Active Acoustics, Speech Enhancement and Noise Masking in Multipurpose Venues

F. Kaiser[1], V. Werner[2], Th. Rohde[1]

[1] *Rohde-BeSB Noise + Vibration GmbH, Salzburg, Austria, Email: kaiser@rohde-besb.at*

[2] *PDV Records OG, Vienna, Austria, Email: volker@pdv-records.com*

Abstract

Designing multipurpose rooms living up to their name is still cost-intensive and technically demanding. The requirements for various events may differ greatly. Mechanical or passive measures represent a significant limitation of the architecture, require high investment and deploy personnel and maintenance activities. However, with state-of-the-art signal processing variable acoustics and enhancement of speech intelligibility can also be provided "actively", even without requiring the performers or presenters to be amplified in the traditional manner. This paper presents the challenges still existing for active acoustic systems and examines modern ways of solving these issues. Requirements are carefully defined and special attention is payed to concepts calculating and generating time-accurate and direction-accurate early reflection patterns in 3D space. New approaches of utilizing these reflection patterns to enhance speech intelligibility and to enhance musical envelopment in a way that reflects the actual dimensions and shape of the venue are examined. Further, new approaches to combine noise masking with speech enhancement algorithms to provide seamless transitions between workshop, lecture, conference and debate settings in the same space are discussed. Two examples of installations of Active Acoustics systems are presented.

1. Introduction

When an acoustician or sound designer is asked to address the challenge of designing the acoustics and sound system for a multipurpose venue, the demands can vary quite a bit. On the one hand these demands are defined by the types of events the costumer wants to host, on the other hand by the architecture and size of the venue itself. So, what is a multipurpose venue and what type of events need to be covered?

Two recent examples from the year 2016 are shown in Fig. 1.

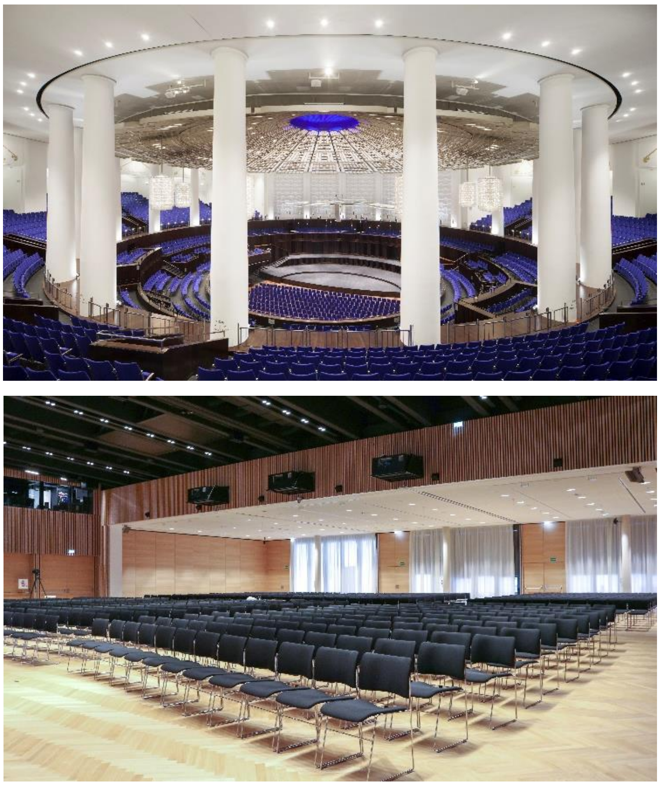

Fig. 1: Top: Hannover Congress Centrum (HCC), Kuppelsaal. Bottom: Congress Centrum Alpbach (CCA), Herz-Kremenak Saal.

Both are multipurpose venues, but even from the picture it is obvious that both are completely different in architecture and therefore also differ greatly in the type of events that can be hosted at least at first glance. But when you get the list of events the users want to host, suddenly they don't look that different any more:

- Presentation
- Conference/ Convention
- Workshop
- Theatre
- Classical (Un-amplified) Concert
- Amplified Concert
- Corporate Events
- Cinema Projection

All of those events vary greatly in their acoustical demands nearly as much as the two venues differ in their architectural characteristics. So, what are the acoustical demands?

For concert hall acoustics, there is a lot of literature to describe the acoustic conditions for a decent performance. The graph in Fig. 2 shows the main criteria following Beranek [1] and Lokki [2].

Most of these factors do not correlate with the characteristics needed for good speech transmission for presentations, conferences etc. Going one step further and looking into the demands for amplified shows or cinema the optimal acoustics would be very low reverberation.

So, the lowest common denominator is a hall with the driest acoustics possible and the addition of certain acoustical qualities for different types of events.

In such a construction, the problem of special architectural ideas like edgy shapes would also be solved since the actual surfaces are more or less absorptive.

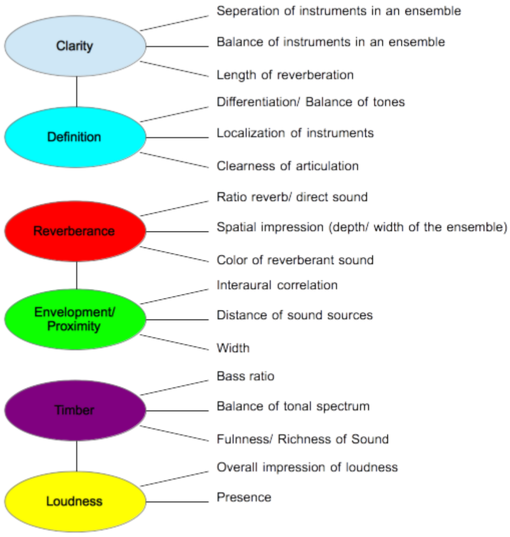

Fig. 2: Subjective criteria of concert hall acoustics (cf. [1], [2]).

To some extent such variability could be realized with structural measures - by adding curtains, reflecting chambers, variable diffusers etc., but this requires high personnel and maintenance demands and is limited by the available space.

On the other hand, there is the "active" approach, that is becoming more and more powerful with the availability of cost-effective high performance computing systems, smaller and better loudspeakers and digital audio networks.

2. Active Acoustics

Active Acoustics is the term comprising several techniques to influence the sound field in rooms in order to achieve goals like speech and acoustic enhancement. This comprises the generation and distribution of early reflections, the generation and shape of late reverberation and the generation and projections of 3D audio scenes.

The idea for such an approach to room acoustic enhancement exists since the 1950's - starting out in the Royal Festival Hall in London – and different variants of such systems are commercially available ever since (cf. [3], [4]).

An Active Acoustics system always consist of microphones, pre-amplifiers, A/D and D/A converters, a signal processing unit, amplifiers and loudspeakers. Fig. 3 shows a schematic drawing of an Active Acoustics system.

The most critical issue is audible feedback. If the transfer function $H_{ML}(\omega)$ exceeds the transfer function $H_{LM}(\omega)$ at some frequency, then feedback becomes audible. The ratio of these transfer functions is called the Loop Gain (cf. [5]):

$$g = \frac{|H_{LM}(\omega)|}{|H_{SM}(\omega)|} \tag{1}$$

Basically, there are two different approaches to active acoustics: *In-Line* and *regenerative* (or *non-in-line*). The difference lies mainly in the way feedback is handled – or, more technically spoken, in the way the transfer function $H_{LM}(w)$ is handled.

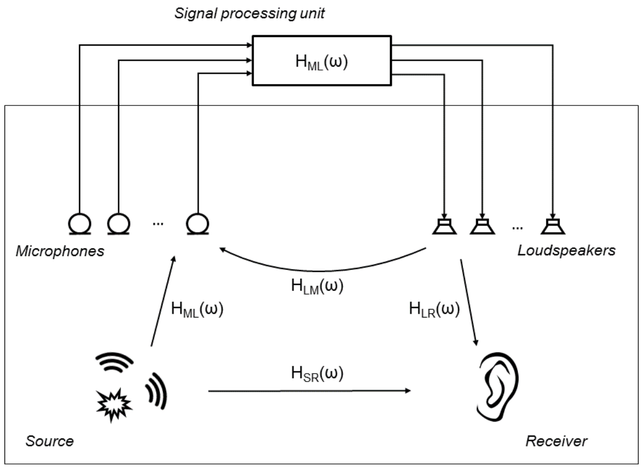

Fig. 3: Schematic drawing of an active acoustic system (cf. [6]).

The In-Line approach uses directive microphones positioned rather closely to the sound sources, usually within their critical distance. The picked-up signals are reverberated either by algorithms or by convolution of synthesized or measured impulse responses. Typically, 4-8 microphones are placed in the stage area and many loudspeakers are placed in the audience area. Feedback is avoided by the directive characteristic and the spatial separation of microphones and loudspeakers. Hence, the loop gain is inherently low and the focus of system design lies on the reverberation generation.

On the other hand, the regenerative approach uses signal loops in between loudspeakers and microphones to generate reverberation. Microphones are placed outside the critical distance of sound sources and loudspeakers. The basic idea is that the picked-up signals are directly given back via the loudspeakers, however, usually a reverberation stage is added to the signal chain. The loop gain is therefore higher and audible artefacts like ringing tones are likely to occur. In order to increase stability a very high number of microphones and loudspeakers has to be used.

A combination of both approaches is typically called a hybrid system (e.g. Amadeus Active Acoustics).

3. Requirements

As stated in the introduction the acoustical requirements for different types of events vary quite a bit. In the following the requirements for different events are discussed.

The basic requirements for presentations, conferences, conventions and partly also theatre performances are a sufficient speech intelligibility (STI $\geq$ 0.6), near constant speech level (S/N $\geq$ 10dB), and a directional reference to the source in the entire audience area. Further, the transmission from sender to receiver must have a flat frequency response in the vocal range and low non-linear distortion. Concerning the room that means a low background noise level from technical equipment and adjacent rooms and no strong echoes or flutter echoes should be prevalent and generally a low reverberation time is desired.

In conference centers workshop situations often are used for group work. In these cases it is desired to increase privacy of groups by decreasing speech intelligibility of other groups. Some sort of masking of noise is necessary to mask one group from another while maintaining good speech intelligibility inside the workshop groups.

For acoustical concerts a certain loudness and reverberation time suited for the musical material is required. A high binaural quality index, good bass ratio (bass strengths), low initial time delay gap, high lateral fraction, high diffusion, low background noise level and no echo effects are other examples from the list of requirements.

Amplified Concerts/Cinema Screenings in contrast require a low reverberation time. There should be no echo effects. The compliance to cinema standards is desirable, therefore virtualization of different speaker setups may be necessary as well as 3D Audio for amplified shows.

All these diverse acoustical requirements lead to distinguished technical requirements.

To achieve a constant speech level and good localization a high-density, high-quality, full-range speaker grid with axis-symmetric directionality of the speaker systems is required to reproduce the necessary reflection patterns. A high-channel count improves gain before feedback, sharpens localization and prevents speaker localization.

Further, a high-density microphone grid with high-quality omni or directional microphones in the entire venue is required for pick-up of the sound sources. This way, the reflection patterns can be created for increased loudness (reduction of the natural drop in sound level per doubling of distance), to provide accurate localization of the sound sources and high gain before feedback.

A high channel count low latency audio processor is needed as well as a high channel count low latency audio network for interconnection of the whole system.

For workshops noise masking with 3D Soundscapes should be implemented to decrease disturbance from one workshop group to the next one.

For acoustic concerts, instead of STI optimized reflection patterns, an adaption of the reflection patterns to customize the virtual venue size and sound characteristics to the musical content is required (always with the background of the actual venue size and shape to avoid "dual slope" acoustics).

A subwoofer system needs to be installed to provide a proper bass ratio (bass strength). The microphone setup in the stage/ performance area needs to be optimized to provide an acceptable initial time delay gap ITDG technical latency (processor, audio network, A/D, D/A) and acoustical latency from source to microphone An increase in reflection density achieves a good binaural quality index, and a high lateral fraction. In addition to the reflection patterns a high-channel-count decorrelated reverberation system needs to be implemented to adjust reverberation time, diffusion and improve gain before feedback.

For amplified concerts the foremost requirement is the possibility to turn the Active Acoustics "off". Also, a 3D Audio system with minimal "sweet spot" behavior to provide comparable localization and sound level distribution for the entire audience area ("WFS Light") is desirable. For movie screenings, a virtualization of different loudspeaker setups might be necessary and up-mixing algorithms might be desirable.

Another part of the technical requirements is the human interface. Most venues require for a "caretaker interface" on the one hand (to turn system on/off, check status, select presets and playback preproduced content) and a "Tonmeister interface" on the other hand (to adjust presets and produce 3D Audio Content).

4. Towards a modern approach to Active Acoustics

4.1. Reverberation control

The main challenge of every Active Acoustic system is the control of reverberation inside a room. As mentioned earlier the environment the system is installed in, i.e. the room, always exhibits some kind of reverberation. The first step of reverberation control therefor is always a profound analysis of the prevalent situation. In case of new constructions this also means careful planning of room acoustic measures.

The reverberant field inside rooms is usually separated into two parts of time: early and late reverberation. The early reverberation consists of rather discrete reflections from floor, walls and ceiling reaching a listener in concert halls usually inside the first 120ms after the direct sound. The loudness, intimacy and apparent source width is strongly influenced by the early reverberation. The late reverberation seamlessly connects to the early field and exhibits diffuse characteristics, i.e., no specific direction of reverberation can be perceived. The late reverberation is mainly responsible of envelopment and time smoothing of musical notes. The balance of early to late reverberation determines the clarity and definition in sound. Two measures help to discriminate the two parts, i.e., the mixing time (cf. [7]), determining the transition, and the critical distance, defining the distance from a sound source at which the direct and diffuse energy densities are equal.

4.1.1. Time-accurate early reverberation

As a basic reverberant sound field in rooms for active acoustics is always prevalent and also desired, the control of reverberation needs to especially take care of the formation of early reflections. The time-structure of added reflections should be chosen according to the prevalent architecture. A so called "double sloping", the perception of two different reverberant sound fields (cf. [8]), can be avoided and more natural acoustic enhancement is achieved. Hence, the early reverberation is mainly shaped by the level of added reflections, the density and the spectral characteristics.

Fig. 4 shows an example of early reverberation shaping. On the left side, the natural decrease of sound pressure over distance is rebuilt. On the right side the shape parameter is adapted so that the sound pressure increases over distance. Additionally, by shaping the spatial distribution, a so called spatial shaping, an emphasis of certain directions, e.g. the side

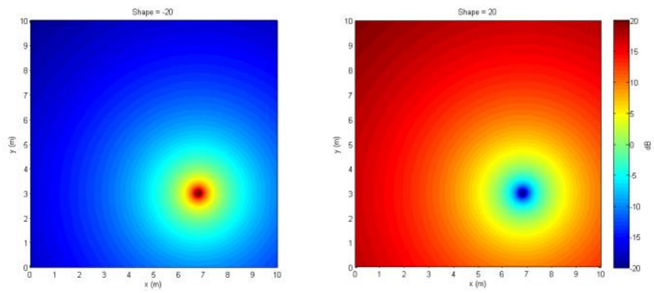

Fig. 4: Distribution of relative sound level from a microphone to loudspeakers across a 10x10m room. Top: Natural decay of level over distance (Shape = ±20). Bottom: Less decay of level over distance (Shape = ±10).

walls are made stronger compared to the ceiling, can be achieved.

4.1.2. Late reverberation

As mentioned earlier there are two main approaches in Active Acoustics: In-line and regenerative. In case of in-line systems artificial reverberation is key to the reverberation control. However, for regenerative approaches it is beneficial to use artificial reverberation to increase the possibilities of control. As described in [9] it is possible to use so-called Feedback Delay Networks (FDN) to recreate a measured impulse response by using convolution of the early part only and using FDNs to recreate the late reverberation. The key is to adapt the parameters of the FDN to match the spectral and temporal characteristics of the measured impulse response (cf. [10]).

This approach can be transferred to active acoustics. By analysing the prevalent reverberation, the parameters of the reverberation unit can be set to spectrally approximate the existing reverberation. The control is done by changing the temporal parameters based on the basic set of parameters. Further, the time-accurate reflections created in the early reverberation stage can be used to feed the late reverberation stage. Therefore, a natural transition between early and late part is achieved.

4.2. 3D Audio and content creation

In order to achieve precise sound localization in 3D space and full immersion of the audience there are a few different approaches to 3D audio (cf. [11]): The so-called stereophonic techniques, i.e., Vector-Base Amplitude Panning (VBAP) and Ambisonics Panning, and the so-called sound-field synthesis techniques, i.e., Wave-Field Synthesis (WFS) and Higher-Order Ambisonics (HOA).

Stereophonic techniques exhibit the benefit of low complexity in understanding and implementation. Usually a few loudspeakers are used to position a virtual sound source and if a listener is located equally distant to the loudspeakers then the localization works well. However, if a listener is located closer to one side of the listening area, virtual sound sources tend to be perceived from the closest loudspeakers.

Sound-field synthesis techniques on the other hand exhibit a significant higher complexity for the involved signal processing. However, localization and motion of sound sources is generally of better quality and the sweet spot issue is reduced (HOA) or avoided (WFS).

When it comes to Active Acoustics in multi-purpose venues a typical application is cinema or playback of video content with 3D Audio. The challenges are, that the listening area can be very large with an audience of 500 people and higher. The goal is to achieve an even distribution of sound level while keeping the sound quality high for every seat. Hence, amplitude panning and also HOA seem not appropriate, as the localization and coloration outside a sweet area is not of the same quality as inside. A WFS approach on the other side has the property of reducing the sweet spot issue. Further, the number of active loudspeakers can vary while keeping localization stable.

Generally, it has to be mentioned that the quality of the loudspeakers in use determines the quality of a 3D Audio reproduction. Hence, the choice of loudspeakers and their position is highly critical.

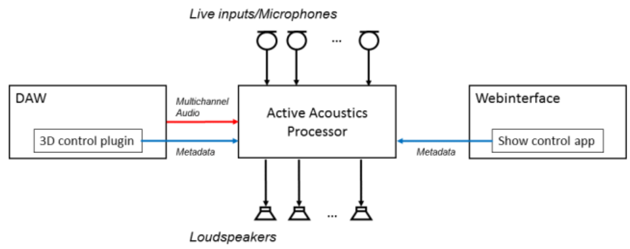

Fig. 5: 3D Audio production chain for recorded and live inputs.

4.3. Signal processing framework

The signal processing blocks in an active acoustic system are the early reverberation stage, the late reverberation stage and the 3D Audio stage. These stages provide the basic signal processing for the different acoustic aspects plus the described shaping of spectrum, level and density. Fig. 6 shows a schematic overview of the proposed framework.

The implementation of these signal processing blocks can be done on DSP chips or using standard CPUs. The performance of modern CPUs is already so high, that the performance of DSPs for signal processing tasks is achieved or even exceeded. Hence, it is possible to have all these units run in parallel while still providing very low latency at an inexpensive price. Also, using standard software development tools and methods the development cycles can be kept quite short compared to the implementation on DSP platforms.

4.4. Audio Network

The main and still rather new challenges for an audio network for Active Acoustics are, as described above, the high channel count (min. 64 bidirectional @ min. 48kHz/24Bits), low latency, ease of planning, installation and configuration and interoperability of products. As Audio over IP solutions become more available the audio world is facing a paradigm shift. Traditional analogue or digital point-to-point solutions are replaced by IP-based (Layer 3) networks. The advantages are high channel counts, routing flexibility, low-cost (devices and cable are essential to the computer industry), connection of office PCs. However, the typical merits are the complexity

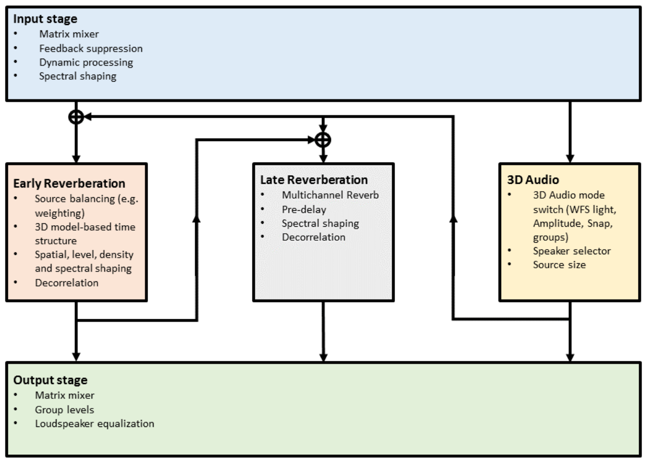

Fig. 6: Signal processing framework for an Active Acoustic system.

of planning, installing and configuring a network and latency is not always the lowest possible.

As Active Acoustics installations, usually are standalone systems, similar to lighting or air-control systems, the need for flexible routing and interfacing with other audio components is not necessary. Point-to-point networks are appropriate for most installations.

However, the amount of copper cabling for an installation of up to 100 single microphones or loudspeakers can be excessive. An approach reducing this is to use sub-stations. The sub-stations are interconnected using only a few copper or optical cables, which can transmit a high number of channels (typical 64). The electro acoustic components are connecting to the sub-stations using far less copper cabling. The installation of such a system is much simpler using Layer 3 Networks compared to legacy digital solutions like MADI or AES-3.

An Audio Network based on MADI seems appropriate at first. However, as the characteristics of audio networks yield advantages for many applications the use of such increased over the last year's. Especially Layer 3 Networks become very popular, as the involved hardware is very cheap, well-known and offers better interoperability compared to layer 1 or 2 networks. With the outcome of the AES67 standard [12] one can expect the different systems like Dante or Ravenna to become interoperable.

5. Case studies

5.1. New construction – CCA Alpbach

The Congress Centrum Alpbach (CCA) is located in the middle of the Tyrolean mountains in Austria. Every year it hosts the well-known European Forum Alpbach, which is a platform to discuss current topics of politics, economy and sciences.

The CCA decided to build a second building to be able to host larger congresses and events. The hearth of the new building is the Herz-Kremenak-Hall, which has a ground floor of 685m² hosting 750 people. In case of the CCA, all requirements mentioned earlier applied and hence an active acoustic system was the right choice.

The system should be able offer a uniform performance across the whole hall. There are no preferred spots for the stage, performers and musicians can be positioned at any point in the hall. A special feature of this hall is that the rear third of the hall can be separated from the main hall and be split in three small conference rooms. This needed to be taken into account in the process of developing the system.

Fig. 7 shows the layout of loudspeakers and microphones. In the hall (the upper part) a hexagonal grid was used at the ceiling (17 microphones, DPA, 18 loudspeakers + 4 subwoofer, Renkus-Heinz). The microphones had to be hung from the ceiling ranging a height from 4-6m above floor level and following the line of projection. On the walls there are 22 loudspeakers (Renkus-Heinz) installed in a height of 2m.

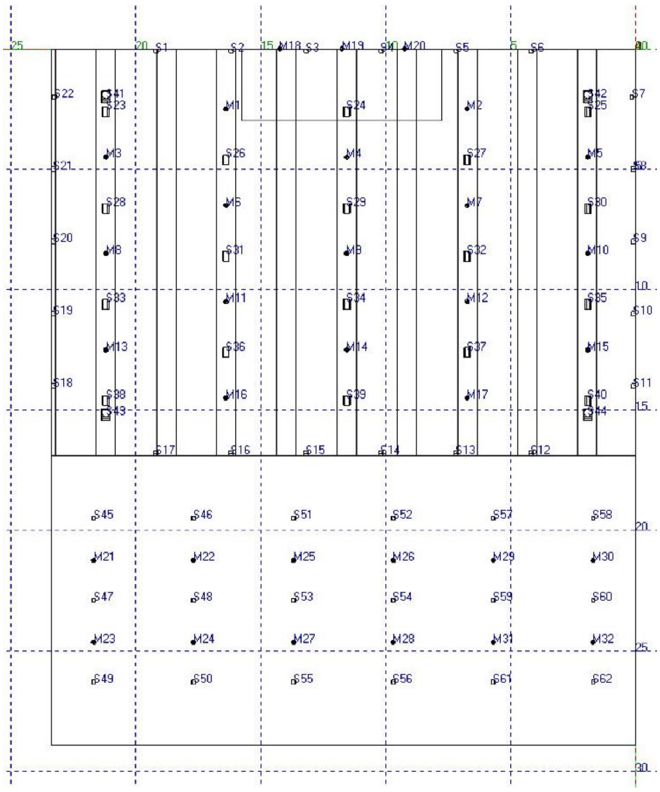

Fig. 7: Layout of loudspeakers and microphones in the CCA Alpbach.

The system was tuned to five different acoustic presets dedicated to different applications like solo concert, chamber music, orchestra or choir music. A special challenge was that every acoustic preset had to be tuned for two cases, the open and the separated hall.

The system can be remotely controlled via a web interface using an iPad or a regular computer. Using the web interface the technical team is able to turn the system on/off, select presets, get status information about the components, and playback 3D audio-content or select virtual speaker setups for motion picture playback.

Fig. 8 shows the measured reverberation times of the different acoustic presets.

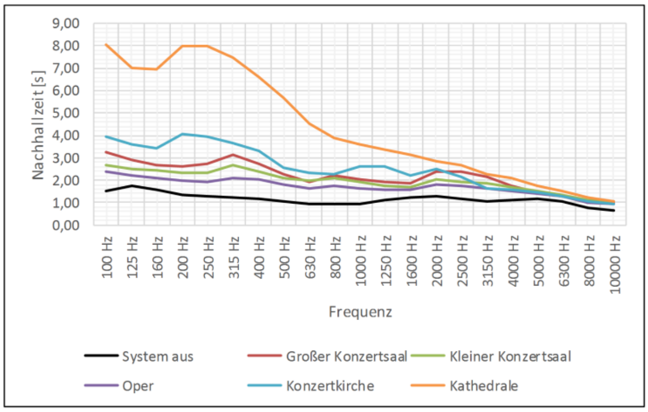

Fig. 8: Measured reverberation time of different acoustic presets of the active acoustic system in the CCA Alpbach.

5.2. Renovation – Kuppelsaal Hannover

The Kuppelsaal of the Hannover Congress Centrum (HCC) is the largest concert hall in Germany hosting up to 3500 people. Every year the opening of the Hannover Messe takes place here.

The HCC decided to profoundly renovate the whole hall and also to adapt the acoustics. However, the requirements differed quite a bit. The problem of the existing hall was a very long reverberation time, while having quite a low loudness, especially on the main floor. The enormous size and unfavorable shape creates nearly no early reflections. The feeling for a musician was like performing in a quite dead space until suddenly a huge reverb tail occurred.

Hence, the system needed to be designed to provide strong early-reflections for a decent loudness level in the hall.

A grid of 24 microphones (Sennheiser) were installed in 3m height above stage covering the whole area. Fig. 9 shows the layout of loudspeakers: 12 loudspeakers (Kling&Freitag) were integrated into the balustrade to cover the main floor, 15 speakers (d&b) were installed on the ceiling to cover the second rank, and two speakers (d&b) were installed in every niche in the back of the hall. Fig. 9 shows the loudspeakers layout.

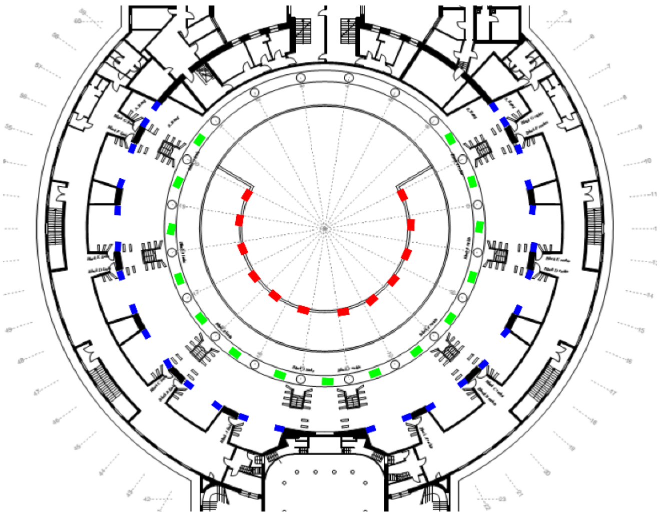

Fig. 9: Layout of loudspeakers in the Kuppelsaal Hannover.

The amplification for conference/presentation situations were implemented with a traditional P.A. system and therefore not integrated in the active acoustics processing.

Different acoustic presets were tuned to provide early reverberation following different setups of ensembles ranging from full orchestra to chamber music ensemble. The system can also be remotely controlled via a web interface using a regular computer.

Acoustic measurements were carried out using the IRIS 3D measurement systems [13]. Fig. 10 shows the reflection pattern in the first 20 dB (normalized to direct sound). On the top left side, the system is turned off, on the ride side the system is turned on. The source is positioned on stage and the microphone in the middle of the audience area at ground floor level. The bottom figure shows the reflection patterns for the second floor. It can be seen that especially the early energy increases substantially.

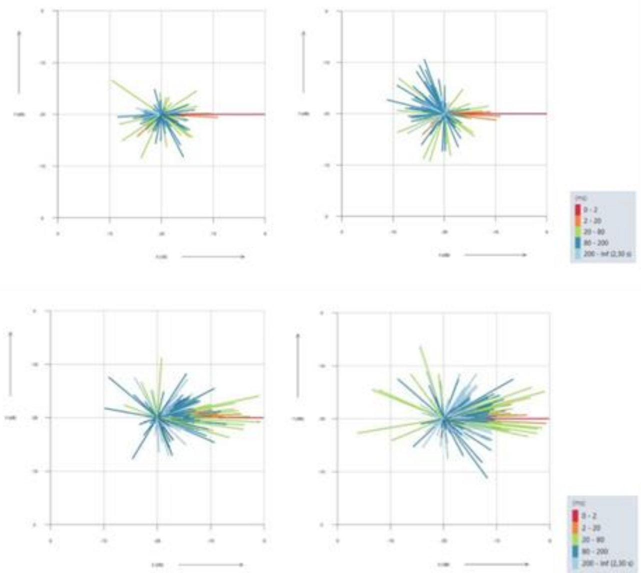

Fig. 10: Reflection patterns in the Kuppelsaal Hannover. Top: main floor, system off (left) and on (right). Bottom: second rank system off (left) and on (right).

6. Conclusion

Active Acoustics is a hot topic nowadays, as multichannel audio technologies become cheaper and easier to integrate. Further owners and managers of multi-purpose halls become more and more open to and aware of such solutions. The main challenges in planning are still the integration of loudspeakers and microphones into the architecture. Also, the price of an installation is still in a range, where persuasive efforts are necessary to show the benefits.

Acoustically, it is key to provide time-accurate early reverberation, in order to ensure a natural sounding reverberation. Also, spectrum-accuracy of late reverberation contributes to a natural sounding system. When it comes to 3D Audio for events it is critical to make sure the whole audience area is provided with near-equal sound quality. Therefore, approaches that do not provide even sound level distribution, independent of virtual source location, and stable localization across the audience area are not suitable for such applications.

The described approaches to active acoustics were realised in two projects, one a renovation (Kuppelsaal Hannover) and one a new construction (CCA Alpbach). In both cases, the acoustic requirements for acoustic enhancement and variable acoustics could be excellently fulfilled.

7. References

[1] Baranek, L.: Concert Halls and Opera Houses. Springer, New York, 2004

[2] Lokki, T., et al.: Disentangling preference ratings of concert hall acoustics using subjective sensory profiles. J. Acoust. Soc. Am. 132 (5), November 2012

[3] Kaiser, F.: Acoustic Enhancement Systems. Bachelor Thesis, TU Graz, 2009

[4] Werner, V.: Der Virtuelle Konzertsaal, Elektroakustische Raumklangerzeugung im Überblick. Diplomarbeit, Universität für Musik und darstellende Kunst Wien, 2008

[5] Schroeder, M.R.: Improvement of acoustic feedback stability in public address systems. Proc. 3rd Int. Congress on Acoustics, 897-901, 1959

[6] Svensson, P., On reverberation enhancement in auditoria, Phd. Thesis, Department of Applied Acoustics, Göteborg, 1994

[7] Lindau, A., et al.: Perceptual evaluation of physical predictors of the mixing time in binaural room impulse responses. AES 128th Convention, London, UK, 2010 May 22-25

[8] Poletti, M.A.: Active Acoustics System for the Control of Room Acoustics. Proceedings of the International Symposium on Room Acoustics, ISRA 2010

[9] Carpentier, Th., et al.: Hybrid Reverberation Processor with Perceptual Control. Proc. Of the 17th Int. Conference on Digital Audio Effects (DAFx-14), Erlangen, Germany, September 1-5, 2014

[10] Jot, J-M., et al.: Digital delay networks for designing artificial reverberators. Presented at the 90th Convention 1991 February 19-22 Paris

[11] Spors, S., et al.: Spatial Sound With Loudspeakers and Its Perception: A Review of the Current State. *Proceedings of the IEEE*, vol. 101, no. 9, pp. 1920-1938, Sept. 2013.

[12] AES67-2015: AES standard for audio applications of networks – High-performance streaming audio-over-IP interoperability

[13] Protheroe, D.: 3D impulse response measurements of spaces using an inexpensive microphone array, ISRA 2013