Monday 8/20 Wednesday 8/22 back to program

Schedule for Tuesday, August 21

Time	Room	Plenary Session 2
9:00 - 9:10	Plato	Opening Remarks, AVAR Conference Organizing Committee
9:10 - 10:10	Plato	Keynote Speaker: Ivan Tashev, Microsoft Research *Capture, representation, and rendering of 3D audio for VR/AR*
10:15 - 11:00	Plato	Panel Discussion: *Personalized vs. Generic HRTFs for VR/AR audio*
11:00 - 11:25	Cafe	Coffee Break

Time	Room	Paper Sessions	Time	Room	Workshop Sessions
11:30 - 1:00	Michelangelo	HRTF Personalization	11:30 - 1:30	Plato	HRTF and Ambisonics
1:00 - 2:00	Cafe	Lunch 1	1:30 - 2:30	Cafe	Lunch 2
2:00 - 3:30	Michelangelo	Ambisonics	2:30 - 3:30	Plato	XR Audio in Diverse Industries
3:30 - 4:00	Cafe	Coffee Break			
4:00 - 5:30	Michelangelo	Binaural Rendering of 3D Sound Fields	4:00 - 5:00	Plato	XR Audio and the Music World

Time	Place	Event
5:15, 5:45	DigiPen Front Entrance	Bus rides to Hollywood Schoolhouse
6:00 - 7:00	Hollywood Schoolhouse and Wine Cellars	Mixer with Bar and Jazz
7:00 - 10:00		AVAR 2018 Banquet

Special thanks to our sponsors!

Plenary Session for Tuesday, August 21

back to top

9:00 - 9:10 Opening Remarks, AVAR Conference Organizing Committee

9:10 - 10:10 Keynote Speaker: Ivan Tashev, Microsoft Research (speaker bio)

Capture, representation, and rendering of 3D audio for VR/AR

The design of virtual and augmented reality devices requires realistic rendering of 3D audio, which creates the need for technologies and algorithms in the entire chain of capture, processing, and representation of spatial audio. In this talk, we will discuss techniques for capturing 3D audio using specialized microphone arrays and its representation with high order ambisonics or as sound objects and diffuse component. Rendering of spatial audio can be done using headphones with generic or personalized HRTF's, or through a set of loudspeakers. We will discuss the advantages and disadvantages of both approaches and share some of our experience in generating personalized HRTF's.

10:15 - 11:00 Panel Discussion:

Personalized vs. Generic HRTFs for VR/AR audio

Rendering spatial audio through headphones requires using Head-Related Transfer Functions (HRTFs). Using generic HRTFs provides good experience for people with similar spatial hearing, but causes front-back and up-down confusion in people with different head size/pinna shape/ear positions. The way to increase the percentage of satisfied listeners has been to use personalization of the HRTFs. This introduces additional complexity and questions about how far we should go, and whether the extra efforts/price/complexity pay off. Can we make improved HRTFs that suit a higher percentage of listeners? Can we make HRTFs that are perceived to be better than even one's own HRTFs? Which approach is better for augmented reality scenarios, and which one for virtual reality scenarios? These and more related issues will be discussed during the panel with the help of conference attendees.

Panel Members:

- **Ivan Tashev,** *Partner Architect, Microsoft Research*
- **Edgar Choueiri (Moderator),** *Professor of Applied Physics, Princeton University*
- **Ramani Duraiswami,** *Professor, University of Maryland*
- **Hannes Gamper,** *Audio and Acoustics Research Group, Microsoft Research*
- **Agnieszka Roginska,** *Associate Professor of Music Technology, New York University*
- **Nicolas Tsingos,** *Director, Sound Technology Research, Dolby Laboratories*

About the Keynote Speaker:

back to top

Ivan Tashev, *Partner Architect, Microsoft Research*

Dr. Ivan Tashev received his Masters degree in Electronic Engineering (1984) and PhD in Computer Science (1990) from the Technical University of Sofia, Bulgaria. He was Assistant Professor in the same university when he joined Microsoft in 1998. Currently Dr. Tashev is a Partner Architect and leads the Audio and Acoustics Research Group in Microsoft Research Labs in Redmond, WA, USA. He has published two books as a sole author and has written chapters in two other books, authored more than 100 scientific papers, and is listed as inventor of 40 granted US patents. Dr. Tashev created audio processing technologies incorporated in Windows, Microsoft Auto Platform, and the Round Table device. He served as the audio architect for Kinect for Xbox and for Microsoft HoloLens. His team designed the algorithms for all four audio subsystems in HoloLens. Ivan Tashev is a member of AES and served as AES Pacific Northwest Committee member (2006-2016), is a senior member of IEEE Signal Processing Society, and serves as member or affiliated member of the Audio and Acoustics Technical Committee and Industrial DSP Standing Committee of this society. Dr. Tashev is affiliate professor in the Department of Electrical Engineering of the University of Washington in Seattle.

About the Panelists:

back to top

Edgar Choueiri (Moderator), *Professor of Applied Physics, Princeton University*

Dr. Edgar Choueiri, is a professor of applied physics at the Mechanical and Aerospace Engineering Department of Princeton University, and Associated Faculty at the Department of Astrophysical Sciences, Program in Plasma Physics. He is also Director of Princeton University's Engineering Physics Program and Chief Scientist at the university's Electric Propulsion and Plasma Dynamics Lab, a recognized center of excellence in research in the field of advanced spacecraft propulsion. He is also the director of Princeton's 3D Audio and Applied Acoustics (3D3A) Lab.

Professor Choueiri is a world-renowned leader in the field of plasma physics and plasma propulsion for spacecraft. He is the author of more than 150 scientific publications, and encyclopedia articles on plasma rockets, plasma physics, instabilities and turbulence in collisional plasmas, plasma accelerator modeling, space physics and applied mathematics. He has been the Principle Investigator (PI) ion more than 30 competitively selected research projects (including two space experiments), funded by NASA, the US Air Force, the National Science Foundation, and other governmental and private institutions. He is Fellow of the American Institute of Aeronautics and Astronautics and the recipient of many awards and honors including a knighthood.

An avid audiophile, acoustician and classical music recordist, his decades-long passion for perfecting the realism of music reproduction has led him to work on the difficult fundamental problem of designing advanced digital filters that allow the natural 3D audio to be extracted from stereo sound played through two loudspeakers, without adding any spectral coloration to the sound (i.e. without changing its tonal character). He was able to solve this problem mathematically by applying analytical and mathematical tools he uses in his plasma physics research.

Ramani Duraiswami, *Professor, University of Maryland*

Ramani Duraiswami is founder and President of VisiSonics (www.visisonics.com), a startup devoted to spatial audio. He is also a Professor in the Department of Computer Science at the University of Maryland, where he also directs the Perceptual Interfaces and Reality Laboratory. He has broad research interests, including spatial audio, computer vision, machine learning and scientific computing. He has a Ph.D. from Johns Hopkins and a B.Tech. from IIT Bombay. See www.umiacs.umd.edu/ ramani for more on his research.

Hannes Gamper, *Audio and Acoustics Research Group, Microsoft Research*

Hannes Gamper is a researcher in the Audio and Acoustics Research Group at MSR Redmond. He received his PhD in Media Technology/Computer Science from Aalto University in Helsinki, Finland in 2014, and a Master's degree in Sound Engineering from the University of Technology and the University of Music and Performing Arts in Graz, Austria, in 2010.

Agnieszka Roginska, *Associate Professor of Music Technology, New York University*

Dr. Agnieszka Roginska is Music Associate Professor, and Associate Director, of Music Technology at the Steinhardt School, at New York University. In 1996 she received a Bachelors degree in music from McGill University in Montreal, Canada, with a double major in Piano Performance and Computer Applications in Music. After receiving an M.M. in Music Technology from New York University in 1998, she pursued doctoral studies at Northwestern University where she obtained a Ph.D. in 2004. At NYU, Dr. Roginska conducts research in the simulation and applications of immersive and 3D audio including the capture, analysis and synthesis of auditory environments, auditory displays and its applications in augmented acoustic sensing. She is the author of numerous publications in the areas of the acoustics and psychoacoustic of spatial sound, immersive audio and auditory displays, and is the co-editor of the book titled "Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio". Dr. Roginska is an AES Fellow and Governor, and is the faculty sponsor of the Society for Women in TeCHnology (SWiTCH) at NYU.

Nicolas Tsingos, *Director, Sound Technology Research, Dolby Laboratories*

Nicolas Tsingos leads the virtual and augmented reality exploration group at Dolby Laboratories. Previously, he designed and prototyped the authoring and rendering tools for the Dolby Atmos cinema technology and the spatial coding tools that brought Dolby Atmos to the home. He also developed audio rendering technology for game engines and for Dolby Axon, a massively multiplayer voice chat system. In addition to his current position at Dolby Laboratories, Nicolas holds a tenure research position at INRIA the French National Institute for Computer Science and previously was a member of the technical staff at Bell Laboratories. He received a MSc and PhD in computer science from the Joseph Fourier University in Grenoble, France and a Habilitation/accreditation to supervise research (HDR) in computer science from Univ. of Nice/Sophia Antipolis, France.

Paper Session 4

back to top

HRTF Personalization

Room: Michelangelo

Session Chair: Matt Klassen

11:30 - 11:55 Tuesday, August 21

AUTHORS: **David Poirier-Quinot and Brian Katz,** *Sorbonne Université, CNRS, Institut Jean Le Rond d'Alembert, France*

TITLE: **Impact of HRTF individualization on player performance in a VR shooter game II**

We present the extended results of a previous experiment to assess the impact of individualized binaural rendering on player performance in the context of a VR "shooter game". Participants played a game in which they were faced with successive enemy targets approaching from random directions on a sphere. Audio-visual cues allowed for target localization. Participants were equipped with an Oculus CV1-HMD, headphones, and two Oculus Touch hand tracked devices as targeting mechanisms. Participants performed six sessions alternatively using their best and worst-match HRTFs from a "perceptually orthogonal" optimized set of 7 HRTFs [Katz2012]. Results suggest that the impact of the HRTF on participant performance (speed and movement efficiency) depends both on participant sensitivity and HRTF presentation order.

12:00 - 12:25 Tuesday, August 21

AUTHORS: **Rishi Shukla, Rebecca Stewart, Mark Sandler,** *Queen Mary University of London, UK*, **Agnieszka Roginska,** *New York Univeristy*

TITLE: **User selection of optimal HRTF sets via holistic comparative evaluation**

If well-matched to a given listener, head-related transfer functions (HRTFs) that have not been individually measured can still present relatively effective auditory scenes compared to renderings from individualised HRTF sets. We present and assess a system for HRTF selection that relies on holistic judgements of users to identify their optimal match through a series of pairwise adversarial comparisons. The mechanism resulted in clear preference for a single HRTF set in a majority of cases. Where this did not occur, randomised selection between equally judged HRTFs did not significantly impact user performance in a subsequent listening task. This approach is shown to be equally effective for both novice and expert listeners in selecting their preferred HRTF set.

12:30 - 12:55 Tuesday, August 21

AUTHORS: **Ramani Duraiswami, Nail Gumerov, Adam O'Donovan and Dmitry Zotkin,** *VisiSonics Corporation*, **Justin Shen, Matthias Zwicker,** *University of Maryland*

TITLE: **Large Scale HRTF personalization**

Many applications in creating relistic audio for augmented and virtual reality require individual head-related transfer functions. Typically individual HRTFs are measured typically via a slow and relatively tedious procedure, making them unusable in practical applications. We discuss two approaches that allow the creation of HRTFs at large scale for applications. The first is a fast reciprocal measurement approach which allows fast measurement. We have applied the technique to measure the HRTF of several individuals. The second approach is based upon computation of the HRTF in the cloud using meshes in a few minutes. We have shown this approach works with 3D scans, and are now extending it to photograph based processing.

Paper Session 5

back to top

Ambisonics

Room: Michelangelo

Session Chair: Joseph Tylka

2:00 - 2:25 Tuesday, August 21

AUTHORS: **Fernando Lopez-Lezcano,** *CCRMA/Stanford University*

TITLE: **The *SpHEAR project update: the TinySpHEAR and Octathingy soundfield microphones**

This paper in an update of the *SpHEAR (Spherical Harmonics Ear) project, created with the goal of using low cost 3D printers to fabricate Ambisonics microphones. The initial four-capsule prototypes reported in 2016 have evolved into a family of full-featured high quality microphones that include the traditional tetrahedral design and a more advanced eight capsule microphone that can capture second-order soundfields. The project includes all mechanical 3d models and electrical designs, as well as all the procedures and software needed to calibrate the microphones for best performance. A fully-automated robotic arm measurement rig is also described. Everything in the project is shared through GPL/CC licenses, uses Free Software components, and is available on a public GIT repository (https://cm-gitlab.stanford.edu/ambisonics/SpHEAR/)

2:30 - 2:55 Tuesday, August 21

AUTHORS: **Michael Goodwin,** *Xperi, Inc / DTS*

TITLE: **A hybrid beamforming framework for B-format encoding with arbitrary microphone arrays**

In this paper we consider the problem of B-format encoding of live audio scenes captured using practical compact microphone arrays with unconstrained microphone locations. We formulate the problem and provide a mathematical framework for a hybrid adaptive beamformer which combines active beamforming based on frequency-domain spatial analysis and synthesis for estimated directional audio sources and least-squares passive beamforming for residual audio scene components. The necessary calibration measurements are described and algorithmic details for efficient implementation are given.

3:00 - 3:25 Tuesday, August 21

AUTHORS: **Calum Armstrong, Damian Murphy and Gavin Kearney,** *University of York, UK*

TITLE: **A Bi-RADIAL Approach to Ambisonics**

This paper introduces Binaural Rendering of Audio by Duplex Independent Auralised Listening (Bi-RADIAL), a new scheme for the reproduction of 3D sound over headphones. Principles, considerations and methods of Bi-RADIAL decoding and their application within the binaural rendering of Virtual Ambisonics is discussed. Three methods of delivering Bi-RADIAL Ambisonics are compared and the advantages of exploiting a Bi-RADIAL scheme over traditional binaural-based Ambisonic playback are highlighted. Simulation results for standard and Bi-RADIAL Ambisonic decoders are given using a perceptual based comparison of their frequency spectra. Analysis is made for 1st, 3rd and 5th order decoders considering both basic and maxrE weighting schemes. Results show a clear improvement in spectral reconstruction when using a Bi-RADIAL decoder

Paper Session 6

back to top

Binaural Rendering of 3D Sound Fields

Room: Michelangelo

Session Chair: Agnieszka Roginska

4:00 - 4:25 Tuesday, August 21

AUTHORS: **César Salvador, Shuichi Sakamoto, Jorge Trevino and Yôiti Suzuki,** *Tohoku University, Japan*

TITLE: **Enhancing binaural reconstruction from rigid circular microphone array recordings by using virtual microphones**

Spatially accurate binaural reconstruction from rigid circular arrays requires a large number of microphones. However, physically adding microphones to available arrays is not always feasible. In environments such as conference rooms or concert halls, prior knowledge regarding source positions allows for the prediction of pressure signals at positions without microphones. Prediction is performed by relying on a physical model for the acoustically rigid sphere. Recently, we used this model to formulate a surface pressure interpolation method for virtual microphone generation. In this study, we use virtual microphones to enhance the high-frequency spatial accuracy of binaural reconstruction. Numerical experiments in anechoic and reverberant conditions demonstrate that adding virtual microphones extends the frequency range of operation and attenuates the time-domain artifacts.

4:30 - 4:55 Tuesday, August 21

AUTHORS: **Axel Plinge,** *Frahunhofer IIS, Germany*, **Sebastian Schlecht, Oliver Thiergart, Thomas Robotham, Olli Rummukainen and Emanuel Habets,** *International Audio Laboratories, Erlangen, Germany*

TITLE: **Six-Degrees-of-Freedom Binaural Audio Reproduction of First-Order Ambisonics with Distance Information**

First-order Ambisonics (FOA) recordings can be processed and reproduced over headphones. They can be rotated to account for the listener's head orientation. However, virtual reality (VR) systems allow the listener to move in six-degrees-of-freedom (6DoF), i.e., three rotational plus three transitional degrees of freedom. Here, the apparent angles and distances of the sound sources depend on the listener's position. We propose a technique to facilitate 6DoF. In particular, a FOA recording is described using a parametric model, which is modified based on the listener's position and information about the distances to the sources. We evaluate our method by a listening test, comparing different binaural renderings of a synthetic sound scene in which the listener can move freely.

5:00 - 5:25 Tuesday, August 21

AUTHORS: **Thomas McKenzie, Damian Murphy and Gavin Kearney,** *University of York, UK*

TITLE: **Directional Bias Equalisation of First-Order Binaural Ambisonic Rendering**

The human auditory system is more accurate at localising sounds in front than at lateral, rear and elevated directions. In virtual reality applications, where Ambisonic audio is presented to the user binaurally (over headphones) in conjunction with a head-mounted display, it is imperative that audio in the frontal direction is as accurate as possible. This paper presents a method for improving frontal high frequency reproduction of binaural Ambisonic rendering through a novel adaptation of the diffuse-field equalisation technique to exploit the non-uniform directional sensitivity of human hearing. The method is evaluated via spectral difference and a height localisation model, and results show improved frontal reproduction at the expense of lateral fidelity.

Workshop 6 & 7

back to top

HRTF and Ambisonics

Room: Plato

11:30 - 12:25 Tuesday, August 21

Michel Henein, *VisiSonics*

TITLE: **The Basics of HRTF Personalization**

VR/AR strives to convince the human brain that the virtual world is real. Since sound in the real world is critical to the human experience, virtual sound, based on HRTFs, must be convincing. Due to anthropometric differences between individuals, a generic HRTF set does not provide convincing 3D effect for all humans, which is a major problem for mass VR/AR adoption. Research conducted by the VisiSonics team, first at the University of Maryland, and then at the company for its RealSpace 3D engine has resulted in a few techniques enabling individualized HRTFs to be created quickly, thus helping to provide the most accurate spatial audio experience – as good as being there – for all VR/AR users.

12:35 - 1:30 Tuesday, August 21

Markus Zaunschirm, Daniel Rudrich and Christian Schörkhuber, *IEM, University of Music and Performing Arts, Graz, Austria*

TITLE: **Binaural Ambisonics rendering for headphones**

With the widespread adoption of the Ambisonics format for headphone-based playback of spatial audio the binaural rendering of order-limited Ambisonic signals became of great interest in the fields of virtual reality and 360-degree multimedia productions. Classic binaural renderers adopt decoding strategies from loudspeaker-based rendering (e.g. virtual loudspeaker approach). They either yield strongly direction-dependent coloration (dull in front) when using a dense set of virtual loudspeakers, or they yield strongly sampling-dependent spatial cues and spatial aliasing with a sparse layout which only uses those few virtual loudspeaker required for the Ambisonic order. This workshop compares classic binaural decoding with new cutting-edge binaural Ambisonic renderers. We discuss differences in coloration and localization, and present approaches for improving externalization.

Workshop 8

back to top

XR Audio in Diverse Industries

Room: Plato

2:30 - 3:25 Tuesday, August 21

Joel Douek and Benedict Green, *EccoVR, Inc.*

TITLE: **Riding the XR Sound Wave into Diverse Industries**

The nascent technologies of Virtual, Augmented and Mixed reality, although originally driven from the world of entertainment and video games, are extending into an untold number of industries. What this means for anyone working in XR, including on the audio side, is that all these industries are becoming part of our world by extension. It means that career opportunities are multiplying manifold in fascinating new areas where sound never played an obvious role. This expansion of reach has real implications for our professional lives, not only in terms of job opportunities and earning potential, but also the possibility of being at the forefront of new technological and creative developments that will influence the shape of the future.

Workshop 9

back to top

XR Audio and the Music World

Room: Plato

4:00 - 4:55 Tuesday, August 21

Joel Douek, *EccoVR, Inc.*

TITLE: **Can XR Rekindle the Music World?**

Against the often bleak picture of music industry and the diminishing value of music, XR offers a ray of hope. With the personal presence, proximity and free agency the VR listener has, a sense of intimacy, wonderment and privilege can be restored to the musical experience. That magical connection can be regained as we stand (albeit virtually) in the presence of our musical heroes or be present at the birth of a song. AR, MR & Social VR can further enrich musical experiences with opportunities for interaction and deeper information. This talk will showcase some recent music-driven XR experiences and explore the future potential of these new platforms and technologies to breath life into the creation and appreciation of music.

AVAR 2018 SPONSORS

Microsoft

facebook
Reality Labs

amazon devices

VALVE

magic leap