DOI: 10.22032/dbt.39974
DOI (proceedings): 10.22032/dbt.39936

vdt Verband Deutscher Tonmeister e.V.

Spatial audio production for immersive fulldome projections

Presented * by VDT.

Johannes Ott[1], Anca-Stefania Tutescu[1], Niklas Wienböker[1], Jan Rosenbauer[1], Thomas Görne[1]

[1] *Hamburg University of Applied Sciences (HAW Hamburg)*
Department of Media Technology
Email: johannes.ott@haw-hamburg.de

Abstract

Fulldome is an immersive half-spherical video format utilized mainly in planetariums which is often combined with spatial audio playback. This combination on one hand offers new ways of perceiving sound in space and on the other hand helps enhancing fulldome productions with audiovisual synergy. However, audio production for fulldome video poses some technical and artistic challenges. Limited time slots and resources seldom allow to work on sound productions inside a planetarium directly. Likewise, the various spatial audio technologies provide the user with fairly different approaches to create, position and move sounds in space.

This paper investigates three different approaches to create spatial audio content for fulldome productions in remote studios and to present them in a planetarium: object based proprietary Fraunhofer "SpatialSound Wave" (SSW) system, scene based Higher Order Ambisonics (HOA), and channel based production will be compared. Technical challenges and potentials of storytelling in spatial audio will be discussed.

1. Introduction

The terms "fulldome" or "fullspace" refer to half spherical video projections designed to present the audience with an immersive multi-media experience in venues like planetariums, resembling illusionistic Renaissance ceiling paintings (Gerling, 2013). Today fulldome productions form an increasing part of multi-media shows in planetariums worldwide. In the wake of this new immersive content, planetariums have started to recognize the potential of spatial audio and have installed suitable playback systems in their domes.

Although the workflow for producing spatial audio content for fulldome is very similar to other immersive formats such as fixed media concerts, sound installations or VR, playing it back in a planetarium brings up a variety of challenges:

- The size of the room is very likely much larger than the listening room in which the content was produced. Even though their size is similar to cinemas, no guidelines or standards exist for audio playback in planetariums, which makes it difficult to judge sound characteristics (such as loudness) in the production stage.
- Because of the floor area occupied by the star projector – the essential piece of technology in a planetarium that is always installed in the very center of the dome – seats are usually grouped around the center, leaving open the area where the ideal listening positions would be.
- To maximize seating capacity, seats in planetariums generally go right up to the walls of the room, giving the audience seated there necessarily a different auditory experience (see Fig. 2, p. 5).

Dealing with these challenges, this paper investigates workflows when producing with generically different spatial audio representations (object based, scene based, channel based), discussing their assets and drawbacks.

2. Production formats

2.1. Object based production (SSW)

The proprietary "SpatialSound Wave" (SSW) System developed at Fraunhofer IDMT has been installed in several planetariums in Germany, such as Hamburg, Jena, Berlin and Bochum (Fraunhofer IDMT, [n. d.]). Producing in this format is therefore an obvious choice for fulldome production.

SSW is an object based / hybrid system based on the works of Brandenburg et al. (2013) with its origin in Wave Field Synthesis (WFS). The core of the System is a real-time renderer that spatializes monophonic sound sources for multichannel loudspeaker arrays. The audio content is fed into the renderer from an external playback machine.

Spatialization in SSW is controlled wirelessly via a web interface in real time. To produce SSW content, a workstation for playback of the mono sound sources, the SSW renderer and a loudspeaker array are needed. Playback through headphones is not possible. The movement of sound sources is synchronized with the playback machine via timecode.

SSW productions can be delivered to other SSW systems in the form of 32 mono tracks containing the audio information of the sound sources and the SSW session containing the positional meta-data. If the playback venue has a SSW system installed the studio production translates easily to the playback system. To compensate for the difference in size of studio and venue, a scaling factor can be applied. It is also possible to pre-render the loudspeaker signals for a specific playback system by recording the outputs of the rendering unit.

2.2. Scene based production (HOA)

A convenient way of producing sound for a multi-directional medium like fulldome is to use auditory scenes, in this case achieved by utilizing Higher Order Ambisonics (HOA). This technology, based on the works of Gerzon (1973), has become very popular in the spatial audio scene, not least due to format specific advantages of scene based coding like e.g. the possibility of rotating the soundfield. Lots of solutions exist today for producing Ambisonics content and many of them are open-source. For linear content such as fulldome, working with tools like VST plugins in a DAW differs the least from conventional linear audio production, and the resulting signal takes full advantage of the spatial resolution of the playback venue given that the utilized Ambisonics order is high enough.

Further advantages of producing in HOA are that no specific hardware except for the loudspeaker array is needed, and that the signal can be rendered for virtually any playback situation including stereo or surround, i.e. the Ambisonics format can be regarded as system-agnostic. This is helpful since the number of planetariums equipped with spatial audio playback systems is still small. The content can also easily be adapted for binaural playback and other applications. It is theoretically even possible to produce on headphones, which reduces the technical demands for producing for planetarium to a minimum compared to the other approaches discussed.

Ambisonics and HOA

Higher Order Ambisonics (HOA) should not be confused with first order Ambisonics, created e.g. with the now popular tetrahedral microphones. First order Ambisonics is known for suffering greatly from sweet spot issues, as the spatial auditory scene tends to collapse when the listener is situated at a non-ideal position. We therefore would not recommend it as a format for fulldome productions. In contrast, HOA productions – specifically from 5th order upwards – are rather robust regarding the listening position, as shown by Frank and Zotter (2017). The authors' experiences conform with these findings.

Ideally the playback facility would be equipped with an Ambisonics renderer configured for the playback system. This renderer could easily be implemented with an ordinary computer powerful enough for HOA rendering, running e.g. a DAW software with a dedicated plugin like the AllRADecoder of the IEM Ambisonics plugin suite (Zotter and Frank, 2012). Setting up a HOA renderer like this would be a low-cost yet very valuable addition to any planetarium's technical setup.

In case the HOA signal cannot be decoded in real time, i.e. if no Ambisonics renderer is available in the venue, it is also possible to pre-render for the specific playback system of the venue, creating a multi-channel loudspeaker feed. In this case the loudspeaker signals would be played back directly by short-cutting a potentially installed 3D audio processor like SSW. However, this direct access to the loudspeakers might not be available in every venue.

In case these two "best options" are not realizable, a less optimal workaround is possible, assuming the venue is equipped with SSW: the HOA signal can be decoded to a virtual loudspeaker array created within the SSW system. The obvious way would be recreating the production venue's setup as a virtual loudspeaker dome in SSW. The alternative way would be to render the HOA in a last step of the production process to a virtual setup matching the loudspeakers in the playback venue, as if it would be rendered for direct HOA playback. Unfortunately, either way requires two layers of virtualization which can result in a significant degradation of localization compared to other workflows. Although the latter solution might appear needlessly complicated (as virtual point sources would be positioned at actual loudspeaker positions), this might be worth trying for the reward of a less colored and more stable soundfield.

A factor to be taken into account for scene based audio is the listening position, which can vary greatly depending on the size of the venue and space occupied by projectors or similar. While in theory the sweet spot – the "perfect reconstruction area" – for Ambisonics is very small, listener feedback suggests a satisfying listening experience and stable scene perception in an extended radius outside the sweet spot, specifically when working with Higher Order Ambisonics (Frank and Zotter, 2017).

2.3. Channel based production

The production workflow with a channel based approach differs significantly from those discussed previously. It is assumed that the production will be played back on a known loudspeaker array. Sounds might be distributed over several speakers using e.g. Vector Base Amplitude Panning (VBAP) after Pulkki (1997), or they might be static using dedicated loudspeakers for the different sonic objects (which in fact means utilizing the playback system as a "loudspeaker orchestra" in the tradition of spatialized electro-acoustic music; cf. e.g. Brech and Paland 2015; Voit 2014). Either way such a production benefits greatly from a High Density Loudspeaker Array (HDLA) with 20+ loudspeakers.

Advantages of a HDLA channel based approach are the very precise localization of sounds with clearly localizable auditory objects, virtually no spatial blur, no sweet spot restriction (if no panning is applied), and perfect control of spatialization assuming there is a sufficient number of loudspeakers available.

Of course, in less avant-garde environments, "channel based production" usually refers to material produced in and for a well-defined comparatively sparse array, like Auro-3D 9.1. The latter or similar 3D Audio systems can be classified as "surround with height", with one sole height layer above ear level, and with the height loudspeakers typically situated vertically above the ear-level loudspeakers, which complicates triangular panning like VBAP and leads to unstable phantom sources / auditory objects at the sides, rear and height (for a list of channel based systems see ITU-R BS.2051 2018). The advantage of a 3D 9.1 or similar approach would be at least the availability of production facilities.

However, since there is no standardized loudspeaker array for planetariums, playback for channel based fulldome productions is difficult. Ideally a channel based production would mean to produce directly at the location, i.e. in the planetarium dome. This is very unlikely because of time constraints of a running planetarium, and furthermore this would lead to a production playable just in one particular planetarium. So it comes back to the task of matching a channel based studio production – be it common 3D 9.1 or fancy 33.2 – to a different loudspeaker array.

A rather pragmatical way of dealing with non-matching loudspeaker setups would be choosing the "nearest to perfect" loudspeakers from the array. Here the task would be – similar to playing back decoded HOA – to get direct access to the loudspeakers. And the significance or insignificance of spatial errors and coloration introduced by the utilization of a non-ideal setup are hard to predict.

A practical approach to adapt a channel based spatial production to different playback systems is to render the channels within the respective playback system (like SSW or HOA) as virtual loudspeakers. In theory this would preserve the spatial and sonic properties of the production, but in practice some advantages of channel based production might get lost as spatial blur and coloration of the virtual loudspeaker system are introduced.

3. Storytelling and formal considerations in spatial audio

A common assumption in content production for so called "immersive media" is that they facilitate a deeper experience of immersion and stronger emotional impact compared to conventional audiovisual media (see e.g. Hahn 2018; Uhrig 2015). While it is widely believed that the immersive experience provided by media like fulldome, 360° video or VR might be explained by an enhanced "realism", it should be noted that this technically mediated realism – specifically with fictional content – is basically different from the real-world experience, as the audience is always aware of its fictional nature however strong its emotional impact might be; Carroll coined the term *paradox of fiction* (Carroll, 1990; Voss, 2009).

Following Ijsselsteijn et al. (2000), *presence* – a more specific description of the immersive experience – can be understood as feeling present in a mediated environment and losing awareness of the technological medium involved. Thus an objective of media production for fulldome would be enhancing the audience's immersion and experience of presence in a virtual space – realistic or not – by technical means.

But what creates a believable and effective virtual auditory environment, both in fulldome productions and in other media such as VR and cinema? Lennox et al. (2001) argue that realism is based not solely on the directionality of sound and geometrical representation of space, but rather on the relationship between sound objects and their environment, i.e. their sonic context.

Spatial audio systems, for their advanced options to arrange and direct auditory objects in space, bring added spatiality to storytelling, being a powerful tool for better directing the audience's attention and increasing presence, according to Hendrix and Barfield (1996).

Barrett (2010) states that a technically sophisticated spatial audio system – her example being Higher Order Ambisonics – facilitate realism beyond sheer listener envelopment and might be capable of even *"breaking the paradox between reality and fiction"*.

And Ijsselsteijn et al. (2000) note spatial sound as a conveyor of sensory information that increases presence depending on fidelity and extent of auditory events. These sound properties do not necessarily mean real-world accuracy.

According to Grimshaw (2014), immersion is further improved via referencing other senses (touch, smell, taste) through sound: *"A visually rendered corpse alone may not trigger an olfactory sensation of decay, but the sound of buzzing flies and wriggling maggots around it has a much greater potential to do so. [...] Although a realistic virtual soundscape may (virtually) reflect its reality counterpart, the lack of audio input compensating for other sensory modalities creates an incomplete experience, lacking immersion and, ironically, appearing unrealistic."* (Grimshaw 2014: 368)

Lennox et al. (2001) talk about creating a sound hierarchy: "my space", "adjacent space" and "distant space", with the first containing the closest and most urgency-filled sounds, the second being less relevant in terms of urgency or threat to the listener, but having the potential to move into their close space, and distant space providing the least amount of threat and requiring the least amount of localization to be believable (cf. Hall's well-established concept of social distances, see Hall 1969). Context is hereby a requirement for the ability to employ selective attention to distinct objects and features, and the authors speak of "selective inattention" with respect to background information in the environment, required in order to detect consistency or inconsistency in space. Likewise, Dalton and Fraenkel (2012) investigate "inattentional deafness", describing the effect of not consciously perceiving a sonic event if the attention is captured by a different part of the auditory scene. Nonetheless, one can expect that the "overheard" elements of the auditory scene have an emotional impact as well (Bargh, 1988).

Movement is also an important source of information and can be expressed not only by change in location, but also through spectral changes like high-frequency comb filtering (Lennox et al., 2001). Consequently, movement should be considered as a key element of a compelling spatial audio production (cf. Karadoğan and Görne 2019).

Spatial sound productions in planetariums allow for novel approaches to storytelling. Cinematography for such a medium brings challenges such as the audience's orientation in space or the risk of disorienting and making the viewers sick through jump-cuts and sudden panning because of the fact that the viewers' field of view is enveloped, unlike with traditional screens (Yu et al., 2007).

The role sound plays and its relationship to the visuals are also in question. One of the tasks sound can fulfill is directing the audience's attention via auditory cues. Although this is a common stylistic device in film sound design (Görne, 2017) it is crucial for sound design in fulldome video (and comparable to other immersive media like VR or 360° video) because of the visual envelopment, even though the viewers are rather stationary in this case, encouraged to lean back their chairs and being set up in a way that creates roughly the same preferred spot of viewing (even if from different directions).

An experiment done by Sheikh et al. (2016) employed 360° video and stereo sound to divert attention from the main characters to a third person in the scene, employing a variety of means (gestures from characters, the main characters addressing the viewer directly, another person walking to the target): *"Audio cues have the advantage that no assumption is made about the viewer's focus of attention at the time of the cue. Even without fully spatialised audio, the use of sound also alerts the viewer that there is something to see; with the visual cue alone, participants sometimes followed the cue, but not as far as the target. When both audio and visual cues were used, all participants saw the target."* (Sheikh et al., 2016)

Sound can be tied to the picture in a straightforward (hyper-)realistic fashion, or it can act independent from it, which French (2018) claims to increase extended presence in fulldome productions through viewer participation and exchange with the medium and its different sound layers and objects (for an extended discussion on sound/image relations in film see Görne 2017).

An example of an approach that tries to go one step further, taking advantage of the capabilites of a planetarium, is an audio-visual piece produced by some of the authors in 2018. It was attempted to employ a unique storytelling approach: half of the production was meant to be experienced with the eyes closed, to communicate events from the story via flashes of light and color and basic geometric shapes that simulate what we normally perceive visually in that state. As this approach has its limitations due to the default brightness of the video projection when displaying pure black (which makes it hard to simulate directional light and shapes), spatial sound was the principal medium conveying the story, but the light did have an interesting function in maintaining a visual connection to the medium. The second half of the piece was realized as a conventional fulldome video with spatial sound.

The concept was to put the audience in the shoes of someone "blindly" going through his or her last day before suffering a car accident, at which point an experience of the "afterlife" or an altered reality, reassembled from bits and pieces of the last day's memories, was introduced. In between, the "transitioning" was marked by a star projection using the planetarium's Carl Zeiss Universarium Projector.

4. Case studies

In 2018 and 2019, two evening programs under the title "Equinox" were produced at HAW Hamburg's Immersive Audio Lab (IAL) for screening at Planetarium Hamburg. They consisted of several distinct fulldome productions with spatial audio. Most tracks were produced in 7th order Ambisonics while two were produced using a channel based approach.

Fig. 1: Situation at IAL (without ring 0 / floor loudspeakers).

The production system at IAL is an 33.2 array arranged in 5 height layers with ring 1 being at ear level (Fig. 1, Table 1). The array layout was designed to accommodate not just HOA beyond 6th order in half-space but also a wide range of channel and object based audio coding formats

Ring	Height	Diameter	Channels
0	−1,0 m	6,5 m	4
1	0 m	8,0 m	8
2	+1.2 m	6,5 m	8
3	+2.4 m	6,5 m	8
4	+3.5 m	4 m	4 + 1 (VoG)

Tab. 1: IAL array layout, height relative to ear level.

including setups commonly used in sound art and electro-acoustic music; for details see Kessling and Görne (2018).

The playback system at Planetarium Hamburg is a 60.4 array arranged in 4 height layers plus one layer solely for the "Voice of God" loudspeaker, with a dense array in the horizon ring (nominally ear level) and rather sparse in the height layers (Fig. 2, Table 2).

Fig. 2: Planetarium Hamburg. Note the dense horizon array (ring 1).

Ring	Height	Diameter	Channels
1	0 m	20.6 m	36
2	+4.0 m	19 m	11
3	+7.3 m	14.5 m	8
4	+9.5 m	7.9 m	4
VoG	+10.3 m	–	1

Tab. 2: Planetarium Hamburg array layout (level 1 at nominal height of 0 m equals the synthesized ear level by the SSW system; actual height is some 1...2 m above the listeners' heads).

In addition to the screenings at Planetarium Hamburg, the productions were also submitted to the 12th and 13th Fulldome Festival Jena Student Competition. Therefore they had to be prepared for two different playback systems.

In 2016 one of the authors produced a fulldome video utilizing a SpatialSound Wave system at Hochschule Darmstadt's Soundscape & Environmental Media Lab (SEM). The SEM-Lab is equipped with a flexible and mobile loudspeaker array which at the time consisted of 24.2 loudspeakers. This production was shown at the 10th Fulldome Festival Jena Student Awards. It was also adapted in 5.1 for screening at a mobile fulldome tent at Hessentag 2017.

4.1. SSW content

This is the most straightforward approach, assuming the planetarium has SSW installed.

The external spatialization of sounds makes the system beginner-friendly. In contrast to technologies like Ambisonics, no deeper understanding of underlining principles is needed to set up a session for spatialization. The user can connect her or his mobile device to the renderer and control all sound sources at once, which makes it very easy to spatialize pre-produced mono tracks – representing the sound objects – from inside the planetarium.

While this approach is very versatile for live applications, it has disadvantages in linear productions such as fulldome video. Movement can only be recorded in real-time, and the automation of movement cannot be edited as it is common in DAW software. This makes operations such as deletion, copy-pasting and adjustments time consuming and imprecise. Time constraints are therefore a big factor if the producer is spatializing her or his pre-rendered sound sources in the planetarium. If the studio used for the production is not equipped with its own SSW renderer, some of the other workflows discussed might be advisable.

The playback of SSW productions in other venues with SSW systems is relatively easy. However different room sizes might make sound adjustments necessary. Since they are transmitted as separate mono channels, each of the different sound objects can still be adjusted in volume and with effects like e.g. equalization, which is not the case for other approaches such as Ambisonics. The transfer in the form of multiple mono files is also susceptible to mistakes by both the producer and the planetarium technicians. It is crucial that a shared naming convention for all files should be discussed beforehand.

If the production is to be shown in a venue without a SSW system it is possible to pre-render the audio channels for the respective playback system. This is achieved by configuring the SSW renderer in the studio with the speaker positions of the venue and recording its outputs. However, this requires changing the routing of the studio and knowledge of the exact speaker positions of the venue. It is also a real time process. Changes during sound check are not possible anymore.

While rendering for other formats such as 5.1 is theoretically possible, in the case of our production playback on this much smaller horizontal array could not achieve a good representation of localization and sound. To preserve the artistic intent, a dedicated manual remix/downmix for this format would be preferred. Further experiments with productions in Ambisonics would be of interest.

4.2. Higher Order Ambisonics (HOA) content

4.2.1. Producing and playing back in HOA

This is the second most straightforward approach, as producing with an open system like HOA is very convenient.

HOA productions were made using Cockos Reaper DAW and the Ambisonics plug-in suite developed by Zotter et al. and

Rudrich at IEM (Institute of Electronic Music and Acoustics) Graz (Frank et al., 2015; Rudrich, 2018; Zotter and Frank, 2012). The IEM StereoEncoder was our primary plug-in to pan channels and write automation for spatial movement. Productions were made in 7th order to take advantage of the full spatial resolution of the IAL loudspeaker dome.

For the presentation in March 2019 we configured IEM's AllRADecoder with the planetarium's speaker positions, rendering the signals for the loudspeaker array in Hamburg planetarium, and connected our playback laptop via Dante, short-cutting the renderer in the venue and directly feeding the pre-produced loudspeaker signals to the playback system. This experiment basically worked quite well, as can be expected with a system-agnostic format like HOA.

However, the perceived locations of auditory objects might be displaced drastically due to precedence effect if single loudspeakers are too close to the audience – a factor we found to be of special importance in planetariums: The seats close to the "rear" wall of a planetarium with the lowest loudspeakers very close above can be considered the worst seats in the venue, where it might happen that the audience perceives everything behind and above the head due to precedence effect. Of the approaches discussed here Ambisonics can be expected to suffer the most from this effect, since all the loudspeakers are contributing equally to the sound field even if just a single point source is rendered.

Disregarding speakers that are too low and too close, thus utilizing just a smaller slice of the system, increases the distance to the nearest speakers for these seats and thus possibly prevents the precedence effect at these positions. Even though the utilization of less speakers might result in a lower effective Ambisonics order and therewith a smaller listening area and lower spatial resolution, we found that trading off Ambisonics order against precedence effect by practically switching off loudspeakers helps to even out the listening experience for all seats of a planetarium. Experimenting with the lowest loudspeakers (i.e. the horizon ring) combined with careful listening at different seats in the venue is greatly encouraged. An alternative might be to keep the seats close to the room boundaries unoccupied.

4.2.2. Adapting HOA to SSW

As easy as it is to render HOA content for the loudspeaker array of the playback venue, as complicated it gets when HOA has to be played back via SSW. This might be necessary if the venue's signal routing cannot be restructured and direct loudspeaker access is not available before the show. We realized the HOA/SSW adaptation in March 2018 at Planetarium Hamburg and for both screenings at Planetarium Jena. Basically two workflows can be implemented:

1. decoding HOA to the loudspeaker array in the planetarium (which then would be substituted with fixed SSW objects),
2. creating SSW sound objects representing the loudspeaker positions of the production venue,

Both options utilize SSW as a static playback system, where movement of auditory objects is implemented in the HOA signal. The first option is similar to rendering directly for the loudspeakers in the planetarium, with an extra SSW processing stage in between, which seemed to be rather inelegant. We therefore opted for the second approach, creating a virtual IAL loudspeaker dome from SSW objects in the planetarium. It turned out that the "bottleneck" of multiple format conversions (sound object → HOA encoding → HOA decoding → SSW encoding → SSW decoding) led to a substantial spatial imbalance, so that the content had to be remixed for the final presentations with a specific focus on loudness of auditory objects above the horizon ring. Another problem is the decreased number of virtual loudspeakers available. Even though SSW can synthesize 32 sound-objects at once, the setup in both planetariums allowed to use only 24 of them. This was a significant degradation of spatial resolution from the 33 loudspeaker array at IAL.

4.3. Channel based content

Here, "channel based production" refers to pre-produced material that has been rendered within a specific loudspeaker array like Auro-3D or similar, with one exception described below. Direct playback of such material is possible only if a subset of the loudspeaker array in the planetarium matches the original production array, at least approximately (which for example is not working for Auro 9.1 playback in Hamburg planetarium).

4.3.1. Live spatialization

A workflow similar to the customs of sonic art and electroacoustic music is the live spatialization of more or less pre-produced tracks, the presentation being a live performance and the loudspeaker array used as a "spatial diffusion system" or "loudspeaker orchestra". If the venue has a system like SSW with its capabilities of live processing installed, then the pre-production would typically be channel based, where during the live event the channels are treated as sonic objects.

4.3.2. Adapting channel based material to SSW

This adaptation is similar to our approach of adapting HOA to SSW, but with a significantly lower number of virtual loudspeakers (e.g. 9 or 11 for Auro-3D content) implemented as SSW objects. Hence the spatial impression is less stable, and possible aberrations of the virtual loudspeakers tend to have a larger impact.

4.3.3. Adapting channel based material to HOA

Adapting a standard 3D Audio production to HOA is similar to adapting to SSW. However, in our experiments with productions for Hamburg planetarium, we also dealt with channel based productions in our in-house 33.2 format, which then were rendered as 7th order HOA to achieve system compatibility. Results were quite satisfactory; there was no obvious difference between generic channel based and generic HOA material when rendered as HOA for the planetarium loudspeaker array.

5. Conclusion

Fulldome productions benefit greatly from spatial audio. Our experiments with dedicated spatial audio compositions and sound designs for fulldome video have been received very positively by the audience, independent of production and playback formats. The extended storytelling options and enhanced immersive experience of spatial audio for fulldome by far outweigh practical issues and more or less complex workflows.

The Fraunhofer SpatialSound Wave system has become a de facto standard in this niche. If a studio equipped with SSW is available, producing in this format makes playback in planetariums with SSW straightforward. However, this requires spatialization outside the DAW with its convenient editing features, and limits the distribution to venues with SSW systems.

For playing back channel based content in planetariums there are several options; however, all of them require some preparation. Channel based material with a rather low spatial resolution and a low number of channels like 9.1 might be presented as "surround with virtual height", but due to the few virtual height loudspeakers issues with sound coloration and with spatial resolution can be expected. Of course, such a "3D" format would still be an advantage over the common stereo or surround playback in the venue, yet immersion, presence and emotional impact might not be as impressive as one would expect from spatial audio.

Producing in HOA provides the producer with a rather familiar workflow similar to conventional audio productions when compared to SSW. Furthermore, HOA leaves all options of distribution open, as the playback venue does not need to have an Ambisonics system installed. Preparing a HOA production for playback in a remote venue is uncomplicated: in case the venue is not equipped with a HOA rendering system, the producer simply needs the geometrical data of the playback array to render loudspeaker signals for the specific venue (and then the only technical obstacle to overcome would be finding a way to feed signals directly to the loudspeakers).

If the production was made in HOA then of course HOA playback is advisable, as the adaptation to a system like SSW indeed is possible, but likely leads to degradation of the content.

A yet still pending experiment for the production of HOA material is the virtual spatialization in the studio utilizing binaural rendering in HOA instead of the loudspeaker dome. A production environment like this could make fulldome productions realizable even for smaller studios. We expect promising results, especially when personalized HRTFs are incorporated.

The choice of technology thus depends on the requirements of the fulldome production itself: on the venues where it will be shown (only one particular planetarium, only planetariums with SSW, other loudspeaker arrays outside of the fulldome niche), and finally if it will be edited for other distribution channels such as 360° video or VR.

One last remark on the production of spatial audio content: in conventional music, audio drama or sound design productions it is very common to think of sound production and "spatialization" (typically in form of a stereo or surround mix) as independent production steps. However, from our experience this does not apply to spatial audio, as the spatialization here becomes a substantial part of the artistic process (see e.g. Karadoğan and Görne 2019). There is a remarkable difference between a workflow where the sounds are produced in a conventional studio and then "mixed" in the loudspeaker dome, and a workflow where the composition is made in and for the immersive environment.

Acknowledgement

Most fulldome video works described in this paper have been produced during winter semester 2017/18 and 2018/19 in the classes of Prof. Almut Schneider (experimental video), Prof. Mareike Ottrand (interactive illustration) and Gloria Schulz (video technology / video mapping) of Hamburg University of Applied Sciences.

The described SpatialSound Wave production for fulldome video was produced during the winter semester 2015/16 under the supervision of Prof. Sabine Breitsameter of Darmstadt University of Applied Sciences.

A special thanks goes to Prof. Thomas W. Kraupe, Principal and Artistic Director of Planetarium Hamburg, and Sascha Kriegel, Technical Director of Planetarium Hamburg, for extensive support and for not just welcoming but stimulating technical and artistic experiments in their house.

Photo credit: Gertje König, Hamburg.

References

John A. Bargh. 1988. Automatic Information Processing: Implications for Communication and Affect. In *Communication, Social Cognition, and Affect*, Lewis Donohew, Howard E. Sypher, and E. Tory Higgins (Eds.). Lawrence Erlbaum Associates.

Natasha Barrett. 2010. Ambisonics and acousmatic space: a composer's framework for investigating spatial ontology. In *Proc. 6th Electroacoustic Music Studies Network Conference*. Shanghai.

Karlheinz Brandenburg, Martin Schneider, Andreas Franck, Walter Kellermann, and Sandra Brix. 2013. Intelligent Multichannel Signal Processing for Future Audio Reproduction Systems. Conference Paper, AES 52nd Int. Conference on Sound Field Control – Engineering and Perception, Guildford.

Martha Brech and Ralph Paland (Eds.). 2015. *Kompositionen für hörbaren Raum – Die frühe elektroakustische Musik und ihre Kontexte*. transcript.

Noël Carroll. 1990. *The Philosophy of Horror. Or: Paradoxes of the Heart*. Routledge.

Polly Dalton and Nick Fraenkel. 2012. Gorillas we have missed: Sustained inattentional deafness for dynamic events. *Cognition* 124, 3 (2012).

Matthias Frank and Franz Zotter. 2017. Exploring the perceptual sweet area in Ambisonics. Convention Paper, AES 142nd Convention, Berlin.

Matthias Frank, Franz Zotter, and Alois Sontacchi. 2015. Producing 3D Audio in Ambisonics. AES 57th Int. Conference, Hollywood.

Fraunhofer IDMT. [n. d.]. IDMT SpatialSound Wave Website. www.idmt.fraunhofer.de/de/institute/projects-products/spatialsound-wave.html. ([n. d.]). Accessed: 2019-08-29.

Michaela French. 2018. Using the layers of presence as a framework for artistic practice in fulldome space. In *IPS 2018 Toulouse Conference Proceedings*.

Winfried Gerling. 2013. Die Kuppel als medialer Raum. In *Fullspace-Projektion. Mit dem 360°lab zum Holodeck*, Gordian Overschmidt and Ute B. Schröder (Eds.). Springer Vieweg.

Michael A. Gerzon. 1973. Periphony: With-Height Sound Reproduction. *Journal of the Audio Engineering Society* 21, 1 (1973), 2–10.

Thomas Görne. 2017. *Sounddesign. Klang, Wahrnehmung, Emotion*. Hanser.

Mark Grimshaw (Ed.). 2014. *The Oxford Handbook of Virtuality*. Oxford University Press.

Ephraim Hahn. 2018. Musical Emotions Evoked by 3D Audio. Conference Paper, AES Conference on Spatial Reproduction – Aesthetics and Science, Tokyo.

Edward T. Hall. 1969. *The Hidden Dimension*. Anchor Books.

Claudia Hendrix and Woodrow Barfield. 1996. The sense of presence within auditory virtual environments. *Presence: Teleoperators and Virtual Environments* 5, 3 (1996), 290–301. https://doi.org/10.1162/pres.1996.5.3.290

Wijnand A. Ijsselsteijn, Huib de Ridder, Jonathan Freeman, and Steve E. Avons. 2000. Presence: Concept, determinants and measurement. In *Proceedings of SPIE - The International Society for Optical Engineering*, Vol. 3959. https://doi.org/10.1117/12.387188

ITU-R BS.2051. 2018. Advanced sound system for programme production. Intern. Telecom. Union Radiocom. Sector Recomm. ITU-R BS.2051-2. (2018).

Can Karadoğan and Thomas Görne. 2019. Auditory Scenography in Music Production: Case Study Mixing Classical Turkish Music in 7th Order Ambisonics. Conference paper, AES Conference on Immersive and Interactive Audio, York.

Philipp Kessling and Thomas Görne. 2018. Studio for immersive media research and production: Immersive Audio Lab at HAW Hamburg. e-Brief EB453, AES 145th Convention, New York.

Peter P. Lennox, John M. Vaughan, and Tony Myatt. 2001. 3D Audio as an Information-Environment: Manipulating Perceptual Significance for Differentiation and Pre-Selection. In *Proceedings of the 2001 International Conference on Auditory Display, Espoo*.

Ville Pulkki. 1997. Virtual Sound Source Positioning Using Vector Base Amplitude Panning. *Journ. Audio Eng. Soc.* 45, 6 (1997), 456–466.

Daniel Rudrich. 2018. Introducing the IEM Plug-in Suite. 'klingt gut' Intern. Symposium on Sound, Hamburg.

Alia Sheikh, Andy Brown, Zillah Watson, and Michael Evans. 2016. Directing attention in 360-degree video. IBC 2016 Conference. https://doi.org/10.1049/ibc.2016.0029

Meike Uhrig. 2015. *Darstellung, Rezeption und Wirkung von Emotionen im Film. Eine interdisziplinäre Studie*. Springer.

Johannes Voit. 2014. *Klingende Raumkunst. Imaginäre, reale und virtuelle Räumlichkeit in der Neuen Musik nach 1950*. Tectum.

Christiane Voss. 2009. Fiktionale Immersion. In *Es ist, als ob. Fiktionalität in Philosophie, Film- und Medienwissenschaft*, Gertrud Koch and Christiane Voss (Eds.). Wilhelm Fink, 127–138.

Ka Yu, Matthew Brownell, Joslyn Schoemer, Daniel Neafus, Thomas Lucas, and Zachary Zager. 2007. Live Action Film Footage for an Astronomy Fulldome Show. *Planetarian* 36 (01 2007).

Franz Zotter and Matthias Frank. 2012. All-Round Ambisonic Panning and Decoding. *Journal of the Audio Engineering Society* 60, 10 (2012), 807–820.