Tuesday 8/18 Wednesday 8/19 back to program

Schedule for Monday, August 17

All Times are PDT = Pacific Time, West Coast USA

Time	**Room**	**Plenary Session 1**
9:00 - 9:15	AVAR Keynotes	Opening Remarks, AVAR Conference Organizing Committee
9:15 - 10:05	AVAR Keynotes	Keynote: Varun Nair, *Facebook* *New Audio Realities at Scale*
10:10 - 11:00	AVAR Panels	Panel Discussion: *Reflections on New Audio Realities at Scale*
11:00 - 11:30	AVAR Lobby	Coffee Break
Time	**Room**	**Parallel Sessions**
11:30 - 1:30	AVAR Papers	Paper Session 1: AR/VR Sound Perception
11:30 - 1:00	AVAR Workshops	Workshop Session 1: Acoustics for XR
1:30 - 2:00	AVAR Lobby	Coffee Break
2:00 - 3:30	AVAR Papers	Paper Session 2: Spatial Sound Reproduction and HRTF's I
2:00 - 4:00	AVAR Workshops	Workshop Session 2: Psychoacoustics and Sound Zones
4:00 - 5:00	AVAR Lobby	Open Chat

Emerald Sponsors:

facebook Reality Labs

VALVE

Silver Sponsors:

GAME SOUND CON
THE ART, TECHNOLOGY AND BUSINESS OF GAME AUDIO

CORE

MAKER OF
OctoMic &
TetraMic

World's Finest Ambisonic Microphones

Plenary Session for Monday, August 17

back to top

9:00 - 9:15 Opening Remarks, AVAR Conference Organizing Committee

9:15 - 10:05 Keynote Speaker: Varun Nair, *Facebook*

New Audio Realities at Scale

What does it take to ship new audio technology at a scale of billions of people? What does it take to productize great research and make it available in products? How can the bleeding edge find impact? The last 1% takes 99% of effort and time. Often the biggest technology blockers are found when scaling solutions that work for many people. The last mile is where a lot of innovation happens. This talk is part retrospective and part future looking on what it takes to create compelling audio technologies and experiences for AR and VR. We'll deep-dive into lessons learned in shipping products over the years and use that as a marker to understand what the future looks like and what needs to be done to improve the state of audio for AR and VR.

Varun Nair, *Facebook*

Varun Nair has spent the past 15 years with audio, working on hundreds of projects including software development, product management, media post-production and interactive content production. He co-founded Two Big Ears and led development of audio tools and technology for VR, AR and emerging immersive media. Two Big Ears was acquired by Facebook in 2016, where he currently leads the AR/VR audio software team working on algorithms, tools, spatial audio technology and scalable audio systems across a wide range of software and hardware products.

10:10 - 11:00 Panel Discussion:

Reflections on New Audio Realities at Scale

This panel will continue the discussion of the keynote topics. Additionally, the panel will focus on individualization/personalization of spatial audio.

Panel Members:

- **Varun Nair,** *Facebook*
- **Scott Selfon (Moderator),** *Facebook Reality Labs*
- **Rémi Audfray,** *Magic Leap*
- **Kaushik Sunder,** *EmbodyVR*
- **Poppy Crum,** *Chief Scientist, Dolby*

About the Panelists:

back to top

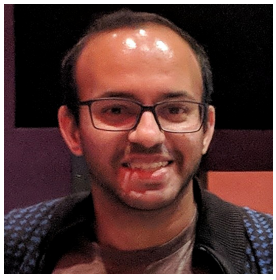
Scott Selfon (Moderator), *Audio Experiences Lead, Facebook Reality Labs*

Scott Selfon is the Audio Experiences Lead at Facebook Reality Labs (formerly Oculus Research), exploring and inventing the technical, creative, and design languages of sound for virtual and augmented reality. Prior to that, he was engaged in developer consultancy, education, and support efforts across three generations of the Xbox game console and broader Windows audio efforts at Microsoft. Scott has composed music for a wide range of media, including film, television, games and live performance, and is a violist with the Puget Sound Symphony in Seattle. A member of the Game Audio Network Guild (G.A.N.G.) advisory board, Scott has been a faculty member of both the Pacific Northwest Film Scoring Program and the University of Southern California, has coordinated and lectured at the Game Developers Conference Audio Boot Camp tutorial sessions for more than fifteen years, and has spoken at the conferences worldwide on audio, interactive entertainment, and real-time sound implementation techniques and technologies. Scott is an alumnus of the University of Southern California, where he obtained dual degrees in music composition (film scoring emphasis) and computer engineering/computer science.

Rémi Audfray, *Magic Leap*

Rémi Audfray is Head of Audio Algorithms R&D at Magic Leap, where he leads the development of new technology and tools for audio in mixed-reality. Before joining Magic Leap, Rémi spent over 7 years at Dolby Laboratories. On the Sound Technology Research team his work was focused on headphone rendering, room acoustics, and content creation tools. He led the development of the Binaural RMU, which became part of the Dolby Atmos authoring tools for VR production. His prior experience includes a short time at Auralex Acoustics, as well as Chief Engineer for Keith Yates Design from 2007 to 2010. In addition to audio system design, calibration, and traditional architectural acoustics simulations, Rémi was responsible for applying numerical simulation methods to solve multi-physics problems. He continues to support KYD as a consultant. Rémi is an author on multiple patents and technical papers. He received his Diplôme d'Ingénieur from the Ecole Centrale de Lyon (France), and his M.Sc. in Music Technology from Indiana University Purdue University Indianapolis (USA) in 2006.

Kaushik Sunder, *EmbodyVR*



Kaushik Sunder currently works as a Principal Research Scientist and leads the Audio and Acoustics Research Team at EMBODY. Kaushik has spent a great deal of his research career in the field of 3D Audio and Psychoacoustics. Over the last few years, his research has focussed on understanding the importance of personalized HRTFs particularly for headphone playback of spatial audio. Prior to working at Embody, Kaushik served as a role of Research Scientist at Ossic and as a Postdoctoral Researcher at the Sound Recording Department, Schulich School of Music at McGill University. He is also a visiting research scholar at the Human factors department at NASA Ames Research Center. Kaushik received his PhD from the Digital Signal Processing Laboratory, NTU Singapore. He has regularly authored articles appearing in the Journal of Acoustical Society of America, IEEE Signal Processing Magazine, Journal of Audio Engineering Society, and AES International Conventions.

About the Panelists:

back to top

Poppy Crum, *Dolby*

Dr. Poppy Crum is Chief Scientist at Dolby Laboratories. She is also an Adjunct Professor at Stanford University in the Center for Computer Research in Music and Acoustics and the Program in Symbolic Systems. At Dolby, Poppy directs the growth of internal science. She is responsible for integrating neuroscience and sensory data science into algorithm design, technological development, and technology strategy. At Stanford, her work focuses on the impact and feedback potential of new technologies including gaming and immersive environments such as Augmented and Virtual Reality on neuroplasticity and learning. Poppy is a U.S. vice-chair to the International Telecommunication Union (ITU) member of the Stanford Research Institute Technical Council, and Consumer Technology Association Board of Industry Leaders. Prior to joining Dolby Laboratories Poppy was Research Faculty in the Department of Biomedical Engineering at Johns Hopkins School of Medicine where her neurophysiological research focused on understanding the neural correlates of how we hear in a complex acoustic environment and the functional circuitry of the auditory cortex. Poppy is a Fellow of the Audio Engineering Society. She is a: 2018 recipient of the Advanced Imaging Societys Distinguished Leadership Award, a 2017 recipient of the Consumer Technology Associations Technology and Standards Achievement Award for work towards the introduction of over-the-counter hearing-aid devices, and has been named to Billboard Magazines 100 most influential female executives in the music industry. She is a frequent speaker including, TED, SXSW, IEEE, TNW, WIRED, on topics related to the intersection of: human experience, artificial intelligence, sensory data-science, and immersive technologies.

Paper Session 1: **AR/VR Sound Perception** back to top

Room: AVAR Papers Session Chair: Johannes Arend

11:30 - 11:55 Monday, August 17

AUTHORS: **Zamir Ben-Hur, David Alon, Philip Robinson and Ravish Mehra**

TITLE: **Localization of Virtual Sounds in Dynamic Listening Using Sparse HRTF's**

Reproduction of virtual sound sources that are perceptually indistinguishable from real-world sounds is impossible without accurate presentation of the virtual sound source location. A key component for such reproduction system is the Head-Related Transfer Function (HRTF), which is different for every individual. In this study, we introduce an experimental setup for accurate localization performance evaluation of a spatial sound reproduction system in dynamic listening conditions with the ability to compare against real-world localization performance. The setup allows testing of different virtual reproduction conditions, such as different HRTF's or different representations and interpolation methods of the HRTF's. Localization experiments for comparing real-world sounds with virtual sounds using high-resolution individual HRTF, sparse individual HRTF's and generic HRTF are investigated.

12:00 - 12:25 Monday, August 17

AUTHORS: **Tim Lubeck, Christoph Porschmann, and Johannes Mathias Arend**

TITLE: **Perception of direct sound, early reflections, and reverberation in auralizations of sparsely measured binaural room impulse responses.**

Binaural auralization applying binaural room impulse responses (BRIRs) requires a high measurement effort and immense computing capacity when rendered with maximum spatial resolution. In this study, we conducted an adaptive ABX listening test to determine the minimum grid resolution of BRIRs (i.e. the spatial resolution) sufficient to achieve an auralization indistinguishable from an auralization of BRIRs with maximum grid resolution. The sparse BRIR sets were calculated by spatial subsampling of dense BRIR sets in the spherical harmonics domain. We determined perceptual thresholds separately for the direct sound, early reflections, and reverberation for four different measured rooms. The results show that the spatial resolution of the early reflections and the reverberation can be reduced significantly to a spatial order of 4.

12:30 - 12:55 Monday, August 17

AUTHORS: **Sebastian Nagel, Daniel Haupt, and Peter Jax**

TITLE: **Coherence-Adaptive Binaural Cue Adaptation**

The recently proposed Binaural Cue Adaptation approach can adapt binaural signals to listener head movements during headphone playback, so that sound sources appear fixed in space instead of to the head. Unlike other methods, it works with pre-rendered signals or binaural recordings acquired using artificial heads or on-ear microphones. This contribution introduces an improved version of the original algorithm which considers signal coherence. Since the human auditory system perceives coherent and incoherent signals differently, the subjective quality of the output signals is significantly improved.

1:00 - 1:25 Monday, August 17

AUTHORS: **Tomasz Rudzki, Damian Murphy, and Gavin Kearney**

TITLE: **On the Measurement of Perceived Lateral Angle Using Eye Tracking**

The quality of binaural reproduction of virtual sound sources is often investigated using auditory localization experiments. An alternative method of collecting listener localization responses to spatial audio stimuli using an eye tracking system built into an AR headset is presented. To validate the method, an experimental procedure is proposed, which includes collection of gaze pattern data in response to binaural signals varying in interaural time and level differences. Additionally, the collected data is compared against data obtained using a visual pointer adjustment as well as data obtained from individual and generic HRTF sets. The preliminary results suggest that the use of eye tracking for auditory localization experiments might be a time efficient and precise testing method for binaural audio.

Paper Session 2: **Spatial Sound Reproduction and HRTF's I** back to top

Room: AVAR Papers Session Chair: David Lou Alon

2:00 - 2:25 Monday, August 17

AUTHORS: **Isaac Engel, David Alon, Kevin Scheumann, and Ravish Mehra**

TITLE: **Listener Preferred Headphone Frequency Response for Binaural Rendering and Stereo Audio**

When spatial audio content is presented over headphones, the audio signal is typically filtered with binaural room impulse responses (BRIRs). An accurate virtual auditory space presentation can be achieved by flattening the headphones' magnitude response. However, when presenting stereo music over headphones, previous studies show that listeners prefer headphones with a magnitude response that simulates loudspeakers in a listening room. It is yet unclear if headphones that are calibrated in such way will be preferred by listeners for spatial audio content as well. This study investigates how listeners' preference for headphone frequency response may differ between stereo audio content and spatial audio content, which was rendered by convolving the same stereo content with in-situ-measured BRIRs of loudspeakers in a room.

2:30 - 2:55 Monday, August 17

AUTHORS: **Steven Crawford, Remi Audfray and Jean-Marc Jot**

TITLE: **Quantifying HRTF Spectral Magnitude Precision in Spatial Computing Applications**

In this paper, an algorithmic approach towards computing quantifiable metrics regarding HRTF spectral magnitude synthesis performance of virtual sound systems, such as those used in VR/AR/MR environments, is presented. Utilizing regularized regression in parallel with a probability and information theory technique, the system provides a detailed analysis of a virtual spatializers spectral magnitude rendering accuracy at a given point in space. Applying the proposed system to the final signal processing stage of a spatial audio rendering pipeline enables the engineer to establish critical performance quantities for benchmarking future modifications to the rendering channel against. The proposed system demonstrates an important step towards automating virtual audio system evaluation and may ultimately act as a participant substitute during critical listening tasks.

3:00 - 3:25 Monday, August 17

AUTHORS: **Md Tamzeed Islam and Ivan Tashev**

TITLE: **Anthropometric Features Estimation Using IntegratedSensors on a Headphone for HRTF Personalization**

Personalization of HRTF is essential for spatial sound rendering, which typically happens based on one or more anthropological measures of the subject. Measuring these anthropometrics accurately and reliably is still a challenge. In this paper, we propose a system for obtaining anthropometric measurements, suitable for HRTF personalization, directly from a high-end headphone. The proposed system is multi-modal and leverages existing sensors to extract features related to listener's head dimensions. We propose three signal processing methodologies for three modalities of sensors and a fusion algorithm to aggregate these extracted features for a robust anthropometry estimation. To verify the design we use a data set, collected from 35 subjects. The proposed algorithm achieves an error (RMSE) of 0.581.21 cm for anthropometry estimation.

Workshop Session 1: **Acoustics for XR**

back to top

Room: AVAR Workshops

11:30 - 1:00 Monday, August 17

Enzo De Sena, Zoran Cvetkovic and Hseyin Hachabibolu, *University of Surrey, UK*

TITLE: **Interactive Room Acoustics Synthesis for XR**

Simulation and rendering of room acoustics enables to make the acoustics of a virtual space audible and is essential for providing a high level of immersion in AR/VR applications–without it, sound sources are perceived inside the head. While it is possible to simulate how sound waves physically propagate, scatter and diffract in an environment, this requires significant computational resources. In many cases, it is possible, and indeed desirable, to simplify the simulation and rendering of room acoustics by leveraging limitations of human auditory perception. This tutorial will provide an overview of the available classes of room acoustics models with a focus on models with low computational requirements that are particularly suitable for XR applications.

Workshop Session 2: **Psychoacoustics and Sound Zones**

back to top

Room: AVAR Workshops

2:00 - 2:55 Monday, August 17

Brian Schmidt, *GameSoundCon, DigiPen*

TITLE: **Beyond Physics: Limitations of Human Hearing and How Source Content and Human Expectations Affect Sound Localization**

By now, we're all familiar with ITD, ILD, HRTF and Reflections (the "Big 4") and the roles they play in our perception of sound in 3D space. However, human sound localization is a complex phenomenon, relying on more than just these physical properties. This workshop will look at human hearing accuracy and examine factors that affect 3D sound perception beyond the "Big 4."

3:00 - 4:00 Monday, August 17

Edgar Choueiri, *Princeton University, Bacch Labs*

TITLE: **Personal Sound Zone: Review of the State of the Art**

We will review the current state of the art of Personal Sound Zone (PSZ) research, where the goal is to create isolated audio sound zones in which separate audio programs (or relative silence) can be delivered to the intended listener in each zone with minimum interference between the zones. We will present an overview of the main promising methods followed so far (Acoustic Contrast Control methods, and Pressure Matching methods), and present a general formalism that covers these methods and that can be used to describe the differences between them. We will focus on the remaining challenges and some of the approaches for meeting these challenges.