Measurement and Visualization of Room Impulse Responses with Spherical Microphone Arrays

(Messung und Visualisierung von Raumimpulsantworten mit kugelförmigen Mikrofonarrays)

Michael Kerscher[1], Benjamin Vonrhein[1], Gunnar Heilmann[1], Sébastien Barré[2], Peter Weigel[3]

[1] *gfai tech GmbH, Berlin, Germany, Email: kerscher@gfaitech.de, vonrhein@gfaitech.de, heilmann@gfaitech.de*

[2] *GFaI e.V., Berlin, Germany, Email: barre@gfai.de*

[3] *ADAM Audio GmbH, Berlin, Germany, Email: peter.weigel@adam-audio.com*

Abstract

Room impulse responses (RIR) contain a vast amount of information about the acoustic properties of a room in a very compact form. Conventional RIR acquisition methods mostly use a single measurement microphone placed to various positions in the room. The use of a spherical microphone array with a high channel number allows for a way more comprehensive analysis as it includes the direction of the incoming soundwaves. Thus a realistic description of the real 3D sound field is possible. The application of a so called delay-and-sum beamforming algorithm permits a precise localization and depiction of the direct sound and the reflections on a 3D model of the room. This paper presents the basic principle and some measurement examples of omnidirectional room impulse responses.

1. Introduction

Multi-array technologies are well-known in the audio industry nowadays. 3D-audio playback systems become more and more popular and give engineers, artists and Tonmeisters the chance to realize a virtual acoustic scenario. On the recording side multi-channel microphone arrays prevail more and more. They have the big advantage of adding a spatial information to the recording by evaluating the run-time delays between the microphones on the grid. With this information sound can be visualized. This technique is already well-established in many industry branches but can also immensely contribute to room acoustics by the visual analysis of room impulse responses and therefore understanding the sound propagation in a room.

Multi-channel microphone arrays can be constructed in many different ways. For room acoustic measurements 3D-arrays are most suitable as they can record the sound field in a three-dimensional way. To minimize the interaction of the array with the sound field the microphones are placed on an acoustically transparent sphere and to ensure a constant sampling of the sound field a uniform microphone distribution is used.

Spherical microphone arrays can replace single channel measurement microphones for room acoustic measurements. Applying deconvolution techniques, which are based on spectral division, directional room impulse responses for all microphone channels can be calculated.

Common methods in array technology are beamforming algorithms. One of the oldest and best-known is the so-called delay-and-sum beamformer which evaluates the run-time delays of the microphone pressure signals. By feeding this beamformer with the directional room impulse responses the locations of the reflections can be determined and be depicted on a 3D-model of the room: room impulse responses become visible!

2. Room Impulse Responses and how to measure them

Room impulse responses are a well-known and comprehensive format to evaluate a room´s acoustic properties. The method is derived from system theory, where impulses are used to identify systems. In our case the system under test is the room, which is excited with an impulsive signal and the way the room “responses” to this excitation is recorded with a microphone.

This response contains all relevant information about the room acoustic properties in a very compact way, e.g. the reverberation time, early decay time, initial time delay gap and pre-delay. Also further room acoustic parameters can be derived from the RIR like definition, clarity, centre time etc.

But there is way more information that can be extracted from the RIR if acquired with directional information: the temporal and spatial development of sound in the room. Therefore a measurement method is needed which realizes a visualisation of sound in general and the locations of the reflections in our case. Such a measurement procedure will be introduced in this paper.

2.1. Structure of Room Sound

But first let´s take a look at the general structure of RIRs. They can be divided into various temporal stages or phases: direct sound, first and early reflections and diffuse sound.

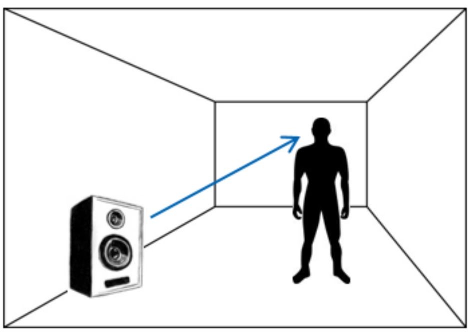

Fig. 1: direct sound path

When sound propagates in a room then of course the first relevant propagation path is the direct sound from the source to the receiver like in figure 1.

The next relevant paths are the first reflections. They are caused by single sound reflections on the floor, ceiling and side walls, the sound does not take any further detours.

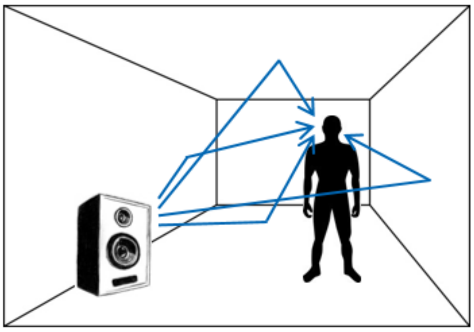

Fig. 2: first reflections

Then multiple reflections start and the reflection pattern gets more and more complicated. Finally single reflections cannot be separated and the sound field becomes diffuse.

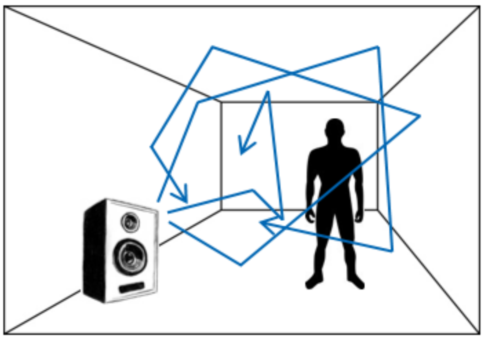

Fig. 3: diffuse sound

2.2. RIR Measurement with Impulse Signals

But how can we get an RIR? The easiest way is to directly excite the room with an impulsive signal, to which the room directly "responses". Common methods are bursting balloons, pistol or gun shots and claps.

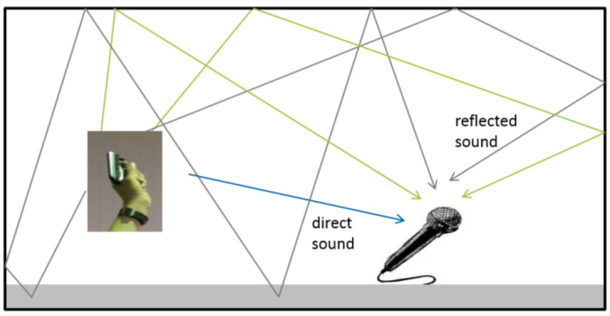

Fig. 4: direct RIR measurement with an impulsive signal

As shown in figure 4 the obtained microphone signal already is the required RIR. This is also the biggest advantage of the direct method, no additional calculations are needed.

An ideal impulse is a theoretically infinitely short time domain signal with a broad and flat frequency response. Any deviation from this ideal leads to a not perfectly flat frequency response, which is important for exciting the room with all frequencies in the audible range. Shots and bursting balloons are very close to this, but cannot be seen as perfect impulse. Thus their frequency response is not perfectly flat and differs from shot to shot. This lack of reproducibility is another big drawback of the direct method.

2.3. RIR Measurement with deterministic Signals

That´s why indirect methods with deterministic measurement signals are mostly more preferable. They use other broadband signals for the excitement of the room, e.g.

- sine sweeps (linear or logarithmic)
- noise
- other measurement signals like MLS etc.

These signals have a flat frequency response and are deterministic. Therefore they can be repeated and reproduced very precisely. To guarantee a uniform excitement of the room omnidirectional loudspeakers are used. Such a situation is depicted in figure 5.

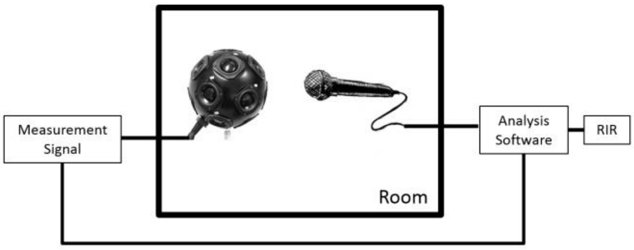

Fig. 5: indirect RIR measurement with deterministic measurement signals

This method does not lead to the RIR directly. Further calculations have to be performed with the microphone signals, which is described in chapter 4.

Since the sound takes different propagation paths through the room depending on the source and receiver position each combination of excitation and microphone position has its own reflection pattern. This must be taken into account when performing the measurements. Hence, it is not sufficient to just use a single source and microphone position, as it represents only one situation. In concert halls acousticians mostly place the sound source on the stage and the microphone at various seats in the audience.

Depending on the main focus of the measurement an averaging over various measurement positions can be made or a single position can be assessed individually.

3. Delay-and-Sum Beamforming

The main focus of this work is the visualization of room impulse responses. Beamforming methods can be used to produce and reproduce directional sound. One of the oldest and best-known algorithm is the so-called delay-and-sum beamforming and it can be used for our visualization task. It is based on the evaluation of run-time delays from a sound source to various receivers, e.g. microphones. A similar concept is partially used for sound localization with the human hearing. The run-time delay between the two ears is interpreted by the human brain to locate sound sources in the environment and tells us, where a sound is actually coming from.

Multi-channel microphone arrays use a much higher channel-count. The microphone data is used to calculate the sound pressure level for a certain position using the following formula:

$$p(\mathbf{x},t) = \frac{1}{M}\sum_{i=1}^{M} p_i(t-\Delta_i). \quad (1)$$

In (1) p denotes the corresponding sound pressure level for a given time t at a position x on a reference plane. M is the number of microphones on the array, p_i the sound pressure of the i-th microphone on the gird and Δ_i its related delay.

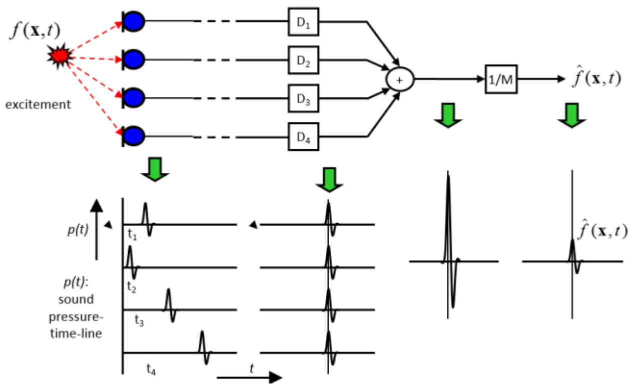

Fig. 6: The principle of time-domain beamforming

Figure 6 shows the situation more in detail. A sound event reaches a number of microphones (in this case four) placed at different positions x at different times t_i, because of the distances between the sound event and the individual microphone positions. The sound reaches the second microphone first, then the first, then number three and finally number four. Using the speed of sound the run-time delays from the source to each microphone can be calculated. These are temporarily equalized by a set of delays D_i. In the next step the delayed microphone signals are summed up, which results in a level that is actually too high. Therefore this level must be divided by the numbers of microphones, in this case four. The end result is then the sound pressure level of the sound event f at the position x and for the time t.

This process is now applied to a number of picture points in a photo plane. There the device under test is located. In our case this is the room and the algorithm works on the points of a point cloud 3D-model or the edges of a mesh 3D-model of the room. For each point a corresponding set of distances r_i and run-time delays t_i can be calculated. This is shown in figure 7.

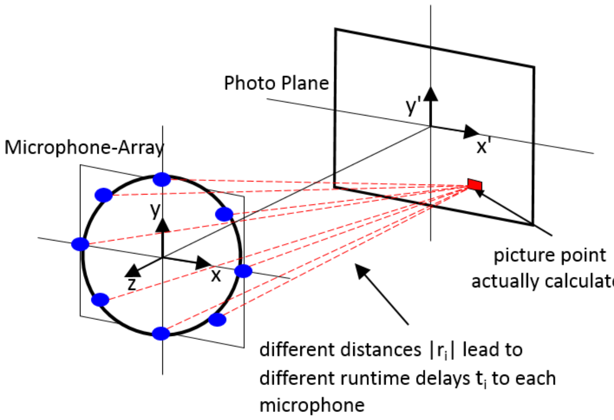

Fig. 7: Photo plane and microphone plane

Hence, the delay-and-sum beamformer allows for the determination of the sound pressure level for all the points of the 3D-model. The points are dyed according to a colour scale, which defines certain sound pressure levels. This so-called acoustic map is overlaid with the 3D-model of the room as in figure 8. This end product is called acoustic image. Here sound pressure levels are represented by colours, similar to heat cameras, where colours represent temperatures: sound becomes visible!

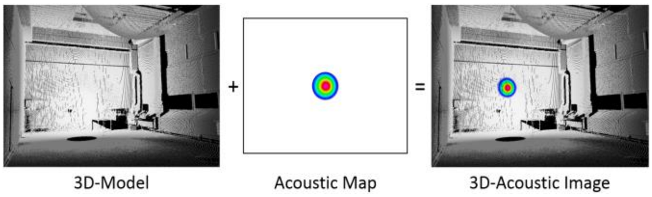

Fig. 8: Overlay of acoustic map and 3D-room model

For our measurements the spherical microphone array in figure 9 with 120 measurement microphones and a diameter of 60 cm was used. This array can localize sound sources three dimensionally and is therefore capable of evaluating the entire sound field in the room.

Fig. 9: The spherical microphone array

4. Directional Room Impulse Responses

The analysis principle which is used for sound visualization is based on the delay-and-sum beamformer and is now well described. But what can be used as input signal for the beamformer? First we replace the single measurement microphone in figures 4 and 5 by our spherical microphone array. Due to the 3D-spatial arrangement of the microphones on the sphere this allows for the recording of directional sound and therefore directional RIRs. There are again two possibilities of obtaining these RIRs:

- Direct measurements with impulsive sound sources
- Indirect measurements with deterministic signals

This was already described in Chapter 2. The same procedures apply for the measurements with the array.

The direct methods lead to the RIRs without any further calculation. The recorded directional RIRs are directly fed into the beamformer which can display them using the algorithm described in the previous chapter. Measurement examples will be shown in the next chapter.

The indirect method needs further calculation to get the directional RIRs. Barré, Jaeckel and Bauer-Diefenbach describe a linear deconvolution method in [2] which is depicted in figure 10:

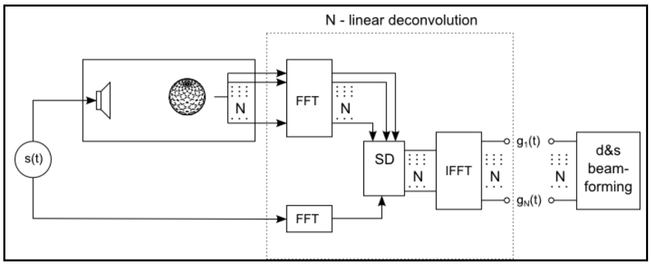

Fig. 10: Linear deconvolution algorithm [2]

The measurement signal *s(t)* is reproduced in the room with an omnidirectional loudspeaker. A reference microphone placed close to the speaker records a reference signal, which also includes its frequency response. The N microphone channels of the array acquire the spatial sound and are then transformed into the frequency domain using an FFT algorithm. The same is done for the reference signal. All N array signals are now spectrally divided by the reference signal. After an inverse Fourier Transform we obtain the N directional RIRs $g_1(t)$ to $g_N(t)$. They can be used as input signal for the delay-and-sum beamformer which realizes the visualisation task.

5. Measurement Examples

5.1. RIR Measurement with Impulse Signals

The following measurement was performed in a quite large room at ADAM Audio speaker factory (figure 11). This room should be optimized for loudspeaker measurements. To evaluate the current acoustic properties of the room measurements with bursting balloons and different kinds of gun shots were performed.

Fig. 11: RIR measurement at ADAM Audio

The array was placed in the centre of the room and for the excitation 3 different positions were chosen. In figure 12 the first 26 ms of all 120 recorded RIRs are depicted.

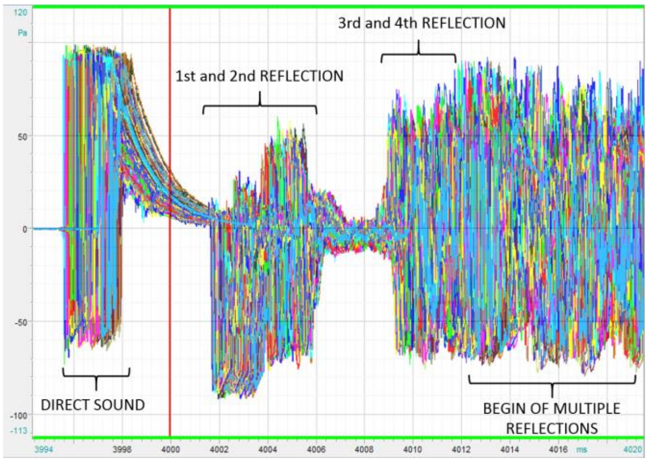

Fig. 12: The first 26 ms of the 120 directional RIRs

For the following analysis the different steps of the RIRs were marked and the beamforming algorithm was applied to them. To obtain the 3D-room model the room was first scanned with a 3D-laser scanner.

Fig. 13: Laser scan of the room

Figure 14 shows the direct sound. The actual shot was made about 2 meters in front of the wall, but as the shooter is not part of the 3D-model the direct sound source is mapped on the nearest points of the 3D-model, which is the back wall. This is another disadvantage of the direct method.

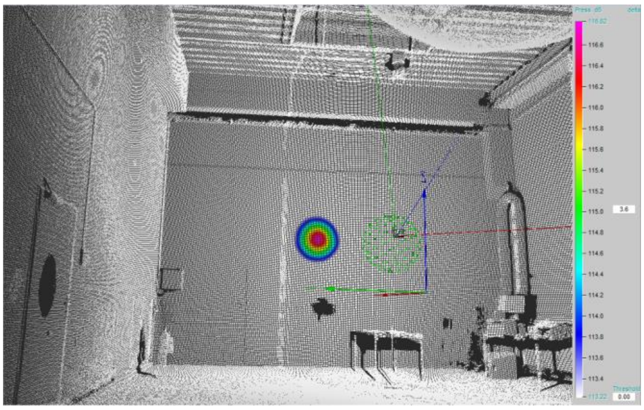

Fig. 14: Direct sound source at measurement position 1 projected on the back wall

Figure 15 shows the first reflection on the floor:

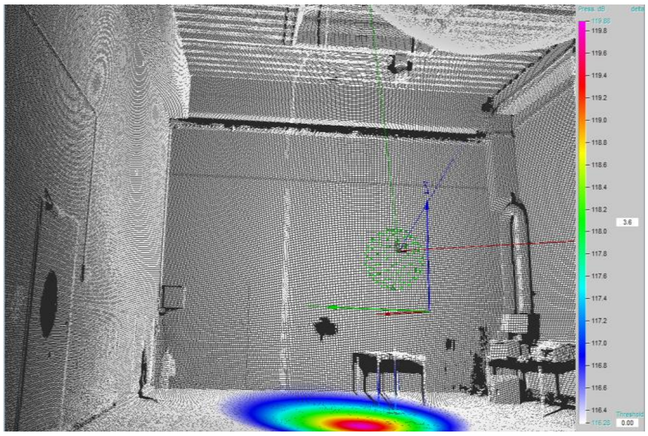

Fig. 15: First reflection

In figure 16 the second reflection on the ceiling can be seen:

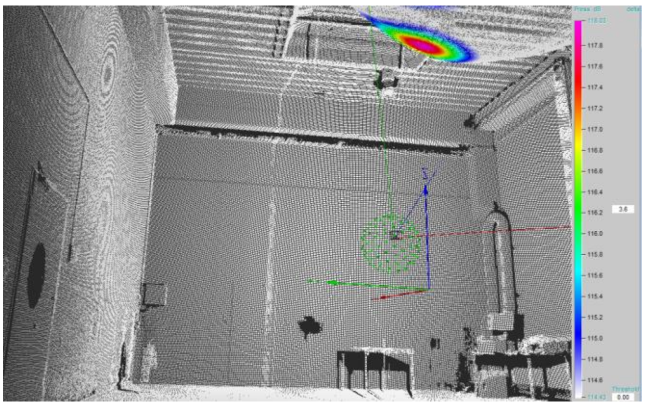

Fig. 16: Second reflection

The third reflection still is located at a discrete position on the side wall. After the fourth reflection multiple reflections begin. This is displayed in figure 17.

The last and longest stage of the RIR is the diffuse field. This can also be visualized with the beamforming algorithm. The sound power is more equally distributed and reflections are less discrete. Such a situation is shown in figure 18.

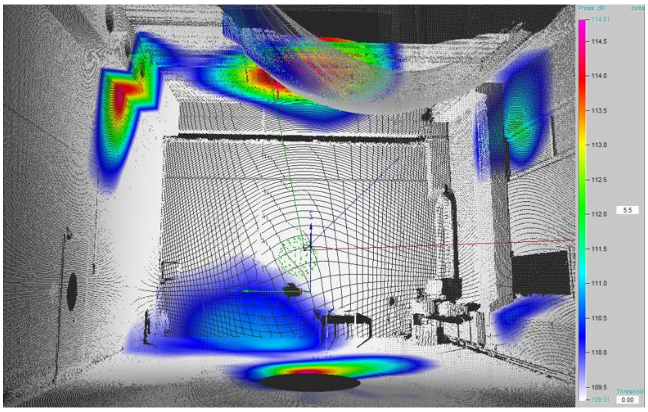

Fig. 17: Begin of multiple reflections

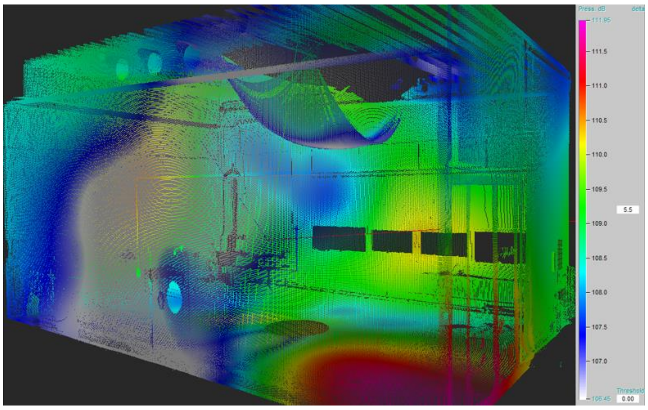

Fig. 18: The diffuse sound field

By looking at the RIR step by step the way the sound travels through the room can be visualized and retraced.

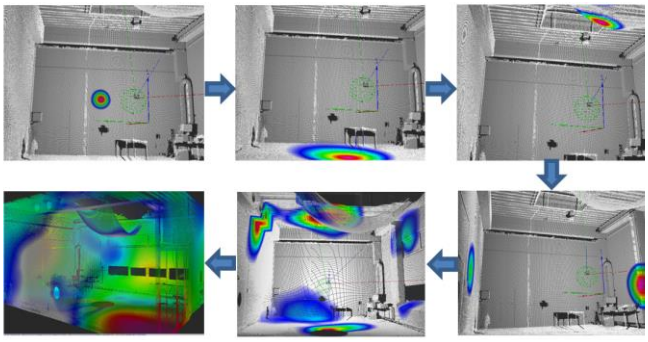

Fig. 19: Retracing the sound in the room

5.2. RIR Measurement with deterministic Signal

The next example explains a measurement that was done in a meeting room at the GFaI e.V. in Berlin. Here a constant power sound source was used as deterministic signal. A single channel measurement microphone was placed close to it to acquire the reference channel. Again here the 120 channel spherical array was used. Figure 20 shows the setup. The objective of this experiment was to determine the influence of absorbing material on room acoustics and its visualization. Therefore Basotect material was put in front of the windows of the room.

Fig. 20: Measurements setup in meeting room with constant power source and spherical microphone array

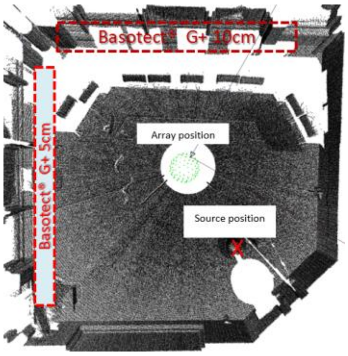

Fig. 21: Positions of array, source and absorbers

Speech intelligibility is of utmost importance in meeting rooms. Poor values lead to fast tiring of the meeting participants. Acoustic parameters like the reverberation time and definition can be derived from the RIRs and evaluate this in an effective way. As table 1 summarizes in this experiment the absorbers could improve all these values in an effective way. Measurement 1 was performed in the plain room without any absorbers. For number 2 the upper windows in figure 21 were covered with 10 cm material and for number 3 both sides were covered with absorbers. RT20 denominates the sound energy decay in the room by 20 dB (similar to RT60), D50 the definition and C80 the clarity.

Measurement	RT20 [sec]	D50 [%]	C80 [dB]
1	0,50	81,5	10,3
2	0,34	88,5	14,8
3	0,29	91,3	16,4

Tab. 1: Room acoustic values without and with treatment

As demonstrated in chapter 5.1 the beamforming algorithm can be used to visualize the phases of the RIR. Looking at the diffuse sound field here reveals the effect of the absorbers. In figure 22 the influence can be clearly seen. The above picture visualizes the diffuse sound field without absorber, the below picture with absorber. In figure 22 the dynamic is the same in both picture, this means, the colours represent the same sound pressure levels. Thus the influence of the absorber becomes obvious. By the colour scale readings on the right a sound attenuation of about 5 dB caused by the absorbers can be detected.

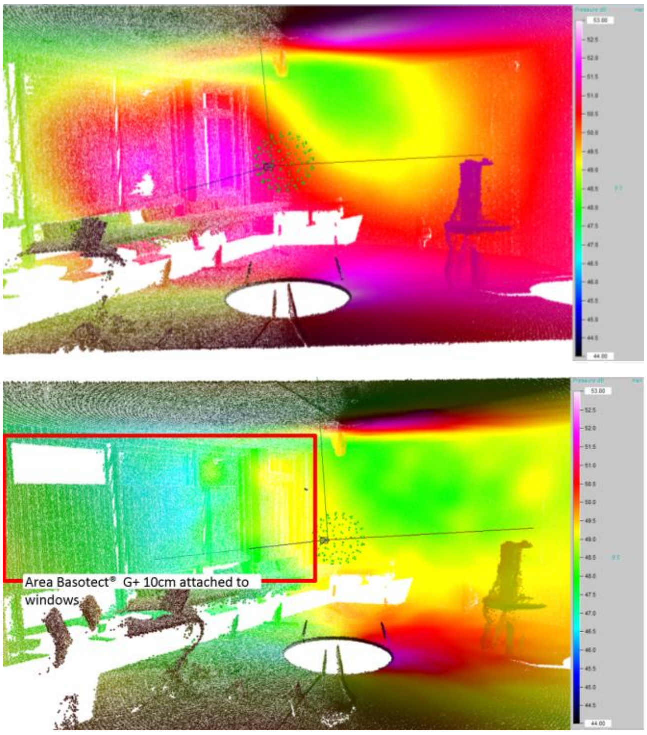

Fig. 22: Diffuse sound field without (above) and with (below) absorbing material

6. Conclusion

Spherical microphone arrays can be used to record directional RIRs. Their evaluation provides a vast amount of information about the real 3D sound field. With the energy decay curve, which is derived from the RIR, parameters like reverberation time, early decay time, definition and clarity etc. can be calculated. But with the application of imaging procedures based on delay-and-sum beamforming a comprehensive understanding of the way the sound takes to travel through a room can be achieved. The visualization helps to understand the complex context in a much easier way.

The performance of this technique was shown in two experiments: locations of reflections can be clearly detected. This can be used for the evaluation of rooms and form the basis for room acoustic improvements e.g. the equal distribution of first reflections over the audience by placing reflectors at the correct positions, the detection of flutter echoes or the efficient mounting of absorbers.

7. References

[1] Gunnar Heilmann, Magdalena Böck, Dirk Döbler: Exploring the limitations and expectations of sound source localization and visualization techniques; InterNoise 2014

[2] Sebastien Barré, Olaf Jaeckel, Ralf Bauer-Diefenbach: Analysis of sound field variations in concert halls via visualization and objective parameter comparison; Berlin Beamforming Conference BeBeC 2016