responses

Philipp Stade
Institute of Communication Engineering, TH Köln, Köln, NRW, 50679, GERMANY; philipp.stade@th-koeln.de

Johannes Arend and Christoph Pörschmann
TH Köln, Köln, NRW, 50679, GERMANY; johannes.arend@th-koeln.de; christoph.poerschmann@th-koeln.de

Binaural room impulse responses (BRIRs) are often applied in spatial audio for the auralization of acoustical environments. In the same field of research, parametric audio coding is an established approach and part of different standards. The presented investigation aims for a parametric description of the sound field to synthesize BRIRs. The model focuses on the main features which characterize a BRIR as well as the acoustical environment. Thereto a spherical microphone array is applied for a spatio-temporal acoustical analysis. Early reflections are determined with sound field decomposition techniques and are described by directional parameters. Diffuse components and the interaural coherence of the late reverberation are characterized with diffuse parameters. In two previous studies, the synthesis of the early and the late part of BRIRs has been elaborated and perceptually evaluated apart from each other. Now both approaches are combined to synthesize entire BRIR datasets using the parametric approach. Fundamentals of the sound field analysis are explained and synthetic BRIRs are compared to their measured counterparts. The approach yielded satisfactorily results and an adequate auralization based on parameters. Surprisingly no relevant enhancement of the auralization due to the additional use of reflections in contrast to diffuse components only could be observed.

Published by the Acoustical Society of America

 https://doi.org/10.1121/2.0000573

1. INTRODUCTION

In the field of architectural and room acoustics, the use of binaural room impulse responses (BRIRs) is a common approach for the plausible auralization of acoustic environments applying 3D audio techniques like dynamic binaural synthesis. Measured as well as simulated or synthesized BRIRs can be used for the convolution process in the binaural renderer to create a virtual acoustic environment (VAE) over headphones. By applying a headtracker and specific BRIRs for different angles, it is possible to adapt the synthesis according to the listener's head-orientation in realtime. By this, a more realistic auralization is achieved with higher localization accuracy, less front-back conversions and an improved externalization of virtual sound sources [1].

The synthesis of reverberation has been in focus of research since years [2,3,4,5]. Furthermore, parametric models for the synthesis of BRIRs with artifical reverberation have been also investigated over decades and are still relevant [6,7,8]. Different models have been proposed and in general the processing is temporally separated in the synthesis of direct sound, early reflections and diffuse reverberation. The parametric analysis and generation of early reflections is usually based on geometric assumptions e.g. the image source model [9] or low order microphone arrays [10,11]. The diffuse reverberation synthesis often applies feedback delay networks [12] or decaying white noise signals using an octave-band frequency resolution [8].

The model presented here focuses on an alternative approach: No temporal segmentation of the BRIRs is applied. Instead, the system differentiates between directional and diffuse components, both features are determined for the whole impulse response and characterized with individual parameters. This enables a discrete generation and adaptation of directional and diffuse components apart from each other for the BRIR synthesis. In two previous studies, the synthesis of the early [13] and the late part [14] of BRIRs has been elaborated and perceptually evaluated. Both approaches are combined in this paper and used for directional and diffuse synthesis to generate entire BRIR datasets based on a parametric description. As a perceptual evaluation, two different rooms and two different sound sources have been synthesized and are compared to measured reference BRIRs in this study using an ABC/HR test design paradigm [15].

BRIR synthesis using a parametric model enables several new features: For example, a BRIR personalization is possible, applying individual head related impulse responses (HRIRs). Changes of room acoustical parameters, like e.g. the frequency dependent reverberation time, are easily to simulate and can be used to modify the auralization. Furthermore, the parametric BRIR synthesis could also be integrated into object-based audio rendering systems, which are currently gaining importance [8].

The paper is organized as follows: In Section 2.A the general idea of our approach is explained and the parametric models for directional and the diffuse components are presented and summarized. The different auralized BRIR datasets based on the investigated rooms, sound sources and gradations of the algorithm are explained in 2.B. Section 2.C deals with the technical environment and the used test design for the perceptual evaluation. Results of the listening experiment are presented and discussed in Section 3. Finally, Section 4 concludes the paper and illustrates possible applications of the approach.

2. METHODS

A. SYSTEM OVERVIEW

The basic idea of the presented approach is to characterize BRIRs and their acoustical and room dependent properties with a few features only. It is aimed for a simplification of measured sound fields to their main characteristics and to describe them with discrete values or mathematical functions. These values are called parameters in the following and are used to generate synthetic BRIR datasets for dynamic auralization. As a consequence, the BRIR synthesis is based on the parametric description only, without any storage of parts of the original measured impulse response.

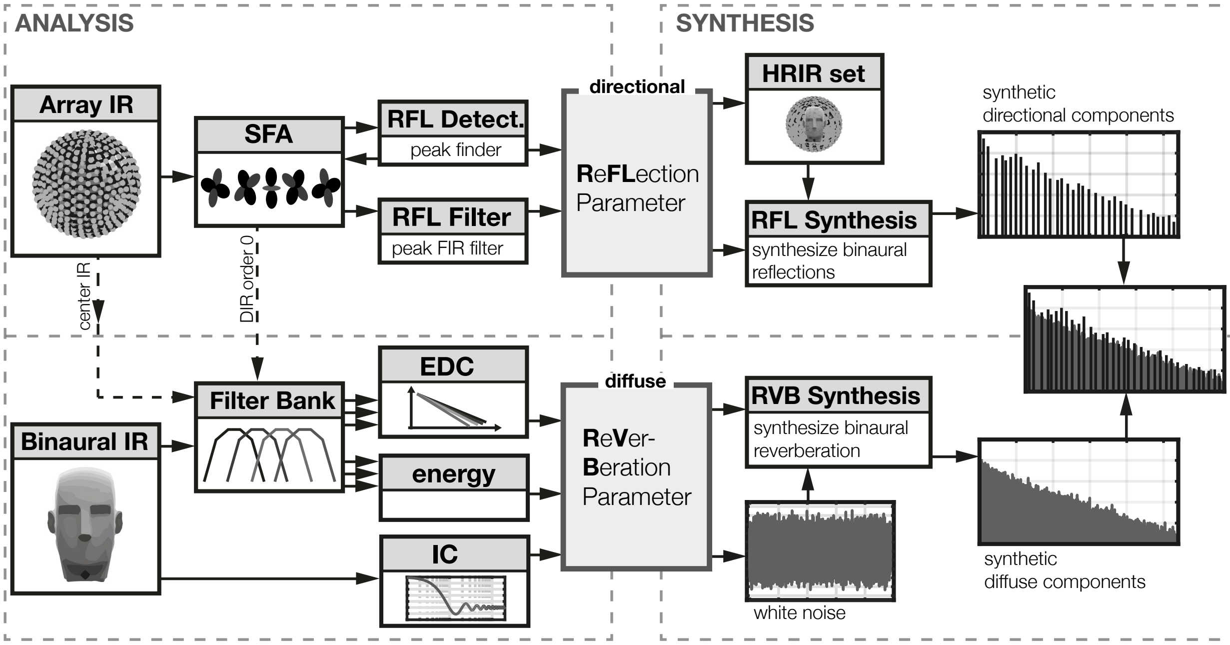

Figure 1: System Overview of parametric BRIR synthesis: Directional reflection parameters (RFL) are extracted from microphone array measurements with sound field analysis (SFA) and used to synthesize reflections with spherical HRIRs. Diffuse reverberation parameters (RVB) are extracted from reference BRIRs or array data and are used to shape white noise as a decaying diffuse masking-signal[13].

The model distinguishes between directional (Figure 1 top) and diffuse components (Figure 1 bottom). Directional reflection parameters (RFL) describe the direct sound and discrete reflections which occur in the room. They are determined with sound field analysis techniques (SFA)[16] based on spherical microphone array measurements. The processing applies sound field decomposition techniques with spherical harmonics[17] and is used to identify the directional components of the sound field. The detected direct sound and the main reflections are stored with their *direction* of arrival (DOA), *time* of arrival (TOA), *level* and spectral coloration (*reflection filter*) in a reflection list. The diffuse reverberation parameters (RVB) contain the frequency dependent energy decay curves (*EDC*), their *mean energy* and the interaural coherence (*IC*) of the late reverberation. An overview of the used parameters is given in Table 1.
In contrast to various other publications in the field of parametric audio coding and also to our previous investigations, the current work does not apply a temporal segmentation of the impulse responses regarding predictors like e.g. the perceptual mixing time[18]. No temporal differentiation between early reflections and late diffuse reverberation is used within the model. Directional components are equally determined for the early and the late part of a BRIR. Thus, it is possible to auralize late reflections which still may occur after a potential mixing time. The diffuse reverberation is synthesized for the total impulse response length as well, because the auralization of synthetic early reflections without diffuse reverberation as a masking signal has shown strong impaired results[13]. Finally, both components are superimposed for the desired listener orientations to achieve synthetic BRIR datasets for dynamic auralization. To avoid impairments of the initial time delay gap (ITDG), the synthetic diffuse reverberation before the first synthetic reflection is truncated. The level of the directional and the diffuse components are both adjusted with reference to the level of the direct sound of the measured BRIR to recreate the original ratio exactly.

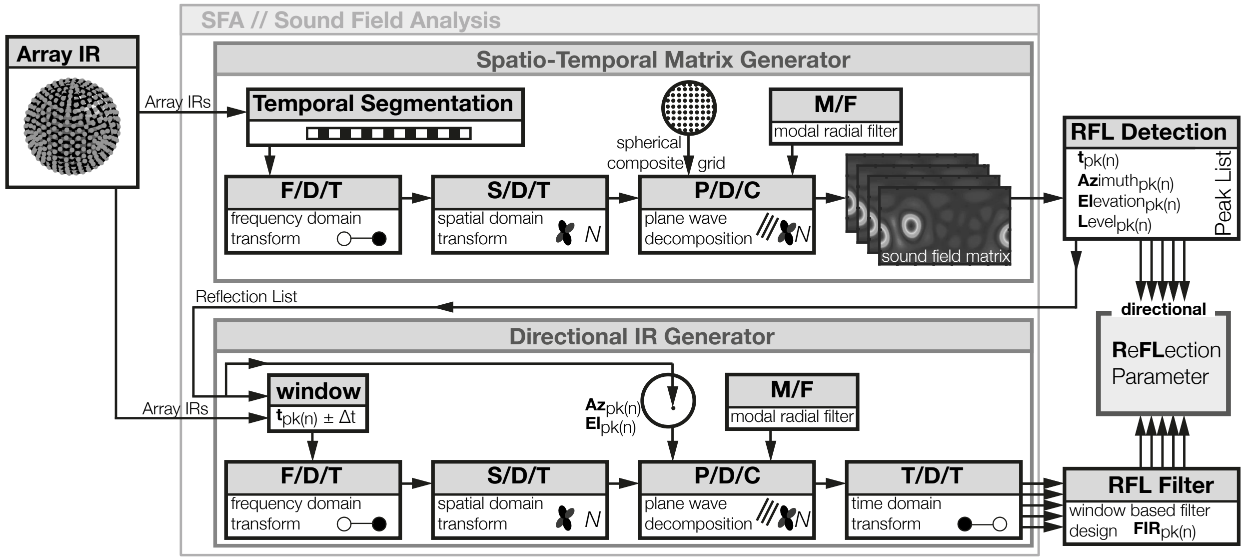

Figure 2: Determination of directional parameters with sound field analysis[16]: spatio-temporal intensity matrix generation for reflection detection and directional impulse responses (DIRs) to calculate FIR filters for each reflection[13].

Table 1: 7 Parameters of the proposed model: Directional components are determined for each reflection ($_{\mathrm{pk}}$), diffuse components are determined for each frequency band ($_{\mathrm{f}}$) of the used filter bank. The interaural coherence (IC) over frequency is also calculated and stored in a parameter.

directional	$\mathrm{direction}_{\mathrm{pk}}$	reflection $\mathrm{filter}_{\mathrm{pk}}$	$\mathrm{level}_{\mathrm{pk}}$	$\mathrm{time}_{\mathrm{pk}}$
diffuse	$\mathrm{EDC}_{\mathrm{f}}$	mean $\mathrm{energy}_{\mathrm{f}}$	IC	

i. Directional model

The synthesis of the directional components is based on a recently presented approach[13]. So far, this system has been investigated for the early part of a BRIR up to the perceptual mixing time. Now, the reflection synthesis is applied for the entire impulse response. Spherical microphone arrays can be used for a spatio-temporal analysis of acoustic environments to detect reflections in the sound field[19,20]. The determination of the RFL parameters in the presented approach is also based on spherical microphone array measurements and the use of sound field decomposition methods. The processing is shortly summarized in the following, for a detailed description of the approach we refer to our previous study[13].

Two main components are implemented: A spatio-temporal intensity matrix of the sound field is calculated for reflection analysis (Figure 2 top) and directional impulse responses (DIRs) are generated for the calculation of specific reflection filters (Figure 2 bottom). Both algorithms are based on functionalities of the SOFiA-toolbox[21]. First, the array impulse responses are temporally segmented and every segment is transformed into the spherical wave spectrum domain applying a conventional Fourier transformation first and afterwards a spatial Fourier transform. By using the respective spatial Fourier coefficients and specific modal radial filters, depending on the chosen composite grid, the spherical decomposition order N and the desired frequency f, the sound field is decomposed in multiple plane waves. This processing calculates the intensities of the sound field for each time segment and each node of the used composite grid. The output is a three-dimensional intensity-matrix which describes the spatio-temporal distribution of the sound field.

Applying a specific classification algorithm, it is possible to detect the main reflections, which are represented in the matrix with local maxima (see our previous study[13] for further information concerning the algorithm and used settings). *Time*, *level* and *direction* of each detected maxima are stored with the RFL parameters. The direct sound is detected equally based on the absolute maxima of the matrix.
Furthermore, for each reflection (and the direct sound) specific hann-windowed linear phase FIR filters, which represent the spectral coloration of the directional components, are determined and stored also within the parametric model. Using Fourier acoustics and microphone array measurements, a very high directivity can be achieved. The *reflection filters* are based on DIRs which are generated for every detected reflection and the direct sound. The processing is similar to the previous explained intensity-matrix generation, but differs in two steps: Instead of a temporal segmentation of the entire impulse responses, only small windows around the TOA of the reflections are generated. After transforming the signals again into the spatial domain, only one plane wave is calculated for each maximum in DOA using the corresponding grid node. This procedure can be compared to in-situ absorption measurements in room acoustics. An inverse Fourier transformation delivers DIRs in the time domain which are used to calculate the reflection filters.
The directional synthesis is based on these four parameters (*direction*, *reflection filter*, *level* and *time*) and arbitrary spherical HRIR datasets. Depending on the detected reflection angle (parameter *direction*), the corresponding HRIR is selected, convolved with the *reflection filter* and attenuated and delayed according to the *level* and *time* parameter. This processing is repeated for each reflection and each desired listener angle to achieve datasets with binaural synthetic reflections which can be used for dynamic auralization with arbitrary motion resolution.

ii. Diffuse model

The synthesis of the diffuse components is based on a recently presented approach[14] and is shortly summarized in the following. So far, this model has been investigated for the late part of a BRIR after the perceptual mixing time, but it also reveals good performance as a diffuse masking signal for synthetic early reflections[13]. Therefore, the diffuse synthesis is now applied for the entire impulse response length.
A reference BRIR is processed with a high-order polyphase filter bank with near perfect reconstruction (32 frequency bands; resolution: f < 8 kHz: 1/3 octave, f > 8 kHz: 1/6 octave). The energy decay curves (*EDC*) are determined for each frequency band and approximated with a polynomial curve fitting algorithm (6-degree polynomial). Furthermore, the *mean energy* of each band is calculated and stored in the parametric model. The interaural coherence (*IC*) of the BRIR is calculated based on a short-time Fourier transform (STFT)[22] using a hamming window with 256 samples and 50 % overlap between the segments. The progression of the IC vs. frequency is approximated with the same curve fitting algorithm as proposed before. A higher-order polynomial is essential ($\geq$ 12-degree) to reproduce the IC precisely.
These three parameters are used to synthesize the diffuse reverberation. A dual channel white noise signal with the length of the reference BRIR is generated and processed with the same filter bank. Each noise band is element-wise multiplied with the correponding frequency-dependent *EDC*[23]. The level of each frequency band is adapted according to the *mean energy* parameter. Afterwards, all noise bands are superimposed to achieve a two-channel broadband decaying noise signal. In a last step, the interaural coherence of the noise signal is adapted to the reference BRIR based on the *IC* parameter and specific filters[22]. If no reference BRIR has been acquired, it is possible to use an omnidirectional room impulse response (measured or generated DIR with order 0) to determine the diffuse parameters. Solely no interaural coherence can be achieved, but the theoretical coherence for a perfectly diffuse sound field[24] can be used and applied to the decaying noise. Our previous investigations have shown, that the adaptation of the coherence of the noise signals and the use of a high-order filter bank are essential for an adequate performance of the diffuse reverberation synthesis[14].

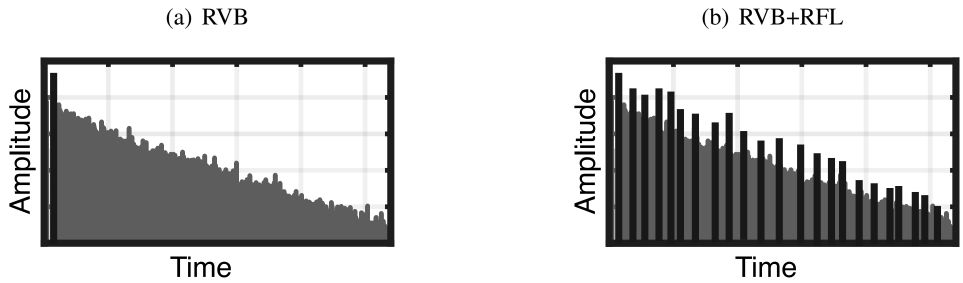

Figure 3: Illustration of the different synthetic BRIRs: (a) direct sound and diffuse reverberation (RVB) and (b) direct sound, reflections and diffuse reverberation (RVB+RFL)

Table 2: Main properties of the auralized rooms, Large Broadcast Studio (LBS) and Small Broadcast Studio (SBS)[25]

Room	**Volume**	**Area**	**RT60**$_{mean}$	$\alpha_{\mathbf{mean}}$
LBS	6098 m^3	480 m^2	1.46 s	0.24
SBS	1247 m^3	204 m^2	0.83 s	0.32

B. BRIR DATASETS

Two different rooms, located at the WDR Cologne radio broadcast studio (*KVB-Saal,* Large Broadcast Studio - "LBS" and *kleiner Sendesaal*, Small Broadcast Studio - "SBS", Table 2 and Figure 4 a/ b) in combination with two different sound sources (directional and omnidirectional, Figure 4 c/ d) have been used to evaluate the presented approach. Using each sound source successively, reference BRIRs were measured in the rooms with an *Neumann KU100* artificial head (see Figure 4 f) in 1° steps on the full azimuthal plane[25]. As directional sound source, a PA stack involving an *AD Systems Stium* Mid/High unit combined with 3 *AD Systems Flex 15* subwoofers was applied. The used omnidirectional sound source *SonicBall* has been developed and manufactured at TH Köln[26]. Spherical microphone array measurements were made with the *VariSphear* (see Figure 4 e), which also has been designed at TH Köln[27]. To avoid influences caused by limitations of the system (like spatial aliasing), a stable array configuration with 1202 sample points on a Lebedev grid and a diameter of 17.5 cm was used[25]. By theory this configuration is stable up to a frequency of approximately 18 kHz and a spherical decomposition order *N* of around 29. A rigid sphere configuration with an *Earthworks M30* omnidirectional microphone was used for the array measurements. The synthesis is based on a spherical HRIR dataset (Lebedev grid with 2702 nodes) measured also with a *Neumann KU100* artificial head[28].

Two different gradations of the approach have been used to synthesize BRIRs and are compared in this paper. The direct sound is synthesized equally in both versions based on the detected peak with the maximum amplitude. In the first version, in addition to the direct sound, only diffuse parameters are used for the synthesis ("RVB", Figure 3 a). The second version is based on the entire parametric model, using directional and diffuse components for the synthesis ("RVB+RFL", Figure 3 b). Approximately 500 respectively 100 reflections were synthesized in LBS and SBS. The evaluation of the synthesis of directional components only has been skipped in this investigation: This approach showed inadequate results in a previous study[13] because reflections were particulately perceived as discrete delays.

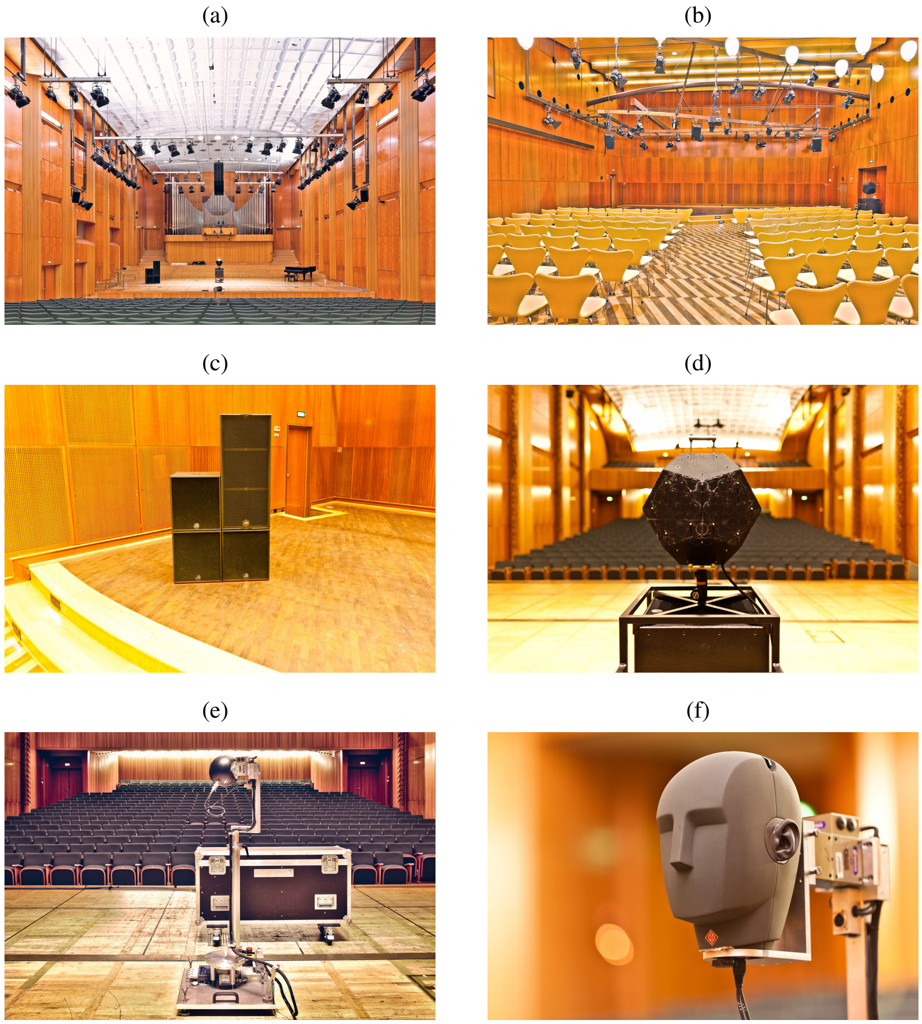

Figure 4: Rooms, sources and measurement systems: (a) Large Broadcast Studio, (b) Small Broadcast Studio, (c) PA stack as directional sound source, (d) SonicBall as omnidirectional sound source, (e) VariSphear spherical microphone array and (f) Neumann KU100 artificial head

C. SETUP AND TEST DESIGN

Dynamic binaural synthesis based on the *SoundScape Renderer*[29], *AKG K601* headphones and a *Polhemus Fastrak* head tracking system has been used for the perceptual evaluation of the presented approach. Azimuthal movements on the full horizontal plane were adapted in 1° steps in the auralization. The experiment was executed in the anechoic chamber of TH Köln to provide a quiet environment. We used a drum loop (kick, snare, hihats) with the length of three seconds and a series of two white noise bursts with 500 ms duration, 1 ms fade in/out (cosine squared ramp) and 1000 ms silence between the bursts as two different test signals for the convolution with the BRIR datasets in the binaural renderer. All datasets were loudness-normalized regarding the respective test signal and in agreement to ITU-R BS.1770-4[30]. The output playback-level was calibrated to about 64.5 dB (A) Leq.
As test design for the perceptual evaluation we used a double-blind triple-stimulus test paradigm ("ABC HR") congruent with ITU-R BS.1116-3[15]. The implementation, controlling and data acquisition of the ex-

periment was made with the MATLAB-based software *Scale*[31] in combination with a tablet computer as a graphical interface and controlling device. The subjects were able to play three stimuli ("A", "B" and "C"), "A" was based on measured BRIRs (reference), stimuli "B" and "C" were randomized with the hidden reference again and a stimulus based on the parametric model (synthesis). Then the subjects were asked to detect the synthesis ("B" or "C") and to rate the difference compared to the reference "A" with two appropriated sliders on the tablet. The faders were labelled with two attributes "identical" and "very different" (in German) only, but the ratings were captured continuously between 1 and 5. Every subject had to complete a training session prior to the grading phase to become familiar with the environment, test design and stimuli. 28 subjects (21 male, 7 female) aged between 19 and 47 (M = 25.8 years, SD = 5.3 years) took part in the investigation. The subjects were instructed to vary their listener orientation during the test. 22 participants can be denoted as experienced listeners since they took part in several listening experiments in virtual acoustics with the used dynamic binaural system before. All other subjects were extensively introduced to get into the topic. Using 8 different BRIR datasets (2_{rooms} x 2_{sources} x $2_{\text{algorithms}}$) with the two test signals, altogether 12 conditions were examined in the listening experiment. Every condition was auralized 8 times, resulting in 128 ratings per subject and a test duration of approximately one hour.

3. RESULTS AND DISCUSSION

Results of listening experiments using the ABC/HR test design are usually calculated as so-called difference grades, subtracting the rating of the reference stimulus ("A") from the rating of the chosen impaired stimulus ("B" or "C"). As a consequence, negative difference grades indicate correct identification of the treated stimulus - in our case the synthetic BRIR datasets. Values close to 0 illustrate a higher similarity of the synthesis in comparison to the given reference (from -4 = "very different" to 0 = "identical"). A Shapiro-Wilk test confirmed normal distribution of the data.

A fully repeated measures ANOVA (2 x 2 x 2 x 2) with the factors *algorithm*, *room*, *source* and *test signal* at a significance level of α = 5 % was used for statistical analysis. We observed significant main effects for *source* ($F(1,27) = 15.69$, $p < .001$, $\eta_p^2 = .37$), *algorithm* ($F(1,27) = 15.61$, $p < .001$, $\eta_p^2 = .37$) and *room* ($F(1,27) = 7.86$, $p = .009$, $\eta_p^2 = .23$). The factor *test signal* was not significant ($p = .25$). Furthermore, the analysis yielded significant interaction effects between a variety of factors. In contrast to the other factors, only the used *test signal* does not influence the performance of the parametric system. We noticed the strongest effect sizes for *source* and *algorithm*, so that especially these factors should be regarded in detail for discussion.

The mean difference grades across subjects and trials with 95 % confidence intervals are presented in Fig. 5 a. The boxplot in Fig. 5 b shows the total distribution of the difference grades of all collected ratings. In general, all ratings are in the upper third of the grading scale and no remarkable variations across the conditions can be observed at first sight. So, an adequate performance can be observed for all investigated conditions. Pairwise t-tests with Hochberg correction were conducted as post hoc analysis for each factor and results are presented in the following. BRIRs based on the parametric model were significantly detected in all conditions. All gradations of the factors *algorithm* (RVB vs. RFL+RVB, Figure 3), *room* (LBS vs. SBS, Figure 4 a/ b), *source* (omnidirectional vs. directional, Figure 4 c/ d) and *test signal* (drums vs. noise bursts) were compared to each other to analyze differences between them.

algorithm: We observed the following results for the post hoc analysis of the factor *algorithm*: In 5 of 8 conditions, the use of directional and diffuse components (RVB+RFL) for the synthesis was rated with a higher similarity than using diffuse components (RVB). But only in 3 conditions (LBS directional noise, LBS omnidirectional drums and LBS omnidirectional noise) we observed significantly higher ratings for the entire parametric model (RVB+RFL). The use of diffuse components only was never rated significantly

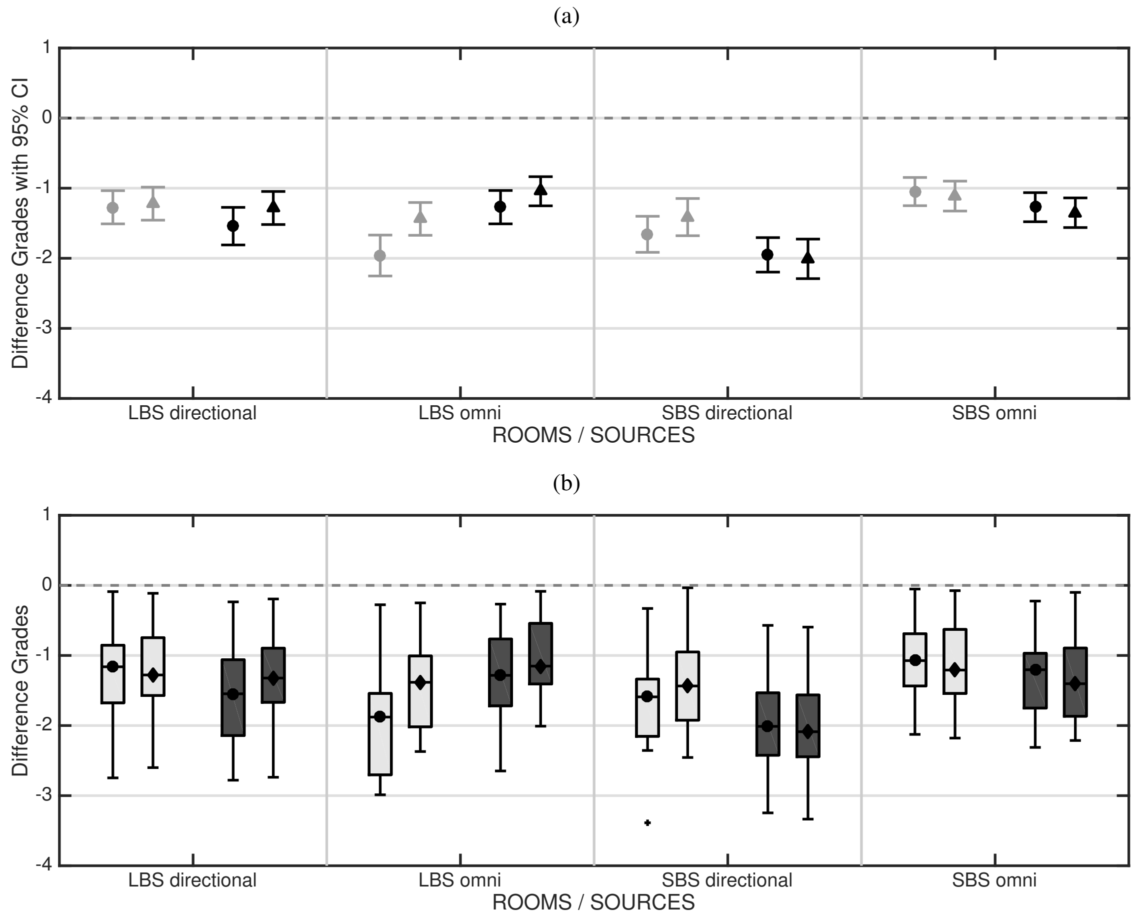

Figure 5: Results of the listening experiment: drums (grey) vs. noise (black), RVB (circle) vs. RVB+RFL (triangle) for different rooms (LBS vs. SBS) and sound sources (omnidirectional vs. directional); (a) Mean difference grades and 95% CI and (b) distribution of the ratings of all subjects. The whiskers are defined as Q±1.5∗IQR (Q: upper/lower quartile, IQR: interquartile range)

higher than the directional and diffuse components. That implies that the decaying diffuse sound field is essential for an adequate performance but depending on the acoustical environment the additional use of directional components can further enhance the auralization. In general, both algorithms were rated with minor differences only compared to each other. Similar results have been observed in our previous publication[13]. So for the majority of the investigated conditions, the synthesis using only diffuse reverberation achieves a good performance as well. But it should be considered, that all gradations include a direction-sensitive synthesis of the direct sound. Informal listening tests evidenced an impaired auralization without dynamic rendering of the direct sound.

source: The post hoc analysis of the factor *source* showed significantly different ratings for all compared conditions. In 6 of 8 conditions, the use of the omnidirectional sound source was rated with a significantly higher similarity, the remaining 2 conditions indicate significantly higher results for the directional sound source (LBS RVB drums, LBS RFL+RVB drums). Although the ANOVA showed a significant main effect for *source*, no trend could be noticed clearly in contrast to the factor *algorithm*, because we observed significant ratings for both sources alternately. Therefore, no precise statement concerning the auralization performance depending on the used sound source can be draft. Furthermore, it has to be remarked that most of the t-tests for the source comparison showed only small p-values.

room: In 6 of 8 conditions, we observed significantly different ratings for the factor *room*. The bigger room (LBS) was rated higher in 4 conditions (RVB directional drums, RVB+RFL directional noise, RVB directional noise, RVB+RFL omnidirectional noise). Using drums and the omnidirectional source, SBS was rated with a significantly higher similarity for both algorithms (RVB omnidirectional drums, RVB+RFL omnidirectional drums). The non-significant comparisons indicate a trend in favour of the bigger room (LBS). As before for the factor *source*, no clear statement can be made for the factor *room* as we observed again alternately significant higher ratings for both rooms and again small p-values.

test signal: Since the ANOVA indicated no significant main effect for *test signal*, the post-hoc analysis of this factor can be disregarded.

4. CONCLUSION

A scalable system for the synthesis of BRIRs for dynamic auralization was presented and perceptually evaluated. The synthesis is based on a parametric model which is determined from measured BRIRs and characterizes the acoustical environment with seven parameters only. The system is separated in two main parts with two processing stages each: The analysis-part deals with the parametrization of the measured data, the synthesis-part generates the artificial BRIRs using the parametric description. The approach distinguishes between directional and diffuse components of the sound field, which are determined, stored and synthesized apart from each other. The processing of the directional analysis is based on sound field decomposition techniques using spherical microphone arrays. The directional synthesis applies spherical HRIRs and a reflection list. Diffuse components are analyzed using a high-order filter bank and a polynomial approximation of the frequency dependent energy decay curves and the interaural coherence. The diffuse synthesis is based on dual channel white noise signals which are shaped according to the original reverberation time, coherence and spectrum using the parametric description.

The approach has been perceptually evaluated in different acoustical conditions: Two rooms (large and small broadcast studio) and two sound sources (directional and omnidirectional) have been compared to each other with two different test signals (drums and noise bursts). Furthermore, two gradations of the algorithm have been observed. For each condition, a synthetic BRIR dataset with diffuse components only (RVB) and a dataset with directional and diffuse components (RVB+RFL) were generated and compared against a measured BRIR as a reference. The direct sound was synthesized in the same way in both versions. In general, the approach works quite satisfying, without big variations across rooms, test signals and sound sources. The majority of the investigated stimuli based on synthetic BRIRs were rated in the upper third of the grading scale. We observed the perceptual highest similarities compared to the measured reference using directional and diffuse components for the synthesis. But surprisingly, the stimuli using diffuse components only were rated nearly equal and this approach also yielded an adequate performance for almost all conditions. Therefore regarding the application, it should be considered if the complex processing for the reflection detection is necessary, or if an even simpler approach performs adequate as well. For example, a fixed reflection pattern which is only adapted to peaks in the impulse response is conceivable[32]. Furthermore, in some applications, the use of diffuse components only could be sufficient.

The presented parametric approach offers several possibilities for the BRIR synthesis. The system can be personalized by using individual HRIRs. For example, a listener can use his own HRIRs for the auralization of rooms where he has never been before. Furthermore, the parameters allow a systematic adaptation of the BRIRs and the acoustical conditions. Reverberation time, coherence, the reflection pattern or the spectral coloration can be easily modified e.g. to simulate room acoustic modifications. Thus, this approach can be applied for a methodical perceptual evaluation of room acoustic parameters. The system can also be used as a parametric convolution reverb which provides a creative tool for sound engineers and producers. The amount of data is heavily decreased due to the parametric description. Indeed for professional environments, this feature can be neglected, but applications for mobile devices could benefit. In future, the resolution of the directional analysis can be scaled down and its impact should be investigated more into detail. Furthermore, the study should be expanded and more rooms have to be synthesized and perceptual evaluated.

ACKNOWLEDGMENTS

The research activities were funded by the Federal Ministry of Education and Research (BMBF) in Germany under the support code 03FH005I3-MoNRa. The authors thank all participants of the listening experiment for their support. Special thanks to Pia Krauss and Tim Lübeck for the help during the tests. Static versions of the presented stimuli can be downloaded on the following webpage:
`http://www.audiogroup.web.th-koeln.de/Boston2017.html`

REFERENCES

[1] Blauert, J., Spatial hearing: the psychophysics of human sound localization, 1997.

[2] Begault, D. R., "Perceptual Effects of Synthetic Reverberation on Three-Dimensional Audio Systems," *Journal of the Audio Engineering Society*, 40(11), pp. 895–904, 1992.

[3] Jot, J. M., Cerveau, L., and Warusfel, O., "Analysis and synthesis of room reverberation based on a statistical time-frequency model," *Audio Engineering Society Convention 103*, 1997.

[4] Valimaki, V., Parker, J. D., Savioja, L., Smith, J. O., and Abel, J. S., "Fifty Years of Artificial Reverberation," *IEEE Transactions on Audio, Speech, and Language Processing*, 20(5), pp. 1421–1448, 2012.

[5] Valimaki, V., Parker, J. D., Smith, J. O., and Abel, J. S., "More Than Fifty Years of Artificial Reverberation," in *AES 60th Conference on Dereverberation and Reverberation of Audio, Music, and Speech*, pp. 1–12, 2015.

[6] Savioja, L., Huopaniemi, J., and Lokki, T., "Creating interactive virtual acoustic environments," *Journal of the Audio Engineering Society*, 47(9), 1999.

[7] Wendt, T., van de Par, S., and Ewert, S. D., "A Computationally-Efficient and Perceptually-Plausible Algorithm for Binaural Room Impulse Response Simulation," *Journal of the Audio Engineering Society*, 62(11), pp. 748–766, 2014.

[8] Coleman, P., Franck, A., Jackson, P., Hughes, R., Remaggi, L., and Melchior, F., "Object-Based Reverberation for Spatial Audio," *Journal of the Audio Engineering Society*, 65(1/2), pp. 66–77, 2017.

[9] Allen, J. B. and Berkley, D. A., "Image method for efficiently simulating small-room acoustics," *Journal of the Acoustical Society of America*, 65(4), p. 943, 1979.

[10] Merimaa, J. and Pulkki, V., "Spatial impulse response rendering I: Analysis and synthesis," *AES: Journal of the Audio Engineering Society*, 53(12), pp. 1115–1127, 2005.

[11] Tervo, S., Pätynen, J., Kuusinen, A., and Lokki, T., "Spatial Decomposition Method for Room Impulse Responses," *Journal of the Audio Engineering Society*, 61(1/2), pp. 17–28, 2013.

[12] Jot, J. M. and Chaigne, A., "Digital delay networks for designing artificial reverberators," *Audio Engineering Society Convention 90*, 1991.

[13] Stade, P., Arend, J. M., and Pörschmann, C., "Perceptual Evaluation of Synthetic Early Binaural Room Impulse Responses Based on a Parametric Model," in *Proc. of the 142nd AES Convention*, pp. 1–10, 2017.

[14] Stade, P. and Arend, J. M., "Perceptual Evaluation of Synthetic Late Binaural Reverberation Based on a Parametric Model ," in *Proc. of the AES International Conference on Headphone Technology*, pp. 80–87, Aalborg, 2016.

[15] ITU-R BS.1116-3, "Methods for the subjective assessment of small impairments in audio systems," 2015.

[16] Bernschütz, B., Stade, P., and Rühl, M., "Sound Field Analysis in Room Acoustics ," in *Proc. of the VDT International Convention 2012*, Cologne, 2012.

[17] Williams, E. G., *Fourier Acoustics*, Sound Radiation and Nearfield Acoustical Holography, Academic Press, 1999.

[18] Lindau, A., Kosanke, L., and Weinzierl, S., "Perceptual evaluation of model-and signal-based predictors of the mixing time in binaural room impulse responses," *Journal of the Audio Engineering Society*, 60(11), pp. 887–898, 2012.

[19] Meyer, J. and Elko, G., "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," *Proc. of ICASSP '02*, 2, pp. II–1781–II–1784, 2002.

[20] Duraiswami, R., Li, Z., Grassi, E., Gumerov, N. A., and Zotkin, D., "Plane-wave decomposition analysis for spherical microphone arrays," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 150–153, 2005.

[21] Bernschütz, B., Pörschmann, C., Spors, S., and Weinzierl, S., "SOFiA sound field analysis toolbox," in *Proc. of the International Conference on Spatial Audio - ICSA 2011*, Detmold, 2011.

[22] Menzer, F. and Faller, C., "Investigations on an Early-Reflection-Free Model for BRIRs*," *Journal of the Audio Engineering Society*, 58(9), pp. 709–723, 2010.

[23] Moorer, J. A., "About this reverberation business," *Computer music journal*, 3(2), pp. 13–28, 1979.

[24] Cook, R. K., "Measurement of Correlation Coefficients in Reverberant Sound Fields," *The Journal of the Acoustical Society of America*, 27(6), pp. 1072–1077, 1955.

[25] Stade, P., Bernschütz, B., and Rühl, M., "A Spatial Audio Impulse Response Compilation Captured at the WDR Broadcast Studios," in *Proc. of the VDT International Convention 2012*, Cologne, 2012.

[26] Meulemann, J., Bernschütz, B., and Pörschmann, C., "Entwurf und Aufbau eines konzentrischen Mehrwegedodekaeders," in *Fortschritte der Akustik - DAGA 2011*, pp. 719–720, Düsseldorf, 2011.

[27] Bernschütz, B., Pörschmann, C., Spors, S., and Weinzierl, S., "Entwurf und Aufbau eines variablen sphärischen Mikrofonarrays für Forschungsanwendungen in Raumakustik und Virtual Audio," in *Fortschritte der Akustik - DAGA 2010*, pp. 717–718, Berlin, 2010.

[28] Bernschütz, B., "A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100," in *AIA DAGA 2013*, pp. 592–595, Meran, 2013.

[29] Geier, M., Ahrens, J., and Spors, S., "The soundscape renderer: A unified spatial audio reproduction framework for arbitrary rendering methods," *Audio Engineering Society Convention 124*, 2008.

[30] ITU-R BS 1770-4, "Algorithms to measure audio programme loudness and true-peak audio level ," 2015.

[31] Vazquez, A., "Scale - Conducting Psychoacoustic Experiments with Dynamic Binaural Synthesis," in *Fortschritte der Akustik - DAGA 2015*, pp. 1128–1130, Nürnberg, 2015.

[32] Pörschmann, C. and Wiefling, S., "Perceptual Aspects of Dynamic Binaural Synthesis based on Measured Omnidirectional Room Impulse Responses," in *International Conference on Spatial Audio - ICSA*, pp. 1–8, Graz, 2015.