# AES 150TH CONVENTION PROGRAM

## MAY 25–28, 2021, STREAMCAST ONLINE

## "GLOBAL RESONANCE"

*AES 150th Convention, "Global Resonance," offered 72-plus hours of scheduled streaming (Stream A and Stream B) content along with a slate of over 60 on-demand Paper and Engineering Brief sessions (presentations were not livestream but were available on demand for the four days). The following is a listing of the events presented. the full schedule can be found at aeseurope.com/program. Default times are Central European Summer Time*

**The Winner of the 150th AES Convention Best Paper Award**

**Delivering Personalized 3D Audio to Multiple Listeners: Determining the Perceptual Trade-Off Between Acoustic Contrast and Cross-Talk**—*Coleman, Philip, Natasha Canter,* Institute of Sound Recording, University of Surrey, UK

*Convention Paper 10452*

* * * * *

The AES has launched an opportunity to recognize student members who author technical papers. The Student Paper Award Competition is based on the preprint manuscripts accepted for the AES convention.

A number of student-authored papers were nominated. The excellent quality of the submissions has made the selection process both challenging and exhilarating.

The award-winning student paper will be honored during the Convention, and the student-authored manuscript will be considered for publication in a timely manner for the *Journal of the Audio Engineering Society*.

Nominees for the Student Paper Award were required to meet the following qualifications:

(a) The paper was accepted for presentation at the AES 150th Convention.

(b) The first author was a student when the work was conducted and the manuscript prepared.

(c) The student author's affiliation listed in the manuscript is an accredited educational institution.

(d) The student will deliver the lecture or poster presentation at the Convention.

* * * * *

**The Winner of the 150th AES Convention Student Paper Award is:**

**Room Geometry Estimation from Higher-Order Ambisonics Signals Using Convolutional Recurrent Neural Networks**—*Nils Poschadel, Robert Hupke, Stephan Preihs, Jürgen Peissig*, Leibniz University Hannover, Institute of Communications Technology, Hannover, Germany

*Convention Paper Paper 10482*

* * * * *

**SPECIAL EVENT**
**Opening Ceremony & Awards**
**Tuesday, May 25, 11:00 – 11:30 am**

Presenters:
**Colleen Harper**, AES Executive Director
**Jonathan Wyner**, AES President
**Ruud Kaltofen**, Spring 2021 Co-Chair
**Bert Kraaijpoel**, Spring 2021 Co-Chair
**Jamie Angus-Whiteoak**, Spring 2021 Papers Co-Chair
**Remy Wenmaekers**, Spring 2021 Papers Co-Chair

**Awards Presentation**

Please join us as the AEs presents special Awards to those who have made outstanding contributions to the Society in such areas of research, scholarship, and publications, as well as other accomplishments that have contributed to the enhancement of out industry. The awardees :are:

**Board of Governors Award**

- Eddy Bøgh Brixen
- Patrick Hegarty
- Matt Klassen
- César Lamschtein
- Piotr Majdak
- Agnieszka Roginska
- Lawrence Schwedler
- Christopher J. Struck
- Alfred J. Svobodnik
- Nadja Wallaszkovits
- Jonathan Wyner

**Fellowship Award**
- Dan Dugan
- Bill Foster
- Richard King

**Distinguished Service Award**
- Bozena Kostek

**PARTNER PRESENTATION: GENELEC**
**Tuesday, May 25, 11:00 – 11:30 am**
**How to Turn Your Spare Room into a Control Room**

Presenters: **Andy Bensley**, Genelec
**Darren Rose**, Genelec

**In this webinar Genelec's Andy Bensley and Darren Rose discuss how Andy converted an "echoey, fluttery" spare room in his new house into a real world studio space.**

Andy's tiny studio—a mere 2.4m x 3m in size—will be a familiar sight to anyone who's watched recent #GenelecProAtHome Instagram Live events. In this session Andy explains how he gradually tackled the acoustic issues in the room, using a combination of acoustic treatment, Genelec Smart Active Monitoring, and GLM calibration software to achieve a truly accurate sonic reference—without breaking the bank!

At the end of the session, Andy and Darren will be answering your questions live. If you're an engineer or creative looking to get the most from your home studio space, this is essential viewing!

**WORKSHOP/TUTORIAL—AUTOMOTIVE AUDIO**
**Tuesday, May 25, 12:00 noon – 1:00 pm**
**Our Roads to the Automotive Industry**

Presenters: **Debrupa Chakraborty**, Fraunhofer IIS
**Eva Hasenburger**, Fraunhoffer IIS
**Sharah Hölscher**, Fraunhofer IIS
**Francesca Nucibella**, Acoustic Systems Architect, Harman
**Hope Sheffield**, Acoustic Engineer, Harman International
**Shelley Uprichard**, Danish Sound

Five female professionals working in the automotive sector get together to present themselves and discuss their experience and the opportunities they've had within automotive. This is an international panel of very talented women with wide-ranging and unique backgrounds . . . which makes for an interesting panel discussion!

**WORKSHOP/TUTORIAL—RECORDING AND PRODUCTION**
**Tuesday, May 25, 12:00 noon – 1:00 pm**
**Deceptive Tracks**

Presenters: **Jim Austin**, Editor, Stereophile
**Thomas Lund**, Senior Scientist, Genelec OY
**Darcy Proper**, Proper Prent Sound LLC

Popular music tracks on 30 year old CDs or vinyl often sound better off the physical medium than if they are streamed, even using one of today's high resolution providers. We consider determing factors in such hidden decline, from track versions over ripping and quality control, to the codecs used for delivery.

Besides from providing technical evidence and listening examples, we discuss content uncertainties from a mastering and a music lover's perspective, including the excellent music listening experience with its variabilities 300 years ago, compared to now.

**PAPER Q & A SESSION: HRTF**

**Tuesday, May 25, 12:15 pm**

- **Individualized HRTF-Based Binaural Renderer for Higher-Order Ambisonics**—*Mengfan Zhang,*[1] *Tianyi Guan,*[1] *Lianwu Chen,*[2] *Tianxiao Fu,*[2] *Dan Su,*[2] *Tianshu Qu*[1]
[1] Key Laboratory on Machine Perception (Ministry of Education), Speech and Hearing Research Center, Peking University, China
[2] Tencent AI Lab, Shenzhen, China

Ambisonics is a promising spatial sound technique in augmented and virtual reality. In our previous study we modeled the individual head-related transfer functions (HRTFs) using deep neural networks based on spatial principal component analysis. This paper proposes an individualized HRTF-based binaural renderer for the higher-order Ambisonics. The binaural renderer is implemented by filtering the virtual loudspeaker signals using individualized HRTFs. We perform subjective experiments to evaluate generic and individualized binaural renderers. Results show that the individualized binaural renderer has front-back confusion rates that are significantly lower than those of the generic binaural renderer. Therefore, we validate that using individualized HRTFs to convolve with those virtual loudspeaker signals to generate virtual sound at an arbitrary spatial direction still performs better than those using generic HRTFs. In addition, by measuring or modeling individual's HRTFs in a small set of directions, our proposed binaural renderer system effectively predict individual's HRTFs in arbitrary spatial directions.
*Paper 10454*

- **Listener-Position Sub-Band Adaptive Crosstalk Canceller Using HRTFs Interpolation for Immersive Audio Systems**—*Valeria Bruschi,*[1] *Stefano Nobili,*[1] *Ferruccio Bettarelli,*[2] *Stefania Cecchi*[1]
[1] Università Politecnica delle Marche, Ancona, Italy
[2] Leaff Engineering Srl, Porto Potenza Picena (MC), Italy

The crosstalk is an undesired phenomenon that occurs in immersive systems when the sound is reproduced over two loudspeakers. In this paper an innovative sub-band adaptive crosstalk canceller considering the listener position is proposed. Starting from a reduced set of measured head related impulse responses (HRTFs), the proposed system is composed of three parts: a listener position tracker, an HRTFs interpolation algorithm, and a sub-band adaptive crosstalk canceller. In particular, the head tracking allows to obtain the listener position and the interpolator is applied to interpolate the HRTFs related to positions different from measurement points. The effectiveness of the proposed approach has been confirmed through several real-world experimental tests.
*Paper 10474*

- **Evaluation of Pinna Point Cloud Alignment by Means of Non-Rigid Registration Algorithms**—*Katharina Pollack,*[1] *Piotr Majdak,*[1] *Hugo Furtado*[2]
[1] Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria
[2] Dreamwaves GmbH, Vienna, Austria

Recently, there have been advances towards providing personalized binaural audio by means of numerical calculations based on a 3D head mesh. Nevertheless, 3D head meshes obtained by non-contact methods such as photogrammetric reconstruction have difficulties in accurately representing the complex biological structure of the pinna. The problems manifest in noisy meshes with

holes, confounding the quality of the calculations. These problems can be tackled by applying a template mesh that is deformed to the personalized but noisy target mesh. In this contribution, we review two state-of-the-art non-rigid registration algorithms applied to the deformation of two high-resolution, high-accuracy pinna point clouds. We evaluate the algorithms by discussing their geometric errors, duration of the alignment task, and their robustness to spatial down-sampling.
*Paper 10475*

• **Global HRTF Personalization Using Anthropometric Measures—***Yuxiang Wang, You Zhang, Zhiyao Duan, Mark Bocko*, University of Rochester, Rochester, NY, USA

In this paper we propose an approach for global HRTF personalization employing subjects' anthropometric features using spherical harmonics transform (SHT) and convolutional neural network (CNN). Existing methods employ different models for each elevation, which fails to take advantage of the underlying common features of the full set of HRTF's. Using the HUTUBS HRTF database as our training set, a SHT was used to produce subjects' personalized HRTF's for all spatial directions using a single model. The resulting predicted HRTFs have a log-spectral distortion (LSD) level of 3.81 dB in comparison to the SHT reconstructed HRTFs and 4.74 dB in comparison to the measured HRTFs. The personalized HRTFs show significant improvement upon the finite element acoustic computations of HRTFs provided in the HUTUBS database.
*Paper 10502*

**SPECIAL EVENTS: KEYNOTE ADDRESS**
**Loudness Normalization, A Personal History**
May 25, 1:00 pm – 2:00 pm

Presenter: **Eelco Grimm**, HKU University of the Arts Utrecht

For almost 17 years, Eelco Grimm has worked on loudness normalization in international broadcast, cinema, and audio streaming. He witnessed one of the greatest revolutions in the audio world from the front row, by taking part in many of the committees involved such as ITU srg3, EBU PLOUD, AES TC-AC, CTA r4wg8 and AES sc-02-12-Q.

In this Keynote he takes us on a personal journey, reflecting on why he was so driven by this topic that he spent a large part of his professional career on it.

**PAPER Q & A SESSION: ACOUSTIC MEASUREMENT 1**
**Tuesday, May 25, 2:00 pm**

• **Real-Time Impulse Response Measurement: Swept-Sine Technique with Minimum Latency for Reflection and Position Fensing—** *Felix Pfreundtner*, Department of Computer Science, ETH Zurich, Zurich, Switzerland; Xilev UG, Schwabach, Germany

State-of-the art impulse response measurement technique imposes a measurement latency that constrains real-time application. This paper proposes a first swept-sine technique that can update subsequent impulse responses within a latency of one recording sample. We verify the technique for various LTI and non-LTI measurement setups by driving a mannequin through a test room. The extracted impulse responses correspond widely to conventional swept-sine technique, while we achieve an eight times higher update rate with our implementation on a conventional laptop. The observed mean error for the LTI case is –60 dB, and builds up to –35 dB with increasing LTI violation. We anticipate, the technique can significantly improve the resolution of ultrasound sensing, monitor quick changes in reflective environment or extract impulse responses for augmented reality applications at immersive update rate.
*Paper 10449*

• **Spatial-Temporal Correlation Based Signals Gathering in WASNs—***Xinwei Guo, Ruijie Meng, Chengshi Zheng, Xiaodong Li*, Chinese Academy of Sciences, Beijing, China

For each node comprising many channels in the wireless acoustic sensor networks (WASNs), the signals from different channels both have the spatial-temporal correlation property. The two-dimensional discrete cosine transform (2D-DCT) is introduced to decorrelate and compute the sparse representation of the signals. Then, only the first K largest 2D-DCT coefficients and the binary sequences consisting of 0 and 1 to indicate the indices of these coefficients are transmitted. The proposed method reduces the number of channels transmitted per node and yet the raw signals can still be recovered accurately at the fusion center. The proposed method is independent of the specific beamformer compared to other distributed algorithms to reduce the communication overhead, and the experimental results demonstrate its validity.
*Paper 10469*

• **A Steered-Beamforming Method for Low-Latency Direction-of-Arrival Estimation in Reverberant Environments Using Spherical Microphone Arrays—** *Jonathan Mathews, Jonas Braasch*, Rensselaer Polytechnic Institute, Troy, NY, USA

This paper introduces a method to estimate the direction of arrival of an acoustic signal based on finding maximum power in iteratively reduced regions of a spherical surface. A plane wave decomposition beamformer is used to produce power estimates at sparsely distributed points on the sphere. Iterating beam orientation based on the orientation of maximum energy produces accurate localization results. The method is tested using varying reverberation times, source-receiver distances, and angular separation of multiple sources and compared against a pseudo-intensity vector estimator. Results demonstrate that this method is suitable for integration into real-time telematic frameworks, especially in reverberant conditions.
*Paper 10493*

• **Direction of Arrival Estimation for Acoustical Sources Using Noise Signals and a Uniform Circular Array—** *Yohann Simon,*[1] *Thierry Heeb*[2]
[1] Storm Audio, Saint Héblain,France
[2] SUPSI-ISIN,Lugano-Viganello, Switzerland

This paper describes an effective and fast solution for estimating speaker direction of arrival in 3D space. The main idea is to use noise as an excitation signal for sources and to determine time of arrivals from recordings of 3D space distributed microphones. Up-sampling and interpolation are applied to improve times of arrivals estimation, which are then used to construct sinusoidal signals. Sine phases are determined from times of arrival estimations and the frequency is optimized from the physical sensors' structure. Hence, speaker direction are predicted by beamforming and direction of arrival algorithms with a precision of

two degrees in azimuth and five in elevation.
*Ebrief 639*

**WORKSHOP/TUTORIAL—ACOUSTICS & PSYCHOACOUSTICS**
**Tuesday, May 25, 2:00 pm – 3:00 pm**
**Perception of Early Reflections in Small Rooms—Psychoacoustic Requiremens for VR/AR Systems with 6DOF**

Presenter: **Annika Neidhardt**, Technische Universität Ilmenau

When a listener walks through a room the spatial-temporal pattern of the early reflections arriving at his ears changes alongside the relative direct sound. Not all of the physical details can be perceived by human listeners.

An in-depth understanding on how a moving listener perceives the early reflections in a room will help to improve the efficiency of auditory AR/VR systems with six degrees-of-freedom

Motivated by the described goal this tutorial will review the literature as well as selected previous studies by the author. The perception of early reflections has been studied in a variety of fields like the precedence effect, speech intelligibility, spatial impression in concert halls, human echolocation, audible effects of distinct reflections in control rooms as well as the excitation of selected early reflections by directed sound sources like beamformers for a controlled shifting of the apparent source location. What can we learn from the results achieved in the different fields so far and how can this knowledge be used to create efficient VR/AR-systems?

Small rooms like living-rooms, offices or seminar rooms are common environments for using AR/VR applications, but their acoustic behavior has not been studied as intensely as that of concert halls. Therefore the tutorial will focus on the perception of early reflections in small rooms.

**WORKSHOP/TUTORIAL—ELECRONIC DANCE MUSIC**
**Tuesday, May 25, 2:15 pm – 3:15 pm**
**Pushing the Envelope: An Introduction to Live Electronic Performance**

Presenter: **Claire Lim**

Over the last century, the need for contemporary creative expression has pushed artists and technologists to the forefront of building new musical interfaces with electronics, from the Theremin, to the TR-808 drum machine, to a plethora of MIDI controllers today. These instruments not only facilitate revolutionary ways of making music, but offer increased access to production, composition, and performance. The possibilities are endless, so how might one even begin to navigate this evolving musical landscape, where change is the only constant?

This performative presentation will survey the expanding field of live electronic music, providing an introduction to electronic performance techniques and examples of their implementation in various performance contexts. References will be made to Berklee College of Music's Electronic Digital Instrument (EDI) program, the first of its kind at an undergraduate level in the United States, which allows students to choose live electronics as their principal instrument at the institution. We will explore a variety of prompts related to incorporating electronic musicians into existing communities, creating new musical opportunities virtually and in-person through live experience design, and offering new pedagogical approaches through the lens of music technology education. The presenter will also demonstrate how electronic musicians perform with their unique systems via a series of short live examples with multiple electronic instruments, notably Ableton Push, and will discuss considerations for getting started in live electronic music performance.

**PAPER Q & A SESSION: GAMES/INTERACTIVE**
Tuesday, May 25, 3:00 pm– 3:45 pm

- **touchEQ: An Eyes-Free Audio Equalizer for a Surface Haptic Interface**—*Jakub Pesek, Brecht De Man*, Royal Conservatoire The Hague, Den Haag, Netherlands

  Over the past three decades, the process of music production moved from operating analogue devices to using software, which leads to accessibility issues for the visually impaired. This paper reviews accessibility features of popular DAWs and explores the use of a surface haptics interface as an eyes-free controller for an audio equalizer. An application prototype was developed to study an alternative human-computer interaction that could be applied to the enhancement of a visually impaired music producer's workflow. The prototype was tested in two usability studies in order to determine if it can be controlled effectively without visual feedback.
  *Paper 10485*

**WORKSHOP/TUTORIAL—GAME AUDIO AVAR SPATIAL AUDIO**
**Tuesday, May 25, 3:00 pm – 4:15 pm**
**Immersive Audio for Live Events**

Presenters: **Etienne Corteel**, L-Acoustics
**Scott Sugden**, L-Acoustics

The live industry is living a transition from traditional left-right, mostly dual mono, to immersive systems. In this workshop specific challenges of immersive audio for live are presented and addressed. These challenges are mostly related to the scale and diversity of audiences and performance spaces, from pre-production, to touring, to post-production. They can be overcome adopting a full system approach. This approach combines specific tools and guidelines for the design of the loudspeaker system, object oriented mixing tools, and specific 3D audio algorithms for loudspeakers and headphones.

**STUDENT & CAREER DEVELOPMENT EVENT**
**Student Delegate Assembly 1**
**May 25, 3:15 pm – 3:45 pm**

**STUDENT & CAREER DEVELOPMENT EVENT**
**Student Recording Competition: Traditional Studio Recording**
**May 25, 3:45 pm – 4:45 pm**

The AES Student Recording Competition is a unique opportunity for student attendees of AES International Conventions to receive feedback and recognition for their audio production work.

Finalists will be announced and prizes awarded during this presentation. **Judge Panelists include:** *Richard King, Darcy Proper, Cesar Lamschtein, Peter Doell*.

**PAPER Q & A SESSION: REPRODUCTION: 3D AUDIO 1**
**Tuesday, May 25, 4:00 pm**

- **The Development of Dummy Head Microphones since**

**1970**—*Martin Schneider*, Georg Neumann GmbH, Berlin, Germany

Recording with dummy heads and reproducing via headphones is the most straightforward way to create immersive environments. Current dummy heads evolved from room acoustic experiments of the late 1960s. Binaural technology found interest especially in radio drama productions. In the late 1970s diffuse-field equalization of dummy heads was adopted to improve on timbral problems. The current generation was further optimized especially regarding diffuse-field equalization. The KU100 now appears to be a recognized standard for binaural recording applications. As a recording microphone, or via its HRTFs for binaural rendering, it is involved in a large percentage of binaural applications. The talk will delve into the interiors and acoustic differences of the dummy head generations.
*Paper 10500*

• **Comparison of Distortion Products in Headphone Equalization Algorithms for Binaural Synthesis**—*Braxton Boren*,[1] *Michele Geronazzo*[2,3]
[1] American University, Washington, DC, USA
[2] University of Udine, Udine, Italy
[3] Imperial College London, London, UK

Headphone design has traditionally focused on creating a frequency response to make commercial stereo audio sound more natural. However, because of the sensitivity of spatial hearing to frequency-dependent cues, binaural reproduction requires headphones' target spectrum to be as flat as possible. Initial attempts to equalize headphones used a naive inversion of the headphone spectrum, which degraded binaural content because the headphone transfer function (HpTF) changes each time headphones are re-seated. Many different algorithms have been proposed to improve binaural equalization, each of which has been tested over a limited sample of HpTFs. The present study gathered 1550 HpTFs from different institutions into a single dataset for large-scale comparisons of equalization algorithms. A numerical metric was designed to quantify auditory perception of spectral coloration from 'ringing' peaks in the post-equalization HpTF. Using this metric, eight of the most prominent equalization methods have been compared over the aggregate HpTF dataset. High-shelf regularization is shown to outperform all other equalization techniques using either individualized or averaged input spectra. In addition, high-shelf regularization without individual measurements gives less average coloration than direct inversion using individualized equalization.
*Paper 10501*

• **Comparison of Spatialization Techniques with Different Music Genres II**—*Shashank Aswathanarayana*, University of California, Santa Barbara, CA, USA

The main characteristics of a good spatialization algorithm are that the resulting sound image is enveloping, stable, clear. Such a sound image would then not only give a good sense of immersion to the listener but will also enable the listener to clearly localize the different sound elements, and ideally move around or at least have freedom to move the head and not have any major distortions in the listening experience. Part 1 of this study compared three algorithms, Ambisonics, VBAP, and DBAP with four different music genres. In part 2, a more in-depth analysis is done of the results found in part 1. Binaural recordings done using a Neumann KU100 dummy head are used to compare the results found in the loudspeaker study with attributes found in the recordings. It was seen how more bright, next spectrally complex and more spectrally flat signals could have resulted in the user ratings.
*Paper 10504*

• **Reproducing the Auditory Width of Individual Acoustic Instruments in Immersive Audio Productions Using a Three-Channel Spot-Microphone Arrangement—System Description and Anechoic Test Recording**—*Leon Hofann*,[1,2] *Frank Melchior*,[1] *Benhamin Johannes Mueller*[2]
[1] Hochschule der Medien, Stuttgart, Germany
[2] Fraunhofer-Institute for Building Physics, Stuttgart, Germany

This e-Brief proposes a spot-microphone arrangement in the context of immersive audio productions for recording acoustic instruments. The presented technique enables reproduction of a single instrument with control of its auditory width in the horizontal and vertical dimension during post production. The geometrical arrangement as well as the required post processing for immersive loudspeaker-based audio productions is discussed. Additionally, a recording of ten individual instruments and two singers with three spot-microphone positions each was carried out in an anechoic chamber using the new method. The audio files and detailed documentation are provided under CC BY-NC-ND license for evaluation, demonstration and further research. An overview of this database as well as the initial findings gathered throughout mixing the 100+ track production for a 7+4 loudspeaker setup are given.
*Ebrief 641*

**STUDENT & CAREER DEVELOPMENT EVENT**
**Non-Audio skills You Need to Succeed as an Audio Freelancer and Business Owner**
**Tuesday, May 25, 4:15 – 5:45 pm**

Moderator: **Ben Gallagher**, Double Shot Audio

Presenters: *Adele Cutting*, Soundcuts Ltd.
*Melissa Pons*, Field Recordist & Sound Designer
*Martin Rieger,* VRTonung
*GeorgeVlad*, sound Designer & Recordist

With the rise of the gig economy, increasingly more audio professionals find themselves operating as freelancers and/or running their own small businesses.

In order to do this effectively, they need not only to understand and master their craft but also the fundamental business concepts that lead to generating reliable income from these skills.

Our panel understands how intimidating it is to start your own business and how stressful it can be when things aren't working. That's why we have gotten together to discuss the underlying non-audio skills (networking, marketing, soft skills, etc.) that helped us to establish successful audio businesses.

Enjoy a wide-ranging talk and discover through the perspectives of our multi-faceted panel the key business concepts that are often the difference between success and failure for many audio freelancers and business owners.

**WORKSHOP/TUTORIAL: BROADCAST & ONLINE DELIVERY**
**The Technology of Streaming**
**Tuesday, May 25, 4:45 – 6:00 pm**

Presenters: *David Bialik*, Consultant, David K. Bialik & Associates
*Tim Carroll*, Dolby Laboratories
*Scott Kramer*
*Robert Minnix*, StreamGuys, Inc.
*Robert Marshall*
*John Schaab*, Modulation Index, LLC

In this panel a number of industry experts will join in a discussion on streaming technologies: the ways we stream video, audio and data in real-time, the latency, quality and scalability of each of those technologies, the applications those particular methods can be applied to, the financial costs and the effort cost, and the current and future developments of the streaming technology itself.

**PAPER Q & A SESSION: EDUCATION 1**
**Tuesday, May 25, 5:00 pm**

- **Teaching a Practical Audio Degree during a Pandemic—** *Paul Doornbusch, Jason Torrens*, Australian College of the Arts, Melbourne, Victoria, Australia

  The Australian College of the Arts moved an intensely hands-on practical audio production course to online delivery during the 2020 COVID-19 pandemic. Great effort went into ensuring the highest quality in student experience. Furthermore, innovative ways were found to give students hands-on experience even in remote settings. This paper discusses the process and techniques used, as well as the (positive) results than came of this change. Thus, informing the development of a new online audio production course which offered greater access and opportunity for students.
  *Paper 10468*

- **Using Multimedia Immersion to Inspire High School Students to Pursue STEM Careers—***Robert Harari*,[1] *Ann Benbow*[2]
  [1] Stevens Institute of Technology, Hoboken, NJ, USA
  [2] Educational Visions, La Plata, MD, USA

  Multimedia Immersion (MI) Inspires STEM Learning is an exploratory DRK-12 project for the National Science Foundation (NSF) in the learning strand. Arts and Technology faculty at Stevens Institute of Technology
  *Paper 10490*

- **The Pandemic Pilot: Are Music Technology Degrees Ready For Online Learning?—***Eli Farnhill, Malachy Ronan*, Limerick Institute of Technology, Moylish, Limerick, Republic of Ireland

  The COVID-19 pandemic has irrevocably changed the educational landscape, forcing institutes of higher education to participate in a global experiment. The pivot to emergency remote teaching and learning in delivery of Music Technology degree programs necessitated changes, as institutional facilities were inaccessible to students. This paper explores the experience of emergency remote teaching from a faculty perspective through semi-structured interviews with three faculty members of a Music Technology degree program. Thematic analysis on these interviews identified five themes that illustrate obstacles to delivering Music Technology degrees in an online medium: (1) authentic assessment, (2) communication, (3) technology as a tool, (4) social distancing, and (5) sharing resources. These barriers negatively affect the student learning experience. If Music Technology Degree programmes are to be delivered online, modules should be divided into onsite and online delivery formats. Furthermore, the development of novel technology tools that facilitate the needs and interactions of Music Technology degrees would be beneficial.
  *Paper 10496*

**WORKSHOP/TUTORIAL: HISTORICAL**
**Rupert Neve Retrospective—Sound Over Specs**
**Tuesday, May 25, 5:45 pm – 7:00 pm**

Moderator: **Alex U. Case**, University of Massachusetts Lowell

Presenters: *George Massenburg*, McGill University
*Ronald Prent*, Valhalla Studios, NY
*Steve Rosenthal*, MARS (MagicShop Archive and Restoration Studios)
*Josh Thomas*, Rupert Neve Designs, LLC.
*Darrell Thorp*, Freelin Inc

Rupert Neve lived life in pursuit of extended frequency response, low distortion, and low noise in service of sound quality. While measurements informed his work and quantified his achievements, his designs evolved because he listened—to sound and to sound engineers. His first “manufacturing” effort began at the age of 13, building radios and amplifiers while with his missionary parents in Argentina. He moved on to England and built his first large format analog console, and the professional audio industry would never be the same—consoles, compressors, equalizers, and amplifiers from Neve Electronics, Focusrite, and AMS Neve. The journey moved next to Texas where his ideas, motivations, and designs continue in the products of Rupert Neve Designs. Join us for this workshop in which we hear from a range of friends, customers, and colleagues discussing the rich life and sonic legacy of the person whose name is unmistakably associated with the highest of audio quality.

**PAPER Q & A SESSION: PSYCHOLOGY**
**Tuesday, May 25 6:00 pm**

- **Mapping Voice Gender and Emotion to Acoustic Properties of Natural Speech—***Eunmi Oh, Jaeeun Lee, Dayoung Lee*, Yonsei University, Seoul, Korea

  This study is concerned with listener’s natural ability to identify an anonymous speaker’s gender and emotion from voice alone. We attempt to map psychological characteristics of the speaker, such as gender image and emotion, to acoustical properties. The acoustical parameters of voice samples were pitch (mean, maximum, and minimum), pitch variation over time, jitter, shimmer, and Harmonics-to-Noise Ratio (HNR). Participants listened to 2-second voice clips and were asked to rate each voice’s gender image and emotion using a 7-point scale. Emotional responses were obtained for 7 opposite pairs of affective attributes (Goble and Ni Chasaide, 2003). The pairs of affective attributes were relaxed/stressed, content/angry, friendly/hostile, sad/happy, bored/interested, intimate/formal, and timid/confident. Experimental results show that listeners were able to identify voice gender and assess emotional status from short utterances. Statistical analyses revealed that these acoustic parameters were related to listeners’ perception of a voice’s gender image and its affective attributes. For voice gender perception, there were significant correlations with jitter, shimmer, and HNR parameters in addition to pitch parameters. For perception of affective attributes, acoustic parameters were analyzed with respect to the valence-arousal dimen-

sion. Voices perceived as positive tended to have higher variance in pitch and higher maximum pitch than those perceived as negative. Voices perceived as strongly active tended to have higher number of voice breaks, jitter, shimmer, and lower HNR than those perceived as passive. We expect that our experimental results on mapping acoustical parameters with voice gender and emotion perception could be applied to the field of Artificial Intelligence (AI) when assigning specific tone or quality to voice agents. Moreover, such psychoacoustical mapping can improve the naturalness of synthesized speech, especially neural TTS (Text-To-Speech), because it can assist in selecting the appropriate speech database for voice interaction and for situations where certain voice gender and affective expressions are needed.
*Paper 10461*

• **Effect of Pleasurable Musical Chills on Driver's Physiological Response**—*Yuki Kameyama,*[1] *Noriyuki Tanimoto,*[1] *Tenchi Murao,*[1] *Kizuna Sudo,*[1] *Shunsuke Ishimitsu,*[1] *Hirofumi Seni,*[2] *Toshihiro Kikuchi,*[2] *Naoko Takahashi*[2]
[1] Hiroshima City University,Hiroshima, Japan
[2] Mazda Motor Corporation, Hiroshima, Japan

In recent years, cars have become increasingly comfortable and attractive to drivers, allowing them to experience the pleasure and excitement of driving. It is now possible to enjoy high-quality music by improving the quietude of car's interior and the sound quality of car's audio system. However, to date no studies have been reported on the effect of "pleasurable musical chills" experienced while listening to high-quality music on the driver's comfort and driving performance. We focus on these musical chills by examining the drivers' physiological responses and investigating the influence of high-quality audio with various sound presentations. Physiological responses, such as brain activity, breathing, and heartbeat, were examined to understand their relationship with sound quality impressions. In the first experiment, the drivers were made to listen to music while in a resting state. The results indicated that the sympathetic nervous system was activated when participants were listening to music, and that they were in a tense or excited state when they felt musical chills. The second experiment focused on participants listening to music while driving. The results indicated that the sympathetic nervous system was again activated, and the $\beta$ wave was stable when the sound quality was higher. Additionally, the results confirmed that the sympathetic nervous system was activated when the driver experienced the musical chills. The results of the semantic differential method indicated that the positive factors and the "familiar" factor were good ratings when the sound quality was high. From the results it can be concluded that high-quality music increases the drivers' excitement and concentration while driving.
*Paper 10472*

• **A Preliminary Study on the Correlation between Subjective Sound Quality Perception and Physiological Parameters**—*Angelica Poli,*[1] *Stefania Cecchi,*[1] *Susanna Spinsante,*[1] *Alessandro Terenzi,*[1] *Ferruccio Bettarelli*[2]
[1] Università Politecnica delle Marche, Ancona, Italy
[2] Leaff Engineering Srl, Porto Potenza Picena (MC), Italy

Subjective listening tests are an important method to evaluate the performance of audio algorithms and of sound reproduction systems. However, listening tests are a costly and complex task whose reliability can be affected by several nuisance variables such as lexicon ambiguity and contextual biases. To help mitigate these aspects, a preliminary investigation on the effects of audio perception on physiological parameters is presented in this paper. In particular, the correlation between the subjective assessment of perceived sound quality and the affective reaction elicited by specific sound stimuli is analyzed and reported.
*Paper 10495*

• **Emotional and Neurological Responses to Timbre in Electric Guitar and Voice**—*Sephra Scheuber,*[1,2] *Mickie Vanhoy*[2]
[1] Oklahoma Christian University, Edmond, OK, USA
[2] University of Central Oklahoma, Edmond, OK, USA

Two types of sounds were created for this study: non-verbal voice sounds and electric guitar sounds. Participants completed categorization of emotion and ratings of intensity and believability. Attack slope was found to be the primary factor in the distinction between emotion categorization. In the second part of the experiment, EEG data were gathered from participants while they made judgements on the emotional similarity of guitar sounds.
*Paper 10505*

• **Implications of Crossmodal Effects and Spatial Cognition on Producing in Spatial Audio**—*Thomas Görne, Kristin Kuldkepp, Stefan Troschka*, Hamburg University of Applied Sciences, Hamburg, Germany

It is quite common to use spatial language in the description of the sensation of sound: A sound can be big or small, it can be edgy, flat or round, a tone can be high or low, a melody rising or falling—all these linguistic metaphors are apparently emerging from the crossmodal correspondences of perception. An auditory object can have a metaphorical size, shape and position in space besides its (perceived) physical size, shape and position in space. The present paper reviews research on crossmodal effects and related findings from different disciplines that might shine a light on the production and aesthetics of spatial audio. In addition, some preliminary results of experiments with complex spatial sonic structures are presented.
*Paper 10506*

**WORKSHOP/TUTORIAL: RECORDING & PRODUCTION**
**Introduction to Copyright Laws for Sound Engineers and Creatives**
**Tuesday, May 25, 6:00 pm – 7:00 pm**

Presenters: **Philipp Lengeling**, RafterMarsh
**Mark A. Pearson**, ARC Law Group

The presenters will go through the common pitfalls, issues, and considerations to watch out for when creating and distributing your recordings/music. Who owns what? Who owns the recording? Who owns the song (composition)? Who owns the rights to publicly perform? What can be contractually waived? What's a must have in your agreements and where can you get help? The basics about music copyright law in North America and Europe and the most important differences between the two will be discussed during this panel.

**WORKSHOP/TUTORIAL: BROADCAST & ONLINE DELIVERY**
**From Live to Virtual and Back: Is Hybrid the Future of Audio?**
**Tuesday, May 25, 7:00 pm – 8:00 pm**

Moderator: **Heather Rafter**, RafterMarsh

Presenters: *Jim Ebdon*, Ebdon Music
*Laura Escudé*, Artist/Entrepreneur
*Antony Randall*, Planet Home
*Frederick Umminger*, Roblox Corporation
*Dave Van Hoy*, Advanced Systems Group, LL

As we ponder both the future of live music and digital streaming, a plethora of questions have arisen. Are the technologies converging? How can we make money going forward? Will live sound return as we know it? Can FOH engineers save money by mixing virtual concerts from their console and obviate the need for pricey broadcast trucks? What are the relevant technologies that currently exist or should be built for the future?

This panel will also explore the monetization of concerts within gaming platforms such as Roblox, as well as the role of AI and animation. Is there a way to go not only from live streaming but from streaming or gaming to live? Join our panel of industry experts, including the engineer who recorded Beck's 3D/360 performance, one of the very first of its kind and Justin Bieber's current engineer, a seasoned live sound engineer, now organizing virtual streams during the pandemic. Other speakers will include a seasoned audio programmer for the Roblox gaming platform, which is providing a new framework for hosting concerts virtually. This panel will explore the future of live sound, along with the audio tools and platforms that are transforming the concert industry.

**SPECIAL EVENT: EDUCATION**
**Ummet Ozcan Sound Design Masterclass**
**Tuesday, May 25, 7:00 pm**
**Genelec Special Education Event**

Presenter: **Ummet Ozcan**

In this session, Ummet will be explaining why sound design is so important for a Producer, before going on to discuss how different synthesis techniques work and the types of sounds you can create with them.

As a superstar DJ Producer, sound designer, software developer, label owner and "scientist spinning records," this is a fascinating insight into Ummet's work and his tools of the trade.

**SPECIAL EVENT**
**SoundGirls.org Mentoring Session: Audio for Music Production**
**Tuesday, May 25, 7:00 pm – 8:00 pm**

Presenters: **Nene Veenman**, Exobia: 3RD Season/Veenman & Morrison Composers
**Audrey Martinovich**, Audio for the Arts
**Petra Randewijk**, Sound Engineer

SoundGirls.org's Netherlands Chapter hosts a mentoring session on music production featuring professionals in the field. Please come prepared to ask questions and gain valuable insight.

**WORKSHOPS/TUTORIALS: ACOUSTICS & PSYCHOACOUSTICS**
**The Sound of Things**
**Wednesday, May 26, 11:00 am – 12 noon**

Presenter: **Agnieszka Oltarzewska**, Siemens Digital Industries

Psychoacoustics has been gradually making its way from lecture halls and academic papers to the everyday life of NVH (Noise, Vibration, and Harshness) and sound quality engineers worldwide. The competitive situation in the market calls for new products to not only function better, but to sound better too. In this tutorial I would like to present how sound quality analysis tools, such as subjective listening tests or objective analyses (including calculations of loudness, sharpness, modulation or tonality) are used during the product development process. The presented examples will link the intricacies of the human auditory system to different acoustic challenges that arise during the design phase. From toothbrushes, through espresso machines and vacuum cleaners up to cars and airplanes, if the thing makes a sound, there's a big chance that someone used the psychoacoustic theory to evaluate and improve it.

**HARMAN PARTNER PRESENTATION**
**Understanding Specification Sheets**
**Wednesday, May 26, 11:00 am – 12:00 noon**

Presenters: **Ross Brett**
**Ed Jackson**
**Christer Lidberg**

Led by the Harman Engineering Team, this course will help you to understand the fundamentals of specification sheets and how to interpret the technical terminology for real life applications.

We will look at loud speakers, amplifiers, consoles, and DSP and will examine the fact that not all specification sheets are created equally and that comparing "apples for apples" is sometime more complicated that it initially appears .

**WORKSHOP/TUTORIAL: GAME AUDIO AVAR SPATIAL AUDIO**
**Loudspeaker Virtualization for Immersive**
**Wednesday, May 26, 12:00 noon – 1:00 pm**

Presenter: **Lasse Nipkow**, Silent work LLC

Immersive Sound offers consumers an extremely immersive listening experience for music through loudspeakers, which is not possible with standard 2CH stereo. Large loudspeaker setups such as Auro 3D 13.1 or Dolby Atmos 9.1.6 are not feasible for most consumers at home. It therefore makes sense to offer the content as headphone productions for the mass market.

A large number of binauralization processors and consumer products have entered the headphone market in recent years. And streaming services have been offering Immersive Sound content for headphones since 2019. But how close does headphone reproduction come to reproduction via loudspeakers? And what needs to be done still to make headphone reproduction equal?

In his video workshop, Lasse Nipkow explains the most important aspects of loudspeaker virtualization, which are necessary for a three-dimensional headphone reproduction. He uses video animations and sound samples to show where we are today on the way to this goal.

**STUDENT & CAREER DEVELOPMENT EVENT**
**Bridging the Technical Divide: Techniques for Effective Communication with Artists**
**Wednesday, May 26, 12:00 noon – 1:00 pm**

Presenters: **Nic Britton**
**Amie Chatterley**, CapsaArx Studios
**Lucy Harrison**, ACM
**Matthew Russell**, ACM

In response to changes in the music industry it has become vital for audio professionals to have the communication skills to work collaboratively with artists while creating a supportive, creative environment. This industry led workshop, aims to provide practical solutions for how producers and audio technicians can communicate effectively with artists in order to provide the best outcome for collaborative, live and studio work.

The workshop will be led by industry expert staff and tutors from The Academy of Contemporary Music (ACM) who have experience both working with artists on high profile tracks and live events and in training the next generation of producers and technicians to work collaboratively in all areas of the music industry. The panel represents experience in artist collaboration through electronic music, studio production, live sound, studio management and sound design for media.

The workshop will provide a discussion and effective, tested, solutions on how to communicate with artists so that you can collaboratively meet your artistic aims while creating a professional and supportive environment.

The session will cover:

• Precision in communication—ensuring that everyone is communicating consistently in order to work towards success.

• Techniques for communicating technical information in a way that can be understood and used by all team members, regardless of technical training.

•Codes of ethics and how to build ethical practices into your collaborative workflow.

•Inclusive communication techniques to ensure that all members of the technical and creative team are able to fully participate in conversations and feel supported to communicate.

•How to build partnerships with artists and ensure that you are easy to work with.

All guidance for this workshop has been developed by industry professionals with technical knowledge of production and audio engineering, ensuring that approaches can be easily included in the modern producer and technician's workflow.

**PAPER Q & A SESSION: AV ART**
**Wednesday, May 26, 1:00 pm**

• **Do You See What I Hear? An Artificial Chromesthetic Experience—** *Michael W. Earle*, Houghton College, Hougton, NY, USA; State University of New York at Fredonia, Fredonia, NY, USA

The scientific and experiential relationship of color and sound frequency has long been a topic of academic interest within the arts. Specifically, color has been recognized as a descriptive mechanism for music. For centuries, composers, musicians, and technical artists have relied, to varying degree, on the use of color as a semantical tool to distinguish emotional intent for a variety of musical events. The aim of this project is to prototype a system for real-time frequency conversion from audible sound to visible light and use such a system to quantitatively enhance the qualitative musical performance. This system will be considered an artificial method of simulating sound to color synesthesia, otherwise known as chromesthesia. This prototype will act as a foundation to plausibly show that a measurable and scientifically definable relationship exists between the aural and visual spectra.
*Paper 10462*

• **An Audio-Driven System for Real-Time Music Visualization—***Max Graf, Harold Opara Chijioke, Mathieu Barthet*, Queen Mary University of London, London, UK

Computer-generated visualisations can accompany recorded or live music to create novel audiovisual experiences for audiences. We present a system to streamline the creation of audio-driven visualizations based on audio feature extraction and mapping interfaces. Its architecture is based on three modular software components: backend (audio plugin), frontend (3D game-like environment), and middleware (visual mapping interface). We conducted a user evaluation comprising two stages. Results from the first stage (34 participants) indicate that music visualizations generated with the system were significantly better at complementing the music than a baseline visualization. Nine participants took part in the second stage involving interactive tasks. Overall, the system yielded a Creativity Support Index above average (68.1) and a System Usability Scale index (58.6) suggesting that ease of use can be improved. Thematic analysis revealed that participants enjoyed the system's synchronicity and expressive capabilities, but found technical problems and difficulties understanding the audio feature terminology.
*Paper 10498*

**STUDENT & CAREER DEVELOPMENT EVENT**
**Starting Your Brand Like a Boss**
**Wednesday, May 26, 1:00 pm – 2:00 pm**

Presenters: **Amie Chatterley**, CapsaArx Studion
**Oisin Lunny**, Galaxy of OM, S.L.
**Helga Osk Hlynsdottir**, Serious.Business

With more and more of our lives taking place in the digital world, particularly during the pandemic, it has never been more important to have a great brand presence online. Human-to-human connections can be a defining factor for your career, but how can you maintain them when more of our lives are becoming digital? All these questions, and more, will be answered in "Starting Your Brand Like A Boss" which will take you from inspiration to creation to branding sensation!

Whether your profile is your website, your portfolio, or your social media activity, this jam-packed expert session will give you a stack of useful tips, tricks, and techniques to successfully manage your online brand like a boss.

Join our three experts for a download of branding awesomeness: daring branding maven Helga Osk Hlynsdottir from leading agency Serious Business, entrepreneur, educator, bass player, and founder of the Power Metal Quest Fest, Amie Chatterley, and journalist, podcaster, and award-winning marketer Oisin Lunny.

Topics covered will include:
• What is branding
• 10 top tips for branding
• The power of the collective
• From competition to collaboration
• How to create a personal brand
• Establishing and maintaining reputation
• The principles of networking
• How to network in a digital world
• The importance of social media
• LinkedIn, Facebook, Instagram, Twitter, WTF?

This session will be packed with great advice, handy checklists, hard-won career lessons, and even a few memes.

**WORKSHOP/TUTORIAL: BROADCAST & ONLINE DELIVERY**
**Next Generation Audio for Advanced Music Creations and Distributions**
**Wednesday, May 26, 1:00 pm – 2:00 pm**

Presenters: **Kimio Hamasaki**, ARTSRIDGE LLC
**Hideo Irimajiri**, WOWWOW Inc.
**Toru Kamekawa**, Tokyo University of the Arts
**Kazuya Nagae**, Nagoya University of the Arts

Immersive audio, high-resolution audio, and high-definition audio have become available for recording, creations, and online delivery of music. While the lossy codec such as MPEG4-AAC was the key technology to enable digital broadcasting and online-delivery, viewers and listeners of digital broadcasting and online-delivery have noticed the difference of audio quality and musical emotion between the legacy 2-ch stereo audio using lossy codec and immersive audio using high-quality audio codec.

This workshop will introduce recent actual use cases of next generation audio such as immersive audio, high-resolution audio, and high-definition audio for the advanced music creations and online-delivery of music.

Kimio Hamasaki will summarize the history and latest status of next-generation audio and introduce examples of his own research and development as well as his musical recording works. Hideo Irimajiri will report on the live experiment of music online-delivery with high-resolution audio and immersive audio in Japan done by WOWOW last October and discuss the future prospects. Toru Kamekawa will report on music creations using next-generation audio at the Tokyo University of the Arts and introduce some actual examples and discuss the future prospects. Kazuya Nagae will report on the recent productions, distributions, and education using immersive audio at Nagoya University of the Arts and discuss the future prospects.

This panel will also discuss the advantages of the next-generation audio for advanced music creations and online-delivery and issues of those they encountered during the actual cases. And finally, they will discuss how to improve the quality and emotion of music recordings and deliveries in the future.

**PAPER Q & A SESSION: ARRAYS/LSP**
**Wednesday, May 26, 2:00 pm**

• **Beamforming Using Two Rigid Circular Loudspeaker Arrays: Numerical Simulations and Experiments—***Yi Ren, Yoichi Haned***a**, The University of Electro-Communications, Chofu, Tokyo, Japan

Beamforming is an important technique in studies involving loudspeaker arrays, and conventional beamforming studies use linear, circular, and spherical arrays. In a previous study the present authors introduced a model involving two circular loudspeaker arrays to reproduce focused sources; the two arrays have rigid baffles, and their multiple scattering offers better performance. The present paper reports on investigations into implementing beamforming with this array model, the performance of which is evaluated using a minimum-variance distortionless-response beamformer. Numerical simulations show that the proposed method outperforms a single circular array at lower frequencies, and the numerical results agree with those from experiments conducted in an anechoic chamber.
*Paper 10450*

• **Enhanced Polygonal Audience Line Curving for Line Source Arrays—***Arne Hölter,*[1] *Florian Straube,*[1] *Frank Schultz,*[2] *Stefan Weinzierl*[1]
[1] Technical University of Berlin, Berlin, Germany
[2] University of Rostock, Rostock, Germany

Line source arrays (LSAs) are used for large-scale sound reinforcement aiming at sound fields that are as homogeneous as possible over the whole audio bandwidth. The deployed loudspeaker cabinets are rigged with different tilt angles and/or are electronically controlled in order to provide the intended coverage of the audience zones and to avoid radiation towards reflective ceilings, sidewalls or residential areas. In this contribution enhancements of the analytical polygonal audience line curving (PALC) approach are presented. PALC was introduced for finding appropriate LSA cabinet tilt angles with respect to the geometry of the receiver area and the intended coverage. The PALC extension includes methods to use discrete sets of inter cabinet tilt angles, to control the target coverage by using weighting factors and to deal with non-continuous audience lines, i.e., zones which are not to be reinforced. The extended PALC is evaluated in comparison with a typical standard LSA curving scheme. An implementation of PALC is provided as an open web application.
*Paper 10451*

• **Propagation Loss of Low Frequency Horn Loudspeakers: Is "Throw" a Real Phenomenon?—** *James Hipperson*, Funktion One Research Ltd., Hoyle, Dorking, UK

Horn loading is frequently used in sound reinforcement to increase efficiency and directivity of high and mid frequency transducers. Low frequency horn loudspeakers are less common due to their large size. Increases in available amplifier power and thermal dissipation in transducers have led to widespread use of high power, low efficiency dual 18" bass reflex loudspeakers. However, some manufacturers and enthusiasts continue to develop and use low frequency horn loudspeakers for their high efficiency and subjective audio quality. In the fields of live event production and noise control, there is sometimes a perception, or "urban myth" that horn low frequency loudspeakers project or "throw" sound a further distance than direct radiating low frequency loudspeakers. This is either considered to be beneficial, or problematic depending on the context. Considering the relevant acoustic theory, it is not immediately apparent why this should be the case, providing the loudspeakers are level matched and of similar physical dimensions. Unfortunately, there is very little investigation of low frequency horns in previous literature to aid in providing a definitive answer. Measurements in this paper demonstrate that horn and direct radiating low frequency loudspeakers and arrays closely follow the theory, and the difference in propagation loss is within measurement uncertainty. The implication for noise control of outdoor events, is that bass loudspeaker size and type are not especially relevant factors, and focus should instead be on system/array design and site layout.
*Paper 10453*

• **Auralizing Concert Venues over Extended Listening Areas Using Wave Field Synthesis—***Jonas Braasch, Samuel Chabot, Evan Chertok, Jonathan Mathews, E.K. Ellington Scott*, Rensselaer Polytechnic Institute, Troy, NY, USA

This paper proposes an efficient method to create auralizations of acoustical landmarks using a 2D ray-tracing algorithm and publicly available floor plans for a 128-channel wave field synthesis (WFS) system with 2.5D approximation. Late reverberation parameters are calculated using additional volumetric data. The approach allows the rapid sonic recreation of historical concert venues with adequate sound sources. The listeners can walk through these recreations over an extended user area (12×10 sqm), and the software suite can be used to calculate room acoustical parameters for various positions directly using a binaural rendering method or via the WFS simulation.
*Paper 10465*

**WORKSHOPS/TUTORIALS**
**Modern Digital to Analogue Converters: HOW Many Bits!?**
**Wednesday, May 26, 2:00 pm – 3:00 pm**

Presenter: **Jamie Angus-Whiteoak**

Recently Digital to Analogue Converters (DACs) that claim to have resolutions of 32 bits have become available. How do they possibly achieve such exalted levels of performance?

Almost all, modern Digital to Analogue convertors (DACs) use oversampled multi-bit convertors with noise-shaping to achieve their high performance. Oversampling and noise-shaping allow one to use a DAC with a small number of levels, which is easier to manufacture. Unfortunately traditional noise shaping does nothing to reduce the effect of component tolerances in the DAC, because the analogue output cannot be fed back to the input.

However, modern DACs do manage to noise shape the output from the DAC without any feedback. This piece of audio alchemy is critical to the exceptional performance of modern Digital to Analogue convertors.

This tutorial will explain how this alchemy is achieved. It will review the problems of component tolerance in DACs and show how they compromise performance. Then noise-shaping, and how it can be applied, without magic, or knowing the actual converted output, to a practical DAC will be explained. It will conclude by discussing how system aspects may limit their performance and discuss how you might prepare audio signals to maximise converter performance.

**SPECIAL EVENT**
**Introducing the TC-MLAI**
**Wednesday, May 26, 2:30 pm – 3:00 pm**

The AES Technical Council has identified that the community of audio engineers working with Machine Learning (ML) and Artificial Intelligence (AI) is underrepresented in the technical committees. The newly formed AES Technical Committee on Machine Learning and Artificial Intelligence (TC-MLAI) intends to answer this need.

The TC-MLAI focuses on applications of machine learning and artificial intelligence in audio, with discussions on topics such as: best practices, data, licensing, social and cultural aspects, technical innovations, and ethics. The goal of the committee is to drive discussion and exchange information by organizing workshops, symposia, tutorials, and technical documents. It will also act as a point of contact and a bridge to other AES technical committees, the AES community at large, and other organizations involved in ML and AI for audio.

In this workshop we will present the committee's mission, values, agenda, and avenues for membership and participation. We will highlight exciting developments and trends as they relate to audio, while at the same time acknowledging topics of controversy, such as data bias, privacy concerns, and when it is appropriate to call an audio technology "artificially intelligent."

This is the introduction only. Please watch on Thursday at 6 pm CEST to view the whole session and panel.

**PAPER Q & A SESSION: REPRODUCTION—3D AUDIO 2**
**Wednesday, May 26, 3:00 pm**

- **Delivering Personalized 3D Audio to Multiple Listeners: Determining the Perceptual Trade-Off Between Acoustic Contrast and Cross-Talk**—*Natasha Canter, Philip Coleman*, Institute of Sound Recording, University of Surrey, UK

3D audio for multiple listeners can be created by combining a personal sound system with cross-talk cancellation to direct binaural content to each listener's ears. However, the relative perceptual importance of controlling these two aspects of the sound field reproduction has not been established. Two headphone-based experiments were carried out in order to understand the trade-offs between acoustic contrast and cross-talk cancellation performance. The first experiment used a method-of-adjustment approach to determine the thresholds at which (a) an interfering program was no longer distracting, with varying cross-talk in the target program, and (b) the threshold at which a target binaural audio program was considered to be enveloping, in the presence of interfering audio. The second experiment used pairwise preference ratings to determine the trade-off in preference between stimuli with different levels of acoustic contrast and cross-talk cancellation. It was found that achieving good acoustic contrast should be prioritized over cross-talk cancellation in a system combining sound zones and binaural technology, but that for a certain level of interference, reducing cross-talk improves listener preference. Moreover, diffuse interferers produced higher thresholds of distraction than localized ones, implying that sound zone systems should consider the spatial characteristics of sound in the dark zone.
*Paper 10452*

- **Comparing immersive Sound Capture Techniques Optimized for Acoustic Music Recording through Binaural Reproduction**—*Will Howie*,[1] *Dennis Martin*,[2] *Toru Kamekawa*,[3] *Jack Kelly*,[2] *Richard King*[2]
[1] CBC/Radio–Canada, Vancouver, BC, Canada
[2] McGill University, Montréal, QC, Canada
[3] Tokyo University of the Arts,Tokyo, Japan

A study was undertaken to compare three immersive sound capture techniques optimized for acoustic music recording, within the context of binaural audio reproduction. 3D audio stimuli derived from 9-channel (4+5+0) recordings of a solo piano were binaurally rendered and presented to listeners over headphones. Subjects compared these stimuli in terms of several salient perceptual auditory attributes. Results of the double-blind listening test found no significant differences between two of the sound capture techniques, "spaced" and "near-coincident," for the perceptual auditory attributes "envelopment," "naturalness of sound scene," and "naturalness of timbre." The spaced technique, however, was shown to create a larger virtual image of the sound source than the near-coincident technique. The coincident technique was found to create an immersive sound scene that occupies a different perceptual space from the other two techniques, delivering less envelopment and naturalness.
*Paper 10455*

- **Spatial Stability Verification of Multichannel Sound Image with Hexagonal Capture for Reduction to Two Channels**—*João Victor Pinto, José Augusto Mannis*, Acoustics and Sound Arts Laboratory (LASom), Arts Institute – University of Campinas (UNICAMP), Brazil

This work is part of a research dedicated to the development of systems and methods of capturing and reproducing multichannel sound for applications in immersive environments for music, sound art, bioacoustics, and environmental monitoring. This paper focuses on the evaluation of hearing quality in a binaural context for headphones reproduction of recordings in six channels with non-coincident hexagonal capture using a device developed at LASom. Several HRTF models were com-

pared through renderings alternatively using Max/MSP and the OpenAir plug-in in DAW. In order to define the HRTF with best performance (greater similarity to the original reproduction), subjective evaluations were implemented. The best result obtained was attributed to the KEMAR model in OpenAir.
*Paper 10477*

• **Perceptual Optimization of Stereo Width Control Methods in Headphones—***Yui Ueno,*[1] *Mitsunori Mizumachi,*[1] *Toshiharu Horiuchi*[2]
[1] Kyushu Institute of Technology, Fukuoka, Japan
[2] KDDI Research, Inc., Saitama, Japan

Legacy 2-ch stereo music sources cause unnatural spatial impressions through earphones and headphones reproduction due to the lack of the crosstalk produced in loudspeakers reproduction. Perceptual widths of the stereo stage in amplitude-based and phase-based stereo width control methods for headphones reproduction were matched with that of the binaural rendering which simulates loudspeakers reproduction. Paired comparison between the stereo width control methods and binaural rendering was carried out to achieve the perceptual optimization for classical, jazz, and pop music sources. Concerning the amplitude-based stereo width control method, the suitable parameter value for each sound source could be determined from the results of the listening test. It is also found that the phase-based stereo width control method should be optimized in different parameter spaces.
*Paper 10480*

• **Mapping Methods of 3D Input Devices and Immersive Audio Environments Using Polar Coordinates for Panning—***Diego Quiroz, Denis Martin,* McGill University, Montreal, Canada; Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Montreal, Canada

This study explores different mapping implementations for various input devices for an immersive mixing task within a polar coordinate system mixing space. A test, in which subjects were asked to pan a monophonic sound object (probe) to the location of a pink noise burst (target), was conducted in a custom 3D loudspeaker array. The objectives were to determine how quickly and accurate the subjects were able to perform the task using each device, which was most appropriate, and which was most preferred overall. Results show significant differences in response time between devices. Input devices with high level of integrality in and direct mapping to spatial parameters were the most accurate and fastest in trial duration when performing the task. Futures studies will involve a more thorough immersive audio mixing task that allows further detail into the variables investigated.
*Paper 10486*

**WORKSHOP/TUTORIAL: EDUCATION**
**Education Panel: Steal My Syllabus**
**Wednesday, May 26, 3:00 pm – 4:00 pm**

Abstract unavailable.

**STUDENT & CAREER DEVELOPMENT EVENT**
**Student Recording Competition: Tradition Acoustic Recording**
**Wednesday, May 26, 3:00 pm – 4:00 pm**

The AES Student Recording Competition is a unique opportunity for student attendees of AES International Conventions to receive feedback and recognition for their audio production work.

Finalists will be announced and prizes awarded during this presentation.

Judge Panelists include: *Martha DeFrancisco, Gary Call Hanley, David Bowles, Ken Blair.*

**PAPER Q & A SESSION: EDUCATION 2**
**Wednesday, May 26, 4:00 pm**

• **Women in Music Industry Roles: A Twenty-Year Analysis—***Chandler Bridges,*[1] *Haley Strong,*[1] *Aaron Overton,*[2] *Justin Berish*[1]
[1] Jacobs School of Music, Indiana University, Bloomington, IN, USA
[2] The Los Angeles Film School, Los Angeles, CA, USA

This research analyzed the gender distribution of the songwriters, producers, and performers of chart-topping music for the 21st Century, over a 20-year period. Some studies have suggested that women are vastly under-represented, however, this amount of data has not been cited nor analyzed previously. Selected from all 21st Century songs in *Billboard* magazine's Hot 100 end-of-year charts (N=210), this research analyzed gender distribution of these participants (N=1624). While the results are indicative of previously reported trends showing women holding a small percentage of creative roles in popular charting music, the data indicate an upward momentum of women holding credited positions on top albums. The discussion includes further results and educational implications as well as reifying how these data are rapidly changing.
*Paper 10497*

• **Design of a Vowel and Voice Quality Indication Tool Based on Synthesized Vocal Signals—***Paul A. Bereuter,*[1,2] *Florian Kraxberger,*[1,2] *Manuel Brandner,*[1,3] *Alois Sontacchi*[1,3]
[1] University of Music and Performing Arts, Graz, Austria
[2] Graz University of Technology, Austria
[3] Institute of Electronic Music and Acoustics, Graz, Austria

Voice disorders due to strenuous usage of unhealthy voice qualities are a common problem in professional singing. In order to minimize the risk of these voice disorders, vital feedback can be given by making aware of one's sung voice quality. This work presents the design task of a vowel and voice quality indication tool which can enable such a feedback. The tool is implemented in form of a VST plug-in. The plugin's interface provides a graphical representation of voice quality and vowel intelligibility by means of two 2D voice maps. The voice maps allow a graphical distinction of three voice qualities (modal, breathy or creaky), and the representation of a sung vowel within the formant space spanned by the first and second formant frequency. The design process includes (i) building a ground truth dataset by using a modified speech synthesizer, (ii) linear prediction analysis, and (iii) the visualisation of the estimated vowel and voice quality by means of the 2D voice maps. The plugin's code is available as open source to enable further development.
*Ebrief 642*

• **Managing Telematic Pain: Migrating a Student Ensemble Online During COVID—***Tom Zlabinger,* York College /CLUNY, Jamaica, NY, USA

During COVID, musical ensembles at high schools and colleges around the globe were forced to develop solu-

tions to rehearsing while students were dispersed from campus or be forced to disband. The concept and required resources of telematic performance (performing with others online in real time) has been in development since the 1980s. The greatest challenge in telematic performance is reducing latency between musicians, enabling musicians to synchronize their performances with one another in real time. Recently, such platforms as JackTrip, Jamulus, and others have allowed musicians to perform together online with very low latency. But the possibility of performing online is dependent on having a stable and clean internet connection, plus the required audio equipment to capture a musician's performance. As was made evident during COVID, the digital divide must be negotiated and mediated as classrooms were migrated online. And musical performance additionally amplifies these challenges. This presentation will discuss and share the technological and unique social challenges of the migration of the York College Jazz Band to an online format.
*Ebrief 644*

**WORKSHOP/TUTORIAL: ACOUSTICS & PSYCHOACOUSTICS**
**Listening Test Data Analysis for Everyone: The A/B Test**
**Wednesday, May 26, 4:00 pm – 4:45 pm**

Presenter: **Darlene Williamson**, Harman International

Listening evaluations are EVERYWHERE. Whether involved in the creative process, design of gear or selection of components, gear or suppliers, audio is being judged by humans! These evaluations might be informal, perhaps one of many taken on a given day. However, oftentimes, formal listening tests are performed to support business decisions: When the stakes are high, it is often prudent to gather scientific, valid data about the test scenario.

In each instalment of the series, HARMAN's data analysis expert Dr. Darlene Williamson will take you through data analytic procedures for a specific type of test. Test design considerations will be covered, but the focus will be on valid data analytic strategies for formal listening tests. Students and professionals alike will benefit from Dr. Williamson's careful and thorough explanations on the pitfalls to avoid and options to navigate when extracting the meaning from the data.

The series begins with probably the most common type of listening test: the A/B Test, which aims to see which of two options are preferable. Analyses from real-world datasets will be showcased, using commonly available software tools. Everyone is invited to come along and give their stats skills a boost!

**SPECIAL EVENT**
**Audio Mythology, Human Bias and How Not to Get Fooled**
**Wednesday, May 26, 4:00 pm – 4:45 pm**

Presenters: **Michael Lawrence**, Rational Acoustics
**Ethan Winer**

As a spiritual successor to his infamous 2009 "Audio Myths Workshop" session, audio engineer, author, and mythbuster Ethan Winer will be joined by *Live Sound International* technical editor Michael Lawrence to discuss what he has learned from a long career dedicated to investigating and debunking widely held audio-related misconceptions and mythology.

Winer and Lawrence will touch on some common audio "truisms" that might not be as true as they seem, along with the perceptual biases that can skew our experiences as listeners, and talk about the "sandbox" approach to designing simple experimentations to test and investigate claims for oneself.

**SPECIAL EVENT: KEYNOTE ADDRESS**
**A Peek Under the Hood of Perceptual Audio Coding:**
**A Review of the History and a Look into the Future**
**Wednesday, May 26, 4:45 pm – 5:45 pm**

Presenter: **Marina Bosi**, CCRMA, Stanford University

Did you ever wonder how your audio files squeeze so much sound into such a small size? Or what is the difference between MP3 and AAC? Or which multichannel audio coding format is best for your application?

The development of perceptual audio coding technologies allowed portable music devices to be launched and "suddenly" these technologies became ubiquitous in our daily lives, residing within mobile devices, DVDs, broad/webcasting, electronic distribution of music, etc. A natural question to ask is: what made all this possible and where is the technology going?

In her presentation, Dr. Bosi will examine major shifts in audio consumption and how they represented new challenges and opportunities in coding audio for entertainment, information, and other purposes. Based upon her deep experience with digital media coding research, and standards, Dr. Bosi will offer unique insights into the widespread use of these technologies in applications ranging from production and distribution of sound to the broader consumer experience, providing the foundation for an informed view of the future of digital media.

**PAPER Q & A SESSION: SYNTHESIS**
**Wednesday, May 26, 5:45 pm**

• **Towards a User-Friendly System for Environmental Sound Synthesis in Games**—*Cezar Floroiu, Ian Gibson*, University of Huddersfield, West Yorkshire, UK

The aim of this project is to explore the creation of an intuitive procedural audio software system for synthesizing environmental effects (natural phenomena sounds) that can be implemented into the Unreal games engine. Current systems are dependent on pre-recorded assets and/or are 'in-house' solutions that are created and implemented by programmers. A need for a system that enables the facile creation of audio assets as well as their real-time manipulation has been identified. The system synthesizes the sounds in real time using physically in-spired models and eliminates the need for pre-recorded assets. The synthesized natural phenomena sounds that are explored are rain, wind, thunder, and fire. The sound synthesis system will offer dynamic interaction with the real-time game environment without the need for the programmer to access lower-level parameters. Topics of exploration include whether the program might improve game development workflow by abstracting low-level controls and having the sound sources intuitively interact with the game environment. The system will present a user-friendly interface thereby minimising the need for knowledge of traditional programming languages. The system's audio processing is achieved using Max-MSP and links to the Unreal Engine 4 (UE4) game engine. Future work will investigate more complex interactions between the sound system and the game engine, as well as developing the system as a plugin or a standalone program.
*Paper 10466*

• **Efficient Synthesis of Violin Sounds Using a BiLSTM Network Based Source Filter Model**—*Yi-Ren Dai, Hung-Chih Yang, Alvin W.Y. Su*, National Cheng-Kung University Tainan, Taiwan

The dynamic changes in playing skills generated from bow-string interaction make synthesizing bowed string instrument sounds a difficult task. Recently, a source filter model incorporating the LSTM predictor and the granular wavetables gives encouraging results. However, the prediction error is still large and the model hasn't caught the nuance caused by the constantly changing characteristics of a playing violin. In this paper the granular wavetable is represented of DCT coefficients and a new training strategy is proposed to reduce the predictor error. In addition, we analyze the difference between the original violin tone and the corresponding synthesis tone. A random pitch perturbation and a DCT coefficient shaping method are proposed to imitate the changing characteristics since results sound regular.
*Paper 10476*

• **Synthesis of Wind Instruments and Their Blowing Noise Using a LSTM Time Varying Source Filter Model**—*Ju-Yen Chen, Hung-Chih Yang, Wen-Yu Su*, National Cheng-Kung University Tainan, Taiwan

Digital Waveguide Filters have been applied to synthesis of wind instruments for years. To design the filter coefficients to synthesize the timbre of a particular instrument usually takes lots of time and effort. A source filter model combined with a long short-term memory(LSTM) recurrent neural network for the synthesis of violin has been successful in this respect, but it lacks the synthesis of the noise part that is important for any playing wind instruments. In this paper we adopt the noise synthesis method of DWF and apply it to the LSTM based source filter model to take the advantages of both methods. The French horn tones recorded in the Real World Computing(RWC) database are used to demonstrate the work. The synthesis tone sounds close to the original tone. Sound files are provided. The proposed method can be efficiently implemented as a VST plugin.
*Paper 10478*

• **A Comparative Perceptual Evaluation of Thunder Synthesis Techniques**—*Joshua D. Reiss*,[1] *Hazar Emre Tez*,[1] *Rod Selfridge*[2]
[1] Queen Mary University of London, London, UK
[2] KTH Royal Institute of Technology, Stockholm, Sweden

The sound of thunder is widely used in game, film, and virtual reality sound design. It is also a phenomenon for which we seek a better understanding of the physics underlying the sound. Though many models of thunder have been proposed, there has not yet been a formal perceptual evaluation of the models to assess their realism and sound quality. Here, we present and evaluate the implementation of several thunder sound effect synthesis models. The models include different physical modeling and signal-based approaches, as well as a recorded sample. Evaluation was with over 50 participants. The results showed that none of the models were close to the recording in terms of realism, though signal-based models slightly outperformed the physical models. This highlights the need for comparative perceptual evaluation in sound synthesis, and identifies the limitations of current thunder simulation approaches.
*Ebrief 640*

**PARTNER PRESENTATION: DESCRIPT**
**AI-Assisted Audio Editing with Descript**
**Wednesday, May 26, 5:45 pm – 6:30 pm**

Moderator: **Jay LeBoeuf**

Presenters: *Harmony Jiroudek*
*Rebecca Kossnick*
*Katrina Lui*
*Sarah Moliner-Roy*

This panel discussion will focus on how studios such as WYNC, NPR, Axios, Al Jazeera, Pushkin Industries, and theSkimm all share a common goal: to streamline media creation and audio editing processes. This panel presentation dives into how artificial intelligence is being used to power the creation & editing process for hundreds of top podcasts.

We'll introduce how producers and audio engineers use technologies such as text-based audio editing, automatic room tone matching, speaker identification, voice cloning, and more to rapidly accelerate their work. Attendees will learn how these features were invented, developed, tested and shipped. We'll conclude with the speculation on the future of AI-driven new media creation.

**SPECIAL EVENT**
**Wednesday, May 26, 5:45 pm – 6:45 pm**
**NFTs Demystified: What Every Audio Engineer (and Investor!) Needs to Know**

Moderator: **Heather Rafter**, RafterMarsh

Presenters: *Jeremy McKane*, OCN/The McKane Organization
*Chris Cooke*, + MD, CMU
*Antony Randall*, Planet Home

Trevor Noah, Clubhouse. . . . NFTs are the hottest three letters on the audio block right now.

This is what you need to know.

What are they? Are they real? What does an audio engineer or artist create to authenticate? Who owns what?

Our dazzling panelists will demystify no one else wants to touch. . . . Wish us luck.

**SPECIAL EVENT**
**Women in Audio Project 2000, A Retrospective**
**Wednesday, May 26, 6:30 pm – 7:15 pm**

Moderator: **Meredith Goldstein**

Presenters: *Laurie Anderson*
*Roma Baran*
*Carol Bousquet*
*Leslie Gaston-Bird*
*Stephanie Hull*

In 1995 at the Audio Engineering Society's 99th Convention in New York, Carol Bousquet invited prominent women to speak about the lack of representation of women in the field of audio. Now, 26 years later, we revisit the event and its impact. Hosting the session is Meredith Goldstein of the *Boston Globe* who leads a discussion with Bousquet, artist and musician Laurie Anderson, producer Roma Baran, author Leslie Gaston-Bird, and Dr. Stephanie Hull of Girls, Inc.

**SPECIAL EVENT**
**Abbey Road Spatial Audio Forum:**
**Audio Production in the Metaverse**
**Wednesday, May 26, 6:45 pm – 8:00 pm**

Presenters: **Stephen Barton**, Afterlight
**Anastasia Devana**, Hear XR

Gavin Kearney, University of York
**Muki Kulhan**, Muki-International, Ltd.
**Ana Monte**, DELTA Soundworks
**Mirek Stiles**, Abbey Road Studios

Join the Abbey Road Studios Spatial Audio Forum in an exclusive round-table discussion about the latest technical innovations in spatial and immersive audio productions, with a focus on virtual creativity and how these metaverses could connect musicians and producers in real-time worlds, anywhere in the world.

**SPECIAL EVENT**
**Women in Audio: Today's Leaders**
**Wednesday, May 26, 7:15 pm – 8:00 pm**

Moderator: **Leslie Gaston-Bird**

Presenters: *Phebean Adedamola Oluwagbemi*
*Erin Barra-Jean*
*Karrie Keyes*
*Ebonie Smith*
*Terri Winston*

During the past decade, organizations focused on women and marginalized genders have been working to uplift these populations with workshops, mentoring, internships, networking, and other opportunities. Learn about the activities and philosophies that are making a measurable impact in the audio world. Moderated by Leslie Gaston-Bird, panelists include Terri Winston of Women's Audio Mission, Ebonie Smith from Gender Amplified, Karrie Keyes from SoundGirls.org, Erin Barra of Beats by Girlz, and Phebean Adedamola Oluwagbemi of Audio Girl Africa.

**SPECIAL EVENT**
**Q&A Session: Women in Audio: Today's Leaders**
**Wednesday, May 26, 8:00 pm – 8:15 pm**

Presenters: **Phebean Adedamola Oluwagbemi**
**Erin Barra-Jean**
**Leslie Gaston-Bird**
**Karrie Keyes**
**Ebonie Smith**
**Terri Winston**

**PARTNER PRESENTATION, NEW AUDIO TECHNOLOGY**
**Dive into a New Immersive and Interactive Audio Dimension**
**Thursday, May 27, 11:00 am – 11:30 am**

Presenter: **Tom Ammermann**

Immersive and interactive audio and music is certainly a new challenging task. But is it just a fashion, what applications does it have and can we create it right from our common individual workflows? The session will give a short overview, shows New Audio Technology approaches, its production tools and strategies to deal with all of this issues quick and efficient.

**WORKSHOP/TUTORIAL: SOUND REINFORCEMENT**
**Can the Same Audience Experience Achieved with a Ground-Based Subwoofer System Be Delivered with a Flown Subwoofer System?**
**Thursday, May 27, 11:00 am – 12:00 noon**

Presenters: **Etienne Corteel**, L-Acoustics
**Adam Hill**, University of Derby
**Michael Lawrence**, Rational Acoustics
**Elena Shabalina**, d&b audiotechnik

Since the advent of the modern line arrays, it is common practice to fly the full-range sources of a main live sound reproduction system. Subwoofers, on the contrary, have remained ground-stacked primarily for practical reasons due to weight and the lack of captive rigging elements. Modern subwoofer designs however have partly alleviated these constraints.

This workshop compares ground-stacked against flown subwoofers in relation to the audience experience: level monitoring of low frequencies, health and safety measures relative to the exposition to high level of low frequencies, tonal balance and level distribution, subwoofer/main system time alignment over the audience, and the acoustical influence of the presence of the audience.

**SPECIAL EVENT: IMMERSIVE AUDIO**
**Genelec Special Webinar: Getting Started with Immersive Audio**
**Thursday, May 27, 11:00 am –12:00 noon**

Presenters: **Eric Horstmann**
**Markus Kahelin**

In this live session, Markus Kahelin covers the fundamentals of immersive audio, before discussing with Eric Horstmann how to actually create an immersive monitoring system.

Markus and Eric will explain topics including hybrid sound formats, Dolby Atmos for Music, channel and object-based audio, ambisonics and VR, and also discuss the fundamental differences between in-room monitoring and headphone listening in the context of immersive audio.

After the discussion there will be a live Q&A, so please bring your questions.

**SPECIAL EVENT**
**"Perceptual Audio coders—What to Listen for"— Launch of the Web Edition**
**Thursday, May 27, 11:30 am – 12:15 pm**

Presenters: **Sascha Dick**, International Audio Laboratories Erlengen, Fraunhofer IIS
**Christof Faller**
**Jürgen Herre**, International Audio Laboratories Erlangen, Fraunhofer IIS

In 2001, the AES Technical Committee on Coding of Audio Signals (TC-CAS) produced the legendary educational CD ROM "Perceptual Audio Coders—What To Listen For." It contains a taxonomy of common types of codec artifacts, as well as tutorial information on the background of each one. Example audio signals with different degrees of impairment illustrate the nature of the artifacts and help in training test listener expertise. Since its initial release, several generations of CD ROMs were sold and found worldwide use for education of the public.

This workshop presents the results of the TC's efforts in producing a second-generation educational package that tutors on new artifact types as they can be typically experienced with advanced audio codec processing, such as bandwidth extension or parametric stereo. Moreover, the format of the material was enhanced for seamless display and playback on PCs, tablets, and mobile phones and includes interactive graphics elements. This makes it an attractive educational package that is now available as an AES publication.

**WORKSHOP/TUTORIAL: ACOUSTICS & PSYCHOACOUSTICS**
**Dimensions of Immersive Auditory Experience**
**May 27, 12:00 noon – 1:00 pm**

Presenter: **Hyunkook Lee**, University f Huddersfield

"Immersive" audio is a popularly used term today, and it is often regarded as synonym of 3D audio. But what does immersive mean exactly? There is currently a lack of consensus in how the term should be defined, and it is not yet clear what techniques are required to make audio content more immersive. This session will first explicate different dimensions of immersion as well as those of related concepts presence and involvement, identifying the source of confusion around the terms and provide a conceptual relationship among them. A universal conceptual model of immersive experience will then be introduced, and various context-dependent factors that might be associated with immersive "auditory" experience will be discussed with practical examples.

**WORKSHOP/TUTORIAL: NETWORKED AUDIO**
**Networked Music Performance for Musicians**
**Thursday, May 27, 12:15 pm – 1:00 pm**

Presenters: **Miriam Iorwerth**, University of the Highlands and Islands
**Rebeka Wilson**, Source Elements

This tutorial will examine some of the challenges and opportunities in Networked Music Performance from a musician's perspective in a home environment. The Covid-19 pandemic has highlighted the importance of playing music with other people in many people's lives, and playing online has allowed this to continue throughout enforced isolation. This tutorial will first look at the two main approaches: asynchronous (or the "virtual ensemble") and synchronous, including the pros and cons of each approach, the particular considerations around how to choose which approach to take, and musical examples.

The tutorial will then go on to focus on the synchronous approach—playing together in (near) real time. The Internet was not designed for real-time transmission of audio, and data packets may arrive late, or not at all, introducing latency and glitches to the received signals. Latency is a major consideration for musicians, and we will discuss the ways it can be used creatively. We will also discuss bandwidth issues, and the trade-offs around this, as well as the impact of different approaches to monitoring. We will give examples of accessible software that musicians can use in their own homes for Networked Music Performance.

**PAPER Q & A SESSION: ACOUSTIC MEASUREMENT 2**
**Thursday, May 27, 1:00 pm**

- **Clean Dialogue Loudness Measurements Based on Deep Neural Networks—***Christian Uhle,*[1,2] *Michael Kratschmer,*[1] *Alessandro Travaglini,*[1] *Bernhard Neugebauer*[3]
  [1] Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
  [2] International Audio Laboratories Erlangen, Germany
  [3] DSP Solutions, Regensburg, Germany

  Loudness normalization based on clean dialogue loudness improves consistency of the dialogue level compared to the loudness of the full program measured at speech or signal activity. Existing loudness metering methods cannot estimate clean dialogue loudness from mixture signals comprising speech and background sounds, e.g., music, sound effects or environmental sounds. This paper proposes to train deep neural networks with input signals and target values obtained from isolated speech and backgrounds to estimate the clean dialogue loudness. Furthermore, the proposed method outputs estimates for loudness levels of background and mixture signal, and Voice Activity Detection. The presented evaluation reports a mean absolute error of 1.5 LU for momentary loudness, 0.5 LU for short-term and 0.27 LU for long-term loudness of the clean dialogue given the mixture signal.
  *Paper 10479*

- **Pyloudnorm: A Simple Yet Flexible Loudness Meter in Python—***Christian J. Steinmetz, Joshua Reiss*, Queen Mary University of London, London, UK

  The ITU-R BS.1770 recommendation for measuring the perceived loudness of audio signals has seen widespread adoption in broadcasting. Due to its simplicity, this algorithm has now found applications across audio signal processing. Here we describe pyloudnorm, a Python package that enables the measurement of integrated loudness following the recommendation. While a number of implementations are available, ours provides an easy-to-install package, a simple interface, and the ability to adjust the algorithm parameters, a feature that others neglect. We outline the design of pyloudnorm and discuss a set of modifications based upon recent literature that improve the robustness of loudness measurements. We perform an evaluation comparing accuracy and runtime with six other implementations, demonstrating that pyloudnorm is both fully compliant and one of the fastest options.
  *Paper 10483*

- **Measurement of the Particle Velocity in Front of Kick Drums in Three States—***Gabe Herman, Christopher Jasinski*, University of Hartford, West Hartford, CT, USA

  The goal of this research is to better understand the precise difference in acoustical output when the front head of a kick drum is ported, removed, or left on. In the studio and live settings, engineers may decide to remove the front head of a drum for microphone access, or because the drummer prefers it to be recorded a certain way. However, no formal study has ever measured how the drum propagates sound differently when the front head is on, off, or on with a dedicated port-hole. To conduct this experiment, a robotic actuator was constructed and multiple measurements were taken in BK Connect Software using an Acoustic Camera system by Brüel and Kjær. Data collected for this project was recorded in an anechoic chamber at the University of Hartford's Acoustics Lab, and in the Hartt School's Recording Studio Live Room. This paper documents the testing methods used to collect preliminary data and presents initial findings. This is an ongoing research project.
  *Paper 10499*

**WORKSHOP/TUTORIAL: RECORDING & PRODUCTION**
**Mentorship in Mastering**
**Thursday, May 27, 1:00 pm – 2:00 pm**

Moderator: **Piper Payne**

Presenters: *Anna Frick*, Airshow
*Margaret Luthar*, Welcome to 1979
*Maria Rice*, Peerless Mastering
*Jessica Thompson*, Jessica Thompson Audio

How would your career look with your mentor/mentee experience were different?

**WORKSHOP/TUTORIAL: GAME AUDIO AVAR SPATIAL AUDIO**

**Audio in Games & Interactive Media**
**Thursday, May 27, 1:00 pm – 2:15 pm**

Presenters: **Sarah Fartuun** Heinze
**Mathilde Hoffmann**, mathildesound.de
**Winifred Phillips**, Generations Productions LLC
**Katja Rogers**, University of Waterloo

Audio in many forms is an important part of the interactive media like games: sound effects and music have a substantial effect on players' experience of games. In this panel, several experts in industry and academia will hold a semi-structured chat. The panel members consist of several experts working in the industry with a background in music composition and sound effects design for games and other interactive media, and a game audio researcher with empirical work in game audio for PC and VR games. The questions will involve the process of audio design for games, how they design or compose for specific experiences and to minimize replay fatigue, how they communicate about the audio they're aiming for, how music can create or break immersion, what factors distract from game audio, and what makes audio in games particularly satisfying.

**PAPER Q & A SESSION: ELAC/AUDIO TECH**
**Thursday, May 27, 2:00 pm**

- **Lumped Parameter Thermal Model for Fast Voice Coil Temperature Prediction**—*Luca Villa, Chiara Corsini, Grazia Spatafora, Emiliano Capucci, Davide Mele, Romolo Toppi*, Faital S.p.A.,San Donato Milanese, Milan, Italy

  Voice coil (VC) temperature prediction is extremely important to make choices on the magnetic circuit during the pre-design and design phases. The aim of this study is to develop a model able to provide a quick response on the steady-state VC temperature. A lumped parameter thermal model describing the relevant heat transfer phenomena occurring in a working loudspeaker was developed. It was validated against measurements of VC, polar plate and magnet temperatures reached after two hours of sinusoidal or pink noise excitation. Differences between predicted and measured temperatures were not significant. By providing a fast and relatively accurate response, this model can be used for VC temperature prediction in the (pre-) design phase.
  *Paper 10458*

- **Temperature Stability of Neodymium Magnets in Voice Coil Transducers**—*Roland Jacques, Claudia Bardt, Meike Faulhaber, Kurt Jürgen Mick*, Sennheiser electronic GmbH & Co. KG, Wedemark, Germany

  Sintered neodymium magnets enjoy continued popularity in electro-acoustic transducers due to their high energy density, which enables the design of lightweight and efficient products. Various, quasi-standardized grades are available which differ in remanence, coercivity, and temperature stability. The latter is depicted in a series of B-H curves in the material datasheet; correctly interpreting them requires a detailed calculation of the complete magnetic circuit. In typical voice coil magnet systems, the presence of the magnetic circuit has quite a favorable effect, preventing thermal demagnetization in a certain temperature range above the nominal temperature limit. Methods and experimental data regarding this question are presented in this paper. Furthermore, the progression of thermal demagnetization over time is researched and presented, and shows interesting characteristics which can be quite relevant in practice, both for cost efficiency and for environmental impact.
  *Paper 10460*

- **Characterizing Non-Linear Behavior of Coupling Capacitors through Audio Feature Analysis and Machine Learning**—*Christopher Johann Clarke, Balamurali B T Jer-Ming Chen*, Singapore University of Technology and Design, Singapore

  Different electrically-equivalent capacitors are known to impact the sonic signature of the audio circuit. In this study the non-linear behavior of five different coupling capacitors of equivalent capacitance (marketed as "audio capacitors"), one at a time, are characterized. A dataset containing the input and output signals of a non-linear amplifier is logged, its audio features are extracted, and the non-linear behavior is analyzed. Machine learning is then applied on the dataset to supplement analysis of the Total Harmonic Distortion (THD). The five capacitors' THD performance seem to fall into two categories: below 200 Hz, there is significant standard deviation of 14.1 dBc; above 200 Hz, the capacitors show somewhat similar behavior, with only 0.01 dBc standard deviation. This separation, however, does not hold at regions below 0.2 V. A support vector machine model is trained and classifies the five capacitors well above chance: the best classification at 84% and worst at 36%. The methodology introduced here may also be used to meaningfully assess the complicated behavior of other audio electronic components.
  *Paper 10463*

- **Tremolo Effect—Optimizing The Use of Handmade Photocouplers**—*Gabriel Celso Kulevicz da Silva*,[1] *Edilberto Costa Neto*,[2] *Sidnei Noceti Filho*[3]
  [1] CELESC: Santa Catarina Central Electric Company, Brazil
  [2] Ciclix E.S.D.E. Ltda., Brazil
  [3] Federal University of Santa Catarina, Brazil

  A circuit suggestion for measurements on handmade photocouplers and its application in a tremolo effect pedal is presented. Design equations are presented, allowing designers to modify the circuits in order to adapt them to their personal taste. These circuits are built with low-cost components easily found in the market. Several practical aspects are discussed.
  *Paper 10487*

- **Comparing the Effect of Different Open Headphone Models on the Perception of a Real Sound Source**—*Christian Schneiderwind, Annika Neidhardt, Dominik Meyer*, Technische Universität Ilmenau, Ilmenau Germany

  The present work investigates the influence of different open headphone models on the perception of real sound sources in augmented audio reality applications. A set of binaural room impulse responses was measured with a dummy head wearing eight different open headphone configurations. A spectral error analysis showed strong deviations between the physical distortions in the real sound field caused by the different headphone models. The resulting perceptual effects were evaluated in a MUSHRA-like psychoacoustic experiment. The results show that all headphones introduce audible distortions. The extra-aural BK211 was found to be the one with the least audible corruption. In contrast, for some of the circumaural headphones strong coloration occurred and the spatial cues of

the real sound sources were seriously affected.
*Paper 10489*

**WORKSHOP/TUTORIAL: GAME AUDIO AVAR SPATIAL AUDIO**
**Innovative Uses of Spatial Audio Technology for Composition and Performing Arts: Wave Field Synthesis and "Sound Holograms"**
**Thursday, May 27, 2:00 pm –3:00 pm**

Moderators: **Bobby McElver**, UC San Diego
**Marcela Rada**, Agonquin College

Presenters: *Wen-Chi Su*, **YiLab**
*Chloe Thompson*, Artist, Sound Designer, Spatial Audio
*Nina Young*, Universityof Southern California

Wave Field Synthesis (WFS) is a spatial audio rendering technique that places virtual sound sources in real space. Using high density arrays consisting of approximately 200–600 discrete loudspeakers, it is possible to place sound sources accurately in physical space in front of the speakers—in short, create sound holograms or "holophones."

While this technology has long been thought of as logistically impossible, there have been a number of systems created in the last few years with the rise of audio-over-IP.

This panel discussion focuses on composers and other creators in the performing arts who have been leading the way in how to use this new spatial audio technology for artistic expression.

The conversation will begin with a brief introduction about the technology, but focus mainly on how artists are using it and why "sound holograms" have caused a fundamental shift in how they think about making artistic work with spatial audio. In the past, everyone in a listening experience hears everything at the same time. Now there can be individual sonic experiences in a live event without headphones. The technology is incredibly flexible and there is enormous room for creativity.

The panelists include composers, sound designers, and a choreographer who have worked closely with WFS. The three projects discussed focus on how the artists are using the technology differently. One is a concert with a roaming audience who walk inside of the sound sources (accompanied by beams of light). One is a seated audience hearing sounds whisper in their ears and moving through them. And the other is a dance in which the sound of the dancer's movement is separated from his body like a ghost, in a conceptual piece about gravity.

**WORKSHOP/TUTORIAL: RECORDING & PRODUCTION**
**Advances in Realism in Sampled Orchestra Performance**
**Thursday, May 27, 2:15 pm – 3:00 pm**

Presenter: **Claudius Bruese**

Ideally an orchestral score should be performed/recorded by an orchestra of musicians playing together. However, due to limitations on budget, time, and perhaps global pandemic conditions, this is sometimes not possible.

The author has worked since the 1990s with orchestras of sampled instruments and has followed and lived through all of the technological changes affecting sampling, sample playback, and advances in computer and musical instrument technology. Where the ultimate deliverable is a sampled orchestra performance, he has always prioritised realism in the performance.

This tutorial will showcase his work through the years, with audio examples to explain how advances in technology have translated into improvements in realism.

**PAPER Q & A SESSION: AUDIO QUALITY/STANDARDS 1**
**Thursday, May 27, 3:00 pm**

• **Dynamic Range Improvement in Digital to Analog Conversion via Multi-Path Topology—***Jon La Grou*, Millennia Media, SPC, Diamond Springs, CA, USA

Conventional digital-to-analog conversion (DAC) is accomplished via a single processing path that must optimize broadband noise level against maximum output level, i.e., dynamic range. By splitting DAC processing into two or more discrete elements or "paths," and passively recombining the analog resultants, order of magnitude improvement in dynamic range and linearity can be realized. Following a brief historical review, this paper will explore design details and experimental results on the author's multi-path DAC prototypes, followed by experiments that assess multi-path design margins which exceed required psychoacoustic delivery parameters.
*Ebrief 645*

• **Microphone Cross-Talk Cancellation in Ensemble Recordings with Maximum Likelihood Estimation—** *Orchisama Das, Julius O. Smith, III, Jonathan S. Abel*, Stanford University, Stanford, CA, USA

While recording an ensemble of musicians, it is often desired to isolate the instruments to avoid interference from other sources. Close-miking and acoustic isolation booths are some techniques for mitigating microphone cross-talk (or "bleed"). In this paper we propose an algorithm for canceling microphone bleed in ensemble recordings in the post-processing stage. We propose a calibration stage to estimate the relative transfer function from each instrument to each mic. Then, we set up an optimization problem to simultaneously estimate the transfer functions and individual sources given the microphone signals and a noisy estimate of the transfer function obtained from the calibration stage. We show that minimizing this cost function gives us the maximum likelihood estimate when we assume the distributions to be normal. Finally, we test our proposed method to cancel microphone bleed in a synthesized environment, and compare our results to an existing multichannel Wiener filter method.
*Paper 10471*

• **Efficient Data Collection Pipeline for Audio Machine Learning of Audio Quality—***Christer P. Volk,*[1] *Jon Nordby,*[2] *Tore Stegenborg-Andersen,*[1] *Nick Zacharov*[1]
[1] FORCE Technology, SenseLab, Hørsholm, Denmark
[2] Soundsensing, Oslo, Norway

In this paper we study the matter of perceptual evaluation data collection for the purposes of machine learning. Well established listening test methods have been developed and standardized in the audio community over many years. This papers looks at the specific needs for machine learning and seeks to establish efficient data collection methods, that address the requirements of machine learning, while also providing robust and repeatable perceptual evaluation results. Following a short review of efficient data collection techniques, including the concept of data augmentation and introduce the new concept of pre-augmentation as an alternative efficient data collection approach. Multiple stimulus presentation style listening tests are then presented for the evaluation of a wide range of audio quality devices (headphones) evaluated by a panel of

trained expert assessors. Two tests are presented using a traditional full factorial design and a pre-augmented design to enable the performance comparison of these two approaches. The two approaches are statistically analyzed and discussed. Finally, the performance of the two approaches for building machine learning models are reviewed, comparing the performance of a range of baseline models.
*Paper 10488*

• **Optimizing the Cinema Sound Experience for a Domestic Setting**—*Ahmed Shalabi*, Edinburgh Napier University, Edinburgh, UK; Overdub Productions, London, UK

Film sound mixing is an iterative process where dubbing mixers combine several stems to form the final soundstage of a motion picture. For cinema: end to end control makes it possible to reproduce near consistent mixes at the playback stage, whereas in a domestic setting: playback hardware and listening levels vary greatly in addition to background sounds or noise in the listening space. While mixers have a general consensus on how their mixes translate in a domestic setting; there are a number of considerations to take into account when making mixing decisions at the post-production stage with little documented on best approaches for the task. This paper describes key considerations and requirements to take into account when mixing film sound to ensure translation in a domestic setting and a listening experiment for the purpose of testing different mixing approaches. Two different film mixing approaches were done.
*Ebrief 638*

**WORKSHOP/TUTORIAL: EDUCATION**
**Education Panel: From Tonmeister to Today**
**Thursday, May 27, 3:00 pm – 4:00 pm**

Presenters: **Jim Anderson**, NYU
**Ulrike Schwarz**

The first Tonmeister program was founded in 1949 at the Detmold Hochschule für Muisik. The unique concept of education in music combined with audio recording continues to this day in Detmold and at schools and universities around the world. How has the tonmeister maintained its tradition and how has the concept changed with continuing evolutions in technology and music itself? In "From Tonmeister to Today," eight prominent international audio educators speak with Tonmeister Ulrike Schwarz and discuss their individual programs. A brief Q and A will follow the video.

**STUDENT & CAREER DEVELOPMENT EVENT**
**Student Recording Competition: Sound for Visual Media**
**Thursday, May 27, 3:00 pm – 4:00 pm**

The AES Student Recording Competition is a unique opportunity for student attendees of AES International Conventions to receive feedback and recognition for their audio production work.

Finalists will be announced and prizes awarded during this presentation. **Judge Panelists include:** *Fei Yu, Scott Hirsch, Luke Klingensmith*.

**PAPER Q & A SESSION: PSYCHOACOUSTICS**
**Thursday, May 27, 4:00 pm**

• **Sound Localization Training and Auditory Adaptation: A Review**—*Federico Lorenzo Martin*, Aural Escuela, Ciudad Autónoma de Buenos Aires, Argentina

The individualization of HRTFs is often considered the most relevant factor in the application of binaural techniques. However, the literature reviewed in this paper demonstrates that not necessarily the degree of HRTF customization determines the plausibility of the system. Multiple contextual factors define the plausibility in binaural audio perception, such as head tracking, the existence of reverberant information in the signal, the degree of divergence between the virtual sound environment and the real listening environment, and auditory adaptation processes. Training in spatial localization has a positive influence as it generates adaptation to spectral cues different from one's own, facilitating the construction of new spatial maps through sensory feedback.
*Paper 10464*

• **Towards an Audio Attribute Framework for Understanding the Perception of Reverberant Spaces Elicitation and Clustering Methods Based on Participant Expectation**—*Luke Child, Natanya Ford, University of the West of England, Bristol, UK*

When used in perceptual audio evaluation, elicitation methods produce a wide variety of raw and unorganized text data. Although at first ambiguous, elicited data can be organized into themes and attributes that are intrinsic to the listener experience. This paper seeks to compare the trends found in descriptions of reverberant locations from memory, isolating key attributes and phrases present in descriptions. These attributes are then cleaned, validated, and clustered to form a series of key parent attributes that encompass the descriptions of the original attributes. Methods for the optimization of each stage are discussed, alongside applications for understanding and utilizing the attributes in future implementations of digital reverberation.
*Paper 10467*

• **Predicting Audio Quality for Different Assessor Types Using Machine Learning**—*Christer P. Volk,*[1] *Jon Nordby,*[2] *Tore Stegenborg-Andersen,*[1] *Nick Zacharov*[1]
[1] FORCE Technology, SenseLab, Hørsholm, Denmark
[2] Soundsensing, Oslo, Norway

In this paper we study how sound quality is evaluated by different groups of assessors, with different levels of hearing loss. Formal listening tests using the Basic Audio Quality scale were designed using 22 headphones spanning a wide range of qualities and sound quality characteristics. The tests were performed with two formally selected listening panels with normal hearing (NH), and mild (N2) or moderate (N3) hearing loss characteristics. It is shown that not only do the two panels evaluate the sound quality consistently within each panel, but also that there are systematic changes in the manner in which hearing loss impacts the evaluation and ranking of the devices under study. Using this data we successfully train machine learning algorithms to predict the sound quality for the two assessor type panels. The prediction performance for each panel is NH: RMSE = 7.1 ± 3.0, PCC = 0.91 ± 0.13; HI: RMSE = 8.7 ± 2.4, PCC = 0.91 ± 0.12. While it may not be practical to run listening tests with multiple panels of assessors, we demonstrate here that machine learning based models can be practically and cost effectively employed to predict the perception of multiple assessor groups rapidly and simultaneously.
*Paper 10494*

• **Evaluation of Hearing Threshold Using a New**

**Response Method for Remote Hearing Profile Detection**—*Maksims Mironovs, Sergejs Sarkovskis*, Sonarworks Ltd, Riga, Latvia

With the rapid technological growth and availability of IOT devices, hearing loss can happen at an accelerated rate and becomes an increasing problem. The current hearing threshold curves, defined in the ISO 389-7 standard, do not take this into account and require a revision. This Engineering Brief will present the results of an ongoing hearing loss study using a novel remote hearing threshold detection method and evaluate its effectiveness. It is a modification of Bekesy's tracking method that is performed remotely on mobile devices to collect results on a large scale and simplify the measurement process. The results showed that the proposed method is effective for hearing loss detection and indicated that new hearing threshold curves can be defined. As the test is being actively distributed, by the time of publication it is expected to have sufficient subject size to represent the general population.
*Ebrief 643*

**WORKSHOP/TUTORIAL: GAME AUDIO AVAR SPATIAL AUDIO**
**360 Ecosystem with Hank Shocklee**
**Thursday, May 27, 4:00 pm – 5:00 pm**

Presenters: **Brian (Bt) Gibbs**, Skyline Entertainment & Publishing
**Paul Womack**, Willie Green Music
**Hank Shocklee**

*What Is the 360 Ecosystem:* **Producers, beat makers, programmers can now begin to imagine the delivery/distribution/presentation of their works at even earlier stages of creation. We'll discuss how artists, arrangers, composers, and songwriters fit into this Ecosystem so they might present their music in a new way to their fans (consumers). Ecosystem Idea: Artist > Producer > Mixing Engineer & Immersive Engineer -then- Producer > Artist > Fans (Consumers through Immersive Playback).**

*How:* Taking close microphone/console audio of a stereo 2-mix and reimagining into Dolby Atmos (Avid Play for Distribution) or Sony RA360 (Orchard for Distribution). Potential future state: using spatial microphones early in the production/writing process to capture the true sonics of what an artist, arranger, composer or songwriter hears in their mind as the musical work develops.

*Why:* Consumers probably already have the technology in their pocket without knowing it's even available to them. Fans can access playback options through streaming services, using speakers/sound bars/headphones (i.e., AirPod Pro2 & AirPod MAX which are now Atmos enabled). Will artists, arrangers, composers & songwriters start to change their thought processes/workflow in the earliest writing stages with the idea of immersing their listeners at inception.

What consumers choose for playback (e.g., speaker(s) vs headphones/binaural) can have impact genre to genre. Every musical style should have access to present their music in these cutting edge formats.

This presentation of the 360 Ecosystem will use a musical work from legendary artist & producer Hank Shocklee to reimagine an original stereo 2-mix into both the Dolby Atmos and Sony RA360 formats.

**WORKSHOP/TUTORIAL: RECORDING & PRODUCTION**
**Creating Podcasts During a Pandemic**
**Thursday, May 27, 4:00 pm – 5:00 pm**

Presenters: **Steve Bone**, Vice Audio
**Jay LeBoeuf**, Descript
**Matthew Shaer**
**Liana Simstrom**, NPR

Over the last year we've seen a huge paradigm shift in how listeners consume content. While many media channels have taken a hit during COVID and will have to continue refocusing their digital efforts, the future of the Podcast looks bright. Creators are adapting to new and updated tools, and listeners are hungrier than ever for fresh content.

In this tutorial we'll uncover how daily and weekly podcasts, reaching millions of listeners, are being created at some of the top media outlets. Guests will include producers and editors from shows at NPR, Campside Media, and VICE. We'll also break down the gear being used including remote recording solutions, editing tools, publishing and collaboration platforms.

**SPECIAL EVENT: HEYSER LECTURE**
**Thursday, May 27, 5:00 pm – 6:00 pm**

Lecturer: **Diana Deutsch**, University of California San Diego

The Heyser Series is an endowment for lectures by eminent individuals with outstanding reputations in audio engineering and its related fields. The series is featured twice annually at both the United States and European AES Conventions. Established in May 1999, The Richard C. Heyser Memorial Lecture honors the memory of Richard Heyser, a scientist at the Jet Propulsion Laboratory, who was awarded nine patents in audio and communication techniques and was widely known for his ability to clearly present new and complex technical ideas. Heyser was also an AES governor and AES Silver Medal recipient.

The Richard C. Heyser distinguished lecturer for the 150th AES Convention is Diana Deutsch.

**Two Perceptual Puzzles: Audio Illusions and Perfect Pitch**

Illusions are often regarded as entertaining anomalies that shed little light on the normal process of perception. In this talk I argue that the contrary is true. Just as the failure of a piece of equipment provides important clues to its successful operation, so illusions provide important information about the brain mechanisms that generally enable us to perceive the world correctly.

Some auditory illusions show that people can differ strikingly in how they hear even simple musical patterns. These differences occur as strongly among expert musicians as among people without musical training. In illusions involving stereo sound—such as the octave illusion, the scale illusion, and the glissando illusion —striking perceptual disagreements tend to arise between right-handers and left-handers, showing that they reflect differences in brain organization. In contrast, perception of the tritone paradox varies with the language or dialect to which the listener has been most frequently exposed.

The speech-to-song illusion demonstrates a strong relationship between speech and music. A spoken phrase is made to morph perceptually from speech to song, without transforming the sounds in any way, or by adding any musical context, but simply by repeating the phrase several times over. The illusion shows that the boundary between music and speech is fragile, and an explanation for the illusion is proposed.

The talk also discusses perfect pitch—the ability to name a musical note when it is presented out of context. This ability is very rare in the Western world, where non-tone language is spoken, but is far more prevalent among speakers of tone languages such as Mandarin, in which the meaning of a word depends on the pitch (or pitches) in which it is spoken. The reasons for this advantage to tone language speakers are discussed.

The talk is accompanied by sound demonstrations.

Diana Deutsch is Professor of Psychology at the University of

California, San Diego, She is internationally known for the musical illusions and paradoxes that she discovered; these include the octave illusion, the scale illusion, the glissando illusion, the tritone paradox, the cambiata illusion, the phantom words illusion, and the speech-to-song illusion, among others. She also explores memory for music and how we relate the sounds of music and speech to each other. In addition she studies absolute pitch—why some people possess it and why it is so rare.

Deutsch has over 200 publications, including *Musical Illusions and Phantom Words: How Music and Speech Unlock Mysteries of the Brain* (2019), *The Psychology of Music*, (1st edition, 1982; 2nd edition (1999), 3rd edition (2013), and the compact discs *Musical Illusions and Paradoxes* (1995) and *Phantom Words and Other Curiosities* (2003). She has been elected a Fellow of the American Association for the Advancement of Science, the Acoustical Society of America, the Audio Engineering Society, the Society of Experimental Psychologists, the American Psychological Society (renamed the Association for Psychological Science), and the American Psychological Association. She received the Rudolf Arnheim Award for Outstanding Achievement in Psychology and the Arts from the American Psychological Association, the Gustav Theodor Fechner Award for Outstanding Contributions to Empirical Aesthetics from the International Association of Empirical Aesthetics, the Science Writing Award for Professionals in Acoustics from the Acoustical Society of America, and the Gold Medal Award from the Audio Engineering Society for "lifelong contributions to the understanding of the human hearing mechanism and the science of psychoacoustics."

**SPECIAL EVENT**
**Introducing the TC-MLAI**
**Thursday, May 27, 6:00 pm – 7:00 pm**

Presenters: **Brecht De Man**, Semantic Audio Labs
**Andy Sarroff**, iZotope, Inc.
**Gordon Wichern**, Mitsubishi Electric Research Laboratories (MERL)
**Christian Uhle**, Fraunhofer IIS

The AES Technical Council has identified that the community of audio engineers working with Machine Learning (ML) and Artificial Intelligence (AI) is underrepresented in the technical committees. The newly formed AES Technical Committee on Machine Learning and Artificial Intelligence (TC-MLAI) intends to answer this need.

The TC-MLAI focuses on applications of machine learning and artificial intelligence in audio, with discussions on topics such as: best practices, data, licensing, social and cultural aspects, technical innovations, and ethics. The goal of the committee is to drive discussion and exchange information by organizing workshops, symposia, tutorials, and technical documents. It will also act as a point of contact and a bridge to other AES technical committees, the AES community at large, and other organizations involved in ML and AI for audio.

In this workshop we will present the committee's mission, values, agenda, and avenues for membership and participation. We will highlight exciting developments and trends as they relate to audio, while at the same time acknowledging topics of controversy, such as data bias, privacy concerns, and when it is appropriate to call an audio technology "artificially intelligent."

The workshop will be composed of four parts: introducing the technical committee, including its mission, values, and membership; providing a brief overview of the state of ML and AI in audio; facilitating a panel discussion about what ML and AI mean in the context of audio engineering; and hosting an open Q&A session with workshop attendees.

**STUDENT & CAREER DEVELOPMENT EVENT**
**How to Get a Job in the Audio Industry**
**Thursday, May 27, 6:00 pm – 7:00 pm**

Presenters: **Joy Lyons**, Harman International Industries
**Cheryl Ottenritter**, Ott House Audio
**Rebeka Wilson**, Source Elements

The question on almost all audio engineering graduates' minds: how to get a job in the audio industry? The panelists will explain good strategies for getting a start and tips for how to stand out in the crowd.

**WORKSHOP/TUTORIAL: GAME AUDIO AVAR SPATIAL AUDIO**
**Master Class on Spatial Audio in Unity for VR**
**Thursday, May 27, 6:00 pm – 7:15 pm**

Presenters: **Jeanine Cowen**, Berklee College of Music
**Marcela Rada**, Algonquin College

This workshop provides a review of the current ecosystem of spatial audio tools while working on Virtual Reality (VR) applications in the Unity interactive game engine. The workshop will begin with the extent and limitations of native Unity audio implementation features. The workshop will then continue to look at outside addons and packages that are purpose built for spatial audio delivery in a VR application. Participants will better understand the various factors that can inform choices to be made while working on audio for VR applications.

The process for working in interactive audio has always included specific and material differences compared to linear media projects. Virtual Reality (VR) audio increases this complexity due to differing workflow and final distribution expectations. This workshop describes and presents the considerations necessary to be efficient and successful while working with audio within the Unity game engine on a VR project.

**WORKSHOP/TUTORIAL: RECORDING & PRODUCTION**
**AI in Audio**
**Thursday, May 27, 7:00 pm – 8:00 pm**

Moderators: **Bobby Lombardi**, PACE Anti-Piracy, Inc.
**Heather Rafter**, RafterMarsh

Presenters: *Sehmon Burnam*, Google Research, Google
*Wisam Reid*, Harvard University
*Alexander Wankhammer*, sonible

A panel of leading experts in Artificial Intelligence (AI) discuss its impact in audio engineering applications.

**SPECIAL EVENT**
**SoundGirls.org Mentoring Session: Audio for Live Production**
**Thursday, May 27, 7:00 pm – 8:00 pm**

Presenters: **Amanda Davis**
**Freyja Lawson**
**Daniela Seggewiss**

SoundGirls.org's Netherlands Chapter hosts a mentoring session on music production featuring professionals in the field. Please come prepared to ask questions and gain valuable insight.

**IRON MOUNTAIN ENTERTAINMENT SERVICES PARTNER PRESENTATION**
**Nurturing the Next Generation of Archivist Talent**
**Friday, May 28, 11:00 am – 12:00 noon**

Presenters: **Bethany Boarts**
**Jen Greenwood**
**Tom Holderness**
**Alex Tomlin**

This session will include four members of the IMES team—Bethany Boarts, Imaging Project Manager at the IMES Boyers digital studio; Jen Greenwood, IMES UK Music Account Manager; Tom Holderness, IMES Junior Audio Engineer; and Alex Tomlin, IMES Senior Manager of Engineering for Western Europe—co-presenting "interview-style," discussing their pathway, chosen field, and what future media archivists should do to prepare for the future of the discipline—all centered around the idea of stressing the importance of safely archiving our cultural heritage assets through music.

**WORKSHOP/TUTORIAL: BROADCAST & ONLINE DELIVERY**
**A Discussion of the Legal Issues of Streaming**
**Friday, May 28, 11:00 am – 12 noon**

Presenter: **Chris Cooke**, + MD, CMU

The streaming music business model has been in the spotlight over the last year as COVID negatively impacted on most of the music industry's revenue streams, except premium streaming.

Streaming services pay fractions of a penny per stream, but there are billions of streams, so the total monies paid over to the music industry are signifciant. But how is that money shared out? Some argue the current business model is unfair. And how does any of this apply the the emerging livestreaming business?

Chris Cooke, Founder + MD of London-based music business consultancy CMU—and author of the book *Dissecting The Digital Dollar*—explains how it all works from a rights and royalties perspective, to help you navigate and understand the bigger debates.

**KLIPPEL PARTNER PRESENTATION**
**Friday, May 28, 12 noon – 12:30 pm**

KLIPPEL introduces the new SCN Near-Field Add On (SCN-NF), which accurately performs all the most relevant transducer measurements using a single hardware setup in a non-anechoic room. Add 2pi acoustical measurement capabilities to the same SCN platform you already use to laser scan diaphragm vibration. The SCN hardware is extended with a microphone in addition to the existing laser sensor. Automated axis control ensures repeatable, precise and fast positioning of microphone and laser sensors. In combination with a round baffle for measuring transducers up to 10" / 30 cm in diameter or compact (smart) speakers, acoustic near-field scanning technology is added to the SCN.

**WORKSHOP/TUTORIAL: ARCHIVING & RESTORATION**
**NEMOSINE—The Future of Media Storage:**
**Follow-Up from the Lab**
**Friday, May 28, 12:15 pm – 1:00 pm**

Presenter: **Nadja Wallaszkovits**, abk Stuttgart

NEMOSINE is an EU founded project for the development of Innovative packaging solutions for storage and conservation of 20th century cultural heritage of artefacts based on cellulose derivatives.

The objective of project NEMOSINE www.nemosineproject.eu is to improve traditional storage solutions by developing an innovative package with the main goal of energy saving and extending the lifetime of cultural objects based on cellulose derivatives. In contrast to conventional film cans or media boxes, the packages will be equipped with the latest sensor technology to monitor decomposition processes and adsorb decomposition products such as acetic acid. The focus is on films, photographs, posters, slides, cinematographic sound, magnetic tapes and discs, based on cellulose acetate and its derivatives. The aim of the four-year project is to achieve more efficient long-term archiving and to increase the life cycle of audiovisual media, as well as other objects of cultural heritage and arts. The complete solution for storage boxes proposed by NEMOSINE is based on multi-nano sensors for different gases (mainly acetic acid and nitric oxide) and a control software platform that simulates degradation processes and then will predict accurate protective treatments.

The project has been introduced already at the last European AES Virtual Vienna Convention. This updated tutorial will now outline the various innovative developments, focussing on the latest insights, such as the development of multi-nanosensors for different gases (mainly acetic acid and nitric oxide) and a control software platform that simulates degradation processes and then predicts accurate protective treatment for the archivist.

**WORKSHOP/TUTORIAL: GAME AUDIO AVAR**
**SPATIAL AUDIO**
**Guiding Audiences with Sound: Techniques for Interactive and Games Audio**
**Friday, May 28, 12:30 pm – 1:15 pm**

Presenter: **Lucy Harrison**, AMC

Within interactive sound and music composers and sound designers are presented with an interesting additional challenge, human behavior.

Interactive structures such as video games and immersive theater events give audiences a chance to explore spaces freely and control their narrative experience. While this free exploration is appealing to audiences, it can lead to narrative fragmentation where audiences miss key parts of storylines that are vital for their understanding of the work.

As a solution, audio can be used to help guide the audience to ensure that they are able to access all relevant areas of the narrative while still retaining exploratory control.

This tutorial will provide practical approaches outlining how sound and music can be used to guide audiences through physical or virtual spaces. It will look at elements such as sound placement, strategic use of frequencies and semiotic associations related to sound and music which build on existing knowledge about audience behaviours in immersive theatre and games. These techniques can be easily adapted to any genre to provide composers and sound designers with effective approaches to take into any of their interactive work.

**PAPER Q & A SESSION: AMBISONICS**
**Friday, May 28, 1:00 pm**

- **Higher Order Ambisonics Compression Method Based on Independent Component Analysis—** *Jiahao Xu, Yadong Niu, Xihong Wu, Tianshu Qu*, Peking University, Beijing, China

  Recently the development of multimedia applications requires a flexible method to represent spatial sound, and higher order ambisonics (HOA) draws more and more attention due to its flexibility between the recording and playback end. To reduce the cost for storage and transmission, some compression methods were developed. However, they result in discontinuity between frames. Here we propose a framework in which independent component analysis (ICA) is used to extract foreground components from HOA signal. We achieve smooth transition by utilizing un-mixing matrices from previous frames. A complete compression system was constructed and a dataset with

simulated and recorded signals was built. Subjective experiments provided evidence for the effectiveness of proposed method.
*Paper 10456*

• **Ambisonic Decoder Test Methodologies Based on Binaural Reproduction—***Enda Bates, William David, Daniel Dempsey*, ADAPT Centre, School of Engineering, Trinity College Dublin, Ireland

The comparative evaluation of the quality of different Ambisonic decoding strategies presents a number of challenges, most notably the lack of a suitable reference signal other than the original, real-world audio scene. In this paper a new test methodology for the evaluation of Ambisonic decoders is presented, using a virtual loudspeaker, binaural rendering approach. A sample study using a MUSHRA test paradigm and three different types of Ambisonic decoders was conducted and the results analyzed using a variety of different statistical approaches. The results indicate significant differences between decoders for some attributes and virtual loudspeaker layouts.
*Paper 10457*

• **Evaluation of Six Degrees of Freedom 3D Audio Orchestra Recording and Playback Using Multi-Point Ambisonics interpolation—***Tomasz Ciotucha,*[1] *Andrzej Ruminski,*[1] *Tomasz Zernicki,*[1] *BartłomiejMróz*[1,2]
[1] Zylia sp. z o. o., Poznan, Poland
[2] Gdansk University of Technology, Gdansk, Poland

This paper describes a strategy for recording sound and enabling six-degrees-of-freedom playback, making use of multiple simultaneous and synchronized Higher Order Ambisonics (HOA) recordings. Such a strategy enables users to navigate in a simulated 3D space and listen to the six-degrees-of-freedom recordings from different perspectives. For the evaluation of the proposed approach, an Unreal Engine-based navigable 3D audiovisual playback system was implemented. Subjective listening tests were conducted which compare the quality of the prepared walk-through.
*Paper 10459*

• **Room Geometry Estimation from Higher-Order Ambisonics Signals Using Convolutional Recurrent Neural Networks—***Nils Poschadel, Robert Hupke, Stephan Preihs, Jürgen Peissig*, Leibniz University Hannover, Institute of Communications Technology, Hannover, Germany

Knowledge of room geometry is a fundamental component for modeling acoustic environments. Since most common methods for room geometry estimation are based on prior knowledge, the generalization to unknown environments is somewhat limited. Deep learning based approaches have delivered promising results for the blind estimation of acoustic parameters considering mainly monaural signals. The purpose of this contribution is to investigate the effect of multichannel higher-order Ambisonics (HOA) signals on the performance of a convolutional recurrent neural network for blind room geometry estimation. Therefore a HOA-dataset of noisy speech signals in simulated rooms with realistic frequency-dependent reflection coefficients is introduced. Results show that for each additional Ambisonics order the estimation performance increases with the fourth-order model achieving a mean absolute error of 1.24 m averaged over all three room dimensions.
*Paper 10482*

• **Optimized Decoders for Mixed-Order Ambisonics—**
*Aaron Heller,*[1] *Eric Benjamin,*[2] *Fernando Lopez-Lezcano*[3]
[1]SRI International, Menlo Park, CA, USA
[2]Surround Research, Pacifica, CA, USA
[3]Stanford University, Stanford, CA, USA

In this paper we discuss the motivation, design, and analysis of ambisonic decoders for systems where the vertical order is less than the horizontal order, known as mixed-order Ambisonic systems. This can be due to the use of microphone arrays that emphasize horizontal spatial resolution or speaker arrays that provide sparser coverage vertically. First, we review Ambisonic reproduction criteria, as defined by Gerzon, and summarize recent results on the relative perceptual importance of the various criteria. Then we show that using full-order decoders with mixed-order program material results in poorer performance than with a properly designed mixed-order decoder. We then introduce a new implementation of a decoder optimizer that draws upon techniques from machine learning for quick and robust convergence, discuss the construction of the objective function, and apply it to the problem of designing two-band decoders for mixed-order signal sets and non-uniform loudspeaker layouts. Results of informal listening tests are summarized and future directions discussed.
*Paper 10507*

**WORKSHOP/TUTORIAL: PRODUCT DEVELOPMENT**
**Career Path as an Audio Engineer in the World of the Audio Software Industry**
**Friday, May 28, 1:00 pm – 1:45 pm**

Presenters: **Angelika Polola**, Native Instruments GmbH
**Maryam Safi**, Steinberg Media Technologies GmbH

The job market in the audio software industry is constantly growing and expanding with new opportunities. It is one of the most dynamic environments to work in! At the same time, it is challenging us to use the new technologies for designing new products or to adapt current ones. Therefore, to succeed in this industry, merely maintaining professional competence is not always the key. It is important to have innovative ideas, stay creative, and keep yourself up-to-date on market demands and the needs of users. In this tutorial we would like to talk about different roles in the audio software industry as well as discussing the strong relationship between companies and their customers; why user-centered methods, as well as continuous research, are useful and necessary for today's software development in professional audio fields.

**WORKSHOP/TUTORIAL: EDUCATION**
**A Complete Guide to Networked Music Performance Using Free and Open-Source Software**
**Friday, May 28, 1:15 pm – 2:00 pm**

Presenter: **Rebeka Wilson**, Source Elements

This workshop will review a thorough representation of the currently active open source and freely available software projects that allow for networked music performance. Even prior to the pandemic, the development of tools that meet the requirements of performing music together online were growing in number; since the last year, we have seen new tools and existing ones have become more sophisticated and powerful, as a unprecedented level of attention is paid to their use and application by those musicians and music groups who traditionally work together offline.

Networked music performance has three particular demands of the technology it uses: first, the audio stream must be suitably high-quality; second, it must minimize interference, for example avoiding echo cancellation algorithms and unnecessary processing; third it must be low-latency, where what defines low-latency exactly depends on the intention and resources of the musicians. Given these factors, the variety of applications and services available that can be used for network performance each have their own specific approach, usually born from the original intended purpose of the developer. Certain tools may, for example, focus on low-latency while others on group usability, and others on solutions where participants do not have access to broadband or 4G.

The outcome of this workshop will be that viewers who are interested in performing together over the internet will receive a comprehensive review of software tools, with the aim to select those that suit their specific needs. In addition, the tools discuss are free-to-use and, in some cases, modify. The availability of such tools leads to an expansive array of music possibilities that extends the core of music practice itself.

**WORKSHOP/TUTORIAL: ARCHIVING & RESTORATION**
**Audio Archiving and Restoration: A Look at Different Facets of Safeguarding Our Musical Heritage**
**Friday, May 28, 1:45 – 3:00 pm**

Moderator: **Nadja Wallaszkovits**, abk Stuttgart

Presenters: *Ilse Assmann*, Apricity Consulting
*Marie del Carmen Ordoño Vidaña*, Fonoteca Nacional de Mexico
*Jessica Thompson*, Jessica Thompson Audio
*Melissa Widzinski*, Library of Congress

The panel leads through various perspectives in the field of audio preservation and restoration. Introducing the work of international communities and networks, the panel will present some glimpses into the variety of related and specialized fields. The presenters will discuss the challenges of in- and outside the world of professional production—from the perspective of a National Audio-Visual Conservation Center to preserving independent musicians' work, small labels and foundations who have historically important recordings but no means or resources for preservation, to safeguarding the unique recordings of Mexican indigenous radio and the related challenges, to the importance of audio archiving and restoration in the context of research accessibility and also in the field of arts preservation.

**PAPER Q & A SESSION: AUDIO QUALITY/STANDARDS 2**
**Friday, May 28, 2:00 pm**

- **Further Insights on the Influence of a Dynamic Binaural Synthesis on Speech Intelligibility in TETRA-coded Voice Communication—** *Nils Poschadel, Mahdi Alyasin, Stephan Preihs, Jürgen Peissig*, Leibniz University Hannover, Institute of Communications Technology, Hannover, Germany

  Within the project VIA2mobiL, we developed signal processing algorithms and methods for a dynamic headphone-based binaural synthesis, with a special focus on radio communication in mobile control centers. Our aim was to achieve a better speech intelligibility in radio communication through the binaural presentation of a conversation scene with several separately locatable interlocutors. In our investigations, a method for determining word recognition rates and 50% speech reception thresholds (SRTs) was developed on the basis of the methodology of the Oldenburg sentence test (OLSA). By means of conducted listening experiments, we examined whether the application of a dynamic binaural synthesis results in a gain in speech intelligibility if compared to monophonic or stereophonic reproduction, in particular for TETRA-coded speech. We present the study design and the results of the experiments that show that considerably higher word recognition rates can be achieved with a binaural synthesis compared to a monophonic and stereophonic reproduction at the same level, for both TETRA-coded and uncoded speech. Furthermore, significantly lower 50% SRTs in the presence of noise were observed for 3D audio compared to a monophonic and stereophonic reproduction.
  *Paper 10484*

- **LC3 and LC3plus: The New Audio Transmission Standards for Wireless Communication—***Markus Schnell,*[1] *Emmanuel Ravelli,*[1] *Jan Büthe,*[1] *Maximilian Schlegel,*[1] *Adrian Tomasek,*[1] *Alexander Tschekalinski,*[1] *Jonas Svedberg,*[2] *Martin Sehlstedt*[2]
  [1] Fraunhofer IIS, Erlangen, Germany
  [2] Telefonaktiebolaget LM Ericsson, Stockholm, Sweden

  The new Low Complexity Communication Codec (LC3) and its sibling Low Complexity Communication Codec Plus (LC3plus) were developed to solve essential shortcomings present in today's short-range wireless communication platforms such as Bluetooth and Digital Enhanced Cordless Telecommunications (DECT). The codec operation modes range from medium bit rates for optimal voice transmission to high bit rates for high-resolution music streaming services. Furthermore, the codecs operate at low latency, low computational complexity, and a low memory footprint.
  *Paper 10491*

- **Scattering Iterative Method Based on Generalized Wave Variables for the Implementation of Audio Circuits with Multiple One-Port Nonlinearities—***Davide Albertini, Alberto Bernardini, Augusto Sarti*, Politecnico di Milano, Milan, Italy

  The Scattering Iterative Method (SIM) is a recently developed fixed-point method relying on Wave Digital (WD) principles for the discrete-time simulation of electrical networks containing multiple one-port nonlinearities. Due to its robustness and efficiency, SIM proved itself to be suitable for the digital emulation of nonlinear audio circuits in Virtual Analog applications. The existent SIM formalization uses voltage wave variables. In this paper we extend such a formalization to accommodate circuit descriptions based on generalized wave variables, including voltage, current, and power-normalized waves, as particular cases. A SIM-based WD implementation of a passive audio compressor employing the newly introduced generalized wave framework is presented, along with an analysis of the SIM convergence speed considering different types of waves and two different initialization strategies.
  *Paper 10492*

- **Automatic Impulse Response Matching for Reverb Plugins—***Andrew Cunningham, Kirk McNally, Peter Driessen*, University of Victoria, BC, Canada

  A system is proposed for automatically tuning the parameters of an algorithmic reverb plugin with the goal of matching the output of another reverb plugin, which can include convolution-based models. The system accepts two VST or VST3 type reverb plugins, and without prior knowledge of the implementation of either plugin, attempts to find a parametrization for the algorithmic reverb plugin that minimizes the difference between the measured impulse responses of both plugins.
  *Paper 10503*

**WORKSHOP/TUTORIAL: GAME AUDIO AVAR SPATIAL AUDIO**
**Immersive Audio Techniques for Beginners**
**Friday, May 28, 2:00 pm – 3:00 pm**

Presenter: **Marcela Rada**, Algonquin College

This masterclass will focus on how to create a spatial audio mix from scratch. It will begin with an overview of the hardware commonly used to create immersive experiences such as ambisonic microphones. It will discuss the state of the art and the type of platforms that support spatial audio. This masterclass will also go over signal flow, and the installation of a variety of spatial audio plugins. The goal of this masterclass is to present beginners with affordable tools such as Reaper and the FB360 Spatial Workstation, so they can easily begin producing content from home.

**PAPER SESSION: MUSIC ANALYSIS**
**Friday, May 28, 3:00 pm**

• **Sequential Modeling of Temporal Timbre Series for Popular Music Sub-Genre Analyses Using Deep Bidirectional Encoder Representations from Transformers—** *Shijia Geng, Gang Ren, Xu Pan, Joel Zysman, Mitsu Ogihara*, University of Miami, FL, USA

The timbral analysis from spectrographic features of popular music sub-genres (or micro-genres) presents unique challenges to the field of the computational auditory scene analysis, which is caused by the adjacencies among sub-genres and the complex sonic scenes from sophisticated musical textures and production processes. This paper presents a timbral modeling tool based on a modified deep learning natural language processing model. It treats the time frames in spectrograms as words in natural languages to explore the temporal dependencies. The modeling performance metrics obtained from the fine-tuned classifier of the modified Deep Bidirectional Encoder Representations from Transformers (BERT) model show strong semantic modeling performances with different temporal settings. Designed as an automatic feature engineering tool, the proposed framework provides a unique solution to the semantic modeling and representation tasks for objectively understanding of subtle musical timbral patterns from highly similar musical genres.
*Paper 10470*

• **Timbre-Based Machine Learning of Clustering Chinese and Western Hip Hop Music—***Rolf Bader, Axel Zielke, Jonas Franke*, Institut of Systematic Musicology, University of Hamburg, Germany

Chinese, Taiwanese, and Western Hip Hop musical pieces are clustered using timbre-based Music Information Retrieval (MIR) and machine learning (ML) algorithms. Psychoacoustically motivated algorithms extracting timbre features such as spectral centroid, roughness, sharpness, sound pressure level (SPL), flux, etc., were extracted form 38 contemporary Chinese/Taiwanese and 38 Western "classical" (USA, Germany, France, Great Britain) Hip Hop pieces. All features were integrated over the pieces with respect to mean and standard deviation. A Kohonen self-organizing map, as integrated in the Computational Music and Sound Archive (COMSAR[6]) and apollon[1] framework was used to train different combinations of feature vectors in their mean and standard deviation integrations. No mean was able to cluster the corpora. Still SPL standard deviation perfectly separated Chinese/Taiwanese and Western pieces. Spectral flux, sharpness, and spread standard deviation created two sub-cluster within the Western corpus, where only Western pieces had strong values there. Spectral centroid std did sub-cluster the Chinese/Taiwanese Hip Hop pieces, where again only Chinese/Taiwanese pieces had strong values. These findings point to different production, composition, or mastering strategies. E.g. the clear SPL-caused clusters point to the loudness-war of contemporary mastering, using massive compression to achieve high perceived loudness.
*Paper 10473*

• **Automatic Audio Source Classification System for Recordings Captured with Microphone Array—***Michał Chrul,*[1] *Andrzej Ruminski,*[1] *Tomasz Zernicki,*[1] *Ewa Łukasik*[2]
[1] Zylia sp. z o. o., Poznan, Poland
[2] Gdansk University of Technology, Gdansk, Poland

The aim of this paper was to create an automatic sound source classification framework for recordings captured with a microphone array and evaluate the sound source separation algorithm impact on the classification results. The preprocessing related to the said evaluation concerned convolving the dataset samples with impulse responses captured with a microphone array, as well as mixing the samples together to simulate their co-presence in a virtual recording scene. This way, the evaluation of the separation algorithm impact on classification results was possible. Furthermore, such approach saved multiple hours of labor that would need to be spent on the recording process itself. Finally, the classification results delivered by different models were evaluated and compared.
*Paper 10481*

**WORKSHOP/TUTORIAL: EDUCATION**
**Education Panel: Coding**
**Friday, May 28, 3:00 pm – 4:00 pm**

Presenters: **Christopher Bennett**, University of Miami
**Eric Tarr**

A discussion on introducing audio coding into our respective curricula and resources for educators to do the same.

**STUDENT & CAREER DEVELOPMENT EVENT**
**Student Recording Competition: Modern Studio Recording & Electronic Music**
**Friday, May 28, 3:00 pm – 4:00 pm**

The AES Student Recording Competition is a unique opportunity for student attendees of AES International Conventions to receive feedback and recognition for their audio production work.

Finalists will be announced and prizes awarded during this presentation. Judge Panelists include: *Magdalena Piotrowska, Daniel Cantor, Marek Walaszek, Mandy Parnell*.

**SPECIAL EVENT: KEYNOTE ADDRESS**
**Creative and Alternative Audio Adventurers**
**Friday, May 28, 4:00 pm – 5:00 pm**

Presenter: **Lenise Bent**, Soundflo Productions

This presentation is to celebrate individuals who currently are using sound and audio technology in creative and alternative ways that enhance the human experience.

Some examples:

Roy Allela, an engineer from Kenya who has created gloves that turn sign language into audible speech.

Ellen Reid, creator of "Soundwalk," a GPS enabled work of public art that uses music to illuminate the natural environment. She has composed musical works to stroll through Central Park and hike in Los Angeles' Griffith Park.

Jonathan Ward, an audio archeologist who collects 78rpm records from around the world including Nigeria, Madacascar, and Panama. Author of *Excavated Shellac, an Alternate History of World Music.*

Brian Bushway, Greatest Blind Mountain Biker and master of Echolocation, how blind people see with sound.

**WORKSHOP/TUTORIAL: PRODUCT DEVELOPMENT**
**The Boring Allpass Filter?**
**Friday, May 28, 5:00 pm – 5:45 pm**

Presenter: **Jayant Datta**

When one thinks of filters, one thinks of classical filters in traditional engineering textbooks—or example lowpass, highpass, bandpass, and bandstop filters. **In the field of audio, one encounters other filters as well.**

In this tutorial we take a deeper look at allpass filters—where phase (instead of frequency shaping) is used as a manipulating tool. We start with the basics of allpass filters and look at how they may be used for audio processing. Then we look at how allpass filters may be used as building blocks for more complicated topologies to create even more interesting filters.

**WORKSHOP/TUTORIAL: BROADCAST & ONLINE DELIVERY**
**Streaming as the Future of High Resolution Audio Distribution**
**Friday, May 28, 5:00 pm – 6:30 pm**

Moderator: **Vicki Melchior**, Consultant, Audio DSP and Software

Presenters:
*Pal Bratelund*, Room Labs LLC
*Chris Horton*, Universal Music Group
*Mike Jbara*, MQA
*Ty Roberts*, Ty Robers Innovation

Streaming is now the dominant method of distributing audio, including high resolution audio (HRA), to consumers. Although HRA has been a small percent of the total, the major labels and the RIAA predicted from 2017 onward that high quality was of strong interest even to young listeners and should thrive given streaming's new affordability and portability.

This workshop looks at current uptake and future directions in high quality audio streaming. Important to both sustainability and growth in this dynamic area are the evolving nature of music distribution infrastructure and the ability for all players in the area—major and indie labels, music streaming services, and designers of platform-spanning software and hardware—to address important issues. Those include growth models expanding to younger listeners, adequate remuneration of artists including indies while limiting costs, worldwide bandwidth constraints, differentiation from one another, and innovative provision for user requests like performance data, music discovery, radio, podcasts, etc. A very exciting new direction is live streaming, that for the first time permits HD video to be combined with HD audio.

**WORKSHOP/TUTORIAL: RECORDING & PRODUCTION**
**Audio for Social Justice**
**Friday, May 28, 5:45 pm – 6:30 pm**

Presenter: **Helen Caddes**, Campaignly Group

In light of recent events in America, we have learned that audio is a powerful catalyst of social justice. As the world heard the last words of George Floyd and the shock spanned the globe, the tragedy of his death in the larger scope of police brutality in America was recognized. Today, high quality audio recording is available to anyone with a cellular telephon—we need to talk about how that is changing the world today and what our role as audio professionals could be.

We are living in a social media driven age where substance is often hidden by external appearances. As a songwriter who learned the craft in order to share my own message through music, I soon learned others had crucial information to share and started helping them to tell their stories through the use of professional audio techniques. Marches, conferences, and direct action are taking place every day across the world, but may not be properly documented without quality audio. I worked on the wrongful conviction case of a woman named Kirstin Blaise Lobato where a single piece of audio was a crucial piece of evidence.

When setting out to help with audio for social justice, there are a number of challenges any audio professional may face:

- Budget
- Audio recorded improperly
- Single channel audio
- Static and mic noise
- Audio overloaded by improper mic placement
- Limited audio knowledge

Though editing tasks may be cumbersome and the availability of equipment onsite for events may be unpredictable, there are many gifts our community has to offer that could make the difference between a successful social media event and a failure for the visionaries who need it most. I will explore these topics in depth and will be happy to do a follow up presentation to expand on this.

**WORKSHOP/TUTORIAL: BROADCAST & ONLINE DELIVERY**
**Importance of Loudness**
**Friday, May 28, 6:30 pm – 8:00 pm**

Presenters:
**David K. Bialik**
**Eelco Grimm**
**John Kean**
**Scott Norcross**
**Robert Orban**
**Jim Starzynski**

A discussion of Loudness for Broadcast and Streaming. We will have a summary of AES73, CTA 2075 and a preview of the Recommendations for Distribution Loudness of Internet Audio Streaming and On-Demand File Playback.

**WORKSHOP/TUTORIAL: RECORDING & PRODUCTION**
**FXpertise: Expansion and Gating**
**Friday, May 28, 6:30 pm – 8:00 pm**

Presenter: **Alex U. Case**, University of Massachusetts Lowell

Compression stole the dynamic effects spotlight. In this tutorial we're stealing it back. Your mixes will benefit from creative applications of that 'other' dynamics processor: the expander/gate. While offering all the virtues of expanded dynamic range, it has the power to create a far wider variety of effects. Expanders are a tool for altering timbre, reshaping sounds, synthesizing new ones, overcoming masking, and fabricating the impossible. Parameters with names like attack, hold, release, decay, range, depth, slope, and side chain filters don't exactly invite creative exploration. This overview of effects strategies brings structure to the sonic possibilities of expansion and gating so that you can quickly find the parameter settings that let you achieve your production goals.

**WORKSHOP/TUTORIAL: GAME AUDIO AVAR SPATIAL AUDIO**

**Spatial Audio Productions for Music & XR Applications**
**Friday, May 28, 6:45 pm – 8:00 pm**

Presenters: **Zoran Cvetkovic**, King's College London
**Enzo De Sena**, University of Surrey
**Huseyin Hacihabiboglu**, METU
**Muki Kulhan**, Muki-International, Ltd.

In AES's first-ever "open viewing" workshop, you'll be able to watch a selected group of participants take part in a LIVE, "hands-on" session producing immersive, spatial sound experiments using the innovative tools of Braud.io software. In under an hour they'll get properly stuck in and learn about new ways to create immersive audio experiences to enhance music and XR storytelling productions. The workshop will be hosted by Muki Kulhan, AES veteran and Executive XR Producer who has created in-depth immersive audio experiences and valuable R&D in the worlds of music and broadcasting (BOSE, MTV, The National Gallery & more), with very special guests from Braud.io, listed below.

Capacity: Spaces limited, up to 10 participants, open to ALL AES delegates on a first-come, first-served basis

Level: Beginner to Intermediate

Workshop Requirements: Available to Mac users only, must have Ableton and latest Zoom software and installed BRAUD.IO software (free license invite via separate email from workshop teacher)

**SPECIAL EVENT**
**Live Streamed Organ Research Event**
**Friday, May 28, 8:00 pm – 9:00 pm**

Presenter: **Rebeka Wilson**, Source Elements

Organists Hans Fidom and Trevor Grahl will stream performances live from the Orgelpark in Amsterdam, Netherlands. This special event is brought to you by Source Elements who is providing the live stream.

**SPECIAL EVENT**
**SoundGirls.org Mentoring Session: Audio Post for Film and TV**
**Friday, May 28, 8:00 pm – 9:00 pm**

Presenters: **Aline Bruijns**, AudioRally Sounddesign
**Twi McCallum**, Formosa Group
**Ana Monte**, DELTA Soundworkd
**Anna Sulley**, Sound As . . .
**Nene Veenman**, 3RD Season/Veenman & Morrison Composers

SoundGirls.org's Netherlands Chapter hosts a mentoring session on post production for film and television featuring professionals in the field. Please come prepared to ask questions and gain valuable insight.

## TC MEETINGS

Monday, May 31, 10 am
**Archiving, Restoration, and Digital Libraries**

Tuesday, June 1, 9:00 am
**Spatial Audio**

Tuesday, June 1, 10:00 am
**Acoustics and Sound Reinforcement**

Tuesday, June 1, 11:00 am
**Automotive Audio**

Tuesday, June 1, 11:00 am
**Coding of Audio Signals**

Tuesday, June 1, 12:00 noon
**Loudspeakers and Headphones**

Wednesday, June 2, 9:00 am
**Machine Learning (ML) and Artificial Intelligence (AI)**

Wednesday, June 2, 11:00 am
**Recording Technology and Practices**

Wednesday, June 2, 12:00 noon
**Perception and Subjective Evaluation of Audio Signals**

Thursday, June 3, 10:00 am
**Audio Forensics**

Thursday, June 3, 10:00 am
**Signal Processing**

Thursday, June 3, 12:00 noon
**High Resolution Audio**

Friday, June 4, 10:00 am
**Fiber Optics for Audio**

Friday, June 4, 11:00 am
**Semantic Audio Analysis**

Friday, June 4, 11:00 am
**Broadcast and Online Delivery**

## STANDARDS COMMITTEE MEETINGS

Wednesday, May 12, 5:00 pm
**SC-02-01 Digital Audio Measurement Techniques**

Wednesday, May 12, 6:00 pm
**SC-02-02 Digital Input Output Interfacing**

Thursday, May 13, 5:00 pm
**SC-03-06 Digital Library and Archive Systems**

Thursday, May 13, 6:00 pm
**SC-03-12 Forensic Audio**

Friday, May 14, 5:00 pm
**SC-05-05 Grounding and EMC Practices**

Friday, May 14, 6:00 pm
**SC-05-02 Audio Connectors**

Monday, May 17, 5:00 pm
**SC-04-08 Measurement of Sound Systems in Rooms**

Tuesday, May 18, 5:00 pm
**SC-02-12 Audio Applications of Networks**

Tuesday, May 18, 6:30 pm
**SC-03-07 Metadata of Audio**

Wednesday, May 19, 5:00 pm
**SC-04-03 Loudspeakers Modeling and Measurement**

Wednesday, May 19, 6:30 pm
**SC-02-08 Audio-File Transfer and Exchange**

Thursday, May 20, 5:00 pm
**SC-04-04 Microphone Measurement and Characterization**

Thursday, May 20, 6:00 pm
**SC-04-09 Loudness and Annoyance**

Friday, May 21, 5:00 pm
**AESSC Plenary**