AES 146TH CONVENTION PROGRAM

MARCH 20–MARCH 23, 2017

THE CONVENTION CENTRE DUBLIN, IRELAND

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 146th 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 146th AES Convention Student Paper Award is:

Factors Contributing to Gender Imbalance in the Audio Industry—
Shelley Ann McCarthy Buckingham, Malachy Ronan, Limerick Institute of Technology, Limerick, Ireland

Convention Paper 10159

To be presented in paper Session P08: Industry Issues on Thursday, March 21

* * ** *

The Winner of the 146th AES Convention Best Peer-Reviewed Paper Award is:

Background Ducking to Produce Esthetically Pleasing Audio for TV with Clear Speech—*Matteo Torcoli,[1] Alex Freke-Morin,[1,2] Jouni Paulus,[1,3] Christian Simon,[1] Ben Shirley[2]*
[1] Fraunhofer IIS, Erlangen, Germany
[2] University of Salford, Salford, UK
[3] International Audio Laboratories Erlangen, Erlangen, Germany

Convention Paper 10175

To be presented on Friday, March 22 in Session 12—Speech

Technical Committee Meeting
Technical Committee Meeting on Spatial Audio
Wednesday, March 20, 09:00 – 10:30
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-02-12-R Task Group on Streaming Audio Metadata over IP
Wednesday, March 20, 09:00 – 10:30
Wicklow Meeting Room 5, Level 2

This group will define a standardized method for transporting metadata associated with audio in an AES67 stream in a separate parallel stream. It shall define synchronization between the audio metadata transport and the associated AES67 transport. The transmission method shall be low latency and have a level of network performance equivalent to AES67. Within the scope is formatting of the streaming audio metadata for transport. Suggested is an open standards based framework that supports both static and dynamic, time synchronous metadata that is optimized for live workflow applications. The standard shall consider all use cases for metadata associated with AES67, support existing AES audio metadata standards, and be extensible for future metadata requirements. The standard will consider binding between the audio metadata transport and the associated AES67 transport.

Tutorial 01 **Wednesday, March 20**
09:15 – 10:45 **Liffey Hall 2**

BLUETOOTH AUDIO AND THE CAR

Chair: **Jonny McClintock**, Qualcomm Technology International Ltd., Belfast, Northern Ireland, UK

Panelists: *Francesco Condorelli*
Jayant Data
Richard Hollinshead

The latest cars have advanced sound systems and many people use their mobile phones to access music while driving, yet relatively little attention has been given to the wireless audio link between the phone and the car sound system. This tutorial will describe the use of Bluetooth with enhanced audio to connect in-car entertainment systems and the benefits this will bring to users. These developments will ensure that drivers and passengers can enjoy CD quality, or better, without using cables, audio and rear seat passengers can enjoy synced gaming audio.

Many drivers are spending more time in the car with some commuters regularly stuck for hours in rush hour traffic. Sound and in-car entertainment systems are therefore very attractive to drivers and a valuable differentiator for car manufacturers. Advanced sound systems, which in some cars cost several thousand dollars and include more than 10 speakers, are a standard feature on some new cars, an important upgrade option on most cars, and represent a significant aftermarket opportunity.

The quality of sound enjoyed by the driver and passengers depends not only on the quality of the sound system but also the quality of the audio source and the connectivity between the audio source and the sound system. Drivers have relied on Satelliteor DAB radio and multidisc CD systems to provide audio in cars. This is changing in a world where Internet radio, streamed music and playlists stored on mobile phones are the audio sources chosen by users at home, on the move and in their cars. Bluetooth connectivity has been available in cars for 15 years and all smart phones support Bluetooth audio. Most new cars now come with Bluetooth connectivity and the rest provide Bluetooth as an option.

Workshop 01 **Wednesday, March 20**
09:15 – 10:45 **Liffey Hall 1**

LOUDSPEAKER RELIABILITY TEST SIGNALS

Chair: **Steven Hutt**, Equity Sound Investments, Bloomington, IN, USA

Panelists: *Laurie Fincham*, THX
Wolfgang Klippel, Klippel GmbH,Dresden, Germany
Phil Knight, Phil Knight Associates, UK
Richard Little, Goertek Electronics Inc., Santa Clara, CA, USA
Roger Schwenke, Meyer Sound Laboratories, Alameda, CA, USA
Alexander Voishvillo, JBL/Harman Professional Solutions, Northridge, CA, USA

Test Signals used for loudspeaker system reliability testing utilize tailored attributes such as band width, duty cycle, and crest factor to evaluate input capacity for loudspeaker systems. While noise can be shaped to emulate music's spectrum and crest factor, music has correlated time/frequency event cycles that impose different thermal and mechanical stress on a loudspeaker. Panelists will discuss correlation of music signals to different test signals as specified in AES and IEC standards or alternatives and comment on how anticipated duress of a loud-speaker system in the field affects useful life.

This session is presented in association with the AES Technical Committee on Loudspeakers and Headphones

Tutorial 02 **Wednesday, March 20**
09:45 – 10:45 **Liffey A**

DANCE MUSIC TRAINING—THE UNREGULATED INDUSTRY AND CPD—ANALYTICAL LISTENING SKILLS

Presenters: **Alexandra Bartles**, Dance Music Production, Wakefield, West Yorkshire, UK
Rick Snoman, Dance Music Production, Barnsley, South Yorkshire, UK

Electronic Dance music has grown into a multi-billion dollar industry but the training remains unregulated. The standards of training material vary greatly, ranging from entertainment to educational. Anyone can set up an online tutorial delivery service and there are no regulations within the industry. This has seen the creation of organizations that have no proven track record. Students find themselves working through a minefield of misinformation, learning and practicing misguided dogmas and not learning how to be innovative but instead how to copy. EMTAS developed a Code of Practice with the help of professionals in the dance music industry.

In this session, we look at the industry and how the code of practice will improve educational standards and help to create a generation of individuals that can produce original electronic dance music. We will then perform an example of continued professional development which is a requirement within the code of practice.

Critical and Analytical Listening (CPD)

The most fundamental skill that all music producers need is the ability to listen both critically and analytically. While there are numerous blogs and software programs, and phone apps, teaching us how to critically listen. Little information is given on how to listen with an analytical ear. Yet without analytical listening, we cannot create an emotional connection with our listening audience.

Technical Committee Meeting
Technical Committee Meeting on Fiber Optics for Audio
Wednesday, March 20, 10:00 – 11:00
Wicklow Meeting Room 4, Level 2

Session P01 **Wednesday, March 20**
10:30 – 12:30 **Meeting Room 3**

LOUDSPEAKERS—PART 1

Chair: **Christof Faller**, Illusonic GmbH, Zurich Switzerland

10:30

P01-1 Large Horns and Small Rooms—Do They "Play Nicely" Together?—*Bjørn Kolbrek*, Celestion, Ipswich, UK

For some audiophiles, having a huge, low-cutoff bass horn built into the wall of the listening room represents the ultimate low frequency solution. Without considering the

practicalities of such an installation, this paper will look at the performance of low frequency horns mounted in the wall of a small room compared to the performance of a typical point source closed box type subwoofer and an array of such subwoofers. Simulation results indicate that in addition to higher efficiency, the horns provide smoother response in the listening position and less seat-to-seat variation.
Convention Paper 10132

11:00

P01-2 Predistortion Technique for Generating Spectrally Clean Excitation Signals for Audio and Electro-Acoustic Nonlinear Measurements—*Antonin Novak, Laurent Simon, Pierrick Lotton, Manuel Melon*, Université du Mans, Le Mans, France

In many audio and electro-acoustic nonlinear measurements we need to excite the nonlinear system under test with an excitation device that is not linear. A typical example is the study of the nonlinear behavior of a loudspeaker mechanical part, where the mechanical part (the nonlinear system under test) is excited externally, either with a shaker or pneumatically using another loudspeaker. We often consider that the excitation device is linear, which is unfortunately not correct. In this paper we present a simple method that corrects the distorted output signal of the excitation device by pre-distorting the input signal. The process is based on harmonic injection and can be applied to any periodic signal that is used for the measurement, e.g., a sine wave to measure the total harmonic distortion (THD), a two-tone signal to measure an intermodulation distortion (IMD), or a multi-tone signal. The experimental results provided on an electrodynamic loudspeaker show that the undesired spectral components of the acoustic pressure inside the sealed box can be suppressed to the level of the background noise.
Convention Paper 10133

11:30

P01-3 Sensory Profiling of High-End Loudspeakers Using Rapid Methods—Part 4: Flash Profile with Expert Assessors—*Irene Arrieta Sagredo,*[1] *Samuel Moulin,*[1] *Søren Bech*[1,2]
[1] Bang & Olufsen, Struer, Denmark
[2] Aalborg University, Aalborg, Denmark

This study is the fourth in a series of papers investigating different rapid sensory profiling methods applied to audio stimuli [1, 2, 3]. In particular, this paper considers Flash Profile, a verbal-based method that allows assessors to use their own vocabulary, for perceptual audio evaluation. A listening test was conducted with expert listeners investigating the ability of Flash Profile to describe and discriminate five sets of high-end loudspeakers. The influence of using different audio-stimuli in order to get a broader perceptual image is supported by doing a track by track analysis, using Multiple Factor Analysis [4, 5]. The results suggest that the differences between loudspeakers lie in two main dimensions related to the timbral and spatial characteristics of the stimuli. Flash Profile seem to be a time-efficient tool for visualization and reduction of perceptual dimensions, being useful for the description and discrimination of a set of audio stimuli with medium to small audible differences.
Convention Paper 10134

12:00

P01-4 Poster Introductions 1

- *Optimized Exciter Positioning Based on Acoustic Power of a Flat Panel Loudspeaker*—Benjamin Zenker; Shanavaz Sanjay Abdul Rawoof; Sebastian Merchel; Ercan Altinsoy
- *Practical Problems in Building Parametric Loudspeakers with Ultrasonic Piezoelectric Emitters*—Antonin Novak; Jose Miguel Cadavid Tobon
- *Time Stretching of Musical Instrument Tones*—Sean O'Leary

Session P02 **Wednesday, March 20**
10:30 – 12:20 **Meeting Room 2**

PERCEPTION

Chair: **Malachy Ronan**, Limerick Institute of Technology, Limerick, Ireland

10:30

P02-1 Comparison of Recording Techniques for 3D Audio Due to Difference between Listening Positions and Microphone Arrays—*Toru Kamekawa, Atsushi Marui*, Tokyo University of the Arts, Tokyo, Japan

The listening experiments comparing three recording techniques for 3D audio, namely Spaced Array, One-point Array, and Ambisonics were executed. First, the evaluation attributes were extracted referring the Repertory Grid Technique. Then participants compared the differences between these microphone techniques including the difference in listening position. From the results, the difference depending on the listening position is the smallest in the Spaced Array. Besides, it is estimated that Ambisonics gives the impression of "hard," One-point Array gives "rich" and "wide," and Spaced Array gives "clear" and "real." Furthermore, "real" was evaluated from the viewpoint of clarity and the richness of reverberation, with a negative correlation with the spectral centroid and a positive correlation with the reflection from lateral and vertical, respectively.
Convention Paper 10136

11:00

P02-2 [Paper moved to Session 8]

11:30

P02-3 Investigation into the Influence of Electromechanical Characteristics of Electrodynamic Transducers on Sound Quality Perception—*Semyung Son,*[1] *Juyoung Jeon,*[1] *Junbae Choi,*[1] *Mikhail Pakhomov*[2]
[1] Hyundai Mobis, Yong-in, Kyung-ki, Korea
[2] SPB Audio R&D Lab,St. Petersburg, Russia

The two most noticeable types of distortion in an audio signal path—frequency and nonlinear—are frequently analyzed by researchers and developers in terms of auditory perception. The effect of transient distortion, though insufficiently studied, is evident in subjective listening tests when comparing loudspeakers with similar frequency response and no audible nonlinear distortions. In the present study we conducted loudspeaker measurements and subjective evaluations to define the critical factors based on the loudspeaker's electromechanical characteristics that affect transient distortion and determined relations

between the factors' values and subjective scores.
Convention Paper 10135

12:00

P02-3 Poster Introductions 2

- *Audio-Driven Multimedia Content Authentication as a Service*—Nikolaos Vryzas; Anastasia Katsaounidou; Rigas Kotsakis; George Kalliris; Charalampos Dimoulas
- *ANC System Using Secondary Path Modeling Based on Driver's Position in Vehicle*—Seyeong Jang; Jongin Jung; Hyungsub Lim
- *Pop and Rock Music Audio Production for 22.2 Multichannel Sound: A Case Study*—Will Howie
- *Sound Recording Studio Renovation at the University of Victoria*—Bezal Benny; Kirk McNally

Standards Committee Meeting
SC-02-12-N Task Group on Media Network Directory Architecture
Wednesday, March 20, 10:30 – 12:00
Wicklow Meeting Room 5, Level 2

This task group is responsible for AES standards and standards-related activities for media network directories. Its scope includes media network device registration, media network device discovery, network membership administration and security activities, and the data models that support them. The group's current activity is AES project X238, a survey that aims to collect and document directory system requirements across the range of professional audio and video applications. The result of X238 will be a report to inform future directory standards activities.

Tutorial 03 — Wednesday, March 20
11:00 – 12:30 — Liffey A

SNARE DRUM STRATEGIES: RECORD AND MIX FOR MAXIMUM IMPACT

Presenter: **Alex Case**, University of Massachusetts Lowell, Lowell, MA, USA

The snare drum demands careful and constant attention—from our ears and our gear. With a dynamic range from whispering brushes to cannon fire, the snare drum challenges us to know our craft. Musical acoustics, room acoustics, and psychoacoustics guide the development of effective microphone techniques and signal processing strategies. Through informed control of spectrum, envelope, and image, the snare can be counted on to drive your music forward.

Tutorial 04 — Wednesday, March 20
11:00 – 12:30 — Liffey Hall 2

DIGITAL FILTERS, FILTER BANKS, AND THEIR DESIGN FOR AUDIO APPLICATIONS

Presenter: **Gerald Schuller**, Ilmenau University of Technology, IImenau, Germany; Fraunhofer Institute for Digital Media technology (IDMT), Ilmenau, Germany

This tutorial will teach you how to design "Finite Impulse Response" and "Infinite Impulse Response" filters for audio applications, in theory and practice and will give examples in the popular Open Source programming language Python. Then it will go on to show you how to design and use filter banks. Examples will be the "Modified Discrete Cosine Transform" (MDCT) filter bank, the (Integer-to-Integer) "IntMDCT", and Low Delay filter banks, which are widely used in MPEG audio coding standards. Further it will show digital filters as predictors for predictive coding, with applications in MPEG Lossless coding standards. Finally it will show how to implement filter banks as convolutional neural networks, which makes them "trainable", and eases the use of GPU's. This has applications for instance in audio source separation.

This session is presented in association with the AES Technical Committee on Signal Processing

Tutorial 05 — Wednesday, March 20
11:00 – 12:00 — Liffey Hall 1

TEACHING THE RECORDING STUDIO BEYOND THE CLASSROOM

Presenter: **Gabe Herman**, The University of Hartford, Hartford, CT, USA

One of the biggest challenges in teaching the recording studio environment is facilitating adequate access to learning labs and studio facilities to practice materials covered in class. This event will explore how innovative software, traditional studio tools, and new interactive programs can provide students new opportunities to continue applied learning outside of the classroom. Included in this presentation are critical listening in headphones and/or non-ideal loudspeaker monitoring environments with the use of innovative acoustical correction software, the use of plugins for learning the esoteric performance characteristics of outboard gear, as well as a proposed new approach to teaching and learning large-format analog mix consoles through HTML code based on research conducted at the Hartt School at the University of Hartford.

Technical Committee Meeting
Technical Committee Meeting on Microphones and Applications
Wednesday, March 20, 11:00 – 12:00
Wicklow Meeting Room 4, Level 2

Special Event
SE01 AWARDS PRESENTATION AND KEYNOTE ADDRESS
Wednesday, March 20, 12:45 – 14:45
Liffey A

Opening Remarks: • Executive Director Colleen Harper
• President Nadja Wallaszkovits
Convention Chairs Enda Bates, Ben Kok, Mariana Lopez
Program: • AES Awards Presentation
• Introduction of Keynote Speaker
• Keynote Address by Stefania Serafin

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 our industry.

BOARD OF GOVERNORS AWARD
- **Paul Gallo**
- **Toru Kamekawa**
- **John Krivit**
- **Valerie Tyler**
- **Christian Uhle**
- **Yuko Watanabe**

FELLOWSHIP AWARD
- **Joel Brito**
- **John Dawson**

SILVER MEDAL

• Jamie A.S. Angus-Whiteoak

Keynote Speaker

This year's Keynote Speaker is **Stefania Serafin**. She is a professor in sonic interaction design at Aalborg University Copenhagen. She received a Ph.D. from CCRMA, Stanford University in 2004 and has been working at Aalborg University since. Serafin is the president of the Sound and Music Computing Association. The title of her Keynote Address is, "Inclusive Sonic Interactions."

Sonic interaction design is a fertile field of investigation at the intersection of sound and music computing and interaction design. At Aalborg University I have been designing, together with several collaborators, sonic interactions for different applications, ranging from physics-based simulations of musical instruments and everyday objects, new interfaces for musical expression, cultural heritage, walking and rehabilitation, learning and training, as well as virtual and augmented reality. In this talk I will present an overview of these activities, with a focus on those who contribute to engage a wider and younger audience in sound, music and technology.

Technical Committee Meeting
Technical Committee Meeting on Audio for Games
Wednesday, March 20, 14:00 – 15:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-04-08 Working Group on Audio Applications of Networks
Wednesday, March 20, 14:00 – 15:30
Wicklow Meeting Room 5, Level 2

The scope of SC-04-08 includes the description, specification, measurement, and calibration of electroacoustic sound systems in rooms and the characteristics of sound presented to an audience.

Tutorial 06 **Wednesday, March 20**
14:15 – 15:15 **Liffey hall 2**

STRATEGIES IN 3D MUSIC PRODUCTION (UPMIXING VS PRODUCTIONS AIMED FOR 3D)

Presenter: **Lasse Nipkow**, Silent Work LLC, Zurich, Switzerland

More and more music productions in 3D / immersive audio are realized native. In contrast to upmix strategies, thereby the potential of 3D / immersive audio can be fully achieved musically,and it can be used to create new forms of creativity. However, it is not trivial to deal with the large number of audio channels meaningfully.

Lasse Nipkow presents his long-term findings and results with 3D audio in this event. He has been making 3D / immersive audio productions, recordings, and presentations for more than a decade at various conventions all over the world. During this session he will explore main aspects of spatial audio psychoacoustics and production strategies. Various sound and video examples will be shown to illustrate the context.

Tutorial 07 **Wednesday, March 20**
14:15 – 15:15 **Liffey Hall 1**

LTI FILTER DESIGN AND IMPLEMENTATION 101

Presenter: **Andrew Stanford-Jason**, XMOS Ltd., Bristol, UK

LTI filter design is a well documented process but the implementation specifics are rarely covered. In this tutorial we will cover how to realize a filter from its specification to the instructions that will execute. We will work through a PDM to PCM conversion example covering how to partition the decimation process to maximize the performance of your hardware while achieving your desired filter performance. The intention is to clearly illustrate the trade-offs a designer has to consider that are typically outside of the specification, how to test for them and how to control them.

Topics include: • Filter design basics: from filter specification to realization; • Designing for floating and fixed point arithmetic; • Optimizing for your hardware; • Characterizing your filter: theory and testing; • Decimation/Interpolation: polyphase filters; • Frequency domain implementation; • Example: PDM to PCM conversion.

This session is presented in association with the AES Technical Committee on Signal Processing

Workshop 02 **Wednesday, March 20**
14:15 – 15:30 **Liffey Hall 2**

METADATA AND WHY EVERY AUDIO ENGINEER NEEDS TO UNDERSTAND IT

Chair: **John Krivit**, Professional Audio Design, Hanover, MA, USA; Emerson College, Boston, MA, USA

Panelists: *Lynne Earls*, NARAS PE Wing, Ireland
Paul Jessop, County Analytics Ltd., Dunstable, Bedfordshire UK
Buddy Judge
Helienne Lindvall, VEVA Sound, London, UK
Drew Waters, VEVA Sound

This panel presents and demystifies the purpose, practical application, and value of capturing metadata at the point of inception—in the studio. We will engage all those with a vested interest in the collection and propagation of quality metadata, including; engineers, producers, production engineers, songwriters, performers, second engineers, and production coordinators across different genres and parts of the industry—thereby creating a cross section of contemporary voices to add insight to this crucial aspect of industry growth.

Session P03 **Wednesday, March 20**
14:30 – 16:30 **Meeting Room 2**

LOUDSPEAKERS—PART 2

Chair: **Alex Voishvillo**, JBL/Harman Professional Solutions, Northridge, CA, USA

14:30

P03-1 Green Speaker Design (Part 1: Optimal Use of System Resources)—*Wolfgang Klippel*, Klippel GmbH, Dresden, Germany

Increasing the efficiency and voltage sensitivity of the electro-acoustical conversion is the key to modern audio devices generating the required sound output with minimum size, weight, cost, and energy. Traditional loudspeaker design sacrifices efficiency for sound quality. Nonlinear adaptive control can compensate for the undesired signal distortion, protect the transducer against overload, stabilize the voice coil position, and cope with time-varying properties of the suspension. The paper presents a new design concept for an active loudspeaker system that uses the new degree of freedom provided by DSP for exploiting

a nonlinear motor topology, a soft suspension and modal vibration in the diaphragm, panel, and in the acoustical systems.
Convention Paper 10138

15:00

P03-2 Green Speaker Design (Part 2: Optimal Use of Transducer Resources)—*Wolfgang Klippel*, Klippel GmbH, Dresden, Germany

Green speaker design is a new concept for developing active loudspeaker systems that generate the required sound output with minimum size, weight, cost, and energy. This paper focuses on the optimization of the transducer by exploiting the new opportunities provided by digital signal processing. Nonlinear adaptive control can compensate for the undesired signal distortion, protect the transducer against overload, stabilize the voice coil position, and cope with time varying properties of the suspension. The transducer has to provide maximum efficiency of the electroacoustical conversion and sufficient voltage sensitivity to cope with the amplifier limitations. The potential of the new concept is illustrated on a transducer intended for automotive application.
Convention Paper 10139

15:30

P03-3 DSP Loudspeaker 3D Complex Correction—
Victor Manuel Catalá Iborra, [1] *Francis F. Li* [2]
[1] DAS Audio, Fuente Del Jarro, Spain
[2] University of Salford, Salford, UK

An advantageous approach to DSP equalization of loudspeakers is proposed in this paper adopting spatial averages of complex responses acquired from 3D balloon measurements. Alignment of the off-axis impulses responses with the on-axis impulse responses are accomplished using a cross-correlation technique prior to spatial averaging to attain meaningful statistics of magnitude and phase responses. This is performed over a pre-defined listening window from the complete loudspeaker response balloons (both magnitude and phase). The resulted average of the complex response within a suitably defined listening window is used to obtain, via the least mean square adaptive technique, an inverse filter that corrects the linear behavior of the loudspeaker.
Convention Paper 10140

16:00

P03-4 Poster Introductions 3

- *Investigation into How Reference Sources and the Experience of Technical Ear Training Work in Mixing through Headphones*—Soohoon Park; Toru Kamekawa; Atsushi Marui
- *Proposal of Power-Saving Audio Playback Algorithm Based on Auditory Masking*—Mitsunori Mizumachi; Tsukasa Nakashima; Mitsuhiro Nakagawara
- *Localization of Natural Sound Sources at Various Azimuth and Elevation Angles*—Maksims Mironovs; Hyunkook Lee

Session P04 **Wednesday, March 20**
14:30 – 16:00 **Meeting Room 3**

SPATIAL AUDIO

Chair: **Jorge Medina Victoria**, University of Applied Sciences Darmstadt, Darmstadt, Germany

14:30

P04-1 Toward Six Degrees of Freedom Audio Recording and Playback Using Multiple Ambisonics Sound Fields—*Eduardo Patricio, Andrzej Ruminski, Adam Kuklasinski, Lukasz Januszkiewicz, Tomasz Zernicki*, Zylia sp. z o.o., Poznan, Poland

This paper describes a strategy for recording sound and enabling six-degrees-of-freedom (6DoF) playback making use of multiple simultaneous and synchronized higher-order ambisonics (HOA) recordings. For the evaluation of the proposed approach a 3D audio-visual navigable playback system was implemented. Subjective listening tests were conducted presenting three distinct scenarios, one using spatialized mono sources and the other two interpolated listening points from 1st and 3rd order multiple ambisonics sound fields. The obtained results demonstrate that HOA recordings are suitable for reproduction of 6DoF immersive audio scenes.
Convention Paper 10141

15:00

P04-2 Recording and Composing Site-Specific Spatial Music for 360 Video—*Enda Bates, Sebastian Csadi, Hugh O'Dwyer, Luke Ferguson, Francis M. Boland*, Trinity College Dublin, Dublin, Ireland

This paper documents the 360 video and audio recording of a newly composed work for saxophone quintet, performed in four distinct locations with differing spatial distributions of performers. The potentially site-specific nature of instrumental spatial music is first discussed via a number of historical examples. A comparative analysis of the recordings of this new work from each location is then performed, and the influence of the acoustic environment on different spatial effects such as mobile performers at varying distances, spill, and spatial trajectories is investigated. The analysis suggests that for exterior locations, localization accuracy in first order Ambisonic recordings is adequately maintained, even when performers are placed at large distances. In addition, the presence or lack of reverberation is shown to strongly influence the effectiveness of spill effects or spatial trajectories in instrumental spatial music compositions.
Convention Paper 10142

15:30

P04-3 3D Ambisonic Decoding for Stereo Loudspeakers with Headtracking—*Dylan Menzies, Filippo Maria Fazi*, University of Southampton, Southampton, Hampshire, UK

Compensated Amplitude Panning (CAP) is a spatial audio reproduction method for loudspeakers that takes the listener head orientation into account. Using CAP it is possible to produce stable images in all directions using only two loudspeakers. In its original formulation CAP is inherently an object-based method, with each image produced separately. Here a natural method is presented for dynamically decoding a first order Ambisonic encoding that is equivalent to using CAP to reproduce the constituents of the encoding. This has the advantage of channel-based methods that complex scenes can be reproduced with little cost, and existing Ambisonic encodings, such as those used in 360° video, can be repro-

duced directly.
Convention Paper 10143

Session P05 **Wednesday, March 20**
15:00 – 17:00 **Liffey B**

POSTERS: SESSION 1

15:00

P05-1 Optimized Exciter Positioning Based on Acoustic Power of a Flat Panel Loudspeaker—*Benjamin Zenker, Shanavaz Sanjay Abdul Rawoof, Sebastian Merchel, Ercan Altinsoy*, TU Dresden, Dresden, Germany

Loudspeaker panels, such as distributed mode loudspeakers (DML), are a promising alternative approach in loudspeaker design. DML have many advantages compared to pistonic loudspeakers. However, the frequency response is mostly associated with higher deviations. The position of the excitation is one parameter to optimize the frequency response. An electro-mechanical-acoustical model is presented that enables the optimization of the exciter location, based on the response of the radiated sound power. A simulation model is presented for different surface areas and aspect ratios of the panel. The appropriated positioning and its excitation are discussed based on a single criterion and finally compared with the State of the Art method.
Convention Paper 10144

15:00

P05-2 Practical Problems in Building Parametric Loudspeakers with Ultrasonic Piezoelectric Emitters—*Jose Cadavid, Antonin Novak*, Université du Mans, Le Mans, France

In this paper we deal with some practical issues that one can encounter when building a parametric loudspeaker with ultrasonic piezoelectric emitters. We measured several of those transducers (with resonance frequency 40 kHz) available on the market, observing a strong nonlinear behavior of many of them. We also tested a hundred of piezoelectric emitters of the same series and studied the influence of the standard deviation of the resonance frequency and the sensitivity on the performance of the parametric loudspeaker. We conclude that, when constructing a parametric loudspeaker with low-cost piezoelectric emitters, the individual behavior of each of them should be considered. This allows to minimize the effect of their differences and, thus, improve the quality of the sound generated.
Convention Paper 10145

15:00

P05-3 Time Stretching of Musical Instrument Tones—*Sean O'Leary*, Dublin Institute of Technology, Dublin, Ireland

This paper will present an approach to time stretching monophonic sounds such as musical instrument and voice samples. While most time stretching algorithms preserve the pitch of signals, typically they distort some aspects of the temporal evolution—such as onset time, vibrato rate, and random variations in amplitude and frequency. The aim of the time stretching algorithm presented in this paper is to preserve such features of the original signal in the transformed signal.
Convention Paper 10146

15:00

P05-4 Audio-Driven Multimedia Content Authentication as a Service—*Nikolaos Vryzas, Anastasia Katsaounidou, Rigas Kotsakis, Charalampos A. Dimoulas, George Kalliris*, Aristotle University of Thessaloniki, Thessaloniki, Greece

In the current paper we present a framework for providing supervisory tools for multimedia Content Authentication As A Service (CAAAS). A double compression method for discontinuity detection in audio signals is implemented and integrated in the provided web service. The user can upload audio/video content or provide links and thereafter, a feature vector is extracted from the audio modality of the selected content for the investigation of discontinuities of the signal via the proposed algorithms. Several visualizations are returned to the user, indicating possible points of forgery in the audio/visual file. Moreover, an audio tampering detection methodology by unsupervised clustering of short-window non-vocal segments, in order to identify differentiations of the acoustic environment of speech signals is presented and evaluated.
Convention Paper 10148

15:00

P05-5 ANC System Using Secondary Path Modeling Based on Driver's Position in Vehicle—*Seyeong Jang, Jongin Jung, Hyungsub Lim*, Hyundai Mobis, Seoul, Korea

In this paper we propose a study of active noise control systems using the concept of Secondary Path modeling based on driver position in the vehicle. The system obtains estimates of the Secondary Path within range of occupant location and applies them to the ANC system to compensate for change depending on the driver's position. We used the Offline Secondary Path modeling method and Fx-LMS algorithm in ANC System. Under assumption of detecting a change in position, the secondary path model is applied according to the occupant position and used as initial value of the ANC system. Therefore, ANC performance is better than a system that does not consider existing changing Secondary Path.
Convention Paper 10149

15:00

P05-6 Pop and Rock Music Audio Production for 22.2 Multichannel Sound: A Case Study—*Will Howie*, CBC/Radio-Canada,
Vancouver, Canada

Advanced sound capture and mixing techniques, optimized for high channel-count three-dimensional audio reproduction systems, are discussed for pop/rock music production. Based on previous research and experimental recordings, newly developed complex close-microphone arrays are designed to deliver realistic sonic images of musical instruments in terms of physical size and timbre. Combined with multiple ambience microphones, these direct sound arrays can be used to create highly realistic or hyper-realistic sound scenes for 22.2 multichannel sound (9+10+3) reproduction, or other 3D audio formats. A specific case study highlights the aesthetic and technical considerations for production of pop/rock music for advanced audio formats such as 22.2 multichannel sound.
Convention Paper 10150

15:00

P05-7 University of Victoria Sound Recording Studio Renovation—*Bezal Benny, Kirk McNally*, University of Victoria, School of Music, Victoria, BC, Canada

A recent renovation of the sound recording studio at the University of Victoria School of Music represents the first major capital project within the school since its opening in 1968. This case study presents an overview of the project from initial briefing to completion, including discussion of the design opportunities and limitations. Acoustical models used in the design process are presented and used to illustrate the challenges faced when attempting to balance control room performance for both research and teaching purposes.
Convention Paper 10151

Tutorial 08 **Wednesday, March 20**
15:15 – 16:15 **Liffey Hall 2**

MODERN MUSIC PRODUCTION FOR IMMERSIVE / 3D FORMATS

Presenter: **Lasse Nipkow**, Silent Work LLC, Zurich, Switzerland

More and more music productions in 3D / immersive audio are realized native. In contrast to upmix strategies, thereby the potential of 3D / immersive audio can be fully achieved musically, and it can be used to create new forms of creativity. However, it is not trivial to deal with the large number of audio channels and with new binaural headphone virtualizations meaningfully.

Lasse Nipkow presents his long-term findings and results with 3D audio in the tutorial. This is a listening session. A series of impressive 3D audio productions will be presented.

Tutorial 09 **Wednesday, March 20**
15:15 – 16:15 **Liffey Hall 1**

MICROPHONES—CAN YOU HEAR THE SPECS? A MASTER CLASS

Presenter: **Eddy Bøgh Brixen**, EBB-consult, Smørum, Denmark; DPA Microphones, Allerød, Denmark

There are numerous microphones available to the audio engineer. It's not easy to compare them on a reliable basis, often the choice of the model is made on the basis of experience or perhaps just habits—or just because it looks nice. Nevertheless, there is valuable information in the microphone specifications. This master class held by well-known microphone experts of leading microphone manufacturers demystifies the most important microphone specs and provides the attendees with up-to-date information on how these specs are obtained and can be interpreted. Furthermore, many practical audio demonstrations are given in order to help everyone to understood how the numbers relate to the perceived sound.

This session is presented in association with the AES Technical Committee on Microphones and Applications

Standards Committee Meeting
SC-02-12-L Task Group on Open Control Architecture
Wednesday, March 20, 15:30 – 17:00
Wicklow Meeting Room 5, Level 2

This task group is responsible for the AES70 network audio system control standard, also known as the Open Control Architecture, or OCA, and for related information documents to help developers of AES70-compliant equipment. AES70 offers a wide range of device control and media stream connection management features for professional media networks of all sizes and includes robust reliability and security features. AES70 became a standard in 2015; the group has recently completed a revision that is available for public comment.

Student and Career Event
SC01 OPENING AND STUDENT DELEGATE ASSEMBLY MEETING – PART 1
Wednesday, March 20, 15:45 – 16:45
Liffey A

The first Student Delegate Assembly (SDA) meeting is the official opening of the Convention's student program and a great opportunity to meet with fellow students from all corners of the world. This opening meeting of the Student Delegate Assembly will introduce new events and election proceedings, announce candidates for the coming year's election for the Europe International Regions Vice Chair, announce the finalists in the Student Recording Competition categories and the Student Design Competition, and announce all upcoming student/education related events of the convention. Students and student sections will be given the opportunity to introduce themselves and their activities, in order to stimulate international contacts. The SDA leaders will then lead a dialog to discuss important issues significant to all audio students.

All students and educators are invited to participate in this meeting. Election results and Recording Competition and Design Competition Awards will be given at the Student Delegate Assembly Meeting—Part 2 on March 23, 12:00-13:30.

Session EB01 **Wednesday, March 20**
16:15 – 17:45 **Meeting Room 3**

SPATIAL AUDIO AND ACOUSTICS

Chair: **Piotr Majdak**, Austrian Academy of Sciences, Vienna, Austria

16:15

EB01-1 Extracting Directional Sound for Ambisonics Mix—*Pei-Lun Hsieh, Tsai-Yi Wu*, Ambidio, Los Angeles, CA, USA

Ambisonics Audio has become the primary format to transmit and reproduce audio in immersive or interactive content including 360 video and virtual reality due to its flexibility to be decoded to various speaker configuration and listener's orientation. However, one of the drawbacks of encoding a sound field to Ambisonics audio is the loss of its spatial precision. Higher order Ambisonics has been developed to use more channels in exchange of better precision. In this brief we present a method to detect and extract directional sound from an encoded Ambisonics mix. Improved extraction of the directional signal can improve the performance of other systems, for example the spatial precision during reconstruction.
Engineering Brief 491

16:30

EB01-2 Height Channel Signal Level in Immersive Audio —How Much Is Enough?—*Richard King,*[1,2] *Brett Leonard,*[3] *Jack Kelly*[1,2]
[1] McGill University, Montreal, Quebec, Canada
[2] The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada

[3] BLP Audio, Richmond Heights, MO, USA

Is there an appropriate level for the height channels in an immersive/3D presentation of recorded music when those channels are used specifically for ambience or spatial information? This paper describes an interactive listening test in which expert listeners were directed to manipulate and set the level for four height channels in the upper ring of a 9.1 channel 3D mix (traditional 5.1 surround sound with the addition of four height channels; front L/R, and rear L/R). Stimuli consisted of three musical excerpts—solo piano, string trio, and orchestra. Results were analyzed for mean level and overall variance as a measure of consistency of level set over multiple trials.
Engineering Brief 492

16:45

EB01-3 MPEG Surround Encoder with Steganography Feature for Data Hiding Based on LSB Method—*Ikhwana Elfitri*,[1] *Doni Nursyam*,[1,2] *Rahmadi Kurnia*,[1] *Baharuddin*[1]
[1] Andalas University, Padang, Sumatera Barat, Indonesia;
[2] Ministry of Information and Communication, Banda Aceh Branch Office, Indonesia

Spatial audio coding becomes more important for future audio technology as the Ultra High Definition TV (UHDTV) is ready to enter the market, which is particularly pioneered by engineers in Japan. In this work a method of audio data hiding (steganography) is proposed to be integrated in an MPEG Surround (MPS) encoder, a standard based on the principle of spatial audio coding. A new Reverse One-To-Two (R-OTT) module is introduced with a capability of hiding a short important and secret data or information. This information is embedded in the least significant bit (LSB) of the spatial parameter of the MPS bit stream. The experiments show that the embedding data do not significantly decrease the quality of the transmitted audio signals.
Engineering Brief 493

17:00

EB01-4 Consideration on the Design of Multi-Zone Control System in a Vehicle Cabin—*Wan-Ho Cho, Ji-Ho Chang*, Korea Research Institute of Standards and Science (KRISS), Daejeon, Korea

The personal audio system to generate different sound conditions for each seat in a vehicle cabin is the representative application of multi-zone sound field control. Here, the effectiveness validation of source positions and the robustness of estimated solutions are investigated for the design of a multi-zone control system in a vehicle cabin. To quantify the efficiency of source position, the linear independency test of transfer matrix between the candidate positions of sources and listener is conducted, and an efficient position is selected by the quantified value estimated by the effective independence method. The dummy head source system is applied to measure the transfer matrix efficiently. With the properly selected source positions, it is observed that the control performance is prominent and robust.
Engineering Brief 494

17:15

EB01-5 Violin Sound Characteristics by its Predominant Formant Frequency Changes—*Ewa Lukasik*, Poznan University of Technology, Poznan, Poland

The goal of this Engineering Brief is to make an insight into the dynamics of violin resonances frequency change while playing the instrument. It was inspired by the experiments of Tai and Chung performed on the individual violin sounds of a scale and associated them with formants of human singing voices. The excerpt of the "Sarabande from Partita d-minor" BWV 1004 by Johann Sebastian Bach has been analyzed from the point of view of its predominant formants within 0–5 kHz band. Violin sounds from AMATI database have been used in experiments.
Engineering Brief 495

17:30

EB01-6 Room & Architectural Acoustics—A New Approach to the Design and Delivery of Critical Acoustic Facilities—*Jim Dunne*, Smart Studio, Dublin, Ireland

There has never been more demand for high quality audio studio facilities than there is today. The growth in music recording and production, sound for picture, and gaming are among the areas that are showing increasing demand for accurate studio acoustics. While equipment has improved in performance terms with significant reductions in cost; the approach to designing and building a professional recording/mixing room is unchanged since the 1970s. However, the demands on facilities to provide fully calibrated and accurate acoustic environments does not allow for the outdated, traditional methods of designing and building critical audio facilities. Time to move on!
Engineering Brief 496

Tutorial 10 **Wednesday, March 20**
16:30 – 18:00 **Liffey Hall 1**

HOW TO RATE THE MAXIMUM LOUDSPEAKER OUTPUT SPLMAX?

Presenters: **Steven Hutt**, Equity Sound Investments, Bloomington, IN, USA
Wolfgang Klippel, Klippel GmbH, Dresden, Germany
Peter Mapp, Peter Mapp Associates, Colchester, UK
Alex Voishvillo, Harman Professional, Northridge, CA, USA

The new IEC standard IEC 60268-21 defines the maximum sound pressure level SPLmax generated by the audio device under specified condition (broadband stimulus, 1-m distance, on-axis under free-field condition). This value can be rated by the manufacturer according to the particular application under the condition that the device can pass a 100 h test using this stimulus without getting damaged. The SPLmax according to IEC 60268-21 is not only a meaningful characteristic for the end-user, marketing, and product development but is also required for cal-ibrating analogue or digital stimuli used for testing modern loudspeaker systems having a wireless input and internal signal processing. The workshop gives an overview on related standards (e.g., CEA 2010) and shows practical ways how to rate a meaningful SPLmax value giving the best sound quality, sufficient reliability, and robustness for the particular application. The methods are demonstrated on passive transducers and active (Bluetooth) speakers.

This session is presented in association with the AES Technical Committee on Loudspeakers and Headphones

Workshop 03 **Wednesday, March 20**
16:30 – 18:00 **Liffey Hall 2**

RECORDING ORCHESTRAL MUSIC AND 3D AUDIO: CHALLENGES, CONSIDERATIONS, AND SOLUTIONS

Co-chairs: **Will Howie**, CBC/Radio-Canada, Vancouver, Canada
Toru Kamekawa, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

Panelists: *Kimio Hamasaki*, ARTSRIDGE LLC, Chiba, Japan
Alex Kosiorek, Central Sound at Arizona PBS, Phoenix, AZ, USA
Hyunkook Lee, University of Huddersfield, Huddersfield, UK

Orchestral music maintains a surprisingly ubiquitous place in modern life, heard regularly in the concert hall, commercial recordings, TV and radio broadcasts, film scores, video game sound-tracks, VR and AR scenarios, and any other number of settings. Capturing an orchestral sound scene for three-dimensional reproduction involves many considerations, including repertoire, venue, setup time, post-production budget, and sonic aesthetics. Currently proposed microphone techniques suitable for 3D orchestral music capture range from largely spaced arrays prioritizing envelopment and sound scene control, to near-coincident techniques designed to combine spaciousness with accurate ensemble imaging, to single-point coincident / ambisonics-based techniques that represent an increase in convenience in terms of setup. In this workshop we compare differences in impression due to differences in design and implementation of 3D recording technique through actual recording examples and discuss how these methods impact technical and aesthetic considerations for orchestra recording.

Workshop 04 **Wednesday, March 20**
17:00 – 18:00 **Liffey A**

INTEGRATING HISTORY INTO THE MODERN AUDIO CURRICULUM

Chair: **Scott Burgess**, University of Colorado Denver, Denver, CO, USA

Panelists: *Fiona Doyle-O'Neill,* University College Cork, Cork, Ireland
Mark Drews, University of Stavanger, Stavanger, Norway; Norwegian Institute of Recorded Sound, Stavanger, Norway
Gabe Herman, The University of Hartford, Hartford, CT, USA
Mariana Lopez, University of York, York, UK

Audio engineers know the value of an old microphone and understand the uses of classic equipment and techniques. However, many current audio students still need to be connected with the rich history of our craft. This panel of experienced educators will discuss how to incorporate history into the curriculum of audio schools. Several approaches to a stand-alone history class will be discussed, as well as methods of including history in survey courses. Among these are the use, maintenance, and repair of historical equipment, examination of documents relating to audio history, preservation, and restoration of older recordings, and utilization of recording techniques from bygone days. Our ultimate goal is to inspire students to take this history to heart by incorporating it into their present-day careers.

Student and Career Event
SC02 STUDENT PARTY
Wednesday, March 20, 19:00 – 23:00
TBD

The AES Student Party is open to any 146th Convention participant with an ALL ACCESS STUDENT BADGE. A great opportunity to meet fellow students from around the world. Check the SDA website/blog for details. It will be hosted at a venue to be announced at the Opening and Student Delegate Assembly Meeting—Part 1.

Session P06 **Thursday, March 21**
09:00 – 11:00 **Meeting Room 3**

LOUDSPEAKERS—PART 3

Chair: **Bjørn Kolbrek**, Celestion, Ipswich, UK

09:00

P06-1 Dynamic Driver Current Feedback Methods—*Juha Backman*, Huawei Technologies, Tampere, Finland; Genelec Oy, Iisalmi, Finland

Current feedback is a versatile method of modifying the behavior of a loudspeaker driver with opportunity for linearization and matching the driver to the enclosure design targets, but depending on the chosen approach a potential risk of increasing the effects of either voice coil impedance variation or driver mechanical parameter nonlinearity, and the current feedback approach needs to be designed to keep these effects well controlled for the intended application. This work compares using a nonlinear simulation model various forms of current feedback, including current drive, finite positive or negative amplifier resistances, negative resistance with reactance. This final part of the work extends the examples given in the earlier papers and presents a feedback approach that would appear to offer benefits in both distortion and thermal compression control.
Convention Paper 10152

9:30

P06-2 Impact of the Coupling Factor on Lossy Voice Coil Impedance —*Isao Anazawa*, NY Works, Toronto, ON, Canada

The voice coil impedance frequency dependence due to Eddy current, skin, and proximity effects (Eddy Losses) becomes more apparent as the frequency becomes higher. The theory is that the magnitude of lossy impedance frequency dependence is $\sqrt{\omega}$. However in the majority of real loudspeakers, the impedance frequency dependence was empirically found to be clearly higher than this. A voice coil blocked impedance model was developed based on a structure that applies a transformer for the voice coil inductance as the primary winding. Surrounding conductive material is treated as an impedance connected to the secondary winding. The model successfully describes the blocked impedance frequency dependence that agreed at a high degree of accuracy with the actual samples. Also the model showed intricate connections between the transformer coupling coefficient k and the magnitude of frequency dependency.
Convention Paper 10153

10:00

P06-3 Compact Stereo Loudspeakers with Dipole Processing— *Christof Faller*, Illusonic GmbH, Uster, Zürich, Switzerland; EPFL, Lausanne, Switzerland

Compact stereo loudspeakers have become increasingly popular. One category of these use side-firing left and right transducers featuring a certain spatial effect due to the transducers' directivity at high frequencies. The presented technique increases the spatial effect by controlling directivity at low/medium frequencies, where the transducers have low directivity. A multi-band filter network is used to increase directivity at these frequencies by partially reproducing the stereo signal with dipole directivity pattern. The problem of interference between left and right direct and dipole reproduced sound is addressed.
Convention Paper 10154

10:30

P06-4 Poster Introductions 4

- *Real-Time Measurement System Detecting Tonal Components and Determining Their Audibility in Environmental Noise*—Magdalena Matys; Kamil Piotrowski; Tadeusz Wszolek; Bartlomiej Kukulski
- *Localization Accuracy of First Order Ambisonics in the Vertical Plane across Multiple Horizontal Positions*—Connor Millns; Hyunkook Lee; Maksims Mironovs
- *A Case Study on the Perceptual Differences in Finite-Difference Time-Domain-Simulated Diffuser Designs*—Julie Meyer; Lauri Savioja; Tapio Lokki
- *Analysis of Polish Web Streaming Loudness*—Piotr Cieslik; Karolina Szybinska

Session P07 Thursday, March 21
09:00 – 10:30 Meeting Room 2

ASSESSMENT

Chair: **Federica Bressan**, Ghent University, Ghent, Belgium

09:00

P07-1 BAQ and QoE: Subjective Assessment of 3D Audio on Mobile Phones—*Fesal Toosy, Muhammad Sarwar Ehsan*, University of Central Punjab, Lahore, Pakistan

With the growing popularity of using cellphones and other handheld electronic devices for surfing the internet and streaming audio and video, it was only a matter of time that technologies like 3D audio would be implemented on such devices and relevant content would start being produced. It is important to know if 3D audio offers an improvement over existing stereo formats in terms of perceived basic audio quality and quality of experience. This paper presents a subjective quality assessment of 3D audio. The results show that 3D audio gives an improvement in perceived basic audio quality and quality of experience over other audio formats.
Convention Paper 10155

09:30

P07-2 Segmentation of Listeners Based on Their Preferred Headphone Sound Quality Profiles—*Sean Olive, Todd Welti, Omid Khonsaripour*, Harman International, Northridge, CA, USA

In previous papers we reported results from two controlled listening tests where both trained and untrained listeners gave sound quality preference ratings for in-ear (IE) and around-ear/on-ear headphones. Both groups of listeners on average preferred headphones with frequency responses that meet the Harman target curves. In this paper we re-analyze the data using cluster analysis to uncover different segments or classes listeners based on their similarity in headphone ratings and explore common demographic (age, gender, listening experience) and acoustic factors associated their headphone preferences.
Convention Paper 10156

10:00

P07-3 Latency Tolerance Range Measurements in Western Musical Instruments—*Jorge Medina Victoria*, Hochschule Darmstadt/CIT, Darmstadt, Germany; Cork Institute of Technology, Ireland

A systematic quantitative listening test was conducted in order to investigate the influence of western musical instruments on the ability to cope with latency. A questionnaire and different control mechanisms, including a predefined score and three different metronomes (aural, visual, and aural-visual), enabled the gathering of data under equal conditions for all participants while performing with self-delay. The influence of the musical instrument was demonstrated with the experimental data. Furthermore, the measurement of the latency tolerance range (LTR) enabled the comparison of different instrument groups and demonstrated the relationship between musical tempo and latency.
Convention Paper 10157

Workshop 05 Thursday, March 21
09:00 – 10:00 Liffey Hall 2

COMPUTATION AND PROCEDURAL LITERACY IN AUDIO ENGINEERING EDUCATION: TEACHING THE ART AND SCIENCE PANEL DISCUSSION

Co-chairs: **Charlie DeVane**, Mathworks/University of Massachusetts Lowell, Lowell, MA, USA
Nyssim Lefford, Luleå University of Technology, Luleå, Sweden

Panelists: *Rebecca Stewart*, Queen Mary University London, London, UK
Jonathan Wyner, M Works Studios/iZotope/ Berklee College of Music, Boston, MA, USA

Most technologies used in professional audio production are digital and involve some form of computational operation in the processing and delivery of an audio signal. Content is distributed in digital formats via digital platforms that utilize computation to search content and facilitate and generate listening experiences. Given this, it is reasonable to expect audio engineering curricula to include, alongside microphone techniques, acoustics and other fundamental concepts, some basic concepts in computer science and digital signal processing. However, within the teaching community, there is little consensus about what concepts are essential and where in the curriculum they might sit. If we mean to prepare students for long careers in our ever-evolving, digital world, these topics deserve deep consideration.

In other areas of media content creation, for example computer graphics, we have seen that innovation comes from technologies that enable low-level scripting of computational procedures. Similarly, technologies such as Arduino, Pd/MAX, etc., are inspiring musicians to make their own interfaces and thereby explore new modes of music composition and interaction. In audio, the enabling potential of computation is growing as well. How do we prepare our students to open new creative doors?

For over a decade, media educators in other domains have promoted "procedural literacy" (Mateas, 2005), and the importance of "[l]earning to become computationally expressive" (Bogost, 2005). This panel considers: where does computational thinking and computer science fit into audio engineering education? How do we or can we teach our students the art and science of computation? What concepts are most pertinent for today's audio creators? How can we best prepare students for today and for what computation will make possible in the future? The panel will explore these questions and will provide some real-world examples of where knowledge of computation and procedural literacy has been important to industry audio engineers.

Workshop 06 **Thursday, March 21**
09:00 – 10:30 **Liffey Hall 1**

MPEG-H 3D AUDIO GOES VR

Chair: **Jürgen Herre**, International Audio Laboratories Erlangen, Erlangen, Germany; Fraunhofer IIS, Erlangen, Germany

Panelists: *Adrian Murtaza*, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
Niles Peters, Qualcomm Technologies, Inc., San Diego, CA, USA

The MPEG-H 3D Audio is a recent MPEG standard that was designed to represent and render 3D audio experiences while supporting all known production paradigms (channel-based, object-based, and Higher Order Ambisonics based audio) and reproduction setups (loudspeaker, headphone/binaural). As the audio production world moves forward to embrace Virtual and Augmented Reality (VR/AR), MPEG-H found considerable adoption and re-use in recently finalized VR standards, such as MPEG-I OMAF (Omnidirectional Media Format), VR Industry Forum (VR-IF) Guidelines as well as 3GPP "VRStream" (Virtual Reality profiles for streaming applications) where it was selected as the audio standard for VR content delivered over 5G networks.

This workshop describes how MPEG-H technology finds its way into the new domain and provides an outlook into the MPEG-I 6DoF VR/AR future.

This session is presented in association with the AES Technical Committee on Coding of Audio Signals

Technical Committee Meeting
Technical Committee Meeting on Network Audio Systems
Thursday, March 21, 09:00 – 10:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-05-02 Audio Connectors
Thursday, March 21, 09:00 – 10:00
Wicklow Meeting Room 5, Level 2

The scope of SC-05-02 includes the usage, description, and contact designation for connectors for audio and ancillary functions used in professional audio recording, reproduction, and reinforcement; and the wiring among such connectors and the circuits to which they connect.

Tutorial 11 **Thursday, March 21**
09:30 – 10:30 **Liffey A**

UNCOVERING THE ACOUSTIC TREASURES OF CATHEDRALS: THE USE OF ACOUSTIC MEASUREMENTS AND COMPUTER MODELING TO PRESERVE INTANGIBLE HERITAGE

Presenter: **Lidia Alvarez Morales**, Department of Theatre, Film and Television, University of York ,York, North Yorkshire, UK

The last decades have seen an increase in interest in the study and preservation of the acoustics of worship sites considering their sound as part of their intangible heritage. This tutorial delves deep into the concepts and the methodology needed to characterize the acoustic field of cathedrals, which are an essential part of Europe's cultural, architectural, and artistic heritage. Its aim is to serve as a guideline on how measurement techniques are applied to register monaural, binaural, and ambisonic room impulse responses, as well as on how simulation techniques can be used to assess the influence of occupancy or to recreate their acoustic field throughout history.

This tutorial is linked to three funded projects: two Spanish national projects on the Acoustics and Virtual Reality in Spanish Cathedrals, and the Marie Sklodowska-Curie Fellowship "Cathedral Acoustics," which highlights the importance of acoustics as an essential element of the intangible heritage of cathedrals, defying the traditional focus on visual heritage.

Technical Committee Meeting
Technical Committee Meeting on Acoustics and Sound Reinforcement
Thursday, March 21, 10:00 – 11:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-05-05 Audio Interconnects & EMC
Thursday, March 21, 10:00 – 11:00
Wicklow Meeting Room 5, Level 2

The scope of SC-05-05 includes all practices affecting usage and performance of analog audio hardware, with respect to the susceptibility of the signals it carries to effects such as noise and cross-talk due to the manner of its connection and construction, and the effects of its signals on other hardware and systems in its vicinity for professional audio recording, reproduction, and reinforcement. It shall not set standards for personal safety with regard to such connections and construction, but shall keep safety considerations in mind in its recommendations.

Tutorial 12 **Thursday, March 21**
10:15 – 11:15 **Liffey Hall 2**

OVERVIEW OF ADVANCES IN MAGNETIC PLAYBACK

Presenters: **John Chester**, Plangent Processes, Nantucket, MA, USA
Jamie Howarth, Plangent Processes, Nantucket, MA, USA

A comprehensive overview of recent advancements in magnetic tape playback, focusing on the minimization of time domain errors ranging from drift, wow, and flutter to phase response. Discussion of standards and measurements and the use of record bias to provide a time reference for time base correction and for precise azimuth adjustment. The advantages of more accurate digital de-emphasis and the implementation of phase equalizers compensate for such errors in the recording process, including modeling in DSP the frequency response and phase characteristics of well-known recorders. We conclude by considering the ethics of preserving artifacts on the recording which can be of service to future restorations. Examples of antique and recent wire, film, and tape recordings will be presented.

Student and Career Event
SC03 STUDENT HANDS-ON WORKSHOPS AND MASTER CLASSES ON IMMERSIVE AUDIO— BUILD A WEARABLE BINAURAL SYSTEM WITH BELA
Thursday, March 21, 10:30 – 13:00
Meeting Room 5

Moderator: **Rebecca Stewart**, Imperial College, London, UK

Bela is an ultra-low latency embedded computer designed for audio applications. This session will introduce how to build sensor circuits and program the Bela board with Pure Data. It will also introduce how to render an interactive binaural audio scene and incorporate head-tracking.

Students should bring along laptops with Pd-Vanilla installed along with headphones. All other kit will be provided for use during the session.

Student and Career Event
SC04 STUDENT RECORDING CRITIQUES
Thursday, March 21, 10:30 – 11:30
Meeting Room 1 (Genelec Demo Room)

Moderator: **Ian Corbett**, Kansas City Kansas Community College, Kansas City, KS, USA; off-beat-open-hats recording & sound reinforcement

Students! Come and get tips, tricks, and advice to push your skills to the next level! The Student Recording Critiques are non-competitive listening sessions in which students get to listen to their recordings and productions on a world-class playback system and then receive feedback from a panel of renowned industry professionals. Students at any stage of their studies can sign up to participate. Sign up at the student (SDA) booth immediately on arrival at the convention and deliver stereo mixes as 44.1 Khz/24 bit AIFF or WAVE files, to the SDA booth when you sign up. (Surround playback may be available, please check upon arrival at the event—but bring both surround and stereo options to be safe.) If you sign up, please make sure you arrive on time at the start of the session, otherwise alternates will be placed on the schedule in your place. Finalists in the Recording Competition are excluded from submitting to these events so that as many students as possible can benefit from this type of feedback. (Recording competition finalists get this feedback as part of the competition process.)

Session EB02 **Thursday, March 21**
10:45 – 12:45 **Liffey B**

E-BRIEF POSTER SESSION 1

10:45

EB02-1 Automatic Mixing Level Balancing Enhanced through Source Interference Identification—*Dave Moffat, Mark Sandler*, Queen Mary University of London, London, UK

It has been well established that equal loudness normalization can produce a perceptually appropriate level balance in an automated mix. Previous work assumes that each captured track represents an individual sound source. In the context of a live drum recording this assumption is incorrect. This paper will demonstrate an approach to identify the source interference and adjust the source gains accordingly, to ensure that tracks are all set to equal perceptual loudness. The impact of this interference on the selected gain parameters and resultant mixture is highlighted.
Engineering Brief 497

10:45

EB02-2 Binaural Rendering of Phantom Image Elevation Using VHAP —*Hyunkook Lee, Maksims Mironovs, Dale Johnson*, University of Huddersfield, Huddersfield, UK

VHAP (virtual hemispherical amplitude panning) is a method developed to create an elevated phantom source on a virtual upper-hemisphere with only four ear-height loudspeakers. This engineering brief introduces a new VST plug-in for VHAP and evaluates the performance of the binaural rendering of VHAP with a simple but effective distance control method. Listening test results indicate that the binaural mode achieves the externalization of elevated phantom images in various degrees of perceived distance. VHAP is considered to be a cost-efficient and effective method for 3D panning in virtual reality applications as well as in horizontal loudspeaker reproduction. The plugin is available for free download in the Resources section at www.hud.ac.uk/apl.
Engineering Brief 498

10:45

EB02-3 A Web-Based Tool for Microphone Array Design and Phantom Image Prediction Using the Web Audio API— *Nikita Goddard, Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

A web-based interactive tool that facilitates microphone array design and phantom image prediction is presented within this brief. Originally a mobile app, this web version of MARRS (Microphone Array Recording and Reproduction Simulator) provides greater accessibility through most web browsers and further functionality for establishing the optimal microphone array for a desired spatial scene. In addition to its novel psychoacoustic algorithm based on interchannel time-level trade-offs for arbitrary loudspeaker angles, another main feature allows demonstration of the phantom image scene through virtual loudspeaker rendering and room simulation via the Web Audio API. The current version of the MARRS web app is available through the Resources section of the APL Website: http://www.hud.ac.uk/apl
Engineering Brief 499

10:45

EB02-4 CityTones: A Repository of Crowdsourced Annotated Soundfield Soundscapes—*Agnieszka Roginska,*[1] *Hyunkook Lee,*[2] *Ana Elisa Mendez Mendez,*[1] *Scott Murakami,*[1] *Andrea Genovese*[1]
[1] New York University, New York, NY, USA
[2] University of Huddersfield, Huddersfield, UK

Immersive environmental soundscape capture and annotation is a growing area of audio engineering and research with applications in the reproduction of immersive sound experiences in AR and VR, sound classification, and environmental sound archiving. This Engineering Brief introduces CityTones as a crowdsourced repository of soundscapes captured using immersive sound capture methods that the audio community can contribute to. The database will include descriptors containing information about the technical details of the recording, physical information, subjective quality attri-

butes, as well as sound content information.
Engineering Brief 500

10:45

EB02-5 Recovering Sound Produced by Wind Turbine Structures Employing Video Motion Magnification —*Sebastian Cygert, Andrzej Czyzewski, Marta Stefaniak, Bozena Kostek*, Gdansk University of Technology, Gdansk, Poland

The recordings were made with a fast video camera and with a microphone. Using fast cameras allowed for observation of the micro vibrations of the object structure. Motion-magnified video recordings of wind turbines on a wind farm were made for the purpose of building a damage prediction system. An idea was to use video to recover sound and vibrations in order to obtain a contactless diagnostic method for wind turbines. The recovered signals can be analyzed in a way similar to accelerometer signals, employing spectral analysis. They can be also played back through headphones and compared with sounds recorded by microphones.
Engineering Brief 501

10:45

EB02-6 Modelling the Effects of Spectator Distribution and Capacity on Speech Intelligibility in a Typical Soccer Stadium—*Ross Hammond,*[1] *Peter Mapp,*[2] *Adam J. Hill*[1]
[1] University of Derby, Derby, Derbyshire, UK
[2] Peter Mapp Associates, Colchester, UK

Public address system performance is frequently simulated using acoustic computer models to assess coverage and predict potential intelligibility. When the typical 0.5 speech transmission index (STI) criterion cannot be achieved in voice alarm systems under unoccupied conditions, justification must be made to allow contractual obligations to be met. An expected increase in STI with occupancy can be used as an explanation, though the associated increase in noise levels must also be considered. This work demonstrates typical changes in STI for different spectator distribution in a calibrated stadium computer model. The effects of ambient noise are also considered. The results can be used to approximate expected changes in STI caused by different spectator occupation rates.
Engineering Brief 502

10:45

EB02-7 Influence of the Delay in Monitor System on the Motor Coordination of Musicians while Performing—*Szymon Zaporowski, Maciej Blaszke, Dawid Weber, Marta Stefaniak*, Gdansk University of Technology, Gdansk, Poland

This paper provides a description and results of measurements of the maximum acceptable value of delay tolerated by a musician, while playing an instrument, that does not cause de-synchronization and discomfort. First, methodology of measurements comprising audio recording and a fast camera is described. Then, the measurement procedure for acquiring the maximum value of delay conditioning comfortably playing is presented. Results of musician's response while playing an instrument along with a delayed signal reproduced from the monitor system are shown. Finally, a presentation of the highest values of delays for musicians playing different instruments is given along with a detailed discussion on the methodology used.
Engineering Brief 503

Tutorial 13 **Thursday, March 21**
11:00 – 12:30 **Liffey A**

RECORDING THE ORCHESTRA—WHERE TO START?

Presenter: **Richard King**, McGill University, Montreal, Quebec, Canada; The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada

With a focus on the orchestra as a complex sound source, this tutorial will cover the topics of how to listen, understanding microphones, concert halls and orchestral seating arrangements. How to set up the monitoring environment and how to approach each section of the orchestra as part of the recording process. Concertos, Vocal and Orchestra, and live concert recording will be discussed. Audio examples will be used to support the discussion.

Student and Career Event
SC05 CREATING AUDIO PLUGINS WITH MATLAB
Thursday, March 21, 11:00 – 12:30
Liffey Hall 1

This tutorial covers the technical foundations needed to enter the competition scheduled to take place at the 147th Convention in New York. After attending, you will be able to build a simple VST plugin using MATLAB. You will learn about structuring code for real-time efficiency, defining clear interfaces to tune parameters via interactive controls, testing generated plugins against original designs, and much more.

The session will make use of practical coding examples—prior programming experience will be beneficial but is not required.

Technical Committee Meeting
Technical Committee Meeting on Archiving, Restoration and Digital Libraries
Thursday, March 21, 11:00 – 12:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-04-03 Loudspeaker Modeling and Measurement
Thursday, March 21, 11:00 – 13:00
Wicklow Meeting Room 5, Level 2

The scope of SC-04-03 includes the specification, modeling, and measurement of the electromechanical and free-field radiation characteristics of loudspeakers, individually and in arrays, intended for propagation of audio signals that are used in professional audio recording, reinforcement, and reproduction applications.

Session P08 **Thursday, March 21**
11:15 – 13:15 **Meeting Room 2**

INDUSTRY ISSUES

Chair: **Róisín Loughran**, UCD School of Business, Dublin, Ireland

11:15

P08-1 Early Causes for Biodegradation of PVA/PVC Tapes for Audio Recording—*Ana Paula da Costa,*[1] *Teresa Rosa,*[1] *Federica Bressan*[2]
[1] Instituto Superior Tecnico, Lisbon, Portugal
[2] Ghent University, Ghent, Belgium

The degradation of magnetic tapes is one of the main threats to the survival of our collective audio heritage. Archives around the world, big and small, are all concerned

with the same challenge, that of counteracting the natural decay of plastic compounds. This study investigates the biodegradation of poly(vinyl alcohol)/poly(vinyl chloride) (PVA/PVC) blends tapes, namely audio magnetic tapes, using the spectrophotometry (FTIR), scanning electron microscopy (SEM), and thermogravimetric analysis (TGA). The tapes (both sides) were studied in the light of their degradation in special conditions. The objective is to obtain more information regarding the polymer degradation of magnetic tapes for audio recording and how it affects the structural composition of the tapes. This study contributes to the long-term goal of building a structured knowledge base about diagnostic tools and recovery methods for magnetic tapes.
Convention Paper 10158

11:45

P08-2 Factors Contributing to Gender Imbalance in the Audio Industry—*Shelley Ann McCarthy Buckingham, Malachy Ronan*, Limerick Institute of Technology, Limerick, Ireland

This paper explores the factors contributing to gender imbalance in the audio industry. The two main goals were: (1) whether the traditional gender-related preference for "agency" or "communal" roles holds in the audio industry, and (2) uncover existing gender-based belief systems in the audio industry. The findings suggest that women in the audio industry possess more agentic personality traits than communal. In a surprising finding, men reported more communal personality traits than agentic. Women reported that they were unsuitable for technical and managerial roles making the need for more visible role models in these areas a critical concern.
Convention Paper 10159

12:15

P08-3 A Psychometric Evaluation of Emotional Responses to Horror Music—*Duncan Williams, Chia-Yu Wu, Victoria Hodge, Damian Murphy, Peter Cowling*, University of York, York, UK

This research explores and designs an effective experimental interface to evaluate people's emotional responses to horror music. We studied methodological approaches by using traditional psychometric techniques to measure emotional responses, including self-reporting and galvanic skin response (GSR). GSR correlates with psychological arousal. It can help circumvent a problem in self-reporting where people are unwilling to report particular felt responses, or confuse perceived and felt responses. We also consider the influence of familiarity. Familiarity can induce learned emotional responses rather than listeners describing how it actually makes them feel. The research revealed different findings in self-reports and GSR data. Both measurements had an interaction between music and familiarity but show inconsistent results from the perspective of simple effects.
Convention Paper 10137

12:45

P08-4 Poster Introductions 5

- *Audio Event Identification in Sports Media Content: the Case of Basketball*—Panagiotis-Marios Filippidis; Nikolaos Vryzas; Rigas Kotsakis; Iordanis Thoidis; Charalampos Dimoulas; Charalampos Bratsas
- *Objective and Subjective Comparison of Several Machine Learning Techniques Applied for the Real-Time Emulation of the Guitar Amplifier Nonlinear Behavior*—Thomas Schmitz; Jean-Jacques Embrechts
- *A Generalized Subspace Approach for Multichannel Speech Enhancement Using Machine Learning-Based Speech Presence Probability Estimation*—Yuxuan Ke; Yi Hu; Chengshi Zheng; Xiaodong Li
- *Road Surface Wetness Detection Using Microphones and Convolutional Neural Networks*—Giovani Pepe; Leonardo Gabrielli; Livio Ambrosini; Stefano Squartini; Luca Cattani
- *Primary Study on Removing Mains Hum from Recordings by Active Tone Cancellation Algorithms*—Michal Luczynski

Tutorial 14 **Thursday, March 21**
11:30 – 12:30 **Liffey Hall 2**

THE SOUND OF WAR: CAPTURING SOUNDS IN CONFLICT ZONES

Presenter: **Ana Monte**, DELTA Soundworks, Germany

In field recording preparation is key—but when preparing to record in a war zone, where do you even start? Ana Monte shares her experiences working on *Picturing War*, a Film Academy Baden-Wüttemberg student documentary directed by Konstantin Flemig.

She and the production team follow journalist and photographer Benjamin Hiller as he captures images of a YPJ all-female fighter unit, a refugee camp in Erbil, the Murambi Genocide Memorial in Rwanda, and Kurdish soldiers fighting in Northern Iraq.

Technical Committee Meeting
Technical Committee Meeting on Semantic Audio Analysis
Thursday, March 21, 12:00 – 13:00
Wicklow Meeting Room 4, Level 2

Tutorial 15 **Thursday, March 21**
12:45 – 14:15 **Liffey A**

RECORDING IRISH MUSIC

Presenter: **Sean Davies**

This session traces the recording activity that took place at first in the USA, where immigrant Irish musicians found good work playing to the large Irish communities in America. Recordings had a ready market (of course on 78-rpm discs), many of which were also pressed in UK and Ireland and that are often the only extant evidence of some legendary artists. We will play some of these, including the Sligo fiddler Michael Coleman, Leo Rowsome's pipe band, and later recordings of, e.g., Michael Gorman, the Donegal fiddler John Doherty, and the family band of the McPeakes. Additionally we must remember that the classical music was enriched by John Field, who invented the "Nocturne" later popularized by Chopin. On the vocal front we cannot ignore Count John McCormack plus the many balladeers of popular songs. It may be only a slight exaggeration to say that preservation of the Irish Music tradition would be justification enough for the invention of recording.

Tutorial 16 **Thursday, March 21**
12:45 – 14:15 **Liffey Hall 2**

PSYCHOACOUSTICS OF 3D SOUND RECORDING AND REPRODUCTION (WITH 9.1 DEMOS)

Presenter: **Hyunkook Lee**, University of Huddersfield, Huddersfield, UK

3D multichannel audio aims to produce an immersive auditory experience by adding the height dimension to the reproduced sound field. In order to use the added height channels most effectively, it is necessary to understand the fundamental psychoacoustic mechanisms of vertical stereophonic perception. This tutorial/ demo session will provide sound engineers and researchers with an overview of important psychoacoustic principles to consider for 3D audio recording and reproduction. The talk will first introduce recent research findings on topics such as 3D sound localization, phantom elevation, vertical image spread, and 3D listener envelopment. It will then show how the research has been used to develop new 3D/VR microphone array techniques, 2D–3D upmixing techniques and a new 3D panning technique without using height channels. The talk will be accompanied by 9.1 recording demos of various types of music, including orchestra, choir, organ, EDM, etc.

Tutorial 17 **Thursday, March 21**
12:45 – 14:15 **Liffey Hall 1**

ALMOST EVERYTHING YOU EVER WANTED TO KNOW ABOUT LOUDSPEAKER DESIGN

Presenter: **Christopher Struck**, CJS Labs, San Francisco, CA, USA

This tutorial will walk the audience through an entire loudspeaker design as well as introducing the basic concepts of loudspeakers. Equivalent circuits, impedance, and Thiele-Small Parame-ters are shown. Inherent driver nonlinearities are explained. The effects of modal behavior and cone breakup are demonstrated. Closed Box and Ported Box systems are analyzed and several design examples are meticulously worked through, both with hand calculations and using CAD. Issues with multiple drivers and cabinet construction are discussed. Directivity and diffraction effects are illustrated. Crossover network design fundamentals are presented, with a specific design example for the previously shown ported enclosure design.

This session is presented in association with the AES Technical Committee on Loudspeakers and Headphones

Student and Career Event
SC06 SAUL WALKER STUDENT DESIGN COMPETITION
Thursday, March 21, 13:00 – 15:00
Liffey B

All accepted entries to the AES Student Design Competition are given the opportunity to show off their designs at this poster/tabletop exhibition. The session is free and open to all convention attendees and is an opportunity for aspiring student hardware and software engineers to have their projects seen by the AES design community. It is an invaluable career-building event and a great place for companies to identify their next employees. Students from both audio and non-audio backgrounds are encouraged to participate. Few restrictions are placed on the nature of the projects, which may include loudspeaker designs, DSP plug-ins, analog hardware, signal analysis tools, mobile applications, and sound synthesis devices. Attendees will observe new, original ideas implemented in working-model prototypes.

Technical Committee Meeting
Technical Committee Meeting on Signal Processing
Thursday, March 21, 13:00 – 14:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-02-12-Q Task Group on Streaming Loudness
Thursday, March 21, 13:00 – 14:00
Wicklow Meeting Room 5, Level 2

The group is developing recommendations for loudness normalization of streaming and network file playback content. Streaming is rapidly becoming a major vehicle for media delivery.Audio quality suffers as a result of loudness differences among and within streams, as well as some very high loudness targets, resulting in distortion. Streaming therefore requires a leveling solution based on loudness, with an appropriate loudness target. These recommendations are primarily intended for radio-like mono and stereo streams as opposed to very dynamic stereo and surround sound streams with content such as movies or video specials.

Session P09 **Thursday, March 21**
14:00 – 15:30 **Meeting Room 2**

MACHINE LEARNING—PART 1

Chair: **Konstantinos Drossos**, Tampere University of Technology, Tampere, Finland

14:00

P9-1 Feature Selection and its Evaluation in Binaural Ear Acoustic Authentication—*Masaki Yasuhara,*[1] *Shohei Yano,*[1] *Takayuki Arakawa,*[2] *Takafumi Koshinaka*[2]
[1] Nagaoka College, Nagaoka City, Niigata, Japan
[2] NEC Corporation, Tokyo, Japan

Ear acoustic authentication is a type of biometric authentication that uses the ear canal transfer characteristics that show the acoustic characteristics of the ear canal. In ear acoustic authentication, biological information can be acquired from both ears. However, extant literature on an accuracy improvement method using binaural features is inadequate. In this study we experimentally determine a feature that represents the difference between each user to perform a highly accurate authentication. Feature selection was performed by changing the combination of binaural features, and they were evaluated using the ratio of between-class and within-class variance and equal error ratio (EER). We concluded that a method that concatenates the features of both ears has the highest performance.
Convention Paper 10160

14:30

P09-2 Deep Learning for Synthesis of Head-Related Transfer Functions—*Sunil G. Bharitkar*, HP Labs., Inc., San Francisco, CA, USA

Ipsilateral and contralateral head-related transfer functions (HRTF) are used for creating the perception of a virtual sound source at a virtual location. Publicly available databases use a subset of a full-grid of angular directions due to time and complexity to acquire and deconvolve responses. In this paper we compare and contrast subspace-based techniques for reconstructing HRTFs at arbitrary directions for a sparse dataset (e.g., IRCAM-Listen HRTF database) using (i) hybrid-based (combined linear and nonlinear) principal component analysis (PCA)+fully-connected neural network (FCNN), and (ii) a fully nonlinear (viz., deep learning based) Autoencoder (AE) approach. The results from the AE-based approach show improvement over the hybrid approach, in both objective and subjective tests, and we validate the AE-based approach on the MIT dataset.
Convention Paper 10161

15:00

P09-3 Bayesian Optimization of Deep Learning Techniques for Synthesis of Head-Related Transfer Functions—
Sunil G. Bharitkar,[1] *Timothy Mauer*,[2] *Teresa Wells*,[2] *David Berfanger*[2]
[1] HP Labs., Inc., San Francisco, CA, USA
[2] Prism Lab, HP Inc., Vancouver, WA, USA

Head-related transfer functions (HRTF) are used for creating the perception of a virtual sound source at horizontal angle ø and vertical angle θ. Publicly available databases use a subset of a full-grid of angular directions due to time and complexity to acquire and deconvolve responses. In this paper we build up on our prior research [5] by extending the technique to HRTF synthesis, using the IRCAM dataset, while reducing the computational complexity of the autoencoder (AE)+fully-connected-neural-network (FCNN) architecture by ~ 60% using Bayesian optimization. We also present listening test results, demonstrating the performance of the presented approach, from a pilot study that was designed for assessing the directional cues of the proposed architecture.
Convention Paper 10162

Technical Committee Meeting
Technical Committee Meeting on Hearing and Hearing Loss Prevention
Thursday, March 21, 14:00 – 15:00
Wicklow Meeting Room 4, Level 2

Tutorial 18 **Thursday, March 21**
14:30 – 15:30 **Liffey Hall 2**

CASE STUDIES IN JAZZ AND POP/ROCK MUSIC PRODUCTION FOR 3D AUDIO

Presenter: **Will Howie**, CBC/Radio-Canada, Vancouver, Canada

Music recording and mixing techniques for 3D audio reproduction systems will be discussed through several case studies covering jazz, pop/rock, and new music genres. Complex multi-microphone arrays designed for capturing highly realistic direct sound images are combined with spaced ambience arrays to reproduce a complete sound scene. Mixing techniques are aesthetically and technically optimized for 3D reproduction. Initially developed for use with large-scale loud-speaker-based formats, such as 22.2 Multichannel Sound (9+10+3), it will be shown how these sound capture techniques can be scaled to smaller, more common standardized channel-based formats, such as 4+5+0 and 4+7+0. Numerous corresponding audio examples have been prepared for 9.1 (4+5+0) reproduction. The focus will be on practical, aesthetic, and technical considerations for 3D music recording.

Workshop 07 **Thursday, March 21**
14:30 – 16:00 **Liffey Hall 1**

IMPACT AND AUDIBILITY OF DISTORTION IN AUTOMOTIVE AUDIO APPLICATIONS

Chair: **Alfred Svobodnik**, MVOID Group, Karlsruhe, Germany; Vienna, Austria

Panelists: *Rafael Kassier*, Harman Becker, Karlsbad, Germany
Joachim Schlechter, Klippel GmbH, Dresden, Germany
Roger Shively, JJR Acoustics, LLC, Seattle, WA, USA
Steve Temme, Listen, Inc., Boston, MA, USA

Distortion effects, either linear or nonlinear, are an important aspect of audio quality. Especially for professional audio applications there are lots of research results available. However, the field of car applications is very specific, and not much research has been performed on this topic so far. This workshop will focus on automotive applications and discuss the current state of science on this topic. While the focus will be on audibility, we will especially tackle on active noise control applications and the specific aspects of loudspeaker distortion.

This session is presented in association with the AES Technical Committee on Automotive Audio

Special Event
SE02 AES DIVERSITY AND INCLUSION COMMITTEE TOWN HALL: FOCUS ON GENDER EQUALITY
Thursday, March 21, 14:30 – 16:00
Liffey A

Moderator: **Mariana Lopez**, University of York, York, UK

Panelists: *Leslie Gaston-Bird*, Mix Messiah Productions, Brighton, UK; Institute of Contemporary Music Performance, London, UK
Cesar Lamschtein, Kaps Audio Production Services
Agnieszka Roginska, New York University, New York, NY, USA
Nadja Wallaszkovits, Phonogrammarchiv, Austrian Academy of Science, Vienna, Austria; NOA GmbH, Vienna, Austria
Jonathan Wyner, iZotope, MA, USA

The AES Diversity and Inclusion Committee recently received approval from the Board of Governors to support the United Nations HeForShe campaign for gender equality. Panel moderator Mariana Lopez will bring members up to date on what this means for the AES and along with invited panelists discuss this and other matters related to diversity and inclusion in the AES and audio industry.

Workshop 08 **Thursday, March 21**
14:45 – 15:45 **Meeting Room 3**

FORENSIC AUDIO—WHAT'S THAT?

Presenter: **Eddy Bøgh Brixen**, EBB-consult, Smørum, Denmark; DPA Microphones, Allerød, Denmark

Working with audio forensics is serious business. Depending on the work of the forensics engineer, people may eventually end up in prison. This event will present the kind of work related to the field. This covers fields as acoustics, when audio analysis can be a part of the crime scene investigation. Voice acoustics: Who was speaking? Electroacoustics: Checking data on tapes, discs or other data storage media: Did anyone tamper with this? Recording techniques: Is this recording an original production or is it a copy of others' work. Even building-acoustics and psychoacoustics, when the question is raised: Who could hear what? However, the most important "everyday work" of the audio forensics engineers is cleaning of audio recordings and providing transcripts.

This session is presented in association with the AES Technical Committee on Audio Forensics

Session P10 **Thursday, March 21**
15:15 – 17:15 **Liffey B**

POSTERS: SESSION 2

15:15

P10-1 Investigation into How Reference Sources and the Experience of Technical Ear Training Work in Mixing through Headphones —*Soohoon Park, Toru Kamekawa, Atsushi Marui*, Tokyo University of the Arts, Tokyo, Japan

This paper reports an investigation into how reference sources and the experience of technical ear training work in mixing through headphones. In the experiment, participants were asked to adjust the EQ of the stimulus source while monitoring by using five different types of headphones respectively. There were significant differences between the two groups based on the experience of ear training and in the EQ adjustment results of the high-frequency region depending on whether or not the reference was provided. Based on the experimental results of the experiments, the mixing result has been shown to be influenced by the existence of the reference source and the experience of ear training.
Convention Paper 10163

15:15

P10-2 Proposal of Power-Saving Audio Playback Algorithm Based on Auditory Masking—*Tsukasa Nakashima,*[1] *Mitsuhiro Nakagawara,*[2] *Mitsunori Mizumachi*[1]
[1] Kyushu Institute of Technology, Fukuoka, Japan
[2] Panasonic Corporation, Yokohama City, Kanagawa, Japan

Power consumption is an important issue while listening to music using portable audio devices. The authors have previously proposed a power-saving audio playback algorithm, which has adjusted filter-bank outputs according to our auditory characteristics. It succeeds in reducing power consumption but causes perceptual distortion. In this paper the power-saving audio playback algorithm is improved based on auditory masking, which attenuates audio components below the masking threshold. As a result of a listening test, it is confirmed that the proposed method is subjectively superior to the previous method with the same power consumption.
Convention Paper 10164

15:15

P10-3 Localization of Natural Sound Sources at Various Azimuth and Elevation Angles—*Maksims Mironovs, Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

A bird recording was compared against an airplane takeoff sample at various azimuth and elevation angles in this study. A total of 33 source positions were tested, ranging from 0° to 180° azimuth and –30° to 90° elevation angles with 30° intervals. The results showed that both perceived azimuth and elevation are significantly affected by the source frequency content. Furthermore, a significant azimuth shift towards the lateral plane was observed on the off-center axis. This effect was stronger for the elevated positions on the rear hemisphere. Additionally, the pitch-height effect was present and was most dominant on the median plane and frontal hemisphere. Last, confusion errors were present for both stimuli; however, they were significant only on the median plane.
Convention Paper 10165

15:15

P10-4 Real-Time Measurement System Detecting Tonal Components and Determining Their Audibility in Environmental Noise—*Magdalena Matys, Kamil Piotrowski, Tadeusz Wszolek, Bartłomiej Kukulski*, AGH University
of Science and Technology, Kraków, Poland

The presence of tonal components in a sound signal usually increases its annoyance but their detection and proper qualification is not always unambiguous. Despite the relatively easy recognition of a tonal noise, its objective identification and tonality measurement is much more difficult. The identification and classification of tonal components presence in measured noise is described in the standard ISO/PAS 20065:2016(E). In this paper authors introduce a system that was created in LabVIEW environment. The main objective was to develop the easy to use system running in real-time, which is capable to perform automatic calculations based on ISO/PAS 20065:2016(E).
Convention Paper 10166

15:15

P10-5 Vertical Localization Accuracy of Binauralized First Order Ambisonics across Multiple Horizontal Positions—*Connor Millns, Hyunkook Lee, Maksims Mironovs*, University of Huddersfield, Huddersfield, West Yorkshire, UK

This presents a systematic investigation into the localization accuracy of two First Order Ambisonics (FOA) decoding methods for head-static binaural reproduction: the magnitude least squared method used in the IEM Binaural Decoder and the basic decoding method for the cube virtual loudspeaker layout. The two decoding methods were compared against a directly binauralized reference for five vertical positions (–45° to 45° at 22.5° intervals) for eight horizontal positions (0° to 315° at 45° intervals). A train of pink noise bursts were used as stimuli. Results indicate that little elevation was perceived across the tested azimuths for all three reproduction methods. The lack of elevation has implications for FOA microphone placement in terms of microphone height.
Convention Paper 10167

15:15

P10-6 A Case Study on the Perceptual Differences in Finite-Difference Time-Domain-Simulated Diffuser Designs—*Julie Meyer, Lauri Savioja, Tapio Lokki*, Aalto University, Espoo, Finland

This paper presents a method to determine if differences between the scattering created by geometrically-similar diffuser designs are perceivable. Although there exist standards to measure the scattering and diffusion coefficients, the perceptual evaluation of the scattering created by diffusing surfaces has previously been scarcely examined. In the context of the optimization of a diffuser design, such audibility study can be used to assess the relevance of optimized geometries from a perceptual point of view. The proposed approach uses finite-difference time-domain (FDTD) numerical simulations to generate impulse responses (IRs) from which diffuser responses of geometrically-close designs are extracted. For each diffuser geometry, a set of three such time-domain responses convolved with a click-like signal, white Gaussian noise, and a male speech, are used as stimuli in an ABX listening test. Percentage of correct answers show

that subjects are able to perceive differences for the click stimulus for all tested conditions (geometries and receiver positions), while discrimination rates are mitigated across conditions for the white Gaussian noise and are not significant for the speech signal. Results also indicate that subjects' performance depends on the receiver location.
Convention Paper 10168

15:15

P10-7 Analysis of Polish Web Streaming Loudness—*Piotr Cieslik*,[1] *Karolina Szybinska*[2]
[1] AGH University of Science and Technology, Krakow, Poland
[2] Jagiellonian University, Krakow, Poland

The aim of the study was to identify the problem related to the lack of sound normalization and law regulations in the online streaming. The method was based on analysis of samples of recorded materials from PC's web browser players and Android and iOS apps. Samples were taken from the Polish Internet streaming stations. The data were analyzed and the results were compared. The results showed that the loudness of Polish web streaming was very differentiated. There is significant discrepancy in the loudness between the stations. Moreover, in some cases, there are substantial loudness differences between advertisements, music, and programs.
Convention Paper 10169

15:15

P10-8 Primary Study on Removing Mains Hum from Recordings by Active Tone Cancellation Algorithms—
Michal Luczynski, Wroclaw University of Science and Technology, Wroclaw, Poland

In this paper the method of removing the mains hum has been presented. This method is based on active tone reduction. Active tone reduction is active noise reduction, where the secondary signal is a signal synthesized based on tonal components detected in the primary signal. The author of the paper has made tests of his own algorithm. The tested signals are the mains hum and hum with the guitar sound. The effect of the work is to indicate the advantages and disadvantages of the algorithm comparing with commonly used solutions.
Convention Paper 10147

Session P11 **Thursday, March 21**
16:00 – 18:00 **Meeting Room 2**

MACHINE LEARNING—PART 2

Chair: **Bezal Benny**, University of Victoria, Victoria, Canada

16:00

P11-1 Audio Inpainting of Music by Means of Neural Networks—*Andrés Marafioti*,[1] *Nicki Holighaus*,[1] *Piotr Majdak*,[1] *Nathanaël Perraudin*[2]
[1] Austrian Academy of Sciences, Vienna, Austria
[2] Swiss Data Science Center, Switzerland

We studied the ability of deep neural networks (DNNs) to restore missing audio content based on its context, a process usually referred to as audio inpainting. We focused on gaps in the range of tens of milliseconds. The proposed DNN structure was trained on audio signals containing music and musical instruments, separately, with 64-ms long gaps and represented by time-frequency (TF) coefficients. For music, our DNN significantly outperformed the reference method based on linear predictive coding (LPC), demonstrating a generally good usability of the proposed DNN structure for inpainting complex audio signals like music.
Convention Paper 10170

16:30

P11-2 A Literature Review of WaveNet: Theory, Application, and Optimization—*Jonathan Boilard, Philippe Gournay, Roch Lefebvre*, Universite de Sherbrooke, Sherbrooke, QC, Canada

WaveNet is a deep convolutional artificial neural network. It is also an autoregressive and probabilistic generative model; it is therefore by nature perfectly suited to solving various complex problems in speech processing. It already achieves state-of-the-art performance in text-to-speech synthesis. It also constitutes a radically new and remarkably efficient tool to perform voice transformation, speech enhancement, and speech compression. This paper presents a comprehensive review of the literature on WaveNet since its introduction in 2016. It identifies and discusses references related to its theoretical foundation, its application scope, and the possible optimization of its subjective quality and computational efficiency.
Convention Paper 10171

17:00

P11-3 Sparse Autoencoder Based Multiple Audio Objects Coding Method—*Shuang Zhang, Xihong Wu, Tianshu Qu*, Peking University, Beijing, China

The traditional multiple audio objects codec extracts the parameters of each object in the frequency domain and produces serious confusion because of high coincidence degree in subband among objects. This paper uses sparse domain instead of frequency domain and reconstruct audio object using the binary mask from the down-mixed signal based on the sparsity of each audio object. In order to overcome high coincidence degree of subband among different audio objects, the sparse autoencoder neural network is established. On this basis, a multiple audio objects codec system is built up. To evaluate this proposed system, the objective and subjective evaluation are carried on and the results show that the proposed system has the better performance than SAOC.
Convention Paper 10172

17:30

P11-4 Poster Introductions 6

• *jReporter: A Smart Voice-Recording Mobile Application*—Lazaros Vrysis; Nikolaos Vryzas; Efstathios Sidiropoulos;
Evangelia Avraam; Charalampos Dimoulas
• *Two-Channel Sine Sweep Stimuli: A Case Study Evaluating 2-n Channel Upmixers*—Laurence J. Hobden; Christopher Gribben
• *A Rendering Method for Diffuse Sound*—Akio Ando

Session EB03 **Thursday, March 21**
16:00 17:15 **Meeting Room 3**

MICROPHONES AND CIRCUITS

Chair: **Joerg Panzer**, R&D Team, Salgen, Germany

16:00

EB03-1 Analysis of Beam Patterns of Super Directive Acoustic Beamformer—*Adam Kupryjanow*, Intel Technology Poland, Gdansk, Poland

In this brief beam patterns of a super directive acoustic Beamformer were presented. The analysis was done based on the recordings made in diffuse far field environment. State of the art MVDR (Minimum Variance Distortionless Beamformer) was utilized as representative of super directive Beamformer. Two types of uniform microphone arrays were investigated: linear and circular. Experiments were performed for various number of microphones in the arrays, i.e., two, four, six, and eight.
Engineering Brief 504

16:15

EB03-2 Configuration for Testing Intermodulation of Ultrasonic Signals in the Microphone Path—*Dominik Stanczak, Jan Banas, Jedrzej Prysko, Pawel Trella, Przemek Maziewski*,
Intel Technology Poland, Gdansk, Poland

The paper presents a comparison of five configurations used to test intermodulation of ultrasonic signals. All five require the use of ultrasonic loudspeakers in an anechoic, low-noise environment. Settings vary in the number and type of digital-to-analog converters and loudspeakers, as well as connection types. Best configuration introducing a small amount of self-intermodulation for the high power of the ultrasonic signals is identified.
Engineering Brief 505

16:30

EB03-3 Considerations for the Next Generation of Singing Tutor Systems—*Behnam Faghih, Joseph Timoney*, Maynooth University, Maynooth, Kildare, Ireland

Recently software systems have been proposed to accelerate the progress of singing beginners. The basics of these systems are: the pitch of the sung notes is detected and algorithmic errors removed. Then, an alignment is made with a melodic ground truth, often as a midi representation, using techniques including Dynamic Time Warping and Hidden Markov Models. Although results have been reasonable, significant drawbacks to these alignment schemes include how a "musically acceptable" alignment can be identified, dynamic singer behavior, multiple repeated notes, and dealing with omitted or extra notes. To this end an improved singing analysis system structure is proposed that includes psychoacoustic models and intelligent decision making. Justification is given along with a description of a structured evaluation procedure.
Engineering Brief 506

16:45

EB03-4 Control Techniques for Audio Envelope Tracking—*Robert Bakker, Maeve Duffy*, NUI Galway, Galway, Ireland

One of the main applications for class-D audio amplifiers is portable or battery-operated devices such as Bluetooth speakers, smartphones, car stereos, etc. These devices often use a boost converter to increase the battery voltage to a suitable level to achieve the desired power output. The use of envelope tracking (ET) has been shown to significantly improve the efficiency of a class-D audio amplifier, particularly at lower power levels. However, modulating a boost converter to provide envelope tracking at a high bandwidth is complicated due to the right half-plane zero in its transfer function. This paper discusses the effect of envelope bandwidth on the overall system performance, and how it affects the control of the boost converter. It also discusses the different control methods for boost converters, and variations of this type of DC/DC converter.
Engineering Brief 507

17:00

EB03-5 Sound Synthesis Using Programmable System-on-Chip Devices—*Larry Fitzgerald, Joseph Timoney*, Maynooth University, Maynooth, Ireland

An approach to building analog synthesizers may be found by exploiting a new mixed-signal technology called the Programmable System-on-Chip (PSoC), which includes a CPU core and mixed-signal arrays of configurable integrated analog and digital peripherals. Another approach is to exploit a System on Chip (SoC) comprising an ARM-based processor and an FPGA. Two synthesizers were built and evaluated for sound quality and difficulty of implementation. Each of the approaches produced a synthesizer of good sound quality. The mixed-signal approach was cheaper in both component costs and development time compared to the FPGA-based approach.
Engineering Brief 508

Student and Career Event
SC07 STUDENT RECORDING COMPETITION—PART 1
Thursday, March 21, 16:00 – 18:00
Liffey Hall 2

The Student Recording Competition is a highlight at each convention. A distinguished panel of judges participates in critiquing finalists of each category in an interactive presentation and discussion. The top three finalists in each category, as identified by our judges, present a short summary of their production intentions and the key recording and mix techniques used to realize their goals. They then play their projects for all who attend. Meritorious awards are determined here and will be presented at the closing Student Delegate Assembly Meeting (SDA-2) on Saturday. The competition is a great chance to hear the work of your fellow students at other educational institutions. Everyone learns from the judges' comments, even those who don't make it to the finals, and it's a great chance to meet other students and faculty.

Technical Committee Meeting
Technical Committee Meeting on Automotive Audio
Thursday, March 21, 16:00 – 17:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-02-12 Audio Applications of Networks
Thursday, March 21, 16:00 – 17:30
Wicklow Meeting Room 5, Level 2

The scope of SC-02-12 includes the use of various network types for audio and audio-related applications in professional recording and broadcasting.

Tutorial 19 **Thursday, March 21**
16:15 – 17:30 **Liffey Hall 1**

PRACTICAL DEEP LEARNING INTRODUCTION FOR AUDIO PROCESSING ENGINEERS

Presenter: **Gabriele Bunkheila**, MathWorks, Madrid, Spain

Are you an audio engineer working on product development or DSP algorithms and willing to integrate AI capabilities within your projects? In this session we will walk through a simple Deep Learning example for speech classification. We will use MATLAB code and a speech command dataset made available by Google. We will cover creating and accessing labeled data, using time-frequency transformations, extracting features, designing and training deep neural network architectures, and testing prototypes on real-time audio. We will also discuss working with other popular Deep Learning tools, including exploiting available pre-trained net-works.

This session is presented in association with the AES Technical Committee on Semantic Audio Analysis

Tutorial 20 **Thursday, March 21**
16:30 – 18:00 **Liffey A**

SOFTWARE COMPLEMENTS LOUDSPEAKER HARDWARE

Presenters: **Wolfgang Klippel**, Klippel GmbH, Dresden, Germany
Joachim Schlechter, Klippel GmbH, Dresden, Germany

Digital signal processing, amplification, and the electroacoustical conversion converge to one active transducer, providing more sound output at higher quality by using less size, weight, energy, and manufacturing cost. Adaptive nonlinear control technology based on monitoring the electric input current cancels undesired signal distortion, actively protects the transducer against mechanical and electrical overload, and stabilizes the voice coil at the optimum rest position to ensure maximum excursion over product life. This control software opens new degrees of freedom for the passive transducer and system development. The tutorial presents a new design concept that optimizes efficiency and voltage sensitivity while using a minimum of hardware resources. The design steps are illustrated through practical examples using new simulation, measurement, and diagnostic tools that analyze the performance of the audio device with nonlinear control.

This session is presented in association with the AES Technical Committee on Loudspeakers and Headphones

Technical Committee Meeting
Technical Committee Meeting on Loudspeakers and Headphones
Thursday, March 21, 17:00 – 18:00
Wicklow Meeting Room 4, Level 2

Special Event
SE03 MIXER PARTY
Thursday, March 21, 18:00 – 19:00
Liffey B

Join us for the AES Dublin Networking Mixer, sponsored by Tyrell, taking place this Thursday at 6pm in the Liffey B Hall, within the Product Showcase.

Special Event
SE04 THE RICHARD C. HEYSER MEMORIAL LECTURE
Thursday, March 21, 19:15 – 20:15
Liffey A

Lecturer: **David Griesinger**

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 Heyser Lecturer this year is **David Griesinger**. The title of his talk is "Learning to Listen."

This talk describes a few of the experiences I have had working with artists and musicians that taught me to hear what makes their art work. These experiences changed my life, and I hope that describing them will help others understand some of the amazing abilities of our ears and brains to detect, separate, and ultimately decode sonic information. Just as important, insight into these listening experiences can illuminate the possible physical processes by which the ear and brain achieve these feats. Such analysis often contradicts current dogma. For example, the phases of sounds higher than 1500 Hz are often considered unimportant but are actually critical to our ability to remember information. Working with these musicians and with acoustic enhancement taught me that many popular acoustic design goals can be counterproductive, and popular practices in sound reinforcement often reduce the ability to remember what was said. Accurate acoustic memory is short. Finding new models and methods for improving sound and acoustics requires reproducing live sound in a laboratory with high accuracy. We will describe an inexpensive technique that uses non-individual binaural recording and non-individual crosstalk cancellation for reproduction. The reproductions are startlingly real, and allow for instant comparisons between spaces. When used with measurement data the effects of individual reflections can be determined. Contrary to widespread practice early lateral reflections are often harmful, not helpful. We find that a reduction in the direct to reverberant ratio of only one dB, or the presence of a single reflection, can change sound from clear to muddy. These findings are being put into practice in old and new venues with exceptional results. We can reproduce live recordings made this way from all over the world. I will demonstrate this system to interested listeners during this conference as time and space allows.

Session P12 **Friday, March 22**
09:00 – 11:00 **Meeting Room 3**

SPEECH

Chair: **Yuxuan Ke**, University of Chinese Academy of Sciences, Beijing, China

09:00

P12-1 Background Ducking to Produce Esthetically Pleasing Audio for TV with Clear Speech—*Matteo Torcoli,*[1] *Alex Freke-Morin,*[1,2] *Jouni Paulus,*[1,3] *Christian Simon,*[1] *Ben Shirley*[2]
[1] Fraunhofer IIS, Erlangen, Germany
[2] University of Salford, Salford, UK
[3] International Audio Laboratories Erlangen, Erlangen, Germany

In audio production background ducking facilitates speech intelligibility while keeping the background track enjoyable. Technical details for recommendable ducking practices are not currently documented in literature. Hence, we first analyze common practices found in TV documentaries. Second, a subjective test investigates the preferences of 22 normal-hearing listeners on the Loudness Difference (LD) between commentary and background during ducking. Highly personal preferences are observed, highlighting the importance of object-based personalization. Statistically significant difference is found between non-expert and expert listeners.

On average, non-experts prefer LDs that are 4 LU higher than the ones preferred by experts. Based on the test results, we recommend at least 10 LU difference between commentary and music and at least 15 LU between commentary and ambience.
Convention Paper 10175

09:30

P12-2 Factors Influencing the Spectral Clarity of Vocals in Music Mixes—*Kirsten Hermes*, University of Westminster, Middlesex, UK

Vocal clarity is one of the most important quality parameters of music mixes. The clarity of isolated sounds depends heavily on spectral factors and can therefore be manipulated with EQ. Spectrum is also an important factor in determining vocal timbral and quality parameters. An experiment where listeners rate the spectral clarity of equalized vocals within a noise backing track can provide insight into spectral predictors of vocal clarity. Overall, higher frequencies contribute to vocal clarity more positively than lower ones, but the relationship is program-item-dependent. Changes in harmonic centroid (or dimensionless spectral centroid) correlate well with changes in clarity and so does the vocal-to-backing track ratio.
Convention Paper 10174

10:00

P12-3 High-Resolution Analysis of the Directivity Factor and Directivity Index Functions of Human Speech—*Samuel Bellows, Timothy Leishman*, Brigham Young University, Provo, UT, USA

The detailed directivity of a sound source is a powerful tool with broad applications in modeling of sound radiation into various acoustic environments, ideal microphone positioning, and other areas. While the directivity of human speech has been assessed previously, the results have lacked the necessary resolution to accurately model radiation in three dimensions. In this work high-resolution measurements were taken using a multiple-capture spherical-scanning system. The frequency-dependent directivity factors and indices of speech were then calculated from the data and their spherical-harmonic expansions. Although past models have represented these measures in simple terms, high-resolution measurements demonstrate that over the audible range they have more variation than previously known, with important ramifications for three-dimensional modeling and audio.
Convention Paper 10173

10:30

P12-4 Poster Introductions 7

• *Quantitative Analysis of Streaming Protocols for Enabling Internet of Things (IoT) Audio Hardware*—Marques Hardin; Rob Toulson
• *Automatic Detection of Audio Problems for Quality Control in Digital Music Distribution*—Pablo Alonso Jiménez; Luis Joglar Ongay; Xavier Serra; Dmitry Bogdanov
• *A High Power Switch-Mode Power Audio Amplifier* — Niels Ekljær Iversesen, Jóhann Björnsson, Patrik Boström, Lars Petersen

Session P13 **Friday, March 22**
09:00 – 10:30 **Meeting Room 2**

DSP—PART 1

Chair: **Emmanouil Theofanis Choudakis**, Queen Mary University of London, London, UK

09:00

P13-1 Applying Modern Sampling Methods to the Mastering Process for Digitally Recorded Material—*Jamie Angus-Whiteoak*, University of Salford, Salford, UK

Mastering often involves a change in sampling rates from a higher sampling rate to the sampling rate required by the distribution medium such as CD, etc. This rate change implicitly implies a resampling process that can introduce artifacts into the output. Modern sampling theory gives useful insight into how to improve this process. This paper introduces modern sampling theory to highlight both the problems, and possible solutions, sample rate changing tof recorded digital audio at the highest quality possible. Possible methods for changing the rate are discussed and means of reducing the huge computational cost are described. The paper will show that by using modern sampling methods it is possible to change sample rates with near perfect to perfect fidelity.
Convention Paper 10176

09:30

P13-2 Application of a Resonance-Based Signal Decomposition to the Analysis of Subtractive Synthesizer Filter Resonances—*Joseph Timoney,*[1] *Kemal Avci,*[2] *Victor Lazzarini*[1]
[1] Maynooth University, Maynooth, Kildare, Ireland
[2] Izmir Demokrasi Üniversitesi, Karabaglar/Izmir, Turkey

This paper investigates the analysis of resonant filters as they appear in subtractive synthesizers. These filters and their properties are a key component in the synthesis chain. The work investigates the application of a new wavelet-like signal decomposition for examining the components that make up the filter output. It produces a pair of "low" and "high" components. The results will examine these components spectrally with the intention that they might lead to new insights into synthesis and modeling.
Convention Paper 10177

10:00

P13-3 An Automatic Mixing System for Multitrack Spatialization for Stereo Based on Unmasking and Best Panning Practices—*Ajin Tom,*[1] *Joshua D. Reiss,*[2] *Philippe Depalle*[1]
[1] McGill University, Montreal, Quebec, Canada
[2] Queen Mary University of London, London, UK

One of the most important tasks in audio production is to place sound sources across the stereo field so as to reduce masking and immerse the listener within the space. This process of panning sources of a multitrack recording to achieve spatialization and masking minimization is a challenging optimization problem, mainly because of the complexity of auditory perception. We propose a novel panning system that makes use of a common framework for spectral decomposition, masking detection, multitrack sub-grouping and frequency-based spreading. It creates a well spatialized mix with increased clarity while complying to the best panning practices. Both real-time and

off-line optimization-based approaches are designed and implemented. We investigate the reduction of inter-track auditory masking using the MPEG psychoacoustic model along with various other masking and spatialization metrics extended for multitrack content. Subjective and objective tests compare the proposed work against mixes by professional sound engineers and existing auto-mix systems.
Convention Paper 10178

Tutorial 21 **Friday, March 22**
09:00 – 10:00 **Liffey Hall 1**

WHAT'S BETWEEN LINEAR MOVIES AND VIDEO GAMES?

Presenters: **Lidwine Ho**, France télévisions, Paris, France

360° video is increasingly used in many areas. Broadcasters whose video is the core business cannot ignore it. The question that they have to seize now is: how to reinvent television through new narrative formats without losing their identity and their objectives.

Former television aims to reconstruct reality by editing, recreating emotions with music while the 360° video shows a sequence in which the user can explore the universe. Beyond the technical question of the tools and formats we used, we will explain, step by step what keys we used to build a coherent and immersive sound universe while guiding the user in his axes of vision.

The other question we will consider is: what is the next step? Do we have to build video games to reach our objectives? Is there a place between linear movies and video game for broadcasters? Which tools or standards do we need?

Tutorial 22 **Friday, March 22**
09:00 – 10:00 **Liffey Hall 2**

AMBISONICS OR NOT?—COMPARISON OF TECHNIQUES FOR 3D AMBIENCE RECORDING

Presenter: **Felix Andriessens**, Ton und Meister, Germany

With the spreading of head-tracked audio in games and VR-applications and 3D-sound systems for cinema like Dolby Atmos and Auro3D on the rise, the demand for 3D-audio recordings and recording setups keeps growing. This event shows a comparison between different recording techniques from coincident to spaced setups and discusses whether and where Ambisonic microphones are the best option for field recordings, especially considering ambience recordings.

Technical Committee Meeting
Technical Committee Meeting on Audio for Telecommunications
Friday March 22, 09:00 – 10:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-02-01 Digital Audio Measurement Techniques
Friday March 22, 09:00 – 10:00
Wicklow Meeting Room 5, Level 2

The scope of SC-02-01 includes measurement methods for equipment used for the recording, reproduction, and transmission of digital audio signals for professional recording and broadcasting. It includes effects of perceptually based coding algorithms on audio signals for professional recording and broadcasting. It includes psychophysical and electrical analysis under all operational and listening conditions. It includes ranking of codecs and test methods to determine presence of coders or their proper operation.

Workshop 09 **Friday, March 22**
09:45 – 10:45 **Liffey A**

AUDIO, ACCESSIBILITY, AND THE CREATIVE INDUSTRIES

Chair: **Mariana Lopez**, University of York, York, UK

Panelists: *Emilie Giles*, The Open University
Sarah McDonagh, Queen's University Belfast, Belfast, UK
Ben Shirley, University of Salford, Salford, Greater Manchester, UK; Salsa Sound Ltd., Salford, Greater Manchester, UK

Audio is a powerful tool in the provision of accessible experiences for people with disabilities. Advancements in production and broadcasting technologies means that now more than ever we should be thinking about how those can help provide inclusive experiences within the creative sector. This workshop brings together experts in the field of audio and accessibility and will focus on new research on Audio Description for visually impaired audiences, the use of object based broadcasting to better cater for the needs of audiences with hearing loss as well as the use of e-textiles to provide accessible experiences within creative contexts.

Session EB04 **Friday, March 22**
10:00 – 12:00 **Liffey B**

E-BRIEF POSTER SESSION 2

10:00

EB04-1 A Study in Machine Learning Applications for Sound Source Localization with Regards to Distance—*Hugh O'Dwyer, Sebastian Csadi, Enda Bates, Francis M. Boland*, Trinity College Dublin, Dublin, Ireland

This engineering brief outlines how Machine Learning (ML) can be used to estimate objective sound source distance by examining both the temporal and spectral content of binaural signals. A simple ML algorithm is presented that is capable of predicting source distance to within half a meter in a previously unseen environment. This algorithm is trained using a selection of features extracted from synthesized binaural speech. This enables us to determine which of a selection of cues can be best used to predict sound source distance in binaural audio. The research presented can be seen not only as an exercise in ML but also as a means of investigating how binaural hearing works.
Engineering Brief 509

10:00

EB04-2 Setup and First Experimentation Over an AES67 Over 802.11 Network—*Mickaël Henry*,[1,2] *Willy Aubry*[2]
[1] University of Grenoble, Grenoble, France
[2] Digigram, Montbonnot, France

In this paper the AES67 standard tackles the transport of audio data over IP technology. This standard was originally created for audio transmission over local area ethernet networks. However, other types of IP links exist with behavior different from ethernet. This paper investigates the setup and constraints to put on WiFi links to support AES67 transmission. We will show the limits to employ an AES67 stereo audio stream on a Wireless Local Area Network. We will detail the difficulties encountered when setting up an AES67 wireless network. Then we will ana-

lyze disruptions brought by wireless networks on device PTP offset, audio packets, and audio PSNR compared to an AES67 ethernet network. We will show the gains brought by the SMPTE 2022-7 redundancy technique.
Engineering Brief 510

10:00

EB04-3 Computational Complexity of a Nonuniform Orthogonal Lapped Filterbank Based on MDCT and Time Domain Aliasing Reduction—*Nils Werner, Bernd Edler*, International Audio Laboratories Erlangen, Erlangen, Germany

In this brief we investigate the computational complexity of a non-uniform lapped orthogonal filterbank with time domain aliasing reduction. The computational complexity of such filterbank is crucial for its usability in real-time systems, as well as in embedded and mobile devices. Due to the signal-adaptive nature of the filterbank, the actual real-world complexity will be situated between two theoretical bounds and has to be estimated experimentally by processing real-world signals using a coder-decoder pipeline. Both the bounds and the real-world complexity were analyzed in this brief, and a median 14–22% increase in complexity over an adaptive uniform MDCT filterbank was found.
Engineering Brief 511

10:00

EB04-4 Research on Reference-Processed Pair Differential Signal Objective Parameters in Relationship with Subjective Listening Tests Results in Digital Audio Coding Domain—*Krzysztof Goliasz, Sylwia Prygon, Mikolaj Zwarycz*, Dolby Poland, Wroclaw

This engineering brief presents results of the research on reference-processed pair differential signal objective parameters in relationship with subjective listening tests results in digital audio coding domain. Authors encoded set of encoder-stressful test signals using two different digital audio encoders. Various bitrates have been used in order to cover many types of audio coding artifacts. Listening tests were conducted against reference signals. Differential signals were created from reference and processed signals pairs. Set of objective parameters of differential signals have been analyzed in order to find relationship between those signals parameters and listening tests results. Conclusion was made based on found dependencies.
Engineering Brief 512

10:00

EB04-5 360° Binaural Room Impulse Response (BRIR) Database for 6DOF Spatial Perception Research—*Bogdan Ioan Bacila, Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

This engineering brief presents an open-access database for 360° binaural room impulse responses (BRIR) captured in a reverberant concert hall. Head-rotated BRIRs were acquired with 3.6° angular resolution for each of 13 different receiver positions, using a custom-made head-rotation system that was automated and integrated with the Huddersfield Acoustical Analysis Research Toolbox. The BRIRs are provided in the SOFA format. The library also contains impulse responses captured using a first-order Ambisonic microphone and an omnidirectional microphone. The database can be downloaded through the Resource section of the APL website: http://www.hud.ac.uk/apl. It is expected that the database would be useful for studying the perception of spatial attributes in a six degrees-of-freedom context.
Engineering Brief 513

10:00

EB04-6 Vertical Localization of Noise Bands Pairs by Time-Separation and Frequency Separation—*Tao Zhang, Toru Kamekawa, Atsushi Marui*, Tokyo University of the Arts, Tokyo, Japan

This study investigates how the vertical localization of the low-frequency is released from the domination of the high-frequency sound. Four different noise band pairs were used. They consist of a low-frequency band and a high-frequency band with different cut off frequency, and the low-frequency band was reproduced with a delay time of 25 ms to 200 ms to the high-frequency band. These stimuli were presented from one of five full-range speakers with different elevation set in the median plane. The speakers cover an upper vertical angle of 30 degrees and a lower vertical angle of 30 degrees. The subjects answered both the vertical sound image localization and the sound image width of the low-frequency band and the high-frequency band, respectively. As a result, while the high-frequency band was localized at the actual reproduction position, the low-frequency band showed a tendency of different vertical localization depending on the reproduction position. This tendency of the low-frequency band depending on the reproduction position was shown: in 30°, 15°, and 30°. As the delay time increased, the offset (deviation from the reproducing position) also increased, but the direction of deviation was different.
Engineering Brief 514

10:00

EB04-7 Development of a 4-pi Sampling Reverberator, VSVerb—Application to In-Game Sounds—*Masataka Nakahara,*[1,2] *Akira Omoto,*[1,3] *Yasuhiko Nagatomo*[4]
[1] ONFUTURE Ltd., Tokyo, Japan
[2] SONA Corp., Tokyo, Japan
[3] Kyushu University, Fukuoka, Japan
[4] Evixar Inc., Tokyo, Japan

The authors develop a 4-pi sampling reverberator named "VSVerb." The VSVerb restores a 4-pi reverberant field by using information of dominant reflections, i.e., dominant virtual sound sources, which are captured in a target space. The distances, amplitudes, and locations of virtual sound sources are detected from measured x, y, z sound intensities at the site, and they are translated into time responses. The generated reverb-effect by the VSVerb provides high flexibility in sound design of post-production works. In order to verify its practicability, reverb-effects at several positions in a virtual room of a video game are re-generated from a VSVerb data that is sampled at one position in a real room. Re-generated reverbs for in-game sounds are implemented into a Dolby Atmos compliant production flow, and their spatial impressions and sound qualities are listened to be checked.
Engineering Brief 515

Technical Committee Meeting
Technical Committee Meeting on Recording Technology and Practices
Friday March 22, 10:00 – 11:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-02-02 Digital Input/Output Interfacing
Friday March 22, 10:00 – 11:00
Wicklow Meeting Room 5, Level 2

The scope of SC-02-02 includes synchronization and the specification of configurations and operating limits for digital interfaces carrying audio, labeling, and control data for professional recording and broadcasting.

Tutorial 23 **Friday, March 22**
10:15 – 11:15 **Liffey Hall 1**

LIVE PRODUCTION OF BINAURAL MIXES OF CLASSICAL MUSIC CONCERTS

Presenters: **Matt Firth**, BBC R&D, Salford, UK
Tom Parnell, BBC R&D, Salford, UK

In recent years the BBC has produced binaural mixes of the BBC Proms, a series of classical music concerts, and has made them available to the public online. In this event BBC engineers will present these binaural mixes and discuss the approach used in their production; this includes custom production tools for a live broadcast environment and the training of sound engineers in spatial audio mixing. There will be an opportunity to listen to these binaural mixes on headphones.

Workshop 10 **Friday, March 22**
10:15 – 11:45 **Liffey Hall 2**

MIX IT! ARE THERE BEST MIXING PRACTICES?

Chair: **Ian Corbett**, Kansas City Kansas Community College, Kansas City, KS, USA; off-beat-open-hats recording & sound reinforcement

Panelists: *Nadezhda Rakhmanova*, Mariinsky Theatre, Russia
Paul Thompson, Leeds Beckett University, Leeds, West Yorkshire, UK

This workshop, ideal for students, new engineers, and more experienced engineers looking to speed up their workflow and productivity in studio mixing, live sound, and event production will discuss topics including:

- Starting a mix, and the key elements of a mix.
- General approaches to mixing music.
- Improving the workflow of mixing, and whether there are best practice workflows or procedures that can be followed in order to make the process of mixing together a large number of audio tracks or channels into a cohesive mix an easier, faster, and more predictable experience, and less of a mysterious black-art.
- Common mistakes to avoid.
- What do you really need, and how do you select different tools and techniques for use in your mix?
- Getting the job done and finishing the mix given real-world time and budget constraints and limitations.

Session P14 **Friday, March 22**
11:00 – 13:00 **Meeting Room 2**

DSP—PART 2

Chair: **Thomas Schmitz**, University of Liege, Liege Belgium

11:00

P14-1 Prediction of Least Significant Bits from Upper Bits in Linearly Quantized Audio Waveform—*Akira Nishimura*, Tokyo University Information Sciences, Chiba-shi, Japan

Bit-depth expansion of digital audio is essential for enhancing the quality of digital contents in re-mastering and up-conversion processes. The current study predicts the least significant bits for the bit-depth expansion from upper bits in linearly quantized samples of a framed audio waveform. A simulated annealing technique is applied to minimize the effective power of the residual signal derived from linear prediction of the framed waveform by localizing positions of the least significant bit (LSB) to be added in the frame. The results of computer simulation using various genres of 100 mono and 10-s music signals exhibit that the mean correct rate of the predicted LSB is 72% using 8-bit quantized waveforms. Measurements of the objective sound quality degradation reveal that the mean objective difference grade (ODG) of the 8-bit signals improved from –2.96 to –2.56 after addition of the predicted LSB.
Convention Paper 10179

11:30

P14-2 B-Format Decoding Based on Adaptive Beamforming—*Alexis Favrot, Christof Faller*, Illusonic GmbH, Uster, Switzerland

B-Format signals can be decoded into signals with first order directivity. For stereo and multichannel decoding it would be desirable to have more channel separation than what is achievable by first order. DirAC (directional audio coding) and HARPEX (high resolution plane wave expansion) achieve higher channel separation by means of using a parametric B-Format model to estimate plane waves and diffuse sound, and adaptively rendering those. A limitation is that plane wave and diffuse models are too simple to represent complex B-Format signals. We propose a B-Format decoder, where each channel is generated by an independent adaptive B-Format beamformer. Each beam is generated independently of the other beams, circumventing the limitation when using a single B-Format signal model.
Convention Paper 10180

12:00

P14-3 Optimizing Wide-Area Sound Reproduction Using a Single Subwoofer with Dynamic Signal Decorrelation—*Adam J. Hill, Jonathan Moore*, University of Derby, Derby, UK

A central goal in small room sound reproduction is achieving consistent sound energy distribution across a wide listening area. This is especially difficult at low-frequencies where room-modes result in highly position-dependent listening experiences. While numerous techniques for multiple-degree-of-freedom systems exist and have proven to be highly effective, this work focuses on achieving position-independent low-frequency listening experiences with a single subwoofer. The negative effects due to room-modes and comb-filtering are mitigated by applying a time-varying decorrelation method known as dynamic diffuse signal processing. Results indicate that spatial variance in magnitude response can be significantly reduced, although there is a sharp trade-off between the algorithm's effectiveness and the resulting perceptual coloration of the audio signal.
Convention Paper 10181

12:30

P14-4 Poster Introductions 8

- *Investigation of an Encoder-Decoder LSTM Model on the Enhancement of Speech Intelligibility in Noise for Hearing Impaired Listeners*—Iordanis Thoidis; Lazaros Vrysis;
Konstantinos Pastiadis; Konstantinos Markou; George Papanikolaou
- *Noise Exposure of PC Video Games Players*—Gino Iannace; Giuseppe Ciaburro; Amelia Trematerra
- *Key Benefits and Drawbacks of Surrounding Sound when Wearing Headphones or Hearing Protection*—Oscar Kårekull; Magnus Johansson
- *The Assessment of Maximum and Peak Sound Levels of F3 Category Fireworks*—Kamil Piotrowski; Adam Szwajcowski; Bartlomeij Kukulski

Workshop 11 **Friday, March 22**
11:00 – 12:30 **Liffey A**

IRISH INNOVATORS

Chair: **Alex Case**, University of Massachusetts Lowell, Lowell, MA, USA

Panelists: *Kevin Killen*, Producer/Mixer/Engineer, New York, NY, USA
Brian Masterson, Windmill Lane, Dublin, Ireland
Bill Whelan, Producer, Composer Arranger, Dublin, Ireland

A panel of engineers and producers whose audio careers began in Ireland will share the insights that have enabled them to find decades of success, even as the art, technology, and industry changed continuously around them. How did they get their start? How do they stay cutting edge? What were some of the key decision points in their career that helped them find and maintain a healthy and successful life in audio? Your career is sure to benefit from some Irish insight and inspiration. .

Technical Committee Meeting
Technical Committee Meeting on Coding of Audio Signals
Friday March 22, 11:00 – 12:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-02-08 Audio-File Transfer and Exchange
Friday March 22, 11:00 – 12:00
Wicklow Meeting Room 5, Level 2

The scope of SC-02-08 includes the specification, user implementation, and adoption of technologies for the exchange of audio data files and editing information among systems, by either transfer over a network or by exchange of physical media, for professional recording and broadcasting.

Tutorial 24 **Friday, March 22**
11:30 – 12:45 **Liffey hall 1**

MICROPHONE TECHNIQUES FOR LIVE MUSIC PRODUCTIONS (45´)

Presenter: **Cesar Lamschtein**, Auditorio Nacional del Sodre, Montevideo, Uruguay; Mixymaster, Montevi-deo, Uruguay

It is quite common to be involved in live music production where musical instruments with a significant difference in the sound level they produce share the same stage. Under those circumstances, natural acoustic musical balance is lost or cannot exist.

This matter together with instrument distribution within a stage design that is ruled by visual guidelines (video or audience) rather than musical/acoustic concerns jeopardizes the signals we get from the microphones. Signal to noise ratio falls as noise only rises (bad bleed from other instruments, stage monitoring, stage noise, ambience noise, etc.). This poor S/N ratio may lead to a loss of cla rity and continuity in the program mix and also may impeach some signal processing that may be necessary during postproduction.

We will discuss techniques borrowed from pop music production, which when applied to acoustic instruments lead to a better control of these parameters through the production process that helps not only in live music production mixing but also in acoustic music sound reinforcement situations.

Workshop 12 **Friday, March 22**
12:00 – 13:30 **Liffey Hall 2**

SPATIAL AUDIO FOR CONTEMPORARY MUSIC PERFORMANCES AND THEATER

Chair: **Enda Bates**, Trinity College Dublin, Dublin, Ireland

Panelists: *Fergal Dowling*, Dublin Sound Lab, Dublin, Ireland
Jimmy Eadie, Trinity College Dublin, Dublin, Ireland
Ji Youn Kang, Institute of Sonology, the Hague
Gráinne Mulvey, Dublin Sound Lab, Dublin, Ireland
Alexis Nealon, Quiet Music Ensemble, Dublin, Ireland

This panel discussion will focus on practical issues relating to the use of spatial audio in theater and contemporary music performances and will concentrate on the practical issues that arise in such works and how engineers can adapt artistic requirements to the particular features of different venues and loudspeaker systems.

The panel will be chaired by Enda Bates, a composer and engineer whose research frequently explores the practical issues of delivering spatial audio to distributed audiences.

Jimmy Eadie is a founding member and sound engineer for the Crash Ensemble, and also an award winning theater sound designer. His work frequently employs spatial audio, most notably for the theatrical presentation of radio plays by Samuel Beckett such as 500+ loudspeakers used for Embers, and the binaural production of Cascando, both in collaboration with PanPan Theatre.

Fergal Dowling, Gráinne Mulvey, and Alexis Nealon have composed and engineered numerous works of spatial contemporary music for composers and ensembles such Dublin Sound Lab, the Quiet Music Ensemble, Kaija Saariaho, Karlheinz Essl, and Jonathan Harvey. They will discuss the importance of communication between composers and engineers, and the challenge of implementing varying spatial strategies within a single concert programme, particularly in terms of their work on the Music Current Festivals in Dublin.

Ji Youn Kang is a composer and member of the Institute of Sonology, the Hague. Most of her music pieces have been composed based on the rites of Korean Shamanism, and many of them were written for Wave Field Synthesis System (192 loudspeakers) playback, exploring the relationship between musical and physical spaces.

Student and Career Event
SC08 STUDENT RECORDING CRITIQUES
Friday, March 22, 12:00 – 13:00
Meeting Room 1 (Genelec Demo Room)

Moderator: **Ian Corbett**, Kansas City Kansas Community College, Kansas City, KS, USA; off-beat-open-hats recording & sound reinforcement

Students! Come and get tips, tricks, and advice to push your skills to the next level! The Student Recording Critiques are non-competitive listening sessions in which students get to listen to their recordings and productions on a world-class playback system and then receive feedback from a panel of renowned industry professionals. Students at any stage of their studies can sign up to participate. Sign up at the student (SDA) booth immediately on arrival at the convention and deliver stereo mixes as 44.1 Khz/24 bit AIFF or WAVE files, to the SDA booth when you sign up. (Surround playback may be available, please check upon arrival at the event - but bring both surround and stereo options to be safe.) If you sign up, please make sure you arrive on time at the start of the session, otherwise alternates will be placed on the schedule in your place. Finalists in the Recording Competition are excluded from submitting to these events so that as many students as possible can benefit from this type of feedback. (Recording competition finalists get this feedback as part of the competition process.)

Technical Committee Meeting
Technical Committee Meeting on Perception and Subjective Evaluation of Audio Signals
Friday March 22, 12:00 – 13:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
SC-07-01 Metadata for Audio
Friday March 22, 12:00 – 13:00
Wicklow Meeting Room 5, Level 2

The scope of SC-07-01 includes formal descriptions and metadata sets for audio and audio-related elements useful to support audio operations based on the use of computers and computer networks.

Student and Career Event
SC09 EDUCATION AND CAREER/JOB FAIR
Friday, March 22, 12:30 – 14:30
Lobby Level 1

The combined AES 146th Education and Career Fair will match job seekers with companies and prospective students with schools.

Companies

Looking for the best and brightest minds in the audio world? No place will have more of them assembled than the 146th Convention of the Audio Engineering Society. Companies are invited to participate in our Education and Career Fair, free of charge. This is the perfect chance to identify your ideal new hires!

All attendees of the convention, students and professionals alike, are welcome to come visit with representatives from participating companies to find out more about job and internship opportunities in the audio industry. Bring your resume!

Schools

One of the best reasons to attend AES conventions is the opportunity to make important connections with your fellow educators from around the globe. Academic Institutions offering studies in audio (from short courses to graduate degrees) will be represented in a "table top" session. Information on each school's respective programs will be made available through displays and academic guidance. There is no charge for schools/institutions to participate. Admission is free and open to all convention attendees.

Tutorial 25 **Friday, March 22**
12:45 – 14:15 **Liffey A**

ON-LOCATION/SURROUND CLASSICAL RECORDING FOR BROADCAST WITH CENTRAL SOUND AT ARIZONA PBS

Presenter: **Alex Kosiorek**, Central Sound at Arizona PBS, Phoenix, AZ, USA

Live classical recording for broadcast distribution over a variety of delivery systems can have challenges, even more so when they are broadcast in surround. Among the concerns can be maintaining a consistent aesthetic across broadcasts, dealing with varying venue acoustics, discretion in microphone placement for minimal visual interference and/or performer/patron access, redundancy, codec, and more. Central Sound at Arizona PBS, a premier multi-award-winning provider of audio-media services, has dealt with many of these challenges, producing over 120 live classical and acoustic music productions annually for local, national and international broadcast. Whether radio, television, on-demand and/or for "live" streaming/broadcast, over 50% of the events are in surround (and soon surround w/height). Manager of Central Sound and Executive Producer of Classical Arizona PBS, Alex Kosiorek, will include microphone techniques and practical solutions that result in high-quality productions. Audio examples will be included.

Workshop 13 **Friday, March 22**
12:45 – 14:15 **Liffey Hall 1**

VIDEO CREATIONS FOR MUSIC, VIRTUAL REALITY, SIX DEGREES OF FREEDOM (6DOF) VR, AND 3D PRODUCTIONS–CASE STUDIES

Chair: **Tomasz Zernicki**, Zylia sp. z o.o., Poznan, Poland

Panelists: *Przemyslaw Danowski*, Frederic Chopin University of Music, Warsaw, Poland
Hans-Peter Gasselseder, Aalborg University, Gramastetten, Austria
Maria Kallionpää, Hong Kong Baptist University, Hong Kong
Eduardo Patricio, Zylia Sp. x o.o., Poznan, Poland

The goal of this workshop is to present spatial audio-video creations in practice. Professional audio engineers and musicians will talk about their 360°, 3D, and ambient productions combining the sound and the vision. Speakers will tell about the process of making 3D audiovisual footage displayed in the 360° dome as well as spatial recordings of the concert music. The workshop will focus especially on the usage of spherical microphone arrays that enable to record the entire 3D sound scene as well as six degrees of freedom experience (6 DoF VR). The separation of individual sound sources in post-production and Ambisonics give creators unlimited possibilities to achieve unique audio effects.

Technical Committee Meeting
Technical Committee Meeting on Broadcast and Online Delivery
Friday March 22, 13:00 – 14:00
Wicklow Meeting Room 4, Level 2

Session P15 **Friday, March 22**
13:30 – 15:30 **Meeting Room 3**

PRODUCTION AND SYNTHESIS

Chair: **Joe Timoney**, Maynooth University, Maynooth, Ireland

13:30

P15-1 Investigating the Behavior of a Recursive Mutual Compression System in a Two-Track Environment—*Hairul Hafizi Bin Hasnan, Jeremy J. Wells*, University of York, York, UK

Dynamic range compression is a widely used audio process. Recent trends in music production include the emergence of its use as a creative tool rather than just a corrective device. The control for this process is unidirectional, using one signal to manipulate one or many tracks. This paper examines the behavior of a bidirectional mutual compression system implemented in Max/MSP. Tests were conducted using amplitude-modulated sine waves that highlight different attributes.
Convention Paper 10182

14:00

P15-2 Turning the DAW Inside Out—*Charles Holbrow*, Massachusetts Institute of Technology, Cambridge, MA, USA

"Turning the DAW Inside Out" describes a speculative, internet-enabled sound recording and music production technology. The internet changed music authorship, ownership, and distribution. We expect connected digital technologies to continue to affect the processes by which music is created and consumed. Our goal is to explore an optimistic future wherein musicians, audio engineers, software developers, and music fans all benefit from an open ecosystem of connected digital services. In the process we review a range of existing tools for internet enabled audio and audio production and consider how they can grow to support a new generation of music creation technology.
Convention Paper 10183

14:30

P15-3 Real-Time Synthesis of Sound Effects Caused by the Interaction between Two Solids—*Pedro Sánchez, Joshua D. Reiss*, Queen Mary University of London, London, UK

We present the implementation of two sound effect synthesis engines that work in a web environment. These are physically driven models that recreate the sonic behavior of friction and impact interactions. The models are integrated into an online project aimed at providing users with browser-based sound effect synthesis tools that can be controlled in real time. This is achieved thanks to a physical modelling approach and existing web tools like the Web Audio API. A modular architecture was followed, making the code versatile and easy to reuse, which encourages the development of higher-level models basedon the existing ones, as well as similar models based on the same principles. The final implementations present satisfactory performance results despite some minor issues.
Convention Paper 10184

15:00

P15-4 Reproducing Bass Guitar Performances Using Descriptor Driven Synthesis—*Dave Foster, Joshua D. Reiss*, Queen Mary University of London, London, UK

Sample-based synthesis is a widely used method of synthesizing the sounds of live instrumental performances, but the control of such sampler instruments is made difficult by the number of parameters that control the output, the expertise required to set those parameters, and by the constraints of the real-time system. In this paper the principles of descriptor-driven synthesis were used to develop a pair of software tools that aid the user in the specific task of reproducing a live performance using a sampler instrument by the automatic generation of MIDI controller messages derived from analysis of the input audio. The techniques employed build on existing work and commercially available products. The output of the system is compared to manipulation by expert users. The results show that the system outperforms the human version, despite the latter taking considerably more time. Future developments of the techniques are discussed, including the application to automatic performer replication.
Convention Paper 10185

Student and Career Event
SC10 RECORDING COMPETITION—PART 2
Friday, March 22, 13:45 – 16:45
Lifey Hall 2

The Student Recording Competition is a highlight at each convention. A distinguished panel of judges participates in critiquing finalists of each category in an interactive presentation and discussion. The top three finalists in each category, as identified by our judges, present a short summary of their production intentions and the key recording and mix techniques used to realize their goals. They then play their projects for all who attend. Meritorious awards are determined here and will be presented at the closing Student Delegate Assembly Meeting (SDA-2) on Saturday. The competition is a great chance to hear the work of your fellow students at other educational institutions. Everyone learns from the judges' comments, even those who don't make it to the finals, and it's a great chance to meet other students and faculty.

Session P16 **Friday, March 22**
14:00 – 16:00 **Meeting Room 2**

ROOM ACOUSTICS

Chair: **Ben Kok**, BEN KOK Acoustic Consulting, Uden, Netherlands

14:00

P16-1 Time-Window Differences Evaluation in a Room Acoustic Sound Field Diffuseness Estimation—*Bartlomiej Chojnacki*, AGH University of Science and Technology, Cracow

Diffusion estimation is an unsolved problem, being identified in many papers for years. One of the most common problems in currently known methods is an impulse response time-window used for diffuseness estimation. Different methods for diffuseness estimation will be described, based on estimation of statistical parameters like kurtosis and standard deviation with the discussion on the time-window selection and possible solutions for this problem, considering the so called mixing time problem. This paper will discuss the topic of misunderstanding the

term of diffuseness as room acoustic measure, being the introduction to extended diffuseness estimation in multi-criteria method.
Convention Paper 10186

14:30

P16-2 Modal Decay Times in Ducts and Rooms—*Roberto Magalotti, Valentina Cardinali*, B&C Speakers, Bagno a Ripoli, Italy

In order to model the behavior of environments dominated by modal resonances, it is important to find the relationship between modal decay times and boundary conditions. The paper investigates this relationship in simple systems (rectangular duct and room), with a theoretical approach validated by FEM simulations. In the rectangular room, the classification of modes in axial, tangential, and oblique categories is helpful in assessing how the impedance of walls influences decay times. The results are compared to the Sabine equation for reverberation time. Some hints for exploiting experimentally the results are given.
Convention Paper 10187

15:00

P16-3 How to Prepare a Typical Cinema Theater to Become Multipurpose Music Venue—*Piotr Kozlowski*, Wroclaw University of Science and Technology, Wroclaw, Poland

Many small towns or villages can afford to build and maintain only one cultural facility. Such buildings have one hall that must be used to hold various meetings, concerts, performances, lectures, and film screenings. It is well known that individual stage productions have quite different requirements regarding the room acoustic conditions. In order to be able to correctly perform various stage activities in one room, it is necessary to use solutions that adjust the parameters of the room acoustics. The work presents methods for providing flexible acoustics of multipurpose venues. On the example of existing venues, the possibility of adjusting room acoustics of the cinema hall to become a good space for music and theater shows is presented.
Convention Paper 10188

15:30

P16-4 A Method for Studying Interactions between Music Performance and Rooms with Real-Time Virtual Acoustics—*Elliot K. Canfield-Dafilou, Eoin F. Callery, Jonathan S. Abel, Jonathan Berger*, CCRMA, Stanford University, Stanford, CA, USA

An experimental methodology for studying the interplay between music composition and performance and room acoustics is proposed, and a system for conducting such experiments is described. Separate auralization and recording subsystems present live, variable virtual acoustics in a studio recording setting, while capturing individual dry tracks from each ensemble member for later analysis. As an example application, acoustics measurements of the Chiesa di Sant'Aniceto in Rome were used to study how reverberation time modifications effect the performance of a piece for four voices and organ likely composed for the space. Performance details, including note onset times and pitch tracks, are clearly evident in the recordings. Two example performance features are presented illustrating the reverberation time impact on this musical material.
Convention Paper 10189

Tutorial 26 **Friday, March 22**
14:30 – 15:30 **Liffey Hall 1**

SOUND FOR EXTREME 360° PRODUCTIONS

Presenter: **Martin Rieger**, VRTONUNG, Germany

This event shows various examples of 360-degree video productions under challenging conditions, featuring location recordings and post-production.

The purpose of the talk is to give practical insights of immersive VR-videos and how sound on vision needs to be contemplated, which varies a lot from usual film formats and requires a lot of knowledge additional to audio as such. Different technologies and sometimes even custom solutions are needed on set and in post. There is no use for a boom microphone and its operator, which gets replaced by an immersive microphone array which there is, just like for 360° cameras, no perfect setup for every occasion as people tend to claim that there is.

Workshop 14 **Friday, March 22**
14:30– 16:00 **Liffey A**

CIRCLES OF CONFUSION

Chair: **Thomas Lund**, Genelec Oy, Iisalmi, Finland

Panelists: *Jamie Angus-Whiteoak*, University of Salford, Salford Greater Manchester, UK
David Griesinger, David Griesinger Acoustics, Cambridge, MA, USA
Bob Katz, Digital Domain Mastering, Orlando, FL, USA
Mandy Parnell, Black Saloon Studios, London, UK

30.000 year old cave paintings are among human beings' most impressive cultural heritage, while we are unable to experience how music by excellent composers sounded just 300 years ago. Within the last couple of generations we acquired the technical skills to record sound. However, that asset is degenerating because of cognitive limitations and the circles of confusion, because of which "smart" loudspeakers will have the potential to add even more ambiguity.

Without proper anchoring of spectral balance and level, drifting over time is foreseeable in self-referenced systems, thereby putting legacy recordings at the risk of sounding dated for no good reason, or causing irreversible distortion to be added to pieces of art.

The panel will discuss baseline listening requirements for in-room and headphone spectral balance and level that stand the test of time, putting our interests as a species above commercial trivialities.

Session P17 **Friday, March 22**
15:00 – 17:00 **Liffey B**

POSTERS: SESSION 3

15:00

P17-1 Audio Event Identification in Sports Media Content: The Case of Basketball—*Panagiotis-Marios Filippidis, Nikolaos Vryzas, Rigas Kotsakis, Iordanis Thoidis, Charalampos A. Dimoulas, Charalampos Bratsas*, Aristotle University of Thessaloniki, Thessaloniki, Greece

This paper presents an audio event recognition methodology in the case of basketball content. The proposed meth-

od leverages low-level features of the audio component of basketball videos to identify basic events of the game. Through the process of detecting and defining audio event classes, a sound event taxonomy of the sport is formed. The tasks of detecting acoustic events related to basketball games, namely referee whistles and court air horns, are investigated. For the purpose of audio event detection, a feature vector is extracted and evaluated for the training of one-class classifiers. The detected events are used to segment basketball games, while the results are combined with Speech-To-Text and text mining in order to pinpoint keywords in every segment.
Convention Paper 10190

15:00

P17-2 Objective and Subjective Comparison of Several Machine Learning Techniques Applied for the Real-Time Emulation of the Guitar Amplifier Nonlinear Behavior—*Thomas Schmitz, Jean-Jacques Embrechts*, University of Liege, Liege, Belgium

Recent progress made in the nonlinear system identification field have improved the ability to emulate nonlinear audio systems such as the tube guitar amplifiers. In particular, machine learning techniques have enabled an accurate emulation of such devices. The next challenge lies in the ability to reduce the computation time of these models. The first purpose of this paper is to compare different neural-network architectures in terms of accuracy and computation time. The second purpose is to select the fastest model keeping the same perceived accuracy using a subjective evaluation of the model with a listening-test.
Convention Paper 10191

15:00

P17-3 A Generalized Subspace Approach for Multichannel Speech Enhancement Using Machine Learning-Based Speech Presence Probability Estimation—*Yuxuan Ke*,[1,2] *Yi Hu*,[3] *Jian Li*,[1,2] *Chengshi Zheng*,[1,2] *Xiaodong Li*[1,2]
[1] University of Chinese Academy of Sciences, Beijing, China
[2] Institute of Acoustics, Chinese Academy of Sciences, Beijing, China
[3] University of Wisconsin–Milwaukee, Milwaukee, WI, USA;

A generalized subspace-based multichannel speech enhancement in frequency domain is proposed by estimating multichannel speech presence probability using machine learning methods. An efficient and low-latency neural networks (NN) model is introduced to discriminatively learn a gain mask for separating the speech and the noise components in noisy scenarios. Besides, a generalized subspace-based approach in frequency domain is proposed, where the speech power spectral density (PSD) matrix and the noise PSD matrix are estimated by short-term and long-term averaging periods, respectively. Experimental results show that the proposed method outperforms the existing NN-based beamforming methods in terms of the perceptual evaluation of speech quality score and the segmental signal-to-noise ratio improvement.
Convention Paper 10192

15:00

P17-4 Detecting Road Surface Wetness Using Microphones and Convolutional Neural Networks—*Giovanni Pepe*,[1,2] *Leonardo Gabrielli*,[2] *Livio Ambrosini*,[1,2] *Stefano Squartini*,[2] *Luca Cattani*[1]
[1] ASK Industries S.p.A., Montecavolo di Quattro Castella (RE), Italy
[2] Universitá Politecnica delle Marche, Ancona, Italy

The automatic detection of road conditions in next-generation vehicles is an important task that is getting increasing interest from the research community. Its main applications concern driver safety, autonomous vehicles, and in-car audio equalization. These applications rely on sensors that must be deployed following a trade-off between installation and maintenance costs and effectiveness. In this paper we tackle road surface wetness classification using microphones and comparing convolutional neural networks (CNN) with bi-directional long-short term memory networks (BLSTM) following previous motivating works. We introduce a new dataset to assess the role of different tire types and discuss the deployment of the microphones. We find a solution that is immune to water and sufficiently robust to in-cabin interference and tire type changes. Classification results with the recorded dataset reach a 95% F-score and a 97% F-score using the CNN and BLSTM methods, respectively.
Convention Paper 10193

15:00

P17-5 jReporter: A Smart Voice-Recording Mobile Application —*Lazaros Vrysis, Nikolaos Vryzas, Efstathios Sidiropoulos, Evangelia Avraam, Charalampos A. Dimoulas*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The evaluation of sound level measuring mobile applications shows that the development of a sophisticated audio analysis framework for voice-recording purposes may be useful for journalists. In many audio recording scenarios the repetition of the procedure is not an option, and under unwanted conditions the quality of the capturing is possibly degraded. Many problems are fixed during post-production but others may make the source material useless. This work introduces a framework for monitoring voice-recording sessions, capable of detecting common mistakes, and providing the user with feedback to avoid unwanted conditions, ensuring the improvement of the recording quality. The framework specifies techniques for measuring sound level, estimating reverberation time, and performing audio semantic analysis by employing audio processing and feature-based classification.
Convention Paper 10194

15:00

P17-6 Two-Channel Sine Sweep Stimuli: A Case Study Evaluating 2-n Channel Upmixers—*Laurence Hobden, Christopher Gribben*, Meridian Audio Ltd., Huntingdon, UK

This paper presents new two-channel test stimuli for the evaluation of systems where traditional monophonic test signals are not suitable. The test stimuli consist of a series of exponential sine sweep signals with varying inter-channel level difference and inter-channel phase difference. As a case study the test signals have been used to evaluate a selection of 2-n channel upmixers within a consumer audio-visual receiver. Results from using the new stimuli have been shown to provide useful insight for the improvement and development of future upmixers.
Convention Paper 10195

15:00

P17-7 A Rendering Method for Diffuse Sound—*Akio Ando*, University of Toyama, Toyama, Japan

This paper proposes a new audio rendering method that tries to preserve the sound inputs to both ears instead of the sound direction. It uses a conversion matrix that converts the original sound signal into the converted sound signal with the different number of channels. The least squares method optimizes the matrix so as to minimize the difference between the input signals to both ears by the original signal and those by the rendered signals. To calculate the error function, the method uses the Head Related Impulse Responses. Two rendering experiments were conducted to evaluate the method. In the first experiment, 22 channel signals of 22.2 multichannel without two LFE channels were rendered into three dimensional 8-channel signals by the conventional directional-based method and the new method. The result showed that the new method could preserve the diffuseness of sound better than the conventional method. In the second experiment, the 22 channel signals were converted into 2-channel signals by the conventional downmix method and the new method. The evaluation result based on the cross correlation coefficient showed that there were not so many differences between the downmix method and the new method. However, the informal listening test showed that the new method might preserve the diffuseness of sound better than the downmix method.
Convention Paper 10196

Technical Committee Meeting
Technical Committee Meeting on Audio Forensics
Friday March 22, 15:00 – 16:00
Wicklow Meeting Room 4, Level 2

Standards Committee Meeting
AESSC Plenary
Friday 25 May, 15:30 – 17:00
Wicklow Meeting Room 5, Level 2

Summaries of all the individual working group meetings are presented.

Tutorial 27 **Friday, March 22**
15:45 – 16:45 **Liffey Hall 1**

BINAURAL MULTITRACKING WITH BINAURAL DIRECTIONAL CONVOLUTION REVERB

Presenter: **Matthew Lien**, Whispering Willows Records Inc., Whitehorse, Yukon, Canada; Universal Music Publishing,Taipei, Taiwan

With the majority of today's music delivered over headphones, the time for binaural audio has arrived. But producing binaural music has presented challenges with some producers declaring native binaural unsuitable for popular music production. Binaural tracking often yielded poor results (especially with drums) compared to non-binaural multitracking, and the reliance on recorded room ambiance was not ideal and didn't allow for further wet/dry adjustment/enhancements when mixing.

However, recent research and new recording techniques, and the use of Directional Binaural Convolution Reverb (described in eBrief 277 "Space Explorations—Broadening Binaural Horizons with Directionally-Matched Impulse Response Convolution Reverb") are yielding impressive binaural multitracking results, even with mainstream music styles.

After gathering an extensive collection of directional binaural impulse responses from churches and halls across Europe, Asia, and Canada for the creation of directional binaural convolution reverb; and following experimental "native binaural" recording studio sessions, a new approach to native binaural multitracking has been established.

This tutorial demonstrates the various stages of this new approach to binaural music production with video and audio samples, including early and late reflection directional binaural impulse response capturing (and how each is used for best results), native binaural studio tracking of instruments ranging from drums and bass, to strings, to marimbas, to Hammond B3 and more; to mixing and mastering.

Session EB05 **Friday, March 22**
16:00 – 17:00 **Meeting Room 3**

LOUDSPEAKERS AND ASSISTIVE TECHNOLOGIES

Chair: **Kirsten Hermes**, University of Westminster, Middlesex, UK

16:00

EB05-1 Comparison of Horn Drivers' Nonlinear Distortion Measured by Different Methods—*Alexander Voishvillo, Balázs Kákonyi, Brian McLaughlin*, Harman Professional Solutions, Northridge, CA, USA

Multitone and log-sweep testing signals at progressively increasing levels were applied to a horn driver to obtain a nonlinear response. Musical signals were also applied to the driver. The acoustical signal was received at the throat of the horn and at a 1-meter distance from the horn in a 2-Pi anechoic chamber. The levels of the applied signals were incremented in 3 dB voltage steps. The initial horn driver response was corrected to provide maximum flatness and passed through a high-pass filter. Auralization examples and graphic material are demonstrated. The next stage of the research will involve subjective listening tests with signals obtained from measurements and from nonlinear models of horn driver.
Engineering Brief 516

16:15

EB05-2 New Engineering Method for Design and Optimization of Phasing Plug and Dome-Shaped Compression Chamber of Horn Drivers—*Alexander Voishvillo*, JBL/Harman Professional Solutions, Northridge, CA, USA

In this work an accurate analytical solution is found for the sound field in a dome-shaped compression chamber. This simplifies the design and optimization of the compression chamber's annular exits to suppress high-frequency air resonances. In earlier works by other authors, the solution is also found in spherical coordinates. For low-curvature chambers, an approximation in the form of Bessel function summation was used. For high-curvature compression chambers an analytical approximation did not work and FEA had to be used. The new proposed method is based on Mehler-Dirichlet analytical integral presentation of Legendre functions. This approach handles high-curvature dome chambers and does not require using numerical methods. An evaluation of this new method's applicability to chambers with various different curvatures was implemented.
Engineering Brief 517

16:30

EB05-3 Conception of an IP-based Broadcast Infrastructure Considering SMPTE 2110 Using the Example of an Transmedia Teleshopping TV Station—*Norbert Wilinski*, HSE24, Ismaning, Germany

A future-proof broadcast infrastructure is critical to the success of a live TV channel. In addition to the widespread TV formats, content must also be produced for e-commerce and social media; this is especially true for teleshopping. The production environment has to be designed flexibly in order to adapt it quickly to changing market conditions. The simultaneous production for TV and social media must be just as possible as the successive work for various media, if possible without loss of time due to technical modifications. In addition, it is particularly important in teleshopping to make the course of the programs very variable and to adapt to the interaction of the audience. There is no way around an IP-based solution, but for reasons of transmission security, conventional approaches must continue to be taken into account. SMPTE 2110, Professional Media Over Managed IP Networks, introduces a solution specifically designed for live broadcasting. A concept is presented that further develops a classically designed broadcast infrastructure based on IP.
Engineering Brief 518

16:45

EB05-4 Education and Assistive Technology for Blind and Visually Impaired Musicians, Presenters or Authors of Radio Plays and Broadcast Programs—*Marcus C. Diess*, StudioGuard, Salzburg, Austria

Many visually impaired people are predestined to work in any field of audio-production, in particular musicians with technical ambitions, authors, and presenters. Blind people often own outstanding analytic hearing to orientate in the spatial environment but require special teaching methods and assistive aids, because they cannot read screens, displays, and level-meters of studio hard- and software. Taking first steps into audio work of any kind are difficult for young blind people due to missing opportunities and professional support. Audio hard-and-software is far from barrier-free, but there are first successful attempts made to solve barriers. Such valuable efforts and an approach for teaching methods are revealed in this paper, reporting existing possibilities, and appealing to increase efforts in education opportunities for visually impaired. There is a quite remarkable number of highly professional blind musicians and sound producers, who prove that a blind person can have a career in the performing arts or music industry and take part in society as everybody else does. Methods for barrier free audio lectures for v.i. and blind people will be presented after an introduction with some general facts about visual impairment that have to be known by the reader for better comprehension of the topic. The intention is to encourage and empower visually impaired and blind people to claim participation in audio production, broadcast or even the movie industry´s new departments for audio description for visually impaired. I question the status quo and offer first solutions, improvements and research based on 30 years of practical audio work and 14 years partly awarded work with visually impaired music and radio enthusiasts. This paper is an excerpt and cannot claim being perfect in the solutions offered, as this field of research only just has begun. My intention is to share personal experience and ideas to encourage colleagues and audio professionals to go further and improve what´s already here in that context. To keep the text fluent and comprehensive the phrase "blind" stands for every kind of visual impairment (v.i.).
Engineering Brief 519

Technical Committee Meeting
Technical Committee Meeting on Audio for Cinema
Friday March 22, 16:00 – 17:00
Wicklow Meeting Room 4, Level 2

Session P18 — Friday, March 22
16:30 – 18:00 — Meeting Room 2

MIR

Chair: **Konstantinos Tsioutas**, Athens University of Economics and Business, Athens, Greece

16:30

P18-1 Evaluating White Noise Degradation on Sonic Quick Response Code (SQRC) Decode Efficacy—
Mark Sheppard,[1] *Rob Toulson*[2]
[1] Anglia Ruskin University, Cambridge, Cambridgeshire, UK
[2] University of Westminster, London, UK

With the advent of high-resolution recording and playback systems, a proportion of the ultrasonic frequency spectrum can potentially be utilized as a carrier for imperceptible data, which can be used to trigger events or to hold metadata in the form of, for example, an ISRC (International Standard Recording Code), a website address or audio track liner notes. The Sonic Quick Response Code (SQRC) algorithm was previously proposed as a method for encoding inaudible acoustic metadata within a 96 kHz audio file in the 30–35 kHz range. This paper demonstrates the effectiveness of the SQRC decode algorithm when acoustically transmitted over distance while evaluating the degradation effect of adding ultrasonic banded white noise to the pre and post transmission SQRC signal.
Convention Paper 10197

17:00

P18-2 Tagging and Retrieval of Room Impulse Responses Using Semantic Word Vectors and Perceptual Measures of Reverberation—*Emmanouil Theofanis Chourdakis, Joshua D. Reiss*, Queen Mary University of London, London, UK

This paper studies tagging and retrieval of room impulse responses from a labelled library. A similarity-based method is introduced that relies on perceptually relevant characteristics of reverberation. This method is developed using a publicly available dataset of algorithmic reverberation settings. Semantic word vectors are introduced to exploit semantic correlation among tags and allow for unseen words to be used for retrieval. Average precision is reported on a subset of the dataset as well as tagging of recorded room impulse responses. The developed approach manages to assign downloaded room impulse responses to tags that match their short descriptions. Furthermore, introducing semantic word vectors allows it to perform well even when large portions of the training data have been replaced by synonyms.
Convention Paper 10198

17:30

P18-3 A Custom Integrated Circuit Based Audio-to-CV and Audio- to-MIDI Solution—*Brian Kaczynski*, Second Sound, LLC, Miami, FL, USA

A new synthesizer technology is demonstrated that tracks the fundamental frequency of virtually any acoustic or electric instrument played monotonically. This technology relies on a mixed analog-digital application-specific integrated circuit (ASIC), which contains a very fast frequency-locked loop (FLL) that tracks with the minimum physically achievable latency of one audio cycle. The ASIC also contains a novel fundamental frequency detection circuit composed of two switched-capacitor peak detectors with decay time proportional to the fundamental period of the audio signal and a novel switched-capacitor, zero-ripple envelope follower. This frequency-tracking technology is fast enough to implement an audio-to-CV or even, with the addition of a simple microcontroller, an audio-to-MIDI solution in real time with very high accuracy and negligible latency.
Convention Paper 10199

Tutorial 28 **Friday, March 22**
17:00 – 18:00 **Liffey Hall 1**

CREATING IMMERSIVE SPATIAL AUDIO FOR CINEMATIC VR

Presenter: **Marcin Gorzel**, Google, Dublin, Ireland

In this tutorial we will focus on theory and practice for creating immersive spatial audio for cinematic VR experiences. No knowledge of spatial audio is assumed, however experience in audio technology or sound design would be helpful to understand some concepts covered in this talk.

After completing this tutorial you should have a basic understanding of theory behind spatial audio capture, manipulation, and reproduction as well as practical knowledge that will allow you to create and deploy spatial audio on YouTube. We will cover topics like: what immersive spatial audio is and how it conceptually differs from traditional multichannel stereo audio; what is Ambisonics and the notion of sound field reproduction (theory behind spatial audio); what are spatial audio recording and post-production techniques.

We will also cover some YouTube-specific topics and workflows, including preparing and uploading spatial audio to YouTube, spatial audio formats, and open source codecs supported by YouTube. You'll learn how spatial audio works on YouTube mobile (Android/iOS) and desktop platforms or audio DSP algorithms used by YouTube to render interactive binaural audio. Lastly, we will give some insights into spatial audio loudness and quality considerations as well as propose some good practices for creating spatial audio content.

Workshop 15 **Friday, March 22**
16:30 – 17:30 **Liffey A**

MEDIA PRESERVATION AND RECOVERY

Chair: **Kelly Pribble**, Iron Mountain Entertainment Services

Panelists: *Federica Bressan*, Ghent University, Ghent, Belgium
Brad McCoy, Library of Congress, Culpeper, VA, USA
Alex Tomlin, Bonded Services UK, London, UK; Iron Mountain entertainment Services UK, London, UK

Media Recovery Technology is a skill few have mastered but as our cultural heritage in the form of audio rapidly dissembles, the ability to rescue tapes and disc drives proves invaluable.

Kelly Pribble is a veteran engineer and archivist that has had extensive work in Nashville TN, London England, New York City, and Rio de Janeiro. Kelly leads Media Recovery Technology Program for audio archiving and preservation stalwart Iron Mountain. Alex Tomlin is a veteran audio engineer and archivist based in London, England, who specializes in preservation and remediation of media carriers. Dr. Federica Bressan is a post-doctoral researcher at Ghent University, where she has designed and developed software modules to automatize preservation routines and controls for data integrity verification. She has published her work in international peer-reviewed journals and conference proceedings. She has participated in several international research projects, involving some of the finest sound archives in Europe. Brad McCoy is Senior Studio Engineer in the Audio Preservation Unit at the Library of Congress Packard Campus. Brad has been active in the audio archiving/preservation community for more than 30 years at the Library of Congress. Together, this distinguished group will offer strategies and best practices for preserving and restoring audio.

This session is presented in association with the AES Technical Committee on Archiving Restoration and Digital Libraries

Workshop 16 **Friday, March 22**
17:00 – 18:00 **Liffey Hall 2**

OBJECT-BASED AUDIO PROGRAMS

Co-chairs: **Matthieu Parmentier**, France télévisions, Paris, France
Adrian Murtaza, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Panelists: *Tim Addy*, Dolby
Jacob Smith, Dolby

Object Based Audio for immersive and interactive contents has been tested several times in real conditions during the past year, thanks to the close collaborations between EBU members and NGA industrials.

With the participation of such panelists, this workshop will highlight the lessons learnt from three major use-cases:

• Roland Garros Ultra HD + NGA signal, where object-based audio has been used to feed seven simultaneous versions at the same time

• Eurovision Song Contest, where NGA has been intensively tested to offer multiple languages and musical mix versions at the receiver side

• European Athletics Championship, where a single NGA production mixed channel-based, scene-based and object-based sources to feed three different codec technologies.

Special Event
SE05 SPATIAL ELECTROACOUSTIC MUSIC CONCERT
Friday, March 22, 18:30 – 19:30
Liffey Hall 2

This concert will feature a number of works of spatial electroacoustic music for an octophonic loudspeaker array. The program will include works by panelists on the workshop on Spatial Audio for Contemporary Music Performances and Theatre such as Fergal Dowling, Gráinne Mulvey, & Ji Youn Kang, and the composers of the Spatial Music Collective.

Special Event
SE06 DIVERSITY AND INCLUSION PARTY
Friday, March 22, 20:00 – 23:00
The Brew Dock, 1 Amiens Street

This event is sponsored by iZotope.

Session P19 Saturday, March 23
09:00 – 11:00 Meeting Room 3

AUDIO AND GAMES

Chair: **Dylan Menzies**, University of Southampton, Southampton, UK

09:00

P19-1 A Framework for Understanding and Defining Quality of Musicians' Experience in Network Music Performance Environments—*Konstantinos Tsioutas*,[1] *Ioannis Doumanis*,[2] *George Xylomenos*[1]
[1] Athens University of Economics and Business, Athens, Greece
[2] University of Central Lancashire, Preston, UK

While there is considerable work on network and system level metrics related to Network Music Performance (NMP), assessing the Quality of Musician's Experience (QoME) in NMP sessions must also take into account the emotional and psychological aspects of the participants. We propose a research framework that integrates both subjective and objective aspects of musicians' experience by explicitly considering the psychological state and profile of each musician, the environment acoustic variables, and the performance of the network as the key dimensions that impact QoME. We will use the proposed framework to drive empirical studies designed to explore the QoME of musicians performing musical pieces over the Internet; this paper is a first step in this direction.
Convention Paper 10200

9:30

P19-2 Aretousa: A Competitive Audio Streaming Software for Network Music Performance—*Konstantinos Tsioutas,*[1] *George Xylomenos,*[1] *Ioannis Doumanis*[2]
[1] Athens University of Economics and Business, Athens, Greece
[2] University of Central Lancashire, Preston, UK

Many existing open source systems provide support for Network Music Performance (NMP), with each one catering to a specific system and usage scenario. As our research in evaluating the Quality of Experience (QoE) of NMP systems as perceived by musicians involves widely different scenarios and requires extensive instrumentation of the platform, we built a new NMP system, Aretousa. Our system offers a large number of configuration and monitoring options, without sacrificing latency, the most critical factor for NMP. To show that Aretousa provides flexibility while being competitive with the state of the art in terms of latency, we present measurements comparing it against JackTrip in multiple setups over a high speed research network.
Convention Paper 10201

10:00

P19-3 Measuring the Impact of Level of Detail for Environmental Soundscapes in Digital Games—*Igor Dall'Avanzi, Matthew Yee-King*, Goldsmiths College, University of London, London, UK

The design of sonic environments in digital games poses an unanswered question of believability. How much time and resources should be used to replicate an element that is stochastic and unpredictable in nature, in order to convey a satisfactory experience? We analyze the effect on player's immersion caused by the detail of digital environmental sounds (soundscapes). Two groups of participants are asked to play two different versions of the same game. One processes audio elements on run time for higher levels of detail, while the other one uses looped files. Player's immersion is measured afterwards using the Immersive Experience Questionnaire [1] and qualitative questions. Results showed no considerable difference between the two groups, and we discuss some possible explanations for this.
Convention Paper 10202

10:30

P19-4 Augmented Audio-Only Games: A New Generation of Immersive Acoustic Environments through Advanced Mixing —*Nikos Moustakas, Andreas Floros, Emmanouel Rovithis, Konstantinos Vogklis*, Ionian University, Corfu, Greece

Audio-only games represent an alternative type of gaming genres that continuously evolves following the technological trends that boost the video-games market. Since augmented reality is now a widespread approach for producing new kind of immersive applications, including games, it is expected that audio-only games will be influenced by this approach. This work represents the beginning of an attempt to investigate the process of delivering augmented audio-only games, focusing on specific technical factors that can improve the user interaction and the overall game-play experience. In particular, it focuses on a new augmented reality audio mixing process that is optimized for variable acoustic environments allowing the development of new attractive titles of audio-only games.
Convention Paper 10203

Tutorial 29 Saturday, March 23
09:00 – 10:30 Liffey A

AUDIO MASTERING'S CORE VALUES AND THE INFLUENCE OF MODERN TECHNOLOGY

Presenter: **Jonathan Wyner**, M Works Studios/iZotope/ Berklee College of Music, Boston, MA, USA

There are certain core principles and perhaps "ethics" rolled into traditional mastering practice. As technology changes we accrue opportunities to do things better and challenges when the principles are diluted or obscured. This event will highlight examples of each and offer areas of inquiry worth undertaking for anyone looking to engage in audio mastering. The end of the workshop will examine the influence of technology in recording and audio distribution on the aesthetic of mastered audio, looking ahead to changes that may be on the near horizon.

Workshop 17 Saturday, March 23
09:00 – 10:30 Liffey Hall 1

THE PRESENT AND FUTURE OF AUDIO FOR FILM AND TELEVISION

Chair: **Mariana Lopez**, University of York, York, UK

Panelists: *Howard Bargroff*
Anna Bertmark, Attic Sound Ltd., Brighton, UK
Emma Butt
Fiona Cruickshank, AIR Studios

Film and Television Sound provides an immense richness in terms of roles and specializations within the industry, and with each role comes unique technical and creative challenges. Our panel of experts will be discussing their career paths as well as how they work as part of a filmmaking team to deliver a finished piece. We will also be discussing the relationship between audio technologies and creativity and "what's next" for the world of sound for film and TV.

Session P20 **Saturday, March 23**
10:00 – 12:00 **Liffey B**

POSTERS: SESSION 4

10:00

P20-1 Quantitative Analysis of Streaming Protocols for Enabling Internet of Things (IoT) Audio Hardware—*Marques Hardin*,[1] *Rob Toulson*[2]
[1] Anglia Ruskin University, London, UK
[2] University of Westminster, London, UK

Given that traditional music production techniques often incorporate analog audio hardware, the Internet of Things (IoT) presents a unique opportunity to maintain past production workflows. For example, it is possible to enable remote digital connectivity to rare, expensive, and bespoke audio systems, as well as unique spaces for use as echo chambers. In the presented research quantitative testing is conducted to verify the performance of audio streaming platforms. Results show that using a high-speed internet connection, it is possible to stream lossless audio with low distortion, no dropouts, and around 30 ms round-trip latency. Therefore, with future integration of audio streaming and IoT control protocols, a new paradigm for remote analog hardware processing in music production could be enabled.
Convention Paper 10204

10:00

P20-2 Automatic Detection of Audio Problems for Quality Control
in Digital Music Distribution—*Pablo Alonso-Jiménez*,[1] *Luis Joglar-Ongay*,[1,2] *Xavier Serra*,[1] *Dmitry Bogdanov*[1]
[1] Universitat Pompeu Fabra, Barcelona, Spain
[2] SonoSuite, Barcelona, Spain

Providing contents within the industry quality standards is crucial for digital music distribution companies. For this reason an excellent quality control (QC) support is paramount to ensure that the music does not contain audio defects. Manual QC is a very effective and widely used method, but it is very time and resources consuming. Therefore, automation is needed in order to develop an efficient and scalable QC service. In this paper we outline the main needs to solve together with the implementation of digital signal processing algorithms and perceptual heuristics to improve the QC workflow. The algorithms are validated on a large music collection of more than 300,000 tracks.
Convention Paper 10205

10:00

P20-3 Investigation of an Encoder-Decoder LSTM Model on the Enhancement of Speech Intelligibility in Noise for Hearing Impaired Listeners—*Iordanis Thoidis, Aristotle Lazaros Vrysis, Konstantinos Pastiadis, Konstantinos Markou, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

Hearing impaired (HI) listeners often struggle to follow conversations when exposed in a complex acoustic environment. This is partly due to the reduced ability in recovering the target speech Temporal Envelope (ENV) cues from Temporal Fine Structure (TFS). This study investigates the enhancement of speech intelligibility in HI listeners by processing the ENV of speech signals corrupted by real-world environmental noise. An Encoder-Decoder Long Short Term Memory (LSTM) model is exploited after perceptually motivated processing stages to compensate for the important ENV characteristics of comprehensible speech for hearing impairment. The computational model is evaluated using the Short-Time Objective Intelligibility (STOI) measure for speech intelligibility. Finally, results indicate a 6% improvement in the mean STOI measure across different SNR values.
Convention Paper 10206

10:00

P20-4 Noise Exposure of PC Video Games Players—*Gino Iannace, Giuseppe Ciaburro, Amelia Trematerra*, Universitá della Campania "Luigi Vanvitelli," Aversa, Italy

Video games are a leisure activity that is being practiced by more and more people. Even the average age of the users is gradually increasing, representing a pleasant activity for any age. The literature has widely insinuated the doubt whether such widespread use could have negative consequences for the health of its users. This article describes noise exposure measurement activities for video game users. The damage caused by noise depends on both the acoustic power as well as the exposure time. For this reason, different noise exposure scenarios produced by video games have been simulated. The results of the study show that the daily level of noise exposure is close to the limits imposed by legislation, despite the hours of rest, and were performed in an environment with a low background noise (46.0 dBA).
Convention Paper 10207

10:00

P20-5 Key Benefits and Drawbacks of Surrounding Sound when Wearing Headphones or Hearing Protection—*Oscar Kårekull, Tech Lic, Magnus Johansson*, 3M Peltor, Värnamo, Sweden

Reproduction of sound in headphones or hearing protectors is essentially a trade-off between sound from the signal source, e.g., a cellphone, and environmental sounds. Acceptable signal to noise ratios and the useful noise level range for communication can be determined by already available measurement methods. The attenuation of surrounding noise, e.g., measured according to ISO 4869-1, can determine the signal to noise ratio but also determine the detection threshold of surrounding sound. Speech intelligibility tests can determine the level of surrounding noise where communication with nearby people is possible. In between these limits, a product can be optimized

for different situations. Examples of measured detection levels are presented and the in between performance to the speech intelligibility limit is discussed.
Convention Paper 10208

10:00

P20-6 The Assessment of Maximum and Peak Sound Levels of F3 Category Fireworks—*Kamil Piotrowski, Adam Szwajcowski, Bartłomiej Kukulski*, AGH University of Science and Technology, Kraków, Poland

Fireworks are widely discussed in the aspect of harming people. Apart from their spectacular effects, there is a serious danger of hearing loss caused by too close bursts when attending as an operator and also as a spectator. The unpleasant truth is that the documents regarding fireworks are usually not specific enough in terms of limiting the excessive noise exposure and the sound pressure levels generated by some fireworks can be simply too high. The paper presents the results of the examination on F3 class firecrackers impulsive noise. The authors measured two types of F3 explosive materials and analyzed the obtained maximum, peak, and exposure values in accordance with PN-EN 15947-4:2016-02. Pyrotechnic articles. Fireworks, Categories F1, F2 and F3. Test methods.
Convention Paper 10209

10:00

P20-7 A High Power Switch-Mode Power Audio Amplifier—*Niels Elkjær Iversen,*[1,2] *Jóhann Björnsson,*[2] *Patrik Boström,*[2] *Lars Petersen*[2]
[1] Technical University of Denmark, Kogens Lyngby, Denmark
[2] ICEpower A/S, Søborg, Denmark

Switch-mode power audio amplifiers, also known as class-D, have become the conventional choice for high power applications. This paper presents the considerations for designing a high power audio amplifier power stage. This includes an overview of loss mechanism including reverse recovery losses that are increasingly important at higher output powers. A 4 kW prototype amplifier is implemented. Absolute maximum ratings shows up to +/– 190 V output voltage swing and 6.5 kW for sine wave burst. THD+N levels go down to 0.003% for 100 Hz and are generally below 0.1% up to clipping at 4 kW in a 4 Ω load.
Convention Paper 10215

Tutorial 30 **Saturday, March 23**
10:00 – 11:00 **Liffey Hall 2**

CREATING SOUND EFFECTS LIBRARIES

Presenter: **Benjamin-Saro Sahihi**, SoundBits, Germany

You can hear them in action movies, commercials, documentaries, or video games. They can tell and support a story, define a character, make a world authentic or provide important acoustic feedback on actions or notifications. We're talking about Sound Effects. There are many, many cases where it's impossible for a Sound Designer / Sound Editor to record/edit their own sound effects for a certain project. Mostly this is because of, e.g., lack of budget, time, equipment, or no access to specific props or typical sounds of foreign countries. That is where huge general sound effects libraries and independent sound effect pack creators come into play. In this event professional sound designer and sound effects manufacturer Saro Sahihi, founder of SoundBits (www.soundbits.de) examines the workflows, tools, and specific challenges to craft sounds from scratch for sound effects libraries.

Student and Career Event
SC11 STUDENT HANDS-ON WORKSHOPS AND MASTER CLASSES ON IMMERSIVE AUDIO—COMPOSING AND PRODUCING SPATIAL AUDIO FOR 360 VIDEO USING FREEWARE PRODUCTION TOOLS
Saturday, March 23, 10:00 – 12:15
Meeting Room 5

Moderator: **Rebecca Stewart**, Imperial College, London, UK

Presenter: *Enda Bates*

This session consists of a hands on tutorial on composing and producing spatial audio for 360 video using freeware production tools, with a particular focus on Ambisonics, and spatial audio production using DAWs. The session will cover the theoretical basis of Ambisonics, and how this is put into practice using microphones and freely available, cross-platform production tools such as the DAW Reaper, the Facebook Spatial Workstation, and the ambiX plugin suite, when producing content for online platforms such as YouTube and Facebook.

Students should bring along a laptop and headphones, and download Reaper, the Facebook Spatial Workstation, and the AmbiX plugin suite in advance.

Session P21 **Saturday, March 23**
10:30 – 12:00 **Meeting Room 2**

HEARING/PERCEPTION

Chair: **Matteo Torcoli***,* Fraunhoffer IIS, Erlangen, Germany

10:30

P21-1 Measurement of Bone-Air Differential Transfer Function Based on Hearing Threshold—*Huifang Tang,*[1,2] *Jie Wang,*[3] *Jinqiu Sang,*[1,2] *Xiaodong Li*[1,2]
[1] Institute of Acoustics, Chinese Academy of Sciences, Beijing, China
[2] University of Chinese Academy of Sciences, Beijing, China
[3] Guangzhou University, Guangzhou, China

In this paper the bone-air differential transfer function (BADTF) is defined as the difference between the bone conduction (BC) and the air conduction (AC) transfer functions. This can be equivalent to the bone to air hearing threshold gap when the system is linear. The single ear BADTF was measured at 29 frequencies ranging from 0.5 to 8 kHz with 5 normal hearing subjects. Repeatability of this method was also verified. The results show that there are obvious individual differences but all the curves have similar envelopes. With the BADTF, the individual equalization can significantly improve the performance of the BC reproduction at low frequencies and make the BC sound perceived closer to the target timbre.
Convention Paper 10210

11:00

P21-2 Calibration of Digital Sound Projectors with Scene Uncertainty —*Luke Ferguson, Enda Bates, Hugh O'Dwyer, Sebastian Csadi, Francis M. Boland*, Trinity College Dublin, Dublin, Ireland

This paper addresses the calibration problem for digital

sound projection in the context of uncertain scene geometry. The image method is extended to handle uncertainties in the description of reflectors, specifically the distance parameter, relative to the source and/or receiver. Under the assumption that the source is a linear array of loudspeakers parallel with the back wall of a rectangular room, a novel extended image method is applied to compute probability distributions for the beamforming parameters for digital sound projection. The calibration is enriched with information available from probability distributions of the planar reflectors in the scene. Computer simulations are conducted to validate the calibration accuracy and to evaluate the performance of system. The expected deviation from optimum performance is quantified by analyzing the expected soundfield at the receiver position. This paper also highlights the sensitivity of digital sound projectors to measurement errors under certain constrained conditions.
Convention Paper 10211

11:30

P21-3 Perception of Auditory Events in Scenarios with Projected and Direct Sound from Various Directions—*Tom Wühle, Sebastian Merchel, Ercan Altinsoy*, TU Dresden, Dresden, Germany

Sound projecting audio systems realize the reproduction of sound from different directions via reflection paths using highly focusing sound sources. However, the limited focusing capabilities of real sources, e.g., loudspeaker arrays, cause the perception of the listener in practice to be influenced by direct sound in addition to the projected sound. This study dealt with the separation of auditory events caused by increasing perceptual dominance of the leading direct sound in sound projection. For that, the perception of auditory events for different directions of direct and projected sound and increasing direct sound level was evaluated. The separation varied with the different directions of direct and projected sound. The effect of sound projection, however, was not influenced.
Convention Paper 10212

Workshop 18 Saturday, March 23
10:30 – 12:00 Liffey A

BREAKING DOWN THE STUDIO WALL

Moderator: **Alex Case**, University of Massachusetts Lowell, Lowell, MA, USA

Presenter: *Kevin Killen*

Two iconic albums—U2's *The Unforgettable Fire*, and Peter Gabriel's *So*—were released just two years apart, and Kevin Killen was part of both. These records represented major points of creative inflection for each artist. Most of the production took place outside of any traditional studio, recording U2 in a castle and Gabriel in a cow shed. These special projects placed unique demands on the producers and engineers to deliver. *The Unforgettable Fire* was released in 1984, *So* hit the world in 1986, and the results speak for themselves. The records have become sonic touchstones for many artists and engineers since. Interviewed by Alex Case, Kevin Killen takes us into the sessions so that we might learn from his experiences.

Student and Career Event
SC12 STUDENT RECORDING CRITIQUES
Saturday, March 23, 10:30 – 11:30
Meeting Room 1 (Genelec Demo Room)

Moderator: **Ian Corbett**, Kansas City Kansas Community College, Kansas City, KS, USA; off-beat-open-hats recording & sound reinforcement

Students! Come and get tips, tricks, and advice to push your skills to the next level! The Student Recording Critiques are non-competitive listening sessions in which students get to listen to their recordings and productions on a world-class playback system and then receive feedback from a panel of renowned industry professionals. Students at any stage of their studies can sign up to participate. Sign up at the student (SDA) booth immediately on arrival at the convention and deliver stereo mixes as 44.1 Khz/24 bit AIFF or WAVE files, to the SDA booth when you sign up. (Surround playback may be available, please check upon arrival at the event - but bring both surround and stereo options to be safe.) If you sign up, please make sure you arrive on time at the start of the session, otherwise alternates will be placed on the schedule in your place. Finalists in the Recording Competition are excluded from submitting to these events so that as many students as possible can benefit from this type of feedback. (Recording competition finalists get this feedback as part of the competition process.)

Tutorial 31 Saturday, March 23
10:45 – 12:15 Liffey Hall 1

UNDERSTANDING LINE SOURCE BEHAVIOR FOR BETTER OPTIMIZATION

Presenter: **François Montignies**, L-Acoustics, Marcoussis, France

The key to a good loudspeaker system design is the balance between coverage, SPL, and frequency response performances. It deals with various challenges such as directivity control, auditory health preservation, and sonic homogeneity.

In solutions based on a variable curvature line source, the parameters linked to its physical deployment are often overlooked. The temptation to rely on electronic processing to fix resulting problems may then arise. However, it always compromises other performances to some extent, whether system headroom or wavefront integrity.

Using Fresnel analysis, this tutorial points at important aspects of line source behavior and identify the effect of determinant parameters, such as inter-element angles. It shows how an optimized physical deployment allows for rational electronic adjustments, which just become the icing on the cake.

This session is presented in association with the AES Technical Committee on Acoustics and Sound Reinforcement

Session EB06 Saturday, March 23
11:15 – 12:30 Meeting Room 3

PRODUCTION AND SIMULATION

Chair: **Ajin Tom**, McGill University, Montreal, Quebec, Canada

11:15

EB06-1 The Effect of HRTF Individualization and Head-Tracking on Localization and Source Width Perception in VR—*Hengwei Su, Atsushi Marui, Toru Kamekawa*, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

In this study the effects of head-tracking and HRTF individualization by subjective selection on localization and width perception of widen-processed sources in VR were investigated. Localization test and the perceived width evaluation were conducted under conditions with or without head-tracking and using individualized or non-individual HRTF. For the perceived width evaluation, monophonic signals were processed by a method proposed in previous studies, which aimed to create spatial extent for sound objects in the binaural synthesis. According to the results, head-tracking not only was effective to improve localization accuracies in localization test, but also could help synthesized source widths to be localized more accurately. No difference in perceived width was found under different conditions.
Engineering Brief 520

11:30

EB06-2 Does Spectral Flatness Affect the Difficulty of the Peak Frequency Identification Task in Technical Ear Training? (Part 2)—*Atsushi Marui, Toru Kamekawa*, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

Technical ear training is a method to improve the ability to focus on a specific sound attribute and to communicate using the vocabularies and units shared in the industry. In designing the successful course in a sound engineers' educational institution, it is essential to have the gradual increase of the task difficulty. In this e-Brief the authors report the relation between spectral envelope modifications on the music excerpts and the resulting objective score through the training, with variations on the music excerpts and difficulty levels.
Engineering Brief 521

11:45

EB06-3 Generalized Image Source Method as a Region-to-Region Transfer Function—*Thushara Abhayapala, Prasanga Samarasinghe*, Australian National University, Canberra, Australia

Sound propagation inside reverberant rooms is an important topic of research. This is due to its impact on a plethora of applications especially in spatial audio. Room transfer function (RTF) simulations methods play a major role in testing new algorithms before implementing them. Allen and Berkley's image source method for simulating room transfer function has recently been generalized [1] to incorporate arbitrary source and receiver directivity patterns. In this paper we further illustrate potential applications of the generalized model for recording, reproduction, and manipulating spatial audio within reverberant rooms. We provide detailed mathematical equations of (i) generalized image source method, (ii) its interpretation as a model for region-to-region transfer functions, and (iii) an outline of potential applications of the model.
Engineering Brief 522

12:00

EB06-4 Experimenting with Lapped Transforms in Numerical Computation Libraries Using Polyphase Matrices and Strided Memory Views—*Nils Werner, Bernd Edler*, International Audio Laboratories Erlangen, Erlangen, Germany

In this brief we present a framework for experimenting with lapped linear transforms in modern numerical computation libraries such as NumPy and Julia. We make use of the fact that these transforms can be represented as matrices (and oftentimes as sparse factorizations thereof), and that numerical computation libraries often support strided memory views. This strided memory view very elegantly solves the problem of processing several overlapping frames at once, while simultaneously allowing vectorization.
Engineering Brief 523

12:15

EB06-5 Suitability of Game Engines for Virtual Acoustic Experiments —*Gavriil Kamaris, Eftratios Giannatsis, Konstantinos Kaleris, John Mourjopoulos***,** University of Patras, Patras, Greece

Game engines like Unity [1], are recently becoming widely used not only by the game industry but also in the creation of serious games and for animations by the movie industry. The enhanced audio spatialization assets that have been developed for such platforms [2], may also provide a suitable development environment for interactive acoustic auralization and hence may allow the evolution of controlled acoustic and perceptual experiments within such "virtual laboratory" spaces. Here, the above concept is investigated via some preliminary experiments assessing virtual source localization achieved by using such a virtual platform within a stereo loudspeaker virtual set-up, when compared to the localization achieved using an established auralization method and a listener test. In all cases, the source localization is compared to the ideal (intended—ground truth) image source positions and the estimated differences are quantified via a perceptual binaural model [3].
Engineering Brief 524
[This e-Brief was not presented]

Tutorial 32 **Saturday, March 23**
11:15 – 12:45 **Lifey Hall 2**

EMOTIVE SOUND DESIGN IN THEORY

Presenter: **Thomas Görne**, Hamburg University of Applied Sciences, Hamburg, Germany

The tutorial explores film sound design from the viewpoints of perception, psychology, and communication science starting with the auditory and audiovisual object. A special focus is set on communication through crossmodal correspondences of auditory perception expressed in metaphors like height, brightness or size, as well as on the semantics of sound symbols, on the emotional impact of ambiguous objects, on inattentional deafness due to the limited bandwidth of conscious perception (i.e., the "inaudible gorilla"), and on image–sound relationships. Application of these principles lead to sound design concepts like realism / naturalism, the modern attention guiding "hyperrealistic" approach, and more experimental approaches like expressionism and impressionism.

Student and Career Event
SC13 STUDENT DELEGATE ASSEMBLY MEETING—PART 2
Saturday, March 23, 12:00 – 13:30
Liffey A

At this meeting the SDA will elect a new vice chair. One vote will be cast by the designated representative from each recognized AES student section in the Europe and International Regions. Judges' comments and awards will be presented for the Recording Competitions and Design Competitions. Plans for future student activities at local, regional, and international levels will be summarized.

Workshop 19 **Saturday, March 23**
12:30 – 14:00 **Lifey Hall 1**

NOISE POLLUTION FROM OUTDOOR ENTERTAINMENT EVENTS

Chair: **Adam J. Hill**, University of Derby, Derby, UK

Panelists: *Daniel Belcher*, d&b audiotechnik, Backnang, Germany
Etienne Corteel, L-Acoustics, Marcoussis, France
Christian Frick, Rocket Science GmbH, Zuerich, Switzerland
Marcel Kok, dBcontrol, Zwaag, Netherlands
Elena Shabalina, d&b audiotechnik

The issue of noise pollution due to large outdoor (and sometimes indoor) entertainment events has become increasingly problematic, with a significant number of festivals and other events taking place in centrally-located areas that are often under strict noise regulations. The AES Technical Committee on Acoustics and Sound Reinforcement (TC-ASR) formed a study group in 2018 to examine topics related to this form of noise pollution. This session will be led by members of the AES TC-ASR study group, where the group's initial findings will be presented, covering areas such as noise regulations, guidelines and prediction techniques, audience level regulations, practical measurement methods, and sound system design (primary and secondary) for limited noise spill while maintaining a high-quality audience experience.

This session is presented in association with the AES Technical Committee on Acoustics and Sound Reinforcement

Session P22 **Saturday, March 23**
14:00 – 15:00 **Meeting Room 2**

PHYSICAL SYSTEMS AND CIRCUITS

Chair: **John Robert Emmett**, Nostairway Creative, Hampton, UK

14:00

P22-1 Statistical Analysis of Audio Triode Tube Properties Based on an Advanced Physical Device Model—
Toshihiko Hamasaki, Reo Sasaki, Masaki Inui,
Hiroshima Institute of Technology, Hiroshima, Japan

The timbre of tube amplifier has been considered to depend not only on the tube type but also on the tube manufacture and its individual difference as well. However, quantitative difference of tube properties of manufacture has not been clarified yet. In this study the manufactures differences of triode tube 12AX7/ECC83 are analyzed statistically by an advanced physical device model. The model parameter values are extracted from measured family curves of 60 devices in total for 5 major tube manufactures for a guitar amplifier. The characteristics of each manufacture tube are clarified by an average and a dispersion of respective parameter value set. Furthermore, the root cause of the significant manufacturing process instability is identified based on the correlation of all combination of parameters for the first time. The tube properties of 5 manufactures are divided into 3 groups by clustering analysis using cosine similarity on the vector of parameters.
Convention Paper 10213

14:30

P22-2 A Novel Digital Radio-Frequency Capacitor Microphone with Gain Rangin—*Lars Urbansky, Udo Zölzer*, Helmut-Schmidt-University Hamburg, Hamburg, Germany

Most capacitor microphones use an audio-frequency (AF) implementation.InanAFcircuit,acapacitorischargedwitha constant bias voltage leading to a high-impedance circuit. In contrast, by using a radio-frequency (RF) approach, the capacitor is operated on a higher frequency band which reduces the circuit's impedance. However, state of the art RF microphones are entirely analog. Thus, a novel digital RF condenser microphone system is proposed. Furthermore, it is extended by a corresponding gain ranging approach. The expected advantages are a further improved demodulation linearity due to a digital demodulation and a circumvention of analog disadvantages due to the smaller required analog circuit. Additionally, because of the analog bandpass signal, it is expected to utterly bypass the electrical low frequency 1/f noise.
Convention Paper 10214

Tutorial 33 **Saturday, March 23**
12:45 – 14:15 **Liffey Hall 2**

EMOTIVE SOUND DESIGN IN PRACTICE

Presenter: **Anna Bertmark**, Attic Sound Ltd., Brighton, UK

How can sound give a sense of emotional perspective in cinema? Putting the audience in the shoes of a character in a story can have a powerful impact on empathy and immersion. Sound designer Anna Bertmark gives an insight into her work and the techniques used to depict the point of view of characters. Using clips from films such as *God's Own Country* (for which she won the 2017 BIFA for Best Sound), *Adult Life Skills*, and *The Goob*, among others, she will talk about the reasons and ideas behind these techniques, from contact mic recording through to the final result as heard in the cinema.

Tutorial 34 **Saturday, March 23**
13:45 – 15:15 **Liffey A**

MIX KITCHEN: MIXING TIPS AND TRICKS FOR MODERN MUSIC GENRES

Presenter: **Marek Walaszek**, Bettermaker

In today's modern genres of pop, hip hop, and EDM there is a fine line between production, mixing, and mastering. It is not uncommon that at least two of above are done simultaneously in music creation process. Marek "Maro" Walaszek will take an in-depth look into today's mixing and mastering techniques and how they affect sound.

Tutorial 35 **Saturday, March 23**
14:00 – 14:45 **Liffey Hall 1**

NOISE PREDICTIONS WITH SOUND SYSTEMS USING SYSTEM DATA & COMPLEX SUMMATION (IMPLEMENTED IN SOUNDPLAN & NOIZCALC)

Presenter: **Daniel Belcher**, d&b audiotechnik, Backnang, Germany

As the number of outdoor events in urban areas is increasing, so are the challenges with and the awareness of the accompanied noise emission and therefore the need for accurate noise predictions. A new method for such noise predictions with sound reinforcement systems for outdoor events is introduced that uses system data of actual system designs and applies complex summation.

The system data including all electronic filters is simply imported with a system design file. This procedure eliminates any friction losses because it spares the repeated process of remodeling a sound system in the noise prediction software and therefore ensures that the prediction is done with the actual system design. It also is the only sensible way to include the specific electronic filters (also IIR, FIR) of a sound system.

Noise prediction software did not consider complex summation simply because noise sources are always assumed not to be correlated with each other (e.g., in traffic and industry). This is where sound systems are quite different because the signals sent to the loudspeakers at different positions are usually always correlated. Furthermore, modern sound systems even use coherence effects to influence directivity (e.g., arrays).

Technical Committee Meeting
Technical Council Plenary
Saturday March 23, 14:00 – 16:00
Wicklow Meeting Room 5, Level 2

Workshop 20 — Saturday, March 23
14:30 – 15:30 — Liffey Hall 2

EXPERIMENTAL APPROACH IN 3D AUDIO PRODUCTION

Presenters: **Pawel Malecki**, AGH University of Science and Technology, Krakow, Poland
Szymon Aleksander Piotrowski, Psychosound Studio, Kraków, Poland

This workshop presents several audio projects involving various spatial sound techniques. Our presenters are going to demonstrate implementation of spatial impulse responses and discuss challenges of ambisonics and auralization in music production. Examples include virtual choir project (recording and production of 3D virtual choir in big church), early period classical music, and electronic music. For each production workflow, benefits and problems for each employed method are discussed.

Tutorial 36 — Saturday, March 23
15:00 – 16:30 — Liffey Hall 1

MODERN SAMPLING PART 2: SPARSE / COMPRESSIVE SAMPLING / SENSING

Presenter: **Jamie Angus-Whiteoak**, University of Salford, Salford, Greater Manchester, UK; JASA Consultancy, York, UK

Sparse and Compressive Sampling allows one to sample the signal at apparently less than the Nyquist/Shannon limit of two times the highest frequency without losing any signal fidelity. If some loss of fidelity is allowed the signal can be sampled at an even lower average rate. How can this be? The answer is that the effective information rate is actually lower than the highest frequency.

The purpose of this tutorial is to give a (mostly) non-mathematical introduction to Sparse/Compressive Sampling. We will examine the difference between "Sparse" and "Dense" signals and define what is meant by "rate of innovation" and see how it relates to sample rate. We will then go on to see how we can create sparse signals either via transforms or filters to provide signals that can be sample at much lower rates. We will then show how some of these methods are already used in audio, and suggest other areas of application, such as measurement. Finally we will finish off by showing how a commonly used audio system can be considered to be a form of compressive sensing.

This session is presented in association with the AES Technical Committee on High Resolution Audio