Stop challenging technology! Towards content-based audio production

M. Rumori, D. Pirrò, G. Eckel

University of Music and Performing Arts Graz, Austria, Email: {rumori,pirro,eckel}@iem.at

Abstract

Production tools that hide their technical complexity stimulate an understanding of creation that centres 'content' rather than materiality, or form. What are the consequences when such user-oriented tools become the predominant means for conceiving auditory experience? Many scholarly interpretations highlight as an advantage that 'content' creators do not have to deal with peculiarities of different sound projection methods anymore. However, assuming a universal model for auditory space effectively shadows its exploration. This paper illuminates that the focus on geometric representation may impede experiential sonic practice. The discussion departs from the main arguments of the call for this ICSA's Student 3D Audio Competition. We aim to identify implicit demarcations that contradict the transdisciplinary nature of engineering in an artistic context.

Introduction

Designing the spatial appearance of sound is a desire as old as any conscious sonic utterance. Throughout music history, acoustic properties of rooms, radiation characteristics of instruments, or effects of musicians distributed in space are substantially considered in composition and musical practice, as they are in theatre and many other artistic areas.

A paradigmatic change is marked by the advent of the loudspeaker as a particular and universal means of sound projection, the conversion of sound pressure waves to electrical or numerical representations and the ability to store them. The detachment of a recorded sound from its specific physical origin is even increased by multichannel sound projection techniques, which exploit psychophysical effects of human hearing or seek to reconstruct sound fields from partial components, such as Ambisonics or Wave Field Synthesis. Those techniques aim at establishing an auditory domain that is also spatially detached from the projecting actuators, that is, the loudspeakers.

In contrast to the notion of a transparent medium, the loudspeaker may be also regarded as an instrument with particular acoustic properties, a location and an orientation in space, with an audiovisual and tangible appearance, in short, as an object with a certain physical presence. In the history of electroacoustic music, such an understanding of the loudspeaker as an instrument is notably connected to the concept of the loudspeaker orchestra and to acousmatic performance.

In the past decades, multichannel loudspeaker projection becomes more and more widespread and accessible, starting from consumer-oriented technology such as surround sound formats, binaural rendering, up to high density loudspeaker arrays (HDLA). This trend is accompanied by the development of production tools for an easier access to the potential of such spatial auditory displays. Generally, the consequence of so-called user-oriented tools is a higher level of abstraction from the technical complexity of underlying algorithms, and a tendency towards uniform, iconic representations of space combined with standardised interfaces. Auditory space is mostly visualised as a two-dimensional projection of three-dimensional Cartesian coordinates or, often in case of Ambisonics, of the surface of a hemisphere.

Many implications of such tools are widely featured, such as a much lower threshold for accessing complex algorithms, the democratisation of formerly exclusive or even elitist technology, and the integration of advanced approaches with existing workflows in audio engineering and production. Nevertheless, we believe that some effects are not sufficiently covered in current discourse. Most importantly, we envision the risk of narrowing the reflection and exploration of the manifold unfoldings of sound in space by promoting specific technical approaches as comprehensive means for the spatial in general.

In this paper, we aim to elaborate on our concern. We do so by picking up this ICSA's call for 'Europe's First Student 3D Audio Production Competition' as a prototypical starting point for discussion. We do neither intend to undermine statements of the call by our take on it nor to offend any of its issuers. Instead, we shall be grateful if our contribution was regarded as a catalyser to a so far neglected discourse which, in our opinion, should be pursued.

Call for 'Europe's First Student 3D Audio Production Competition'

A call for 'Europe's First Student 3D Audio Production Competition' has been issued by IEM *(Institute of Electronic Music and Acoustics Graz)* and VDT *(German Tonmeister Association)* for ICSA 2017 [1]. The call invites submissions in several categories such as computer music, audio drama, or recording. It is addressed to students enrolled in universities or comparable programmes of higher education. This call is accompanied by another, more general call for IEM's icosahedral loudspeaker (IKO) and may be understood as a consecution of the call for Ambisonics works at last ICSA 2015, which took place in Graz as well [2, 3].

In our interpretation, the call implies three major claims that we are going to deduce from the call's text and explain them in the comments to each of the following numbered items.

1. *The development of spatial audio technology and standards is about to reach a terminal state that allows to shift the focus in academia away from technology and towards artistic creation. Inversely, this implies that former efforts of sound projection in space were largely dominated by technological concerns rather than musical, artistic, or aesthetic ones.*

   The call puts the above-mentioned shift 'from challenging the technology to a more content-related working style' in a causal connection to the increased relevance of spatial audio technologies and the recent definition of standards for the production and distribution of spatial audio works. '[D]iscussion [. . .] begins to focus on [. . .] the audio content, story, and aesthetics that can be conveyed by the 3D Audio Technologies' [1]. The advent of a widespread standard medium that 'allows for re-performing, dissemination, [and] broadcasting' (ibid.) suggests its maturity for mere application rather than its ongoing critical reflection and further development.

2. *Future attention will be targeted at the production of 'content,' storytelling, and the aesthetics of* 3D *audio technologies. In particular, the next generation of students will not anymore deal with the idiosyncrasies of technology but rather pursue working styles with a specific, universal medium.*

   The same arguments in the call that support the deduction of the first claim positively imply the second claim. The production focus of 'next-generation, enthusiastic students' is 'necessarily' influenced by the mentioned shift (ibid.).

3. *The universal medium in question is the Ambisonics technique of representing sound signals and their spatial direction of origin. Ambisonics is capable of covering all approaches to spatial music, spatial recording, and spatialised sound.*

   'The Student 3D Audio Production Competition requires Ambisonics as a playback format' (ibid.). The call invites 'all students who deal with spatial music, spatial recording' or being 'interested in spatialised sound.' Submissions are organised in the three categories of contemporary or computer music, audio drama or documentary, and music recording. The logical consequence of generally inviting a multitude of approaches to spatial audio and, at the same time, imposing Ambisonics is the attribution of transparent, universal representational capabilities to Ambisonics.

The call involves two other statements that are less general but will be nevertheless included in our discussion:

4. *Two reasons predestine Ambisonics as the required format for submissions: it is used as an internet distribution format by YouTube, and free Ambisonics tools are available.*

   The two reasons are literally provided: 'as both the internet distribution (youtube – virtual reality) technologies focus on Ambisonics [. . .] and because there are free Ambisonic tools' (ibid.).

5. *The provided template session for the Reaper Digital Audio Workstation (DAW) is a comprehensive way of accessing Ambisonics techniques. Its major effect is facilitating the process of submission review while potential aesthetic implications are negligible.*

   The call asks for using the provided Reaper session template for submissions in order 'to simplify the review process' (ibid.). Similar to the argument structure of claim number 3, this implies that Reaper with its paradigm of spatialised source tracks and the *Ambix* plugins involved universally cover all approaches to Ambisonics that might be explored by submitters [4, 5].

We believe that the implicit claims just deduced are not limited to the scope of this call, where they have a reasonable advertising function to a large degree. Rather, they seem to be prototypical as similar statements are often made in the realm of technologically mediated auditory space, both in commercially oriented and in research contexts.

A fictitious submission

In order to prepare the discussion, we would like to pursue a narrative thought experiment along the path of the issued call.

Let us imagine a 'next-generation, enthusiastic student' belonging to the envisioned target audience of the call [1]. She is interested in exploring the immersive qualities of sound projection all around the listener. This is an aesthetic enquiry. She heard about Ambisonics before and remembers some complex maths connected to it. A friend told her about the Ambix plugins which shall make Ambisonics much easier accessible. She is eager to learn more about them and happily takes the ICSA call as a motivation for getting started.

She opens the provided template session in Reaper, plays a bit with the announcements for the spatial orientation, and she imports some of her own recordings. Initially she is a bit confused by the behaviour of the encoder plugin, compared to her prior experience with surround sound panning. She soon understands that the source position actually does not move freely in space, but on the surface of a sphere.

For listening, she resorts to the binaural auralisation of the MUMUTH as recommended. She is amazed about the spatial impression just by listening with headphones, and she wonders how this experience may relate to the real space in Graz.

Our student notices some spectral effects that seem to be introduced by the spatialisation but she knows how to compensate for them by filtering her recordings accordingly. She is not sure though what that will mean for playback in the real MUMUTH or in a YouTube stream.

While she gets more and more acquainted with the newly discovered tools, she finds it particularly challenging to create an impression of varying distances, sound objects that approach the listener or are arranged behind each other. She explores the effects of the more advanced plugins for spatial manipulation, but she also thinks about how to transfer her prior knowledge on working with depth and width in stereophony.

At this moment, let us leave our student with her experiments in order to reflect what she actually did with respect to the medium of Ambisonics.

We believe that she immediately started to explore a specific space, a very peculiar one in addition: a mental extrapolation of a real space that is reconstructed by measured binaural room impulse responses, the MUMUTH Ligeti hall. She does not know the MUMUTH yet so she cannot relate her listening experience to her memories of the real space. This will have changed after her stay during ICSA, and this will influence her future dealing with the auralised MUMUTH to a large extent. We approached this effect under the term 'in-situ composition' in the artistic research project *The Choreography of Sound* and with its auralisation tool, *StiffNeck* [6, 7].

We think that our student also investigated particular properties of Ambisonics, how the technology 'behaves' and how it 'reacts' to what she is doing, but also how it acts on the process of aligning her conceptual ideas, her expectations, and her taste to what she is experiencing. She does not know yet how her study will sound in different Ambisonics systems, with different decoding strategies or as a YouTube stream. But as she is smart enough to already wonder about such effects, she will be developing strategies to cope with them. What she will have gained or embodied then is experiential knowledge on the qualities of a specific medium: that of Ambisonics.

She might be eager to find out more about how her experience relates to the theoretical model and the signal processing taking place behind the scenes. That would mean to leave her role as a content producer as it was assigned to her by the call. Probably she did not even fulfil this role so far. In fact, she conceived a work exploiting the idiosyncrasies of Ambisonics, that is, as an interpretation of the Ambisonics medium. We think that this is far from immaterial 'content production.'

Conclusions

Based on the thought experiment just conducted we are going to draw conclusions regarding above-mentioned widespread claims and their implications. They will be approached in reverse order.

5. *Specific tools – neutral and comprehensive?*

   Specific tools provide specific ways of accessing a certain method or technology. Applied to the analysed call, this means that the Reaper template session, which uses the Ambix plugin suite, does not allow for approaching Ambisonics in all possible, thinkable or yet unthinkable ways. Regardless

of Ambisonics, aesthetic influences of track-based digital audio workstations on electroacoustic music composition is evident, for example, when regarded in the context of algorithmic or generative composition. When dealing with the spatial model of Ambisonics, recording software and plugin architectures favour an object-based understanding of sound in space, that is, the assumption that complex auditory entities may be dissected into or constructed from a number of separate constituent parts and corresponding metadata.

On the other hand, a sound field encoded in Ambisonics, like a B-format recording, poses an irreducible entity, a complete scene, which may be manipulated with special techniques such as beam-forming, spatial up-mixing or blur but nevertheless has to be retained in its specific representation.

A thinkable, experimental approach may operate directly on the level of Ambisonics orders or even single channels based on the direct experience of its auditory consequences, which is something in between – it may neither be described in an object-based nor a scene-based paradigm but it is still and in particular specific to Ambisonics. Such an approach cannot be pursued with common production tools as usually featured.

4. *Ambisonics – ubiquitous and free?*

   Our thought experiment revealed that site-specificity is indeed incorporated in the practical dealing with Ambisonics, despite its claimed ubiquity and independence of actual projection means. Throughout the history of stereo production, ubiquity has been sought – and to a wide extent achieved – by exploiting a sense of media specificity, not by neglecting it, which allowed for gaining both experiential and formal knowledge. Similarly, we believe that it is much more inspiring and promising to investigate the medial idiosyncrasies of Ambisonics with respect to different ways of (re-)production rather than attempting to even them out or to deny them.

   As a side note to the experience of our student in the thought experiment we may add that she discovered, luckily enough, that the Ambix plugins are free and libre software and thus allow for further exploration of their underlying foundations. She did not miss though that, unlike the call implies, this is not true for Reaper, the digital audio workstation obligatory to use for submissions. Strictly speaking, it is not even 'free as in beer,' unless using the evaluation license is regarded as an acceptable way for production.

3. *Ambisonics – the incarnation of spatial audio?*

   The presence of idiosyncrasies when dealing with Ambisonics, as exemplified in our thought experiment, shows that there are principal differences between conceptual approaches to auditory space and its representation in terms of spherical harmonics. Certain conceptual aims have to be adapted from a channel-

based projection (e.g., stereophony) to scene-based soundfield reconstruction, such as constituting the perceptual impression of depth, which can be hardly formalised with the means of either representation. Beyond, many conceptual understandings do not have any representational equivalent in Ambisonics, such as above-mentioned notion of the loudspeaker as an instrumental sculpture.

More generally, the artistic exploration of sound in space with respect to its representation by formal models may be explained in terms of relational ontology [8]. Its basic assumption is that *relations* between entities are more essential than entities themselves. This contrasts with (traditional) substantivist ontology in which entities are primary and relations are derivative.

Building on this postulate, we do not conceive sound as an *object* that could be imagined in isolation. Rather, we treat sound in relation to its performative and technical means as well as to its physical and perceptual qualities or to the acoustical and medial characteristics of its staging. Sound has a relational ontology in the sense that its relative and dynamic relationship to all these diverse aspects constitutes its definition. Spatial sound is an artistic practice that acknowledges and revolves around the radical relational essence of sound: spatial sound is intertwined in a mesh of relations in and around itself which co-determine its production, performance and reception.

That is, all those mutual connections cannot be untangled and they therefore have to enter the composition of spatial sound. Spatial sound artists explore those networks, elicit their resonances and recompose them into aesthetic artefacts. They perceive the sensible effect each relation has on the aesthetic perception of spatial sound and they expose approaches to sound which explicitly or implicitly posit a kind of *absolute* formulation.

As an example in the scope of this discussion, stating that the Ambisonics formulation is capable of covering all possible approaches to spatial sound effectively postulates that this formulation is *absolute* and *invariant* with respect to the particular work and de-coupled from its aesthetic reception. Such statement therefore deprives the relational essence of spatial sound of a fundamental aspect: the conceptual and technical formulation of its spatial appearance in relation to all the other aspects of one particular work.

2. *Content rather than technology?*

   The previous discussion also makes clear that there is neither 'content' independent of the way and means it is conceived, formulated and projected, nor is there a neutral, transparent 'container' independent of what it contains. In both cases, absolute and invariant properties are assigned.

From that follows that there can be no gradual detachment of auditory works from technology. Indeed, the drastically changing appearance of technical means very much influences the ways auditory space is envisioned and approached. Consider, for example, '[t]he times when only nerdy specialists in their MATLAB labs could generate 3D content' [9] compared to recent Ambisonics plugins as discussed above. The increasing emphasis on presenting largely simplified means of access to technology rather than exposing their actual complexity does not mean that technology becomes less influential or negligible. On the contrary: the shift of the perceived relations that is caused by abstracting tools has to be balanced by an opposing shift in reflection, if sonic works shall still explore sonic ontology rather than serving as technical demonstrations within the boundaries of said tools.

1. *Progress towards the end of technology?*

   The implicit notion of technological progress that makes itself vanish triggers subjects well beyond the scope of this paper. In a narrower sense related to spatial sound, the previous discussion concluded that the apparent focus on so-called 'content production' does not diminish the role of the underlying technology. More generally, the notion of progress is always normative and driven by the narrative of finally reaching an envisioned goal. Its normative nature is revealed when the consequence of its actual success is imagined: any quest for improvement, or 'progress,' would immediately cease. On the other hand and again, the normative notion of more and more perfecting the means for actualising spatial audio establishes an absolutism that contradicts the ever relational emergence of sound in space.

   When the desire of designing the appearance of sound in space as expressed, for example, by Karlheinz Stockhausen is called a former 'exotic desideratum' that 'after decades of niche existence has become *en vogue*' due to the development of spatial sound projection technology, this interpretation might ignore the relational role of the 'visionary force of musical avant-garde' [10]. Never will technology fulfil what has been once envisioned without entirely having reshaped both the vision and the conditions of approaching it – just like any vision is never a concrete request towards engineering but a contribution to the superordinate narrative of progress.

Acknowledgements

References

[1] Europe's first Student 3D Audio Production Competition, ICSA 2017, IEM/VDT, URL: `http://ambisonics.iem.at/icsa2017-music`, last retrieved: August 31, 2017.

[2] International Call for Compositions – for and with – the IKO!, URL: `http://iem.kug.ac.at/projects/osil/osil-news/call-for-compositions-201617.html`, last retrieved: August 31, 2017.

[3] PURE Ambisonics Concert & the Night of Ambisonics, Call for Works, ICSA 2015, IEM/VDT, URL: `http://ambisonics.iem.at/icsa2015/pure-ambisonics-concert`, last retrieved: August 31, 2017.

[4] *REAPER* Digital Audio Workstation, URL: `http://reaper.fm`, last retrieved: October 13, 2017.

[5] *Ambix* Ambisonic plug-in suite, URL: `http://www.matthiaskronlachner.com/?p=2015`, last retrieved: October 13, 2017.

[6] Gerhard Eckel, Martin Rumori, David Pirrò, Ramón González-Arroyo: A Framework for the Choreography of Sound, in *Proceedings of* ICMC 2012, Ljubljana.

[7] Gerhard Eckel, Martin Rumori: StiffNeck: The Electroacoustic Music Performance Venue in a Box, in *Proceedings of* ICMC 2014, Athens.

[8] Wesley J. Wildman: An introduction to relational ontology, in John Polkinghorne (ed.): *The Trinity and an entangled world: Relationality in physical science and theology,* Eerdmans 2010, pp. 55–73.

[9] Report from the ICSA 2017, URL: `http://vdt-icsa.de/blog/2017-09-26-report-from-the-icsa-2017/`, last retrieved: October 13, 2017.

[10] inSonic 2017. Immersive Future, URL: `http://zkm.de/event/2017/12/insonic-2017`, last retrieved: October 13, 2017.