Perceptual Aspects of Dynamic Binaural Synthesis based on Measured Omnidirectional Room Impulse Responses

C. Pörschmann, S. Wiefling

Fachhochschule Köln, Institut f. Nachrichtentechnik,50679 Köln, Germany,
Email: Christoph.Poerschmann@fh-koeln.de

Abstract

The auralization of acoustic environments applying dynamic binaural synthesis can be used for multiple applications. Circular sets of binaural room impulse responses (BRIRs) are often acquired by performing measurements with rotated dummy heads. This procedure is rather costly and therefore not always feasible in practice.

Recently, an approach to synthesize circular BRIR sets based on a single omni-directional room impulse response (RIR) was presented [1]. Direct sound, early reflections and diffuse reverberation are extracted from the latter and treated separately. Spatial information is added according to generic geometric assumptions and by exploiting certain psychoacoustic knowledge about diffuse reverberation, like for example the perceptual mixing time.

While it is the highest goal in auralization to archive authentic reproduction of an auditory environment, i.e. the recipient cannot distinguish between a real environment and an equivalent virtual one, a sufficiently plausible representation is actually adequate for many practical applications. This is the basic assumption of the presented approach. Naturally, there are certain physical and perceptual differences between actual measured BRIRs and synthesized ones which are based on a single omni-directional RIR. In order to examine these perceptual differences, a listening experiment based on a SAQI test paradigm [2] for four rooms with different room acoustic properties was conducted. The results show that the proposed solution is promising and has the potential to create circular sets of BRIRs from one single one-channel measurement.

Introduction

Binaural synthesis can be used for a headphone based presentation of auditory virtual environments. This approach can be applied in various areas like audio engineering, telecommunication, or architectural acoustics to create a natural and plausible room impression. By considering the listener's head movements in the auralization, localization accuracy increases and externalization of virtual sound sources improves. Several commercial or scientific rendering engines are available which allow adapting the sound field presented via the headphones according to the head movements in real time (e.g. [3]).

However, applications using this technology in everyday life are still rare. One reason is the complex procedure for the acquisition of spatial sound fields which cannot be conducted with standard equipment and techniques. Two different approaches to collect room data are commonly applied: Either simulated sound fields based on a geometric model are used. This allows for a free positioning of the listener in the virtual room and an adaptation of the sound field according to the listener's movements. Alternatively, measurements of binaural room impulse responses (BRIRs) can be performed. If the reproduction aims at being authentic, i.e. the listener cannot distinguish a presented virtual environment from an equivalent real one, head rotations of the listener need to be considered. For this a circular (or even a full spherical) set of BRIRs needs to be captured. However, such a measurement implies the use of complex devices (e.g. a rotatable artificial head) and an enormous amount of time. Thus for many state-of-the-art applications in the field of spatial audio, the effort is so high that circular sets of BRIRs are not used.

The algorithm introduced in this paper aims at a sufficiently plausible auralization of a room, based on a simple measurement procedure. For many of the applications mentioned above, a plausible presentation is quite sufficient and no authentic reproduction of the sound field is pursued. The approach described here requires only one measured omnidirectional room impulse response to synthesize a complete set of BRIRs.

The presented study investigates to what extent one measured omnidirectional room impulse response can be enhanced in order to deliver a plausible auralization when using dynamic binaural synthesis. The paper is organized as follows: Section 1 describes the proposed algorithm performing this spatialization. Section 2 illustrates the performance and the results of a psychoacoustic evaluation of the algorithm. Measurements of circular sets of BRIRs and of omnidirectional impulse responses of four different rooms were used [4]. Within two psychoacoustic experiments, perceptual differences between the measured sets of binaural impulse responses and the synthesized sets were evaluated. In Experiment 1 the general difference was investigated while in Experiment 2 applying a SAQI test paradigm [2] a large number of perceptual attributes was tested. The results show to what extent a plausible auralization can be obtained and which of the perceptual cues are responsible for the remaining differences.

Basic idea of the algorithm

The aim of the algorithm described here is the synthesis of a circular set of binaural room impulse responses (BRIRs) from a single omnidirectionally measured room impulse response (RIR). Therefore, predictable information from geometrical acoustics as well as knowledge regarding the perception of diffuse sound fields is applied.

The RIR is split into different parts. In the early part the direct sound is identified. According to the direction of incidence of the direct sound, which has been determined by the user in advance, the appropriate head-related impulse response (HRIR) is picked. This HRIR is convolved with the time frame of the omnidirectional RIR which contains the direct sound.

The early reflections reach the listener according to a room and position-specific reflection pattern from various directions. In the algorithm, arbitrary spatial reflection patterns are chosen, which are regarding the directions of incidence not at all adapted to the geometrical shape of the room. The temporal structure of the early reflections is detected from the omnidirectional RIR by analyzing peaks and dips. By convolving the HRIRs of each of the directions of sound incidence with the time section of the respective reflection a binaural synthesis of the early geometric reflected part is applied.

The diffuse reverberation part is considered to reach the listener temporarily and spatially equally distributed. Several studies [5-7] have shown that the part of the measured impulse response after a perceptual mixing time can be simplified. In this investigation, the diffuse part is replaced by binaural noise, which is frequency-dependently shaped according to the decay of the omnidirectional impulse response. Even though an exact reconstruction of the temporal and spatial structure is ideally not required [6] small perceptual differences might remain [8]. Several studies have shown that the value of the mixing time is room dependent and can be chosen according to predictors which are calculated based on geometric room properties [7]. However, as the transition from the early reflection part to the diffuse reverberation part is quite soft, a smooth crossover has been chosen in the design of the algorithm.

Since relevant information about the binaural sound field is absent in the omnidirectional RIR, the algorithm presented here incorporates several inaccuracies and errors. For example, a point source is assumed for all synthetic BRIR sets. Thus, it is not possible to rebuild source width and other properties of the source correctly. Furthermore, the directions of incidence of the synthesized early reflections are not in line with the original ones. Hence, differences in perceptual spatial properties (e.g. envelopment) between the original room and the synthesized room may occur. Finally, the diffusion of the early reflections is not rebuilt precisely. In the algorithm, only a simple model estimating the part of the geometrically reflected and the diffusely reflected energy is applied. In the listening experiments, which complete this study, it is evaluated which perceptual influences these approximations have.

In the literature, several approaches to enhance an omnidirectional impulse response to a binaural one have been proposed [9, 10]. They rely on statistical considerations and perform an adequate reconstruction of the coherence.

Implementation

The algorithm was implemented in Matlab and includes the following steps:

- The algorithm requires the omnidirectional RIR as well as basic geometric information about room volume and the incidence direction of the direct sound as input data. Furthermore, the algorithm has access to a set of HRIRs and to a pre-processed sequence of binaural noise. Both the HRIRs and the binaural noise were acquired based on measurements with a Neumann KU 100 artificial head [11].
- The algorithm only applies to frequency components above 200 Hz. For lower frequencies the coherence of a typical binaural impulse response is nearly one and the omnidirectional impulse response can be maintained.
- The first relevant maximum in the impulse response is identified as direct sound. A frame of 5 ms starting slightly before this maximum is convolved with the HRIR of the direct sound incidence direction.
- The time section from 5 ms to 150 ms, is assigned to the early reflections and the transition towards the diffuse reverberation. In order to determine the early reflections, the energy of a sliding window of 8 ms length is calculated and areas are marked which contain high energy. Peaks which are 6 dB above the RMS are determined and assigned to geometric reflections. Smoothly windowed sections around the peaks are convolved with the HRIRs according to the predefined reflection pattern.
- The omnidirectional RIR excluding the sections of the direct sound and the geometric reflections is split up into 1/6 octave bands by applying a near perfect reconstruction filterbank [12].
- In each of the frequency bands, the binaural diffuse part is synthesized. Therefore, the envelope of the binaural noise is shaped according to the envelope of the omnidirectional impulse response. Thus, the binaural parameters (e.g. coherence) of the noise and the envelope of the omnidirectional room impulse response are both maintained. This method has already been described more detailed in [13].
- The direct sound component, the synthesized geometric reflections, and the binaural diffuse reverberation are appropriately weighted and summed up in a way that the temporal structure of the synthesized BRIR fits the measured omni-directional RIR in all frequency bands.

Figure 1 shows the basic concept of the algorithm.

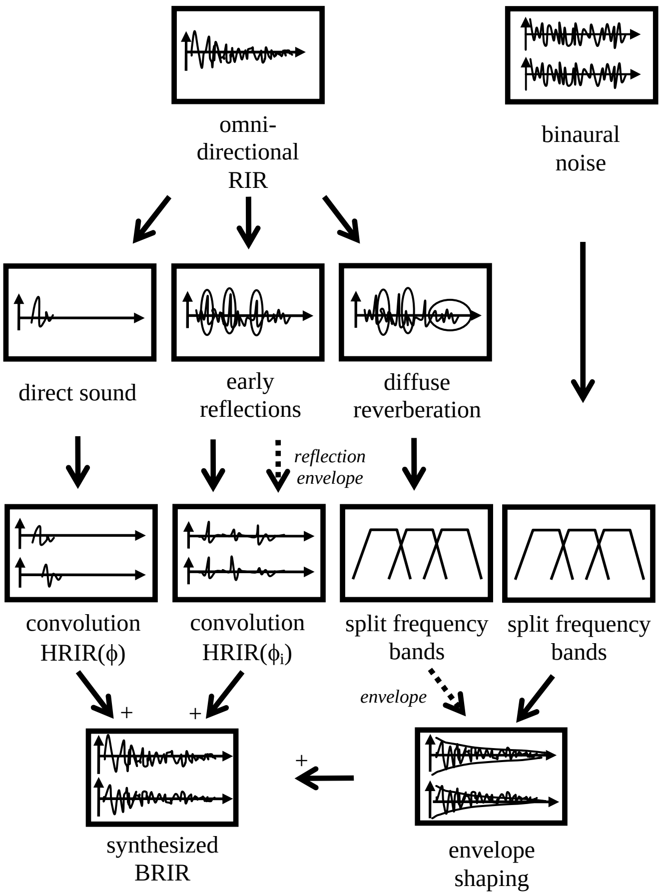

Figure 1: Block diagram of the algorithm for synthesizing a binaural room impulse response based on one single omnidirectional room impulse response.

The calculation of the synthesized BRIRs is repeated for constant shifts in the azimuth angle (e.g. 1°) for the direct and the reflected sound. Thus, a circular set of BRIRs is created, which can be used for dynamic binaural synthesis.

Measured rooms

To evaluate the algorithm and to perform the psychoacoustic experiments, measurement data from four different rooms were used [11]. The measurements comprised both, a circular reference set of BRIRs and an omnidirectional impulse response, measured at the pivot position of the artificial head. The binaural measurements were performed in steps of 1° in the horizontal plane with a Neumann KU 100 artificial head [11]. A Microtech Gefell M296S microphone was used for the TGC room; for the other rooms, an Earthworks M30 microphone was applied. The type of the loudspeaker and the positions of sender and receiver remained unchanged in each series of measurements.

It was considered that the rooms are significantly different regarding their geometrical and acoustical properties. The following rooms were chosen:

- Control Room 7 (WDR Köln):
  volume 168 m³, base area 60 m²,
  T_{60}(500/1000Hz) < 0,25 s
- Large Broadcast Studio (WDR Köln):
  volume 6100 m³, base area 579 m²,
  T_{60}(500/1000Hz) = 1,8 s
- Small Broadcast Studio (WDR Köln):
  volume 1247 m³, base area 220 m²,
  T_{60}(500/1000Hz) = 0,9 s
- TGC Training Room (Köln):
  volume 191 m³, base area 67 m²,
  T_{60}(500/1000Hz) = 2,3 s

Psychoacoustic experiments

To evaluate the performance of the algorithm, a listening experiment using the spatial audio quality inventory (SAQI) [2] was performed. SAQI is a consensus vocabulary which allows the assessment of apparatus-dependent or algorithm-related perceptual differences of a virtual auditory environment. The SAQI test follows the semantic test paradigm where the attributes are rated on bipolar or unipolar scales. In the experiments presented here, the test was used to evaluate the perceptual differences between a synthesized stimulus and a reference.

The psychoacoustic investigations aim at a quantification of the perceptual differences between a measured set of BRIRs and a synthesized one based on a single omnidirectional measurement. Furthermore, the evaluation shall identify the relevant attributes describing the perceptual differences between binaurally measured sets and synthesized ones. Based on the results of the psychoacoustic experiment the algorithm can be optimized further.

The experiments were set up, controlled, and executed running the software SCALE [14]. This software directly accesses the SoundScape Renderer (SSR) [3], which was used as the binaural renderer. The listeners had to enter their response on a touch-screen tablet computer (iPad).

Due to the complexity of the evaluation task, the experiments were run with a small group of expert listeners. The 11 participants were aged between 28 and 63 years (average 42). Five of the participants were sound engineers from the WDR broadcasting corporation. The others were members of the acoustics group at Fachhochschule Köln and were experienced in performing listening tests and in evaluating virtual auditory environments. As all participants were German-speaking and, apart from one subject, their mother tongue was German, the experiments were explained and conducted in German. Thus, the translated set of attributes and scales (SAQI-GER) as suggested in [15] was used.

The audio signals were presented via an AKG K-601 open headphone. To acquire the head movements, a Polhemus Fastrak tracking system was applied. The playback-level was calibrated to 65 dB(A).

A looped drum and a guitar sequence with a length between four and eight seconds were used. These source signals have already been presented in previous listening experiments and were regarded as being suitable to evaluate the perceptual attributes of different rooms.

In each of the tests, the playback order of the stimuli was pseudo-randomized. The subjects were allowed to listen to each of the reference and test stimuli as often and as long as required.

The experiments were separated into two parts. In all presented conditions the measured set of BRIRS was used as reference. In Experiment 1 a test of the global difference between stimulus and reference was assessed. Both, the synthesized BRIRs as well as the original omnidirectional RIR, were tested and the perceptual differences to the reference were investigated. Experiment 2 comprised the detailed evaluation of a broad set of perceptual attributes. Here, only the set of synthesized BRIRs was compared to the reference.

For the evaluation, the average values and the 95% confidence intervals applying a bootstrapping method with 2000 samples [16] were calculated.

Experiment 1

In this experiment 16 comparisons for four rooms and two different source signals were performed. For each of the conditions the perceptual difference of the measured omnidirectional RIRs and the synthesized set of BRIRs to the respective reference set of measured BRIRs were investigated. The subjects rated on a continuous scale from 0 (none) to 3 (very large).

A t-test comparing the differences of the monaural and the synthesized binaural condition showed that for all comparisons the synthesized BRIRs were rated significantly less different to the reference set than the omnidirectional RIRs (α=0.005). Especially for the Small Broadcast Studio, the perceptual differences of the synthesized BRIRs were surprisingly low. The tendencies were similar for both types of source signals. Figure 2 shows the results of the experiment.

However, the analysis showed significant differences between the rooms. The synthesis was rated perceptually closer to the reference, e.g. for the Small Broadcast Studio than for the TGC room (α=0.05).

The generally small differences between the measured BRIRs and the synthesized ones are somehow surprising, because several properties of the spatial sound field were not reconstructed correctly. For example, the properties and the direction of incidence of the early reflections were adapted based on a previously defined reflection pattern. Furthermore, no information about the diffusion of any of the reflecting surfaces of the room was available.

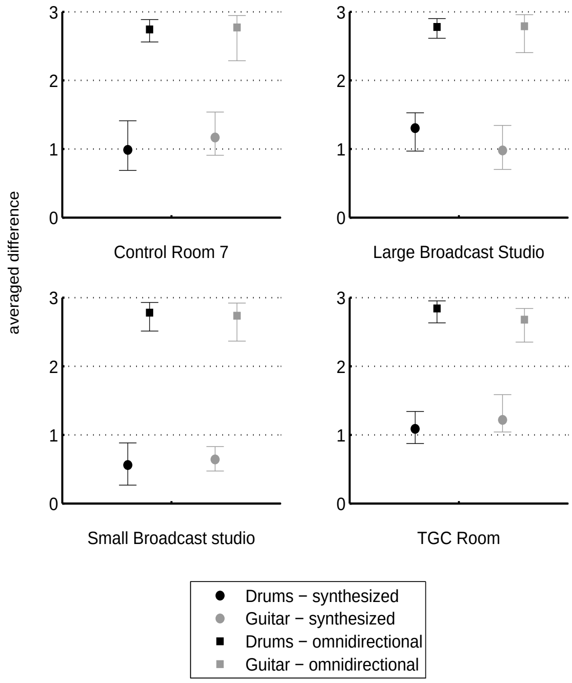

Figure 2: Perceptual differences (mean values and 95 % confidence intervals) between the test stimuli and the reference (measured set of BRIRs) of the following rooms: Control Room 7 WDR, Large Broadcast Studio WDR, Small Broadcast Studio WDR, TGC Training Room (Köln). In each plot, the judgements of the binaurally synthesized rooms and applying the omni-directionally measured room impulse responses are shown. The comparison was performed on a scale from 0 (none) to 3 (very large).

Experiment 1 shows that the algorithm has a performance which is worth evaluating in more detail. Especially in order to enhance the algorithm, an extended knowledge of the different perceptual attributes is of interest. The subsequently performed Experiment 2 focuses on a detailed analysis of these aspects.

Experiment 2

The same four rooms and the same two source signals were presented in the second experiment. The presented stimuli were rated using the spatial audio quality inventory (SAQI). However, due to the complex and time-consuming task of evaluating a large number of perceptual attributes for each stimulus, only the synthesized set of BRIRs was tested here. The detailed evaluation of the omnidirectional RIRs was dropped because it was of no major interest for the further assessment.

The experiment shall answer the question which perceptual cues cause the remaining differences. Furthermore, it shall be analyzed if these perceptual attributes can be explained by inaccuracies due to missing information on the geometric properties of the rooms. Thus, the results can provide information if a more detailed geometric model achieves better results or if other optimizations are required to improve the algorithm's performance.

Concerning the choice of the attributes, the following groups were investigated: General attributes, timbre, tonalness, dynamics, geometry, room, time behavior, and artefacts as suggested in [2]. However, as some of the SAQI attributes were of no importance for this investigation, informal pretests with a small group of listeners were conducted before the actual experiment in order to choose the relevant attributes.

Furthermore, in previous tests, some stimuli caused ambiguous horizontal shifts depending on the head rotation. Thus, in contrast to the original SAQI proposal, an unsigned value for the horizontal shift instead of a signed one was used. Details on the attributes can be retrieved from the Figures 3 – 6.

In Figures 3 – 6, the perceptual differences between the test stimuli (synthesized set of BRIRs) and the reference (measured set of BRIRs) for the different rooms are shown. For each attribute, the values (mean and 95 % confidence intervals) for the drums (black) and the guitar (grey) source signals are given. The transparent circles in the plot indicate that the differences between reference and stimulus are significant at a 5% level.

For the general difference larger values were observed than in Experiment 1. As no omnidirectional room impulse responses were presented in Experiment 2 the scales were used with a bias.

The results show that some of the perceptual attributes are hardly affected by the synthesis. For example, for echos, artefacts and distortions as well as for vertical shift, nearly no significant differences were observed. Sharpness, roughness, coloration due to comb filtering or metallic coloration remained nearly unchanged as well.

A horizontal shift was obvious in all conditions. In more detail, as the interview with the subjects after the test revealed, this shift was in most conditions observed as leading into the same direction. An explanation for this might be that the room reflections were synthesized according to a predefined pattern. Summing localization cannot play a role since the time difference between direct sound and the first reflection is more than 2 ms in all conditions However, the fixed reflection pattern which implies an early reflection from rear right might have caused a slight shift. Similar image shifts due to early reflections have been observed in [17] and [18]. In the following consideration of the different rooms, this influence is not further mentioned.

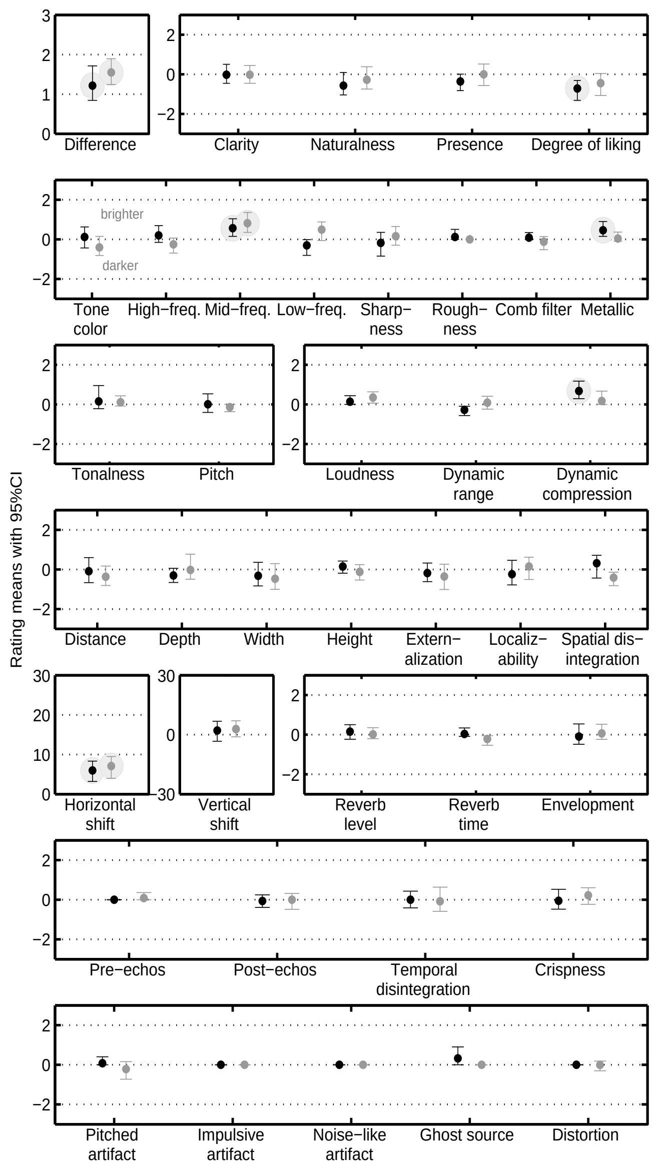

Figure 3: Control Room 7: Perceptual attributes (mean values and 95 % confidence intervals) for the drums (black) and the guitar (grey) source signal.

A closer look at the scores for the different attributes of Control Room 7 reveals that mainly tone color is affected. The scores for the other attributes do not show significant differences between test stimulus and reference. However, the general difference attribute indicates that the subjects very well perceived a difference, which probably could not be manifested in the detailed description.

As the reverberation time of Control Room 7 is very short, the influence of single reflections might be of greater importance. An inadequate reconstruction of spatial properties of the reflections might lead to modifications of the timbre. In contrast, attributes related to reverberation indicate no difference between test stimulus and reference.

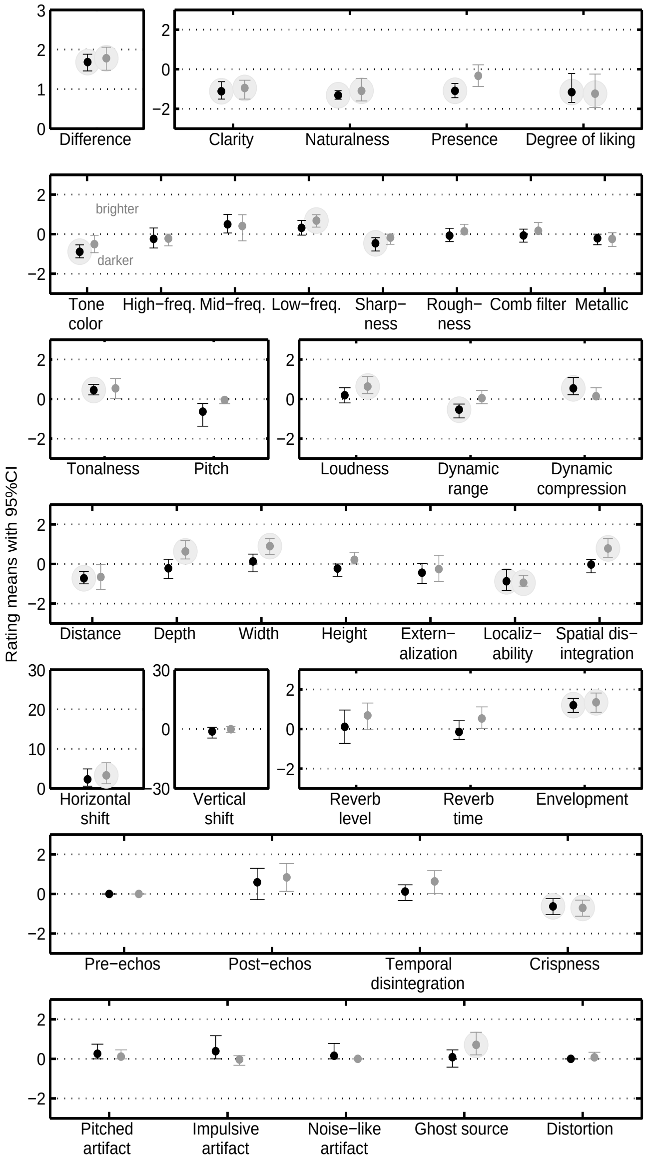

Figure 4: Large Broadcast Studio: Perceptual attributes (mean values and 95 % confidence intervals) for the drums (black) and the guitar (grey) source signal.

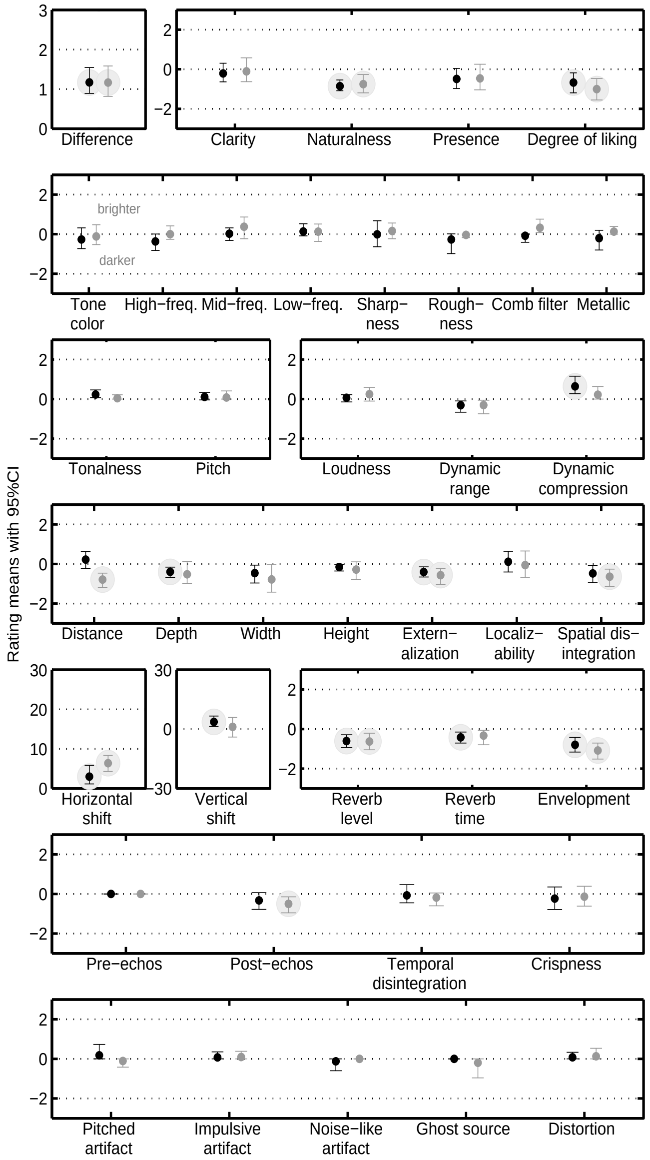

Figure 5: Small Broadcast Studio: Perceptual attributes (mean values and 95 % confidence intervals) for the drums (black) and the guitar (grey) source signal.

Considering the Large Broadcast Studio, the subjects rated many of the reverberation related attributes as significantly different. For example, the envelopment and more general attributes which are influenced by reverberation like clarity, naturalness and presence were affected here. The complex room acoustical properties of the Large Broadcast Studio were probably not rebuilt in the synthetization. Furthermore, the attributes describing shape and position of the sound source were perceived differently. Distance, depth, width and the localizability show at least for one of the source signals significant differences compared to the reference condition. Finally, the listeners identified a perceptually significant change in tone color.

For this room, the use of an improved model, which incorporates more information about the room geometry, could be considered.

The smallest differences between reference and test conditions were detected for the Small Broadcast Studio. Here, mainly geometric attributes and the reverberation attributes cause the dissimilarities: Significant differences for distance, depth and the externalization of the sound source were observed.

Additionally, attributes evaluating the reverberation were affected. However, for this room many perceptual attributes were rated very similar between reference and test stimulus. No significant difference for all attributes describing timbre, tone color, tonalness, time behavior and artefacts could be shown. The only exception is a slight change for post-echo for the guitar source signal.

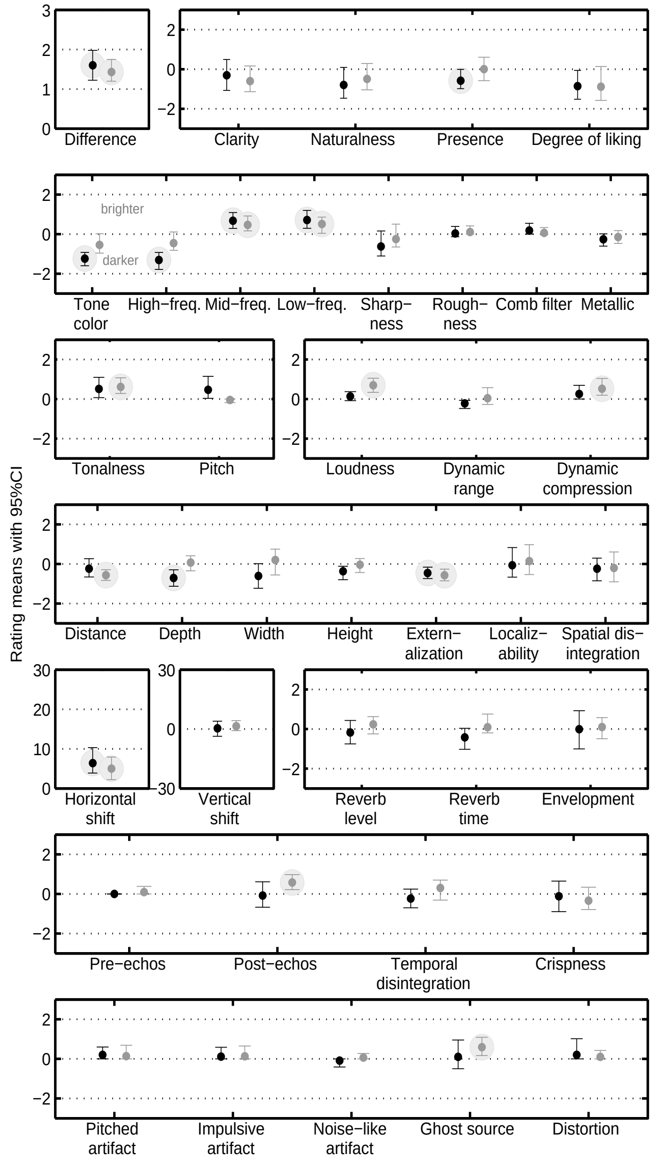

Figure 6: TGC Room: Perceptual attributes (mean values and 95 % confidence intervals) for the drums (black) and the guitar (grey) source signal.

The largest differences of all rooms were observed for the TGC room. A changed perception of timbre and tone color was dominant for this room. Other attributes like distance or depth contributed to the differences as well. Finally, the externalization of the sound sources was affected. In order to reduce the spectral discrepancies for this room, an improved version of the algorithm might be desired. Probably room resonances, which are still present in the reverberation, were not modelled appropriately by the shaped binaural noise.

From the subjects' scores for the four rooms, the following conclusions can be drawn:

First, the synthesis is perceived different concerning the geometry of room and sound source. As spatial information is missing in the omnidirectional RIR, a change of these parameters was expected. The alterations were dominant for the broadcast studios. Second, changes of the spectral attributes contributed to the differences between the rooms as well. These attributes dominated the differences for Control Room 7 and for the TGC Room. For the Control Room 7, single reflections, which were not correctly rebuilt, might have influenced the timbre. The TGC room has a significant amount of room resonances, which might not have been properly recreated in the synthesis. This can lead to a changed spectrum of the reverberation. The spatial characteristics of the reverberation probably are of inferior importance for Control Room 7, which is very dry, and for the TGC Room, which has a very diffuse reverberation tale without strong spatially perceptible components.

Finally, the horizontal shift, which was observed in all scenarios, needs to be evaluated further. Probably, the chosen reflection pattern is not optimal. It needs to be examined if patterns taken from some sample rooms can be generalized or if typical directions of incidence can be estimated based on geometrical assumptions.

Conclusion and Outlook

The presented investigations aim at applying dynamic binaural synthesis based on one measured omnidirectional room impulse response. Due to missing information on spatial aspects, a perfect reconstruction of the sound field is generally not possible. However, it could be shown that generating a full set of BRIRs from a single one-channel room impulse response is feasible as long as a plausible presentation is pursued and as long as perceptual differences to the original room are acceptable. So the presented algorithm can help to acquire full sets of BRIRs in an easy and fast measurement procedure.

The presented method treats direct sound and reflections separated from diffuse reverberation. The early parts of the impulse response are convolved with the HRIR of arbitrary chosen directions of incidence while the reverberation tail is rebuilt from an appropriately shaped binaural noise sequence.

The first psychoacoustic experiment showed that audio signals convolved with the synthesized BRIRs are rated less different to the binaural reference than the ones convolved with the omnidirectional impulse responses. The second psychoacoustic experiment investigated the perceptual attributes contributing to the differences between the synthetized and the reference room.

To some extent, these differences are caused by the assumed room model, which lacks a detailed description of the geometric properties of the original room. A more detailed model, e.g. including directions and times of incidence of the early reflections might help to optimize the synthetization. However, this implies an increased effort when collecting the input data to run such an improved algorithm.

Other perceptual differences might be caused by specific solutions in the implementation of the algorithm. Here, the reconstruction of the reverberation by calculating an appropriate envelope of a binaural noise sequence incorporates a source of error. Room resonances or geometric properties of the real rooms might be reasons that the reverberation tail is not completely diffuse as proposed in the theoretical model.

As next steps some improvements of the algorithm are planned. On the one hand the method for creating the diffuse binaural reverberation shall be optimized. The diffuse reverberation shall be synthesized by convolving the omnidirectional impulse response with varying small frames of binaural noise. Thus the envelope of the omnidirectional impulse response in different frequency bands does not need to be estimated. One the other hand the reflection pattern for the spatialization of the early reflections shall be enhanced. Methods using additional information about the room or applying a statistically improved model shall be investigated here. Finally as suggested in [9] and [10] it shall be investigated to what extent a better matching of the coherence of the binaural noise can improve the algorithm.

Moreover, the approach presented in this paper can be combined with other modifications of the measured room impulse response. In [13], an approach has been discussed which allows a predictive auralization of room modifications by an appropriate adaptation of the BRIRs. Thus the measurement of one omnidirectional RIR is sufficient to obtain a plausible representation of the modified room.

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.

Static versions of the presented stimuli can be accessed via the following webpage:
http://www.audiogroup.web.fh-koeln.de/ICSA2015.html

References

[1] Pörschmann, C., Wiefling, S. (2015). „Dynamische Binauralsynthese auf Basis gemessener einkanaliger Raumimpulsantworten," In: Fortschritte der Akustik – DAGA 2015, DEGA e.V., D – Berlin, pp. 1595 – 1598.

[2] Lindau, A., Erbes, V., Lepa, S., Maempel, H.-J.; Brinkmann, F.; Weinzierl, S. (2014). „A Spatial Audio Quality Inventory (SAQI)," Acta Acustica united with Acustica vol. 100, pp. 984-994.

[3] Geier, M., Ahrens, J., Spors, S. (2008). „The SoundScape Renderer: A Unified Spatial Audio Reproduction Framework for Arbitrary Rendering Methods," 124th Convention of the Audio Engineering Society, Amsterdam.

[4] Stade, P., Bernschütz, B., Rühl, M. (2012). „A spatial audio impulse response compilation captured at the WDR broadcast studios," In: Proceedings of the VDT International Convention, Cologne, Germany.

[5] Pörschmann, C., Bednarzyk, M. (1996). „Erzeugung eines synthetischen binauralen diffusen Nachhalls," in: Fortschritte der Akustik – DAGA '96, DEGA e.V., D – Oldenburg, pp. 400-401.

[6] Pörschmann, C., Zebisch, A. (2012). „Psychoakustische Untersuchungen zu synthetischem diffusen Nachhall," In: Proceedings of the VDT International Convention, Cologne, Germany.

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

[8] Stade, P. (2015). "Perzeptive Untersuchung zur Mixing Time und deren Auswirkung auf die Auralisation," In: Fortschritte der Akustik – DAGA 2015, DEGA e.V., D – Berlin, pp. 1103 – 1106.

[9] Menzer, F. (2010). „Binaural audio signal processing using interaural coherence matching," Thèse No. 4643, EPFL.

[10] Menzer, F., Faller, C., Lissek, H. (2011). „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 (2), pp. 396-405.

[11] Bernschütz, B. (2013). „A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU100," In: Fortschritte der Akustik – AIA-DAGA 2013, DEGA e.V., D – Berlin, pp.592-595.

[12] Lubberhuizen, W. (2007) „Near perfect reconstruction polyphase filterbank," Matlab Central. [Online]. Available: http://www.mathworks.com/matlabcentral/ fileexchange/15813, accessed 15/6/2015.

[13] Pörschmann, C. , Schmitter, S., Jaritz, A. (2013). „Predictive Auralization of Room Modifications by the Adaptation of Measured Room Impulse Responses," In: Fortschritte der Akustik – AIA-DAGA 2013, DEGA e.V., D – Berlin, pp.1653-1656.

[14] Vazquez Giner, A. (2015). „A Solution for Conducting Psychoacoustic Experiments with Real-time Dynamic Binaural Synthesis," In: Fortschritte der Akustik – DAGA 2015, DEGA e.V., D – Berlin

[15] Lindau, A. (2014). "SAQI - Test manual," online http://dx.doi.org/10.14279/depositonce-1.

[16] Bortz, J. (2005). *„Statistik für Human- und Sozialwissenschaftler,"* 6. Auflage. Heidelberg: Springer Medizin.

[17] Barron, M. (1971). "The subjective effects of first reflections in concert halls: The need for lateral reflections," Journal of Sound and Vibration , vol. 15, pp. 475–494.

[18] Olive, S. E., Toole, F. E. (1989), "The detection of reflections in typical rooms," Journal of the Audio Engineering Society, vol. 37, pp. 539–553.