All Around Audio 2024 Abstracts

Session 1

Application of measured room acoustic datasets for modeling spatial auditory perception

Stephan Werner

The objective of audio mixed reality applications is the synthesis of virtual audio objects within a real-world environment. In order to accurately reproduce virtual audio objects, it is essential to consider the local real acoustic environment when creating the virtual room acoustics. Excessive deviations result in perceptible discrepancies between the actual and virtual environments, which can diminish the perceived plausibility and lead to an inaccurate sense of location within the head. Consequently, the position-dynamic binaural synthesis is required to be capable of continuous adaptation in accordance with the local acoustic environment. In addition to the acoustic differences between different rooms, significant discrepancies can also be observed within a single room. Additionally, it would be beneficial to determine at what point perceptible acoustic differences emerge, and whether these differences have a detrimental effect on the listening experience. To address this, a prediction model was developed based on a comprehensive high-resolution acoustic dataset from multiple rooms and a listening test based on this dataset. The used mixed effects model can predict the perceived differences in sound based on acoustic differences. This contribution outlines the conceptual approach of utilizing real and spatially high-resolution acoustic measurements, and provides insight into the modelling methodology employed.

OpenResonance: A FLOSS, extended Higher Order Ambisonics spatialiser for Unity

Martin Rumori, Ludwig Zeller

When it comes to 3D audio spatialiser plugins for Unity, only a few options exist that employ Higher Order Ambisonics (HOA) support and that allow access to the HOA representation of the encoded soundfield, such as the third-order audio bed player of Rapture3D (Blue Ripple Sound). The Resonance Audio project, formerly developed at Google, provided a solid base for further developments as it eventually became Free and Libre Open Source Software (FLOSS). Yet, the project has been discontinued, and its code base was finally archived in 2023. As a result, more recent platforms did not become supported anymore, such as Android 64-bit or Apple Silicon.

OpenResonance departs from a fork of the discontinued Resonance Audio project and maintains its FLOSS license (Apache 2.0). It aims at making the bindings for the Unity game engine working again on common target platforms, such as Apple Silicon and Meta Quest 2 and 3. Yet, as a side effect, other previous targets of Resonance Audio may benefit from the porting efforts as well, such as the Resonance VST plugins running on Apple Silicon.
Furthermore, OpenResonance is being extended by a media framework that enables advanced access to the HOA soundfield representation within 3D scenes while being played. This includes the flexible playback of pre-recorded Ambisonics audio beds as well as in-situ recording of the currently experienced audio scene in up to third order. Apart from documenting soundfields for flexible listening targets (e.g., head-tracked binaural or multi-channel), the latter also enables the creation of three-dimensional audio content from within spatial worlds. The mixed-reality modular synthesizer project OpenSoundLab, for instance, will largely benefit from this functionality and therefore poses a major motivation behind OpenResonance.

The spotlight talk will present the current state of OpenResonance, namely extended platform support and HOA recording, and a roadmap for further development, such as the integration of realtime Ambisonics streaming by means of AOO (Audio over OSC).

Game Audio in 3D: Tools, Techniques, and Cross-Platform Consistency

Christian Schörkhuber & Markus Zaunschirm

This presentation explores the creation of immersive, dynamic, and interactive spatial audio for game development and XR applications. We will provide an overview of the current tools used in game audio, with a particular focus on spatial audio design. Key challenges will be highlighted, including maintaining audio quality, optimizing performance, and achieving consistent cross-platform compatibility. To tackle these complexities, we introduce atmoky trueSpatial—a spatial audio toolset designed for compatibility with various game engines and audio middleware, providing precise rendering to ensure a consistent auditory experience across platforms and playback setups.

Protection vs. situational awareness in very noisy environments

Marian Weger

Spatial hearing is an essential human ability. We rely on it every day, for example, while communicating with multiple persons simultaneously, or while navigating through dangerous environments such as traffic or industrial/construction sites. Dangerous environments, however, are often noisy and thus bear the additional risk of hearing injury. The only means of protecting ourselves from this risk is to wear hearing protection devices, which in turn prevent us from hearing

and communicating properly. Although effective hearing protection is indispensable in any context, some sectors such as heavy industry, emergency services, or the military cannot sacrifice on communication and auditory situational awareness either. This spotlight talk addresses two challenges that arise from the trade-off between protection and perception. The first challenge is to measure and predict the risk of noise-induced hearing injury in order to take the necessary actions: At very high sound pressure levels, beyond the range of standard microphones, hearing protection alone may not be sufficient and the actually consumed noise dose has to be monitored by unconventional methods. The second challenge is to make hearing protection acoustically transparent: While speech communication is easily re-established by radio/intercom systems, the perceptual quality of a natural conversation is only achieved by incorporating spatial sound cues. Likewise, for maintaining natural auditory situational awareness, the outside acoustic environment (except the harmful noise) needs to be recreated binaurally. Finally, this talk provides an insight into acoustics research at the French-German Research Institute of Saint-Louis.

SESSION 2

LTFATPY: Interfacing Octave and Python

Clara Holomey

LTFATPY is a software package for accessing the Large Time Frequency Analysis Toolbox (LTFAT), written in Octave and C, from Python. LTFAT comprising linear transforms for Fourier, Gabor, and wavelet analysis along with computational routines for the specification of invertible constant-Q and auditory-inspired time-frequency representations and phase retrieval algorithms make it particularly suitable for audio machine learning tasks and motivate its addition to Python. We introduce LTFATPY, describe its design considerations and structure, and give guidance on its usage.

Enhanced creativity: How AI can enable new forms of expression without destroying artists' livelihoods.

Mathis Nitschke

In the current debate about artificial intelligence, the fear of losing artistic livelihoods is often at the forefront. But what if AI does not mean the end, but rather an expansion of artistic possibilities? The lecture shows how AI systems can function as tools that enable new forms of interaction and artistic creation - without undermining the value of human work and inspiration. Instead of destroying old structures, the aim is to open up new spaces in which man and machine can create something together that was previously unthinkable

SESSION 3

Sounds of writing
Katharina Groß-Vogt

Sonic interaction design has proven to be useful in learning to write. Even though handwriting is usually a rather silent activity, we have confirmed that writers are very sensitive to non-ecological sound mappings. In our study of writing and drawing, we found that enhanced sound design can lead to a change in the fluency, e.g. when test subjects are asked to copy line drawings. Another possible application of sound in writing is the learning of individual letters. In this case, the users are 6- to 7-year-old schoolchildren who have to memorize and put together the phoneme and grapheme of a letter. During one school year, the children were provided with an interface to experience different sound designs and interaction strategies.

Dark Sonification: Enacting Ecological Environments in Multimodal Display Systems
Miguel Angel Crozzoli

Dark Sonification interactive multimodal display emerges as a phenomena through data and user intra-action, both conceptualized as objects that co-create a relational space for exploration. This space can be perceived as an ecological environment, with dynamic boundaries enacted by the objects agential cuts. Within this environment, interaction occurs in a perceptual field where data and user are entangled. This project investigates how data perceptualization is driffracted through affect based on user intent within this entanglement. The presentation introduces and demonstrates the foundational concepts behind Dark Sonification, discussing a system shaped by new materialism, actor-network theory, and thinking at the edge and focusing in connection with micro-phenomenology.

Sonification as a Composition Technique and Means of Artistic Expression
Maria Kallionpää & Hans-Peter Gasselseder

Despite of the ever-expanding variety of technological means accessible to the composers, performers, orchestrators, music producers, and music pedagogues of our time, the essence of music composing has mostly remained the same for centuries: the key question is, how to form an artistically innovative core idea of a musical piece and to find the most ideal instruments for it?

This paper discusses composers' explorations on establishing their individual voices, as well as finding and selecting the tools and techniques that would best serve their artistic goals in today´s complex and pluralistic network of aesthetics. To facilitate this, various software solutions have come to the market, enabling a wider palette of sounds to be used as musical "raw material". Instead of pre-composed music based on the rhythms and pitches organised by the composer, we will focus on how to translate other kinds of data into "notes", or more generally "sound events", and how to use it in an artistically meaningful manner, resulting into complete musical compositions on their own right. To do this, we will present compositional case studies based on sonification, realised by various methods and technologies. These include, for example, application of the ORCID software in artistic practice. We will present Maria Kallionpaa´s environmentally themed work "El Canto del Mar Infinito" (2020), as well as her composition "The Reef" (2023), the musical material of which is based on the computer-based analysis of the sounds recorded on a coral reef. Moreover, we will discuss Olga Neuwirth´s work "Kloing!", which drew part of its material from the seismic data gathered in the Sumatran area during the seaquake that caused the tsunami disaster on the 26th of December 2004.

Sonification techniques do not produce musically utilisable aesthetics per se. This raises the question for the creative domains in which composers engage when involving sonification techniques in their process. These domains include:

- The choice of the data or subject being sonified.
- The preprocessing of this data (e.g. the quantization of a series of floating point values to integer pitch values and metronomic timings), a specifically musical design of the sonification method, i.e. the transformation or mapping of the data to sound structures.
- A purposeful work with the parameters of the sonification.

A further approach has been explored by the TouchNoise system (2014-2017), developed by Axel Berndt, Nadia Al-Kassab and Raimund Dachselt. It is a sonification of a particle simulation. The compositional domain here is spanned primarily by a palette of interaction techniques with the particles field, including direct manipulations of the particle distribution as well as flow field and flocking algorithms. We will discuss the aesthetics that this approach evokes.

WE ARE EUDRES EUROPEAN UNIVERSITY

fh st.pölten UNIVERSITY OF APPLIED SCIENCES

SESSION 4

DAB+ more than audio broadcast
Gernot Fischer

DAB+ is the standard for terrestrial transmitting audio content and the modern version of FM radio. It offers a set of unique features and numerous advantages over comparable technologies, which are very well worth exploring. The state of the art of its implementation and distribution as well as some further ideas for new services are discussed, with a special focus on innovations for safety, advertisement and – in general – the future of radio.

Automating Speech Audio Post-Production with Auphonic
Christoph Grasser & Lukas Maier

Auphonic develops a tool that intelligently enhances speech recordings. Its capabilities include automatic noise reduction and reverb removal, adaptive leveling, silence and filler word cutting, filtering, auto EQ, and more. All related AI models and signal processing algorithms are developed and trained in our office in Graz.
We aim to provide an in-depth look at how Auphonic works, what the development process entails, and the best ways to use our toolkit.

The Genie is out of the Bottle - Experiences with ADM, the universal object-based audio metadata model
Werner Bleisteiner

For some time now, video and audio production systems can produce Metadata Guided Audio (MGA), also known as "object-based audio", applying ADM - the AudioDefinition Model.
This is a standard jointly developed by BBC R&D, IRT, EBU, NHK, Dolby, Fraunhofer and others that allows audio to become adaptive, immersive and interactive.
We'll have a look at its background, possibilities and applicabilities.