AES 144TH CONVENTION PROGRAM

MAY 23–MAY 26, 2017

NH HOTEL MILANO, CONGRESS CENTRE, MILAN

The Winner of the 144th AES Convention Best Peer-Reviewed Paper Award is:

Real-Time Conversion of Sensor Array Signals into Spherical Harmonic Signals with Applications to Spatially Localized Sub-Band Sound-Field Analysis—*Leo McCormack,*[1] *Symeon Delikaris-Manias,*[1] *Angelo Farina,*[2] *Daniel Pinardi,*[2] *Ville Pulkki*[1]
[1] Aalto University, Espoo, Finland
[2] Università di Parma, Parma, Italy

Convention Paper 9939

To be presented on Thursday, May 24
in Session 7—Spatial Audio—Part 1

* * * * *

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 144th 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 144th AES Convention Student Paper Award is:

Bandwidth Extension Method Based on Generative Adversarial Nets for Audio Compression—Qingbo Huang, Xihong Wu, Tianshu Qu, Peking University, Beijing, China

Convention Paper 9954

To be presented on Thursday, May 24
in Session 10—Audio Coding, Analysis, and Synthesis

* * * * *

Technical Comittee Meeting
Wednesday, May 23, 09:00 – 10:00 **Room Brera 1**

Technical Committee Meeting on Recording Technology and Practices

Tutorial 1 **Wednesday, May 23**
09:15 – 10:15 **Scala 2**

CRASH COURSE IN 3D AUDIO

Presenter: **Nuno Fonseca**, Polytechnic Institute of Leiria, Leiria, Portugal; Sound Particles

A little confused with all the new 3D formats out there? Although most 3D audio concepts already exist for decades, the interest in 3D audio has increased in recent years, with the new immersive formats for cinema or the rebirth of Virtual Reality (VR). This tutorial will present the most common 3D audio concepts, formats, and technologies allowing you to finally understand buzzwords like Ambisonics/HOA, Binaural, HRTF/HRIR, channel-based audio, object-based audio, Dolby Atmos, among others.

Workshop 1 **Wednesday, May 23**
09:15 – 10:15 **Scala 3**

CANCELED

Session P1 **Wednesday, May 23**
09:30 – 12:30 **Scala 4**

LOUDSPEAKERS—PART 1

Chair: **Angelo Farina**, University of Parma, Parma, Italy

09:30

P1-1 Maximizing Efficiency in Active Loudspeaker Systems—*Wolfgang Klippel*, Klippel GmbH, Dresden, Germany

Increasing the efficiency of the electro-acoustical conversion is the key to modern audio devices gen-

erating the required sound output with minimum size, weight, cost, and energy. There is unused potential for increasing the efficiency of the electrodynamical transducer by using a nonlinear motor topology, a soft suspension, and cultivating the modal resonances in the mechanical and acoustical system. However, transducers optimized for maximum efficiency are more prone to nonlinear and unstable behavior. 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 discusses the design of modern active systems that combine the new opportunities provided by software algorithms with the optimization of the hardware components in the transducer and power amplifier.
Convention Paper 9908

10:00

P1-2 How Do We Make an Electrostatic Loudspeaker with Constant Directivity?—*Tim Mellow*, Mellow Acoustics Ltd., Farnham, Surrey, UK

The idea of broadening the directivity pattern of a push-pull electrostatic loudspeaker by partitioning the stators into concentric annular rings, which are connected to tappings along a delay line, isn't new. However, the delay line has traditionally been attenuated to avoid response irregularities due to the finite size of the membrane. An alternative approach is presented here whereby a constant-impedance delay line is configured to imitate an oscillating sphere, which is an ideal constant-directivity dipole source that needs no attenuation. Walker's equation for the on-axis pressure does not account for the effect of the delay line without taking the vector sum of the currents though all the rings, so a simple alternative that does is presented here.
Convention Paper 9909

10:30

P1-3 Analysis of Front Loaded Low Frequency Horn Loudspeakers—*Bjørn Kolbrek*, Celestion International Ltd., Ipswich, UK

The low frequency horn design procedures described by Keele and Leach are extended and generalized to cases where the horn is already specified, or where maximum output or the smoothest response is desired. The impact of finite-length horns is analyzed. A more detailed analysis of the high frequency range is given, where it is shown how the voice coil inductance can be taken into account to create a third order low pass filter of specified shape. A new analysis of reactance annulling is presented that significantly improves the performance above cutoff for a certain class of horns. Large signal behavior is touched upon, and finally, an analysis of the sensitivity of driver and system parameters is given.
Convention Paper 9910

11:00

P1-4 Design and Implementation of a High Efficiency Subwoofer—*Sebastian Tengvall, Niels Elkjær Iversen, Arnold Knott*, Technical University of Denmark, Kgs. Lyngby, Denmark

The demand for battery driven loudspeakers is increasing but the challenge of efficient low frequency reproduction remains. An alternative approach to the conventional 4th order bandpass enclosure design for a subwoofer to achieve a high peak in the passband and increase voltage sensitivity is investigated. The response is corrected with DSP to ensure a flat response in the passband. The results proved that this approach can increase the voltage sensitivity dramatically, reaching an average sensitivity of over 100 dB in the passband from 45 Hz to 90 Hz. It also showed that the design is sensitive to construction errors. Precise assembling is required to achieve satisfactory results while small errors can ruin the purpose of the design.
Convention Paper 9911

11:30

P1-5 Parameterization of Micro Speakers Used in Cell Phones and Earbuds—*Jason McIntosh*, SAATI SPA, Appiano Gentile (C0), Italy

Loudspeaker parameterizations based on measuring a transfer matrix is presented. This approach produces nine complex frequency dependent functions. While Thiele-Small parameters only capture the first piston mode of a moving coil speaker, the transfer matrix approach captures all the linear behavior of the speaker, including the diaphragm modes and internal geometry. Predictions of baffled response of two speakers are presented with the transfer matrix parameters producing better results than the Thiele-Small parameters, especially at high frequencies. The SAATI "Ares" acoustic simulator uses the transfer matrix parameters for effective simulation of complete devices, including proper porting geometry and dampening acoustic meshes for better audio tuning.
Convention Paper 9912

12:00

P1-6 A Fast And Efficient Model for Transmission Line Loudspeakers—*James Hipperson*,[1,2] Jamie Angus,[1,3] Jonathan Hargreaves[1]
[1]University of Salford, Salford, UK
[2]16 Longhurst House, Horsham, UK
[3] JASA Consultancy, York, UK

Transmission Line loudspeakers use a tube behind the driver that is lined, or filled, with absorber to remove the rear radiation. They also use the resonances of the pipe to support the radiation of the driver and reduce displacement at low frequencies. While lumped element models are used for modeling sealed and vented box enclosures, they cannot be used for transmission line loudspeakers because they cannot be accurately modeled as a lumped element. Finite Element and Boundary Element models can be used but they are complex and computationally expensive. A cascaded two port method has been developed that can model varying tube area and absorption. It has been evaluated against acoustic measurements and shown to provide accurate predictions.
Convention Paper 9913

Tutorial 2 **Wednesday, May 23**
09:30 – 11:00 **Scala 1**

REUSING AND PROTOTYPING TO ACCELERATE INNOVATION IN AUDIO SIGNAL PROCESSING

Presenters: **Gabriele Bunkheila**, MathWorks, Cambridge, UK
Jonas Rutstrom, MathWorks, Sollentuna, Sweden

Voice assistants are shifting consumer expectations on performance and capabilities of audio devices and human-machine interfaces. As new products are driven to deliver increasingly complex features, successful manufacturers and IP providers need to reuse more design assets, deliver original innovation more efficiently, and prototype more quickly than ever before. In this session you will learn about different techniques to integrate existing code and IP into early simulations of algorithms and system designs, ranging from embeddable code to cloud-based services. You will also be exposed to quick prototyping workflows, including methods for running in real-time and validating ideas on live real-world signals. The presentation will go through practical worked examples using MATLAB, while discussing some early-stage challenges in the design of voice-driven connected devices..

Workshop 2 **Wednesday, May 23**
10:00 – 11:30 **Lobby**

VIRTUAL REALITY AUDIO: B-FORMAT PROCESSING

Chair: **Christof Faller**, Illusonic GmbH, Uster, Zürich, Switzerland; EPFL, Lausanne, Switzerland

Panelists: *Svein Berge*, Harpex Ltd., Berlin, Germany
Ben Sangbae Chon, Gaudio Lab, Inc., Seoul, Korea
Ville Pulkki, Aalto University, Espoo, Finland
Oliver Thiergart, International Audio Laboratories Erlangen, Erlangen, Germany

B-Format has had a revival in recent years and has established itself as the audio format of choice for VR videos and content. Experts in signal processing and production tools are presenting and discussing latest innovations in B-Format processing. This includes processing on the recording and rendering side and B-Format post-production.

Technical Comitee Meeting
Wednesday, May 23, 10:00 – 11:00 **Room Brera 1**

Technical Committee Meeting on Signal Processing

Tutorial 3 **Wednesday, May 23**
10:15 – 11:15 **Scala 3**

PERCEPTUALLY MOTIVATED FILTER DESIGN WITH APPLICATIONS TO LOUDSPEAKER-ROOM EQUALIZATION

Presenter: **Balázs Bank**, Budapest University of Technology and Economics, Budapest, Hungaryl

Digital filters are often used to model or equalize acoustic or electroacoustic transfer functions. Applications include headphone, loudspeaker, and room equalization, or modeling the radiation of musical instruments for sound synthesis. As the final judge of quality is the human ear, filter design should take into account the quasi-logarithmic frequency resolution of the auditory system. This tutorial presents various approaches for achieving this goal, including warped FIR and IIR, Kautz, and fixed-pole parallel filters, and discusses their differences and similarities. It also shows their relation to fractional-octave smoothing, a method used for displaying transfer functions. With a better allocation of frequency resolution, these methods require a significantly lower computational power compared to straightforward FIR and IIR designs at a given sound quality.

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

Session P2 **Wednesday, May 23**
10:30 – 12:30 **Scala 2**

AUDIO QUALITY—PART 1

Chair: **Thomas Sporer**, Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

10:30

P2-1 **An Auditory Model-Inspired Objective Speech Intelligibility Estimate for Audio Systems—***Jayant Datta, Xinhui Zhou, Joe Begin, Mark Martin*, Audio Precision, Beaverton, OR, USA

Compared with subjective tests, objective measures save time and money. This paper presents the implementation of a new algorithm for objective speech intelligibility, based on the modified rhyme test using real speech. An auditory-model inspired signal processing framework approach gathers word selection evidence in auditory filter bank correlations and then uses an auditory attention model to perform word selection. It has been shown to outperform popular measures in terms of Pearson correlation coefficient to the human intelligibility scores. A real-time version of this approach has been integrated into a versatile audio test and measurement system supporting a number of interfaces (different combinations of devices/channels/systems). Examples and measurement results will be presented to show the advantages of this approach.
Convention Paper 9918

11:00

P2-2 **A Statistical Model that Predicts Listeners' Preference Ratings of Around-Ear and On-Ear Headphones—***Sean Olive, Todd Welti, Omid Khonsaripour*, Harman International, Northridge,
CA, USA

A controlled listening test was conducted on 31 different models of around-ear (AE) and on-ear (OE) headphones to determine listeners' sound quality preferences. One-hundred-thirty listeners both trained and untrained rated the headphones based on preference using a virtual headphone method that used a single replicator headphone equalized to match magnitude and minimum phase responses of the different head phones. Listeners rated seven different headphones in each trial that included high (the new Harman AE-OE target curve) and low anchors. On average, both trained and untrained listeners preferred the high anchor to 31 other choices. Using machine learning a model was developed that predicts the listeners' headphone preference ratings of the headphones based on deviation in magnitude response from the Harman target curve.
Convention Paper 9919
[Paper presented by Todd Welti]

11:30

P2-3 **Comparing the Effect of HRTF Processing Techniques on Perceptual Quality Ratings—***Areti Andreopoulou,*[1] *Brian F. G. Katz*[2]
[1] National and Kapodistrian University of Athens, Athens, Greece
[2] Sorbonne Université, CNRS, Paris, France

The use of Head-Related Transfer Functions for binaural rendering of spatial audio is quickly emerging in today's

audio market. Benefits of individual HRTFs, or personalized HRTF selection, have been demonstrated in numerous previous studies. A number of recent works have examined assisted or automated selection of HRTFs for optimized personalization. Such techniques attempt to rank HRTFs according to expected spatial quality for a given user based on signal, morphological, and/or perceptual studies. In parallel, there exist several HRTF processing methods that are often used to compact and/or smooth HRTFs in order to facilitate real-time treatments. Nevertheless, the potential impact of such processes on HRTF spatial quality is not always considered. This study examines the effects of three commonly used HRTF processing techniques (spectral smoothing in constant absolute bandwidths, minimum-phase decomposition, and infinite impulse response modeling) on perceptual quality ratings of selected HRTFs. Results showed that the frequency and phase-spectra variations introduced in the data by the three processing methods can lead to significant changes in HRTF evaluations. In addition, they highlight the challenging nature of non-individualized HRTF rating tasks and establish the need for systematic participant screening and sufficient task repetitions in perceptual HRTF evaluation studies.
Convention Paper 9920

12:00

P2-4 The Effect of Visual Cues and Binaural Rendering Method on Plausibility in Virtual Environments—*Will Bailey, Bruno Fazenda*, University of Salford, Salford, Greater Manchester, UK

Immersive virtual reality is by its nature a multimodal medium and the use of spatial audio renderers for VR development is widespread. The aim of this study was to assess the performance of two common rendering methods and the effect of the presence of visual cues on plausibility of rendering. While it was found that the plausibility of the rendered audio was low, the results suggest that the use of measured responses performed comparatively better. In addition, absence of virtual sources reduced the number of simulated stimuli identified as real sources and complete absence of visual stimuli increased the rate of simulated audio identified emitted from the loudspeakers.
Convention Paper 9921

Technical Comittee Meeting
Wednesday, May 23, 11:00 – 12:00 Room Brera 1

Technical Committee Meeting on Spatial Audio

Session EB1 Wednesday, May 23
11:15 – 12:45 Arena 2

POSTERS 1

11:15

EB1-1 Experimental Study on Sound Quality of Various Audio Fade Lengths—*Jedrzej Borowski, Krzysztof Bulawski, Krzysztof Goliasz*, Dolby Poland, Wroclaw, Poland

The aim of this paper is learning the shortest possible lengths of audio fades and crossfades that are not audible as audio artifacts. The determined lengths can be utilized in adaptive streaming scenarios during pauses and content switches. Subjective and objective tests were performed, utilizing speech and music signals with various fade-out and fade-in lengths. Subjective evaluation was performed by critical listening tests where listeners were asked to grade the quality of the fade-out or fade-in and listen for unwanted audio artifacts. Basing on the subjective test results, optimal ranges of fade-out and fade-in times were selected—50 to 100 ms for fade-in, and 100 to 200 ms for fade-out. Objective tests were conducted using optimal times chosen by the listening tests. The results confirm that the selected ranges of fade-in and fade-out lengths do not introduce significant har-monic distortion and noise into the signal.
Engineering Brief 404

11:15

EB1-2 3D Sound Intensity Measurement of 1241 Sound Objects on Fine Panning Grids by Using a Virtual Source Visualizer—*Takashi Mikami,*[1] *Masataka Nakahara*,[2] *Kazutaka Someya,*[3] *Akira Omoto*[4]
[1] SONA Co., Tokyo, Japan
[2] ONFUTURE Ltd., Tokyo, Japan
[3] beBlue Co., Ltd. ,Tokyo, Japan
[4] Kyushu University, Fukuoka, Japan;

3D sound intensity measurement of 1241 sound objects rendered by Dolby Atmos on fine panning grids was carried out by using a Virtual Source Visualizer (VSV). The obtained sound localizations were visualized as a 3D panning map. To evaluate properties of reproduced sound fields for several kinds of rendering systems in various rooms, the authors have previously carried out VSV measurements of sound objects on some main panning positions. The results roughly illustrated each acoustic feature of rendered sound fields. This measurement was carried out to find out the relationship between 3D panner's indications and physical sound localizations on finer scale. The visualized sound localizations formed "3D panning position map" that clearly shows the relationship between them.
Engineering Brief 405

11:15

EB1-3 SOFA Native Spatializer Plugin for Unity—Exchangeable HRTFs in Virtual Reality—*Claudia Jenny*,[1,2] *Piotr Majdak*,[2] *Christoph Reuter*[1]
[1] University of Vienna, Vienna, Austria
[2] Austrian Academy of Sciences, Vienna, Austria

In order to present three-dimensional virtual sound sources via headphones, head related transfer functions (HRTFs) can be integrated in a spatialization algorithm. However, the spatial perception in binaural virtual acoustics may be limited if the applied HRTFs differ from those of the actual listener. Thus, SOFAlizer, a spatialization engine allowing to use and switch on-the-fly between listener-specific HRTFs stored in the spatially oriented format for acoustics (SOFA) was implemented for the Unity game engine. With that plugin, virtual-reality headsets can benefit from the individual HRTF-based spatial sound reproduction.
Engineering Brief 406

11:15

EB1-4 PR-VR: Approaching Sound Field Recording in Multi-Reality Environments—*Tiernan Cross*, University of Sydney, Sydney, Australia

This brief will communicate the author's recent research that questions what it is to expand the horizon of field recording beyond the physical sense of sonic immediacy into the simultaneous recording of mixed physical, technological

and network-based realities. In doing so this work proposes a reconstruction to what constitutes a modern sound recordist's immediate sonic environment in today's technologically inundated atmospheres. This brief will discuss the architecture of a PR-VR, software-based audio device capable of recording real-time, multichannel inputs and field recordings from physical, technological, and virtual acoustic spaces concurrently. By blending multi-reality input streams through algorithm this research looks to explore how modern technology can sculpt new variegated sound field recordings and formulate new hybrid, soundscapes.
Engineering Brief 407

11:15

EB1-5 Development and Validation of a Simulation Model for Prediction of Properties of Acoustic Diffusers —*Adam Kurowski, Damian Koszewski, Józef Kotus, Bozena Kostek*, Gdansk University of Technology, Gdansk, Poland

A mathematical simulation is a common approach for early stages of the design process of many devices. In this paper we would like to present a new way to use an open source solution for simulation of scattering phenomena occurring in proximity of acoustic diffusers. The result of our work is a script based on FEniCSÑan open source computing platform for FEM-based solution of differential equations. A ground truth data for validation of our model are ob-tained with the use of sound intensity measurement employing a p-u probe. A visualization and comparison of data obtained from measurement and simulation are presented and discussed.
Engineering Brief 408

11:15

EB1-6 The Immersive Media Laboratory: Installation of a Novel
Multichannel Audio Laboratory for Immersive Media Applications—*Robert Hupke, Marcel Nophut, Song Li, Roman Schlieper, Stephan Preihs, Jürgen Peissig*, Leibniz Universität Hannover, Hannover, Germany

This engineering brief presents the novel multichannel audio laboratory for immersive media applications of the Institut für Kommunikationstechnik (IKT) with its varying multichannel loudspeaker arrangements and acoustical transparent projection screens of nearly 270°. We address the construction process and setup of the laboratory called Immersive Media Lab (IML). It was designed in compliance to the strict recommendations of the ITU-R BS.1116-3 in order to conduct research in 3D audio reproduction. Our brief will first address issues of space and room dimensions as well as the acoustical design of the new listening room. Furthermore, the flexible loudspeaker arrangement consisting of 28 active loudspeakers as well as the projection setup consisting of 3 high definition ultra-short throw video projectors is described.
Engineering Brief 409

11:15

EB1-7 Can Visual Priming Affect the Perceived Sound Quality of a Voice Signal in Voice over Internet Protocol (VoIP) Applications?—*Jack Haigh,[1] Chris Exton,[1] Malachy Ronan[2]*
[1] University of Limerick, Limerick, Limerick, Ireland
[2] Limerick Institute of Technology, Limerick, Ireland

Verbal suggestions of loudness changes have been reported to result in significantly higher loudness ratings than those of a control group [1]. This study seeks to extend these results to VoIP applications by implementing visual priming cues within a VoIP interface and assessing their effect on audio quality ratings. A list of common visual priming cues was compiled and cross-referenced with prevalent design features found in popular mobile VoIP Applications. Fourteen participants were divided into two groups: one received embedded priming cues and one did not. Quality ratings were gathered using a MOS rating scale. The results are presented and their relevance discussed.
Engineering Brief 410

11:15

EB1-8 Communication Through Timbral Manipulation: Using Equalization to Communicate Warmth—Part 1—*Alejandro Aspinwall*, McGill University, Montreal, QC, Canada; Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Montreal, Quebec, Canada

With the advent of new technologies that allow for virtually any modern computer to process high quality audio, many musicians and amateur players are presented with a plethora of sound sculpting tools. Some of these display subjective attributes such as warmth punch and shimmer. If engineers are able to manipulate the timbre of a recorded sound using equalization, are they then able to use this ability to convey specific perceptual intentions (to make a sound "crunchy," "bright," or "warm" for instance). Using Juslin's standard paradigm, this study explores the question: How effective are audio engineers in communicating warmth when applying equalization?
Engineering Brief 411

11:15

EB1-9 An Open Realtime Binaural Synthesis Toolkit for Audio Research—*Andreas Franck,[1] Giacomo Costantini,[1] Chris Pike,[2] Filippo Maria Fazi[1]*
[1] University of Southampton, Southampton, Hampshire, UK
[2] BBC R&D, Salford, Greater Manchester, UK

Binaural synthesis has gained fundamental importance both as a practical sound reproduction method and as a tool in audio research. Binaural rendering requires significant implementation effort, especially if head movement tracking or dynamic sound scenes are required, thus impeding audio research. For this reason we propose the Binaural Synthesis Toolkit (BST), a portable, open source, and extensible software package for binaural synthesis. In this paper we present the design of the BST and the three rendering approaches currently implemented. In contrast to most other software, the BST can easily be adapted and extended by users. The Binaural Synthesis Toolkit is released as an open software package as a flexible solution for binaural reproduction and to foster reproducible research in this field.
Engineering Brief 412

11:15

EB1-10 Measurement of Latency in the Android Audio Path—
Szymon Zaporowski, Maciej Blaszke, Dawid Weber, Gdansk University of Technology, Gdansk, Poland

This paper provides a description of experimental investigations concerning comparison between the audio path characteristics of various Android versions. First, information about the changes in each system version in the context of latency caused by them is presented. Then, a

measurement procedure employing available applications to measure latency is described comparing to results contained in the Internet. Finally, a comparison between tested systems and results of tests are presented along with conclusions on possible audio processing implementations on the Android platform.
Engineering Brief 413

Tutorial 4 **Wednesday, May 20**
11:15 – 12:15 **Scala 1**

BUILD A SYNTH FOR ANDROID

Presenter: **Don Turner**, Developer Advocate, Android Audio Framework, UK

With 2 billion users Android is the world's most popular operating system, and it can be a great platform for musical creativity. In this session Don Turner (Developer Advocate for the Android Audio Framework) will build a synthesizer app from scratch* on Android. He'll demonstrate methods for obtaining the best performance from the widest range of devices, and how to take advantage of the new breed of low latency Android "pro audio" devices. The app will be written in C and C++ using the Android NDK APIs.

*Some DSP code may be copy/pasted

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

Workshop 3 **Wednesday, May 23**
11:15 – 12:45 **Scala 3**

AUDIO REPURPOSING USING SOURCE SEPARATION

Chair: **Philip Coleman**, University of Surrey, Guildford, Surrey, UK

Panelists: *Estefanía Cano Cerón*, Fraunhofer Institute for Digital Media Technology (IDMT), Ilmenau, Germany
Chungeun Kim, University of Surrey, Guildford, Surrey, UK
Jon Francombe, BBC Research and Development, Salford, UK
Jouni Paulus, Fraunhofer IIS, Erlangen, Germany; International Audio Laboratories Erlangen, Erlangen, Germany

Source separation tries to extract sound objects from an existing mixture. In reality, perfect separation is not achievable; the sound quality of single extracted sources is often heavily degraded. Fortunately, if the separated sources are recombined with small alterations in level or position, the degradations are often masked. This workshop discusses using source separation to enable repurposing the original audio content: speech intelligibility can be improved for broadcast listeners or cochlear implant users; sound objects can be extracted from a recording to enable object-based transmission and rendering. It is important to be able to assess the sound quality of the remix and the extent to which audio remixing is possible. We will highlight possible evaluation methods involving listeners and state-of-the-art algorithms.

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

Workshop 4 **Wednesday, May 23**
11:15 – 12:45 **Arena 3 & 4**

MASTERING FOR DIGITAL AND VINYL DISTRIBUTION

Chair: **Magdalena Piotrowska**, Gdansk University of Technology, Gdansk, Poland

Panelists: *Margaret Luthar*, Chicago Mastering Service, Chicago, IL, USA
Mandy Parnell, Black Saloon Studios, London, UK
Jonathan Wyner, M Works Studios/iZotope/ Berklee College of Music, Boston, MA, USA; M Works Mastering

In recent years it has become more common that files destined for digital distribution (iTunes, Tidal, Spotify, etc.) find their way to vinyl release. During this session a brief description of digital distribution release formats and channels, as well as vinyl as a media will be provided. Mastering engineers will discuss differences and specifics of mastering for each of them, as well as commonalities. Discussion will be supported with playback of audio examples.

Workshop 5 **Wednesday, May 23**
11:45 – 12:45 **Lobby**

ASSESSMENT OF SPATIAL AUDIO CONTENTS USING THE MS-IPM METHODOLOGY

Co-chairs: **Matthieu Parmentier**, francetélévisions, Paris, France
Chris Pike, BBC R&D, Salford, Greater Manchester, UK; University of York, Heslington, York, UK
Nick Zacharov, Force Technology, SenseLab, Hørsholm, Denmark

Panelists: *Catherine Colomès*, b<>com Tecnologica, Roma (RM), Italy
Michael Weitnauer, IRT, Munich, Germany

This workshop will explain the MS-IPM methodology that allows the evaluation of several perceptive attributes without anchor, a suitable method to assess spatial audio contents. The discussion will highlight a concrete use-case of this methodology for the assessment of object-based audio contents processed by different production renderers.

Special Event
SE1 AWARDS PRESENTATION AND KEYNOTE ADDRESS
Wednesday, May 23, 13:00 – 14:00 **Arena 3 & 4**

Opening Remarks:
- Executive Director Bob Moses
- President David Scheirman

Convention Chairs Alberto Pinto & Nadja Wallaszkovitzs

Program:
- AES Awards Presentation by Alex Case, Past President
- Introduction of Keynote Speaker
- Keynote Address by Marina Bosi

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. The awardees are:

BOARD OF GOVERNORS AWARD
- Maurycy Kin
- Sascha Spors

• Nadja Wallaszkovits
• Toon van Waterschoot

Keynote Speaker

In her presentation, Dr. Marina Bosi will examine some of the major shifts in audio consumption and how they represent new challenges and opportunities in coding audio for entertainment, information and other purposes, addressing the questions: "Who would have guessed twenty years ago that everybody would be clamoring for devices with MP3/AAC perceptual audio coders that fit into their pockets? As perceptual audio coders become more and more part of our daily lives, residing within mobile devices, DVDs, broad/webcasting, electronic distribution of music, etc., a natural question to ask is: what made this possible, and where is this going?"

Dr. Bosi draws from a rich background of being actively involved in the development of standards for audio coding, video coding, and digital content management, having contributed to the work of ANSI, ATSC, the DVD Forum, DVB, ISO/IEC MPEG, SDMI and SMPTE, as well as the AES. She holds several patents and has authored numerous publications in the field, as well as the acclaimed textbook *Introduction to Digital Audio Coding and Standards* (Kluwer/Springer December 2002). Additionally, Bosi has authored or co-authored over a dozen published AES research papers on perceptual audio coding and related topics, available in the AES E-Library. Dr. Bosi is also a Past-President of the Audio Engineering Society and a senior member of IEEE, and was part of the research team at Dolby Laboratories that developed the AC-2A and AC-3 (Dolby Digital) technologies, where, in addition, she also led the MPEG-2 Advanced Audio Coding (AAC) development. She was Vice President – Technology, Standards, and Strategies at DTS, and later Chief Technology Officer of MPEG LA, an industry leader in the use of patent pools to create a "one-stop shop" for licensing the essential patents required of multimedia technologies, and was a founding Director (together with Leonardo Chiariglione, MPEG Chair), and also the Treasurer, of the Digital Media Project, a non-profit organization created to promote successful development, deployment, and use of Digital Media.

Dr. Bosi has been recognized for her work with such honors as the AES Fellowship Award for her contributions to audio and video standards development and two AES Board of Governors Awards for chairing the 97th AES Convention and again for chairing the first international conference on high quality audio coding.

Session P3 — **Wednesday, May 23**
14:00 – 15:30 — **Scala 2**

AUDIO QUALITY—PART 2

Chair: **Todd Welti**, Harman International, Northridge, CA, USA

14:00

P3-1 Stereo Image Localization Maps for Loudspeaker Reproduction in Rooms—*Gavriil Kamaris, John Mourjopoulos*, University of Patras, Patras, Greece

A novel approach is proposed for maps illustrating the accuracy of image source position reproduced by stereo systems in various scenarios of listening rooms / loudspeakers. Based on previous work by the authors, the maps unify results derived from an image localization classifier and the sweet spot area metric estimating the direction of arrival angles (DOA) of all potential image source positions reproduced by the system via a perceptual model. The statistical analysis of these maps indicates the skewness and kurtosis of the DOA classification error and hence the consistency and robustness of the image definition along the plane of listener positions. Results utilize such parameter mappings to objectively illustrate the robustness of this qualitative aspect of audio reproduction in rooms.
Convention Paper 9922

14:30

P3-2 Method for Quantitative Evaluation of Auditory Perception of Nonlinear Distortion—*Mikhail Pahomov, Victor Rozhnov*, SPb Audio R&D Lab, St. Petersburg, Russia

All loudspeakers, amplifiers, and transmission paths introduce various types of distortion into audio signals. While energy contribution of nonlinear distortion to the distortion signal is comparatively small it has a significant impact on the sound signal quality in terms of the auditory perception. Similar energy characteristics of a nonlinear distortion signal can affect the subjective quality evaluation to different extents. This makes it important to accurately extract the nonlinear distortion signal from a musical signal in the simultaneous presence of significant distortions of other types for the auditory perception relevant objective evaluation of nonlinear distortion. The paper offers a method for quantitative evaluation of a nonlinear distortion signal in terms of its audibility and a method for prediction of subjective ratings of signals with nonlinear distortion.
Convention Paper 9923

15:00

P3-3 On the Effect of Inter-Channel Level Difference Distortions on the Perceived Subjective Quality of Stereo Signals—*Armin Taghipour*,[1,2] *Nadja Schinkel-Bielefeld*,[1,3] *Ankit Kumar*,[4] *Pablo Delgado*,[4] *Jürgen Herre*[1,4]
[1] Fraunhofer Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2] Currently with Empa, Laboratory fo Acoustics/Noise Control, Duebendorf, Switzerland
[3] Currently with Sivantos GmbH, Erlangen, Germany
[4] International Audio Laboratories Erlanged, a joint institution of Universität Erlangen-Nürnberg and Fraunhofer IIS, Erlangen, Germany

Perceptual audio coding at low bitrates and stereo enhancement algorithms can affect perceived quality of stereo audio signals. Besides changes in timbre, also the spatial sound image can be altered, resulting in quality degradations compared to an original reference. While effects of timbre degradation on quality are well-understood, effects of spatial distortions are not sufficiently known. This paper presents a study designed to quantify the effect of Inter-Channel Level Difference (ICLD) errors on perceived audio quality. Results show systematic effects of ICLD errors on quality: bigger ICLD errors led to greater quality degradation. Spectral portions containing relatively higher energy were affected more strongly.
Convention Paper 9924

Session P4 — **Wednesday, May 23**
14:00 – 16:30 — **Scala 4**

LOUDSPEAKERS—PART 2

Chair: **Jamie Angus**, University of Salford, Salford, UK

14:00

P4-1 Live Sound Subwoofer System Performance Quantification—*Adam J. Hill*, University of Derby, Derby, Derbyshire, UK

The general aim of live sound reinforcement is to deliver an appropriate and consistent listening experience across an audience. Achieving this in the subwoofer range (typically between 20–100 Hz) has been the focus of previous work, where techniques have been developed to allow for consistent sound energy distribution over a wide area. While this provides system designers with a powerful set of tools, it brings with it many potential metrics to quantify performance. This research identifies key indicators of subwoofer system performance and proposes a single weighted metric to quantify overall performance. Both centrally-distributed and left/right configurations are analyzed using the new metric to highlight functionality.
Convention Paper 9925

14:30

P4-2 Don't Throw the Loudspeaker Out with the Bathwater! Two Case Studies Regarding End-of-Line Tests in the Automotive Loudspeaker Industry —*Enrico Esposito*,[1] Angelo Farina,[2] Pietro Massini[1]
[1] Ask Industries S.p.A., Monte San Vito, Italy
[2] Università di Parma, Parma, Italy

Mass production of loudspeakers drivers for the automotive market is subjected to the strong requirements dictated by the implementation of the sector Quality System and is heavily conditioned by the low profit margin of what is seen (and actually is) a commodity as many other components of a vehicle; but, differently from other components, a loudspeaker is a complex system made of parts whose performance depends on many factors, including ambient conditions. For these reasons it is quite difficult to impose tight tolerances on loudspeakers and a fair agreement must be found between suppliers and customers to avoid scraping samples that are fine under any aspect, especially considering that the final judgment mainly stays, although not exclusively, in the ears of the end user. In this work two case studies are presented to show how tolerances could be fixed reasonably.
Convention Paper 9926

15:00

P4-3 Fast and Sensitive End-of-Line Testing—*Stefan Irrgang, Wolfgang Klippel*, Klippel GmbH, Dresden, Germany

Measurement time is a crucial factor for the total cost and feasibility of end-of-line quality control. This paper discusses new strategies minimizing the test time for transducers and audio systems while ensuring high sensitivity of defect detection, extracting comprehensive diagnostics information and using available resources in the best possible way. Modern production lines are fully automated and benefit highly from high speed testing. Optimal test stimuli and sophisticated processing in combination with multichannel test design are the key factors for smart testing. Appropriate acoustical, mechanical, and electrical sensors are discussed and suggested. Furthermore, parallel or alternating test schemes reduce the overall test time. Finally, typical concerns and pitfalls when testing at high speed are addressed and illustrated by measurement results.
Convention Paper 9927

15:30

P4-4 Optimal Material Parameter Estimation by Fitting Finite Element Simulations to Loudspeaker Measurements—*William Cardenas, Wolfgang Klippel*, Klippel GmbH, Dresden, Germany

Important characteristics for the sound quality of loudspeakers like frequency response and directivity are determined by the size, geometry, and material parameters of the components interfacing the acoustic field. The higher-order modes after cone break-up play an important role in wideband transducers and require a careful design of the cone, surround, and other soft parts to achieve the desired performance. Finite Element Analysis is a powerful simulation tool but requires accurate material parameters (complex Young's modulus as a function of frequency) to provide meaningful results. This paper addresses this problem and provides optimal material parameters by fitting the FEA model to an existing loudspeaker prototype measured by Laser vibrometry. This method validates the accuracy of the FEA simulation and gives further information to improve the modeling.
Convention Paper 9928

16:00

P4-5 Vision Methods Applied to Measure the Displacement of Loudspeaker Membranes—*Thomas Durand-Texte*,[1] *Manuel Melon*,[1] *Elisabeth Simonetto*,[2] *Stéphane Durand*,[2] *Marie-Hélène Moulet*[3]
[1] Le Mans Université, Le Mans, France
[2] Laboratoire Géomatique et Foncier, Le Mans, France
[3] Centre de Transfert de Technologie du Mans, Le Mans, France

Increased interest has been witnessed over the past decades for techniques measuring the vibration of a loudspeaker membrane. In this proceeding, vision methods have been adapted to measure the displacement of the cones of low-frequency drivers. The movement of the membrane is recorded with industrial or consumer market high-speed cameras with frame rates greater than or equal to 240 fps. The measured displacement shows acceptable to very good agreement with the one measured by a laser vibrometer, depending on the camera model. The displacement of the membrane, coupled to the measured electrical impedance, can be used to retrieve the small signal parameters or some non-linear parameters of the speaker. Finally, the vision methods are used to retrieve the vibration patterns of the membrane.
Convention Paper 9929

Technical Comittee Meeting
Wednesday, May 23, 14:00 – 15:00 Room Brera 1

Technical Committee Meeting on Audio for Games

Session P5 Wednesday, May 23
14:15 – 15:45 Poster Area

POSTERS: APPLICATIONS

14:15

P5-1 Grid-Based Stage Paradigm with Equalization Extension for "Flat-Mix" Production—*Jonathan Wakefield, Christopher Dewey, William Gale*, University of Huddersfield, Huddersfield, UK

In the Stage Paradigm (SP) the visual position of each channel widget represents the channel's level and pan position. The SP has received favorable evaluation but does

not scale well to high track counts because channels with similar pan positions and level visually overlap/occlude. This paper considers a modified SP for creating a "flat-mix" that provides coarse control of channel level and pan position using a grid-based, rather than continuous, stage and extends the concept to EQ visualization. Its motivation was to convert the "overlap" deficiency of the SP into an advantage. All subjects were faster at creating audibly comparable flat-mixes with the novel SP. Subject selected satisfaction keywords were also very positive.
Convention Paper 9930

14:15

P5-2 Real Time Implementation of an Active Noise Control for Snoring Reduction—*Stefania Cecchi,*[1] *Alessandro Terenzi,*[1] *Paolo Peretti,*[2] *Ferruccio Bettarelli*[2]
[1] Universitá Politecnica della Marche, Ancona, Italy
[2] Leaff Engineering, Osimo, Italy

Snoring is a well-known problem in our society. Active noise control systems can be applied to partially solve this annoyance. In this context the presented work aims at proposing a real-time implementation of an active noise control algorithm for snoring reduction by means of a DSP embedded platform and an innovative headboard equipped with microphones and loudspeakers. Several experimental results with different snoring signals have shown the potentiality of the proposed approach in terms of computational complexity and noise reduction.
Convention Paper 9931

14:15

P5-3 Identification of Nonlinear Audio Devices Exploiting Multiple- Variance Method and Perfect Sequences—*Simone Orcioni,*[1] *Alberto Carini,*[2] *Stefania Cecchi,*[1] *Alessandro Terenzi,*[1] *Francesco Piazza*[1]
[1] Universitá Politecnica della Marche, Ancona, Italy
[2] University of Urbino Carlo Bo, Urbino, Italy

Multiple-variance identification methods are based on the use of input signals with different powers for nonlinear system identification. They overcome the problem of the locality of the solution of traditional identification methods that well approximates the system only for inputs with approximately the same power of the identification signal. In this context, it is possible to further improve the nonlinear filter estimation exploiting as input signals the perfect periodic sequences that guarantee the orthogonality of the Wiener basis functions used for identification. Experimental results involving real measurements show that the proposed approach can accurately model nonlinear devices on a wide range of input variances. This property is particularly useful when modeling systems with high dynamic inputs, like audio amplifiers.
Convention Paper 9932

14:15

P5-4 Power Saving Audio Playback Algorithm Based on Auditory Characteristics—*Mitsunori Mizumachi,*[1] *Wataru Kubota,*[1] *Mitsuhiro Nakagawara*[2]
[1] Kyushu Institute of Technology, Kitakyushu, Fukuoka, Japan
[2] Panasonic, Yokohama City, Kanagawa, Japan

Music appreciation can be achieved with a variety of manners such as smartphones, portable music players, car audio, and high-end audio systems. Power consumption is one of the important issues for portable electronic devices and electric vehicles. In this paper a power saving audio playback algorithm is proposed in restraint of perceptual distortion. An original music source is passed through filterbanks and is reconstructed after increasing or decreasing the narrow-band component in each channel. The channel-dependent manipulation is carefully done not to cause perceptual distortion. The feasibility of the proposed method is evaluated by both measuring consumption current while music playback and carrying out a listening test.
Convention Paper 9933

14:15

P5-5 Deep Neural Networks for Road Surface Roughness Classification from Acoustic Signals—*Livio Ambrosini,*[1,2] *Leonardo Gabrielli,*[1] *Fabio Vesperini,*[1] *Stefano Squartini,*[1] *Luca Cattani*[2]
[1] Universita Politecnica delle Marche, Ancona, Italy;
[2] ASK Industries S.p.A., Montecavolo di Quattro Castella (RE), Italy

Vehicle noise emissions are highly dependent on the road surface roughness and materials. A classification of the road surface conditions may be useful in several regards, from driving assistance to in-car audio equalization. With the present work we exploit deep neural networks for the classification of the road surface roughness using microphones placed inside and outside the vehicle. A database is built to test our classification algorithms and results are reported, showing that the roughness classification is feasible with the proposed approach.
Convention Paper 9934

14:15

P5-6 Elicitation and Quantitative Analysis of User Requirements for Audio Mixing Interface—*Christopher Dewey, Jonathan Wakefield,* University of Huddersfield, Huddersfield, UK

Existing Audio Mixing Interfaces (AMIs) have focused primarily on track level and pan and related visualizations. This paper places the user at the start of the AMI design process by reconsidering what are the most important aspects of an AMI's visual feedback from a user's perspective and also which parameters are most frequently used by users. An experiment was conducted with a novel AMI which in one mode provides the user with no visual feedback. This enabled the qualitative elicitation of the most desired visual feedback from test subjects. Additionally, logging user interactions enabled the quantitative analysis of time spent on different mix parameters. Results with music technology undergraduate students suggest that AMIs should concentrate on compression and EQ visualization.
Convention Paper 9935

14:15

P5-7 Real-Time Underwater Spatial Audio: A Feasibility Study—*Symeon Delikaris-Manias, Leo McCormack, Ilkka Huhtakallio, Ville Pulkki,* Aalto University, Espoo, Finland

In recent years, spatial audio utilizing compact microphone arrays has seen many advancements due to emerging virtual reality hardware and computational advances. These advances can be observed in three main areas of spatial audio, namely: spatial filtering, direction of arrival estimation, and sound reproduction over loudspeakers or headphones. The advantage of compact microphone arrays is their portability, which permits their use in everyday consumer applications. However, an area that has

received minimal attention is the field of underwater spatial audio, using compact hydrophone arrays. Although the principles are largely the same, microphone array technologies have rarely been applied to underwater acoustic arrays. In this feasibility study we present a purpose built compact hydrophone array, which can be transported by a single diver. This study demonstrates a real-time underwater acoustic camera for underwater sound-field visualization and a parametric binaural rendering engine for auralization.
Convention Paper 9936

14:15

P5-8 Dual-Band PWM for Filterless Class-D Audio Amplification—*Konstantinos Kaleris, Charalampos Papadakos, John Mourjopoulos*, University of Patras, Patras, Greece

The benefits of Dual-Band Pulse Width Modulation (DBPWM) are demonstrated for filter-less audio class-D amplifier implementations. DBPWM is evaluated in terms of energy efficiency (thermal loading) and reproduction fidelity for direct driving of loudspeaker units by DBPWM signals. Detailed physical models of low-frequency (woofer) and high-frequency (tweeter) loudspeakers are employed for simulation of the coupling between the DBPWM signal and the electro-mechanical and acoustic properties of loudspeaker systems in the broadband PWM spectral range. Derived frequency responses are used to estimate the reproduced sound signal of the loudspeaker. Equivalent impedance of the speaker's voice coil is used to estimate thermal loading by the DBPWM signal's out-of-band spectral energy, comparing standard filtered PWM implementations to the proposed method.
Convention Paper 9937

14:15

P5-9 Low Cost Algorithm for Online Segmentation of Electronic Dance Music—*Emanuel Aguilera, Jose J. Lopez, Pablo Gutierrez-Parera, Carlos Hernandez*, Universitat Politecnica de Valencia, Valencia, Spain

Visual content animation for Electronic Dance Music (EDM) events is an emerging and demanded but costly service for the industry. In this paper an algorithm for automatic EDM structure segmentation is presented, suitable for the automation of video and 3D animation synchronized to the audio content. The algorithm implements low cost time/frequency features smoothed with a novel algorithm, resulting in a low latency output. Segmentation stage is based on a multiple threshold algorithm specifically tuned to EDM. It has been implemented in C performing in real-time and has been successfully integrated in a real-time commercial 3D graphics engine with great results over EDM music sets.
Convention Paper 10027

Tutorial 5 **Wednesday, May 20**
14:15 – 16:15 **Scala 1**

TUNE YOUR STUDIO-ACOUSTIC ANALYSIS OF SMALL ROOMS FOR MUSIC

Presenter: **Lorenzo Rizzi**, Suono e Vita - Acoustic Engineering, Lecco, Italy

The tutorial will talk about acoustic quality issues in small rooms for music. Focusing on internal acoustic quality (time response, frequency response, acoustic parameters) it will be rich in practical examples, useful for sound engineers, musicians and acousticians to focus on the small-room issues.

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

Workshop 6 **Wednesday, May 23**
14:45 – 16:15 **Arena 3 & 4**

THE ART OF VOCAL PRODUCTION

Presenters: **Wes Maebe**, RAK Studios/Sonic Cuisine, London, UK
Barry Marshall, The New England Institute of Art, Boston, MA, USA
Mandy Parnell, Black Saloon Studios, London, UK
Marek Walaszek, Addicted to Music Studio, Warsaw, Poland

Producing great vocals of course requires the best gear chain of microphone, mic pre, etc., as well as a good acoustic space. But what about all of the factors that have nothing to do with gear: atmosphere, drama, mood, psychology, phrasing, coaching, and preproduction? And what about practical considerations like scheduling, comfort level, doubling (or not), harmony, "comping" vs. complete takes, and simply recognizing the peak vocal performance? Our panel will examine these "extra-technical" considerations in getting that great take.

Tutorial 6 **Wednesday, May 23**
15:00 – 16:30 **Lobby**

PSYCHOACOUSTICS OF 3D SOUND RECORDING (WITH 9.1 DEMOS)

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

3D surround audio formats aim to produce an immersive soundfield in reproduction utilizing elevated loudspeakers. In order to use the added height channels most optimally in sound recording and reproduction, it is necessary to understand the psychoacoustic mechanisms of vertical stereophonic perception. From this background, this tutorial/demo session aims to provide an overview of important psychoacoustic principles that recording engineers and spatial audio researchers need to consider when making a 3D recording using a microphone array. Various microphone array techniques and practical workflows for 3D sound capture will also be introduced and their pros and cons will be discussed. This session will play demos of various 9.1 sound recordings, including the recent Auro-3D and Dolby Atmos release for the Siglo de Oro choir.

Technical Comittee Meeting
Wednesday, May 23, 15:00 – 16:00 **Room Brera 1**

Technical Committee Meeting on Archiving, Restoration, and Digital Libraries

Session P6 **Wednesday, May 23**
16:00 – 17:30 **Poster Area**

POSTERS: TRANSDUCERS

16:00

P6-1 Acoustic Power Based Evaluation of Loudspeaker

Distortion Measurements—*Charalampos Ferekidis*, Ingenieurbuero Ferekidis, Lemgo, Germany

Traditionally, drive-unit related distortion figures are evaluated "on-axis." This method leads to straightforward results as long as the frequency of the analyzed distortion component stays below the driver's modal break-up region. When the distortion component is entering the drive-unit's break-up region an on-axis based distortion analysis approach can become difficult if not meaningless. Therefore an alternative method is proposed which makes use of acoustic harmonic distortion sound power values, instead of sound pressure levels acquired in a single observation point (on-axis). These acoustic power values are derived from high resolution directivity data for the fundamental-, 2nd- and 3rd-harmonic-distortion components. Results are shown for different kinds of drive-units and the implications are discussed.
Convention Paper 9914

16:00

P6-2 **Anti-Rattle System Loudspeaker Device—***Dario Cinanni, Carlo Sancisi*, ASK Insustries Spa, Monte San Vito, Italy

On the basis of loudspeaker cabinets and panels vibration problems, this paper deals with a new dynamic loudspeaker device capable to reduce mechanical vibrations transmitted to the panel where it is fixed. Virtual 3D prototype is designed and optimized by simulations. Simulations were carried out using analytical and finite element methods. A working prototype was realized, measured and then tested on a panel, in order to evaluate vibrations reduction.
Convention Paper 9915

16:00

P6-3 **FEM Thermal Model of a Compression Driver: Comparison with Experimental Results—***Marco Baratelli, Grazia Spatafora, Emiliano Capucci, Romolo Toppi*, Faital S.p.A., Milan, Italy

A complete time domain thermal model of a compression driver was developed using COMSOL Multiphysics in order to predict heating phenomena and minimize potential damage. Heat transfer in the model relies on conduction, natural convection, and radiation all together ensuring a rigorous approach. Considerations accounting for power compression are also included in order to provide detail in the temperature prediction through time. Results are satisfactory and represent the outcome of an accurate method to predict operation limits of such devices, together with the change of magnetic induction in the air gap due to thermal effects.
Convention Paper 9916

16:00

P6-4 **Dynamic Driver Linearization Using Current Feedback—***Juha Backman*, Genelec Oy, Iisalmi, Finland; Noiseless Acoustics, Helsinki, Finland

It is well known that the electromechanical parameters of a dynamic driver can be modified through current feedback to suit the requirements of the available enclosure size and desired bandwidth, with a reduced distortion as a commonly observed beneficial effect. However, especially when designing sealed enclosures it is possible to select the loudspeaker parameters from a continuum of combinations of effective moving mass, damping resistance, and compliance. This paper describes optimization of electrical source parameters to improve the linearity based on numerical solution of the nonlinear equation of motion of the driver using the measured driver parameters.
Convention Paper 9917

Standards Comitee Meeting
Wednesday, May, 23, 16:00 – 17:30 Room Castello 1

SC-02-12 Working Group on Audio Applications of Networks

Technical Comitee Meeting
Wednesday, May 23, 16:00 – 17:00 Room Brera 1

Technical Committee Meeting on Coding of Audio Signals

Tutorial 7 Wednesday, May 23
16:30 – 18:00 Scala 1

BENEFITING FROM NEW LOUDSPEAKER STANDARDS

Presenter: **Wolfgang Klippel**, Klippel GmbH, Dresden, Germany

This tutorial focuses on the development of new IEC standards, addressing conventional and modern measurement techniques applicable to all kinds of transducers, active and passive loudspeakers and other sound reproduction systems. The first proposed standard (IEC 60268-21) describes important acoustical measurements for evaluating the generated sound field and signal distortion. The second standard (IEC 60268-22) is dedicated to the measurement of electrical and mechanical state variables (e.g. displacement), the identification of lumped and distributed parameters (e.g. T/S) and long-term testing to assess power handling, thermal capabilities, product reliability and climate impact. The third standard(IEC 63034) addresses the particularities of micro-speakers used in mobile and other personal audio devices. The tutorial gives a deeper insight into the background, theory and practical know-how behind those standards..

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

Tutorial 8 Wednesday, May 23
16:30 – 18:00 Scala 3

MODERN SAMPLING: IT'S NOT ABOUT THE SAMPLING; IT'S ABOUT THE RECONSTRUCTION!

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

Sampling, and sample rate conversion, are critical processes in digital audio. The analogue signal must be sampled, so that it can be quantized into a digital word. If these processes go wrong, the original signal will be irretrievably damaged!

1. Does sampling affect the audio?
2. Can we reconstruct audio after sampling?
3. Does sampling affect the timing, or distort the music?
4. Can modern sampling techniques improve things?

This tutorial will look at the modern theories of sampling, and explain, in a non-mathematical way, how these modern techniques can improve the sampling and reconstruction of audio.

Using audio examples, it will show that sampled audio, when properly reconstructed, preserves all of the original signal. Because it's not the sampling but the reconstruction that matters!

This session is presented in association with the AES Technical Committee on Broadcast and Online Delivery

Tutorial 9 **Wednesday, May 23**
16:30 – 17:30 **Scala 2**

AES67 & ST2110 - AN OVERVIEW

Presenter: **Andreas Hildebrand**, ALC NetworX GmbH, Munich, Germany

In September 2017, the new ST2110 standard on "Professional Media over Managed IP Networks" was published by SMPTE. What is ST2110 covering, how does it relate to AES67, and what is the practical impact to the audio industry? This tutorial describes the basic principles and the commonalities and differences of ST2110 & AES67 differences, and eludes on the constraints defined in ST2110 with respect to AES67. It also includes a brief outlook on how transport of non-linear audio formats (AES3) will be defined in ST2110.

Student Event/Career Development
SC1 OPENING AND STUDENT DELEGATE ASSEMBLY MEETING – PART 1
Wednesday, May 23, 16:30– 18:00 **Arena 3 & 4**

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 Saturday, May 26.

Workshop 7 **Wednesday, May 23**
17:00 – 18:00 **Lobby**

LISTENING SESSION ON LARGE 3D AUDIO SYSTEM

Presenters: **Matthias Frank**, University of Music and Performing Arts Graz, Graz, Austria
Franz Zotter, IEM, University of Music and Performing Arts, Graz, Austria

Advancements of the past years improve our understanding of large 3D-audio sound reinforcement systems, how to make them successfully supply large audiences with plausible direct, diffuse, and distant sounds, also live. Our lab promotes this by offering free Ambisonic tools for a broad usage, own productions, and by organizing European Student 3D Audio Production Competitions. You are warmly invited to this listening session in which we listen to pieces from the 3D Audio Competition, our Al Di Meola's - Elysium and More live production, and selected other productions.

Special Event
SE2 ORGAN CONCERT: SOUND MAGIC AND COUNTERPOINTS IN GERMAN BAROQUE
Wednesday, May 23, 20:00 – 21:30
Chiesa Cristiana Protestante
Via Marco de Marchi, 9, 20121 Milan

Performer: **Alberto Pinto**

AES conventions have entertained delegates with an organ concert for many years,and this year is no exception. Please join organist Alberto Pinto for an evening of music. Program will include works by Pachelbel, Bohm, Buxtehude, Bruhns, and Bach.

The Organ in the Protestant Church of Milan is one of the most interesting instruments of Northern Italy. Designed under the consultancy of Prof. Luigi Ferdinando Tagliavini in 1969 and built by one of the major Italian organ builders, the Tamburini Pontifical Organbuilding firm of Crema (Italy), it represents the first example in Italy of modern organ endowed with a Rückpositiv. His timbers, all very characterised, can be fully appreciated into every details also thanks to the rather dry acoustics of the church.

Session P7 **Thursday, May 24**
09:00 – 12:30 **Scala 4**

SPATIAL AUDIO—PART 1

Chair: **Sascha Spors**, University of Rostock, Rostock, Germany

09:00

P7-1 Continuous Measurement of Spatial Room Impulse Responses on a Sphere at Discrete Elevations—*Nara Hahn, Sascha Spors*, University of Rostock, Rostock, Germany

In order to analyze a sound field with a high spatial resolution, a large number of measurements are required. A recently proposed continuous measurement technique is suited for this purpose, where the impulse response measurement is performed by using a moving microphone. In this paper it is applied for the measurement of spatial room impulse responses on a spherical surface. The microphone captures the sound field on the sphere at discrete elevations while the system is periodically excited by the so-called perfect sequence. The captured signal is considered as a spatio-temporal sampling of the sound field, and the impulse responses are obtained by a spatial interpolation in the spherical harmonics domain. The elevation angles and the speed of the microphone are chosen is such a way that the spatial sampling points constitute a Gaussian sampling grid.
Convention Paper 9938

09:30

P7-2 Real-Time Conversion of Sensor Array Signals into Spherical Harmonic Signals with Applications to Spatially Localized Sub-Band Sound-Field Analysis—*Leo McCormack*,[1] *Symeon Delikaris-Manias*,[1] *Angelo Farina*,[2] *Daniel Pinardi*,[2] *Ville Pulkki*[1]
[1] Aalto University, Espoo, Finland
[2] Università di Parma, Parma, Italy

This paper presents two real-time audio plug-ins for processing sensor array signals for sound-field visualization. The first plug-in utilizes spherical or cylindrical sensor array specifications to provide analytical spatial filters which encode the array signals into spherical harmonic signals. The second plug-in utilizes these intermediate signals to estimate the direction-of-arrival of sound sources, based on a spatially localized pressure-intensity (SLPI) approach. The challenge with the tradi-

tional pressure-intensity (PI) sound-field analysis is that it performs poorly when presented with multiple sound sources with similar spectral content. Test results indicate that the proposed SLPI approach is capable of identifying sound source directions with reduced error in various environments, when compared to the PI method.
Convention Paper 9939

10:00

P7-3 Parametric Multidirectional Decomposition of Microphone Recordings for Broadband High-Order Ambisonic Encoding—*Archontis Politis, Sakari Tervo, Tapio Lokki, Ville Pulkki*, Aalto University, Espoo, Finland

Higher-order Ambisonics (HOA) is a flexible recording and reproduction method, which makes it attractive for several applications in virtual and augmented reality. However, the recording of HOA signals with practical compact microphone arrays is limited to a certain frequency range, which depends on the applied microphone array. In this paper we present a parametric signal-dependent approach that improves the HOA signals at all frequencies. The presented method assumes that the sound field consists of multiple directional components and a diffuse part. The performance of the method is evaluated in simulations with a rigid microphone array in different direct-to-diffuse and signal-to-noise atio conditions. The results show that the proposed method has a better performance than the traditional signal-dependent encoding in all the simulated conditions.
Convention Paper 9940

10:30

P7-4 Adaptive Non-Coincidence Phase Correction for A to B-Format Conversion—*Alexis Favrot, Christof Faller*, Illusonic GmbH, Uster, Switzerland

B-format is usually obtained from A-format signals, i.e., from four directive microphone capsules pointing in different directions. Ideally, these capsules should be coincident, but due to design constraints, small distances always remain between them. As a result the phase mismatches between the microphone capsule signals lead to inaccuracies and interferences, impairing B-format directional responses, especially at high frequencies. A non-coincidence correction is proposed based on adaptive phase matching of the four microphone A-format signals before conversion to B-format, improving the directional responses at high frequencies, enabling higher focus, better spatial image and timbre in B-format signals.
Convention Paper 9941

11:00

P7-5 Advanced B-Format Analysis—*Mihailo Kolundzija,*[1] *Christof Faller*[1,2]
[1] Ecole Polytechnique Féderale de Lausanne (EPFL), Lausanne, Switzerland
[2] Illusonic GmbH, Uster, Switzerland

Spatial sound rendering methods that use B-format have moved from static to signal-dependent, making B-format signal analysis a crucial part of B-format decoders. In the established B-format signal analysis methods, the acquired sound field is commonly modeled in terms of a single plane wave and diffuse sound, or in terms of two plane waves. We present a B-format analysis method that models the sound field with two direct sounds and diffuse sound, and computes the three components' powers and direct sound directions as a function of time and frequency. We show the effectiveness of the proposed method with experiments using artificial and realistic signals.
Convention Paper 9942

11:30

P7-6 Ambisonic Decoding with Panning-Invariant Loudness on Small Layouts (AllRAD2)—*Franz Zotter, Matthias Frank*, University of Music and Performing Arts Graz, Graz, Austria

On ITU BS.2051 surround with height loudspeaker layouts, Ambisonic panning is practice-proof, when using AllRAD decoders involving imaginary loudspea ker insertion and downmix. And yet on the 4+5+0 layout, this still yields a loudness difference of nearly 3 dB when comparing sounds panned to the front with such panned to the back. AllRAD linearly superimposes a series of two panning functions, optimally sampled Ambisonics and VBAP. Both are perfectly energy-preserving and therefore do not cause the loudness differences themselves, but their linear superposition does. In this contribution we present and analyze a new AllRAD2 approach that achieves decoding of constant loudness by (i) superimposing the squares of both panning functions, and (ii) calculating the equivalent linear decoder of the square root thereof.
Convention Paper 9943

12:00

P7-7 BRIR Synthesis Using First-Order Microphone Arrays—*Markus Zaunschirm, Matthias Frank, Franz Zotter,* University of Music and Performing Arts, Graz, Austria

Both the quality and immersion of binaural auralization benefit from head movements and individual measurements. However, measurements of binaural room impulse responses (BRIRs) for various head rotations are both time consuming and costly. Hence for efficient BRIR synthesis, a separate measurement of the listener-dependent part (head-related impulse responses, HRIR) and the room-dependent part (RIR) is desirable. The room-dependent part can be measured with compact first-order microphone arrays, however the inherent spatial resolution is often not satisfying. Our contribution presents an approach to enhance the spatial resolution using the spatial decomposition method in order to synthesize high-resolution BRIRs that facilitate easy application of arbitrary HRIRs and incorporation of head movements. Finally, the synthesized BRIRs are compared to measured BRIRs.
Convention Paper 9944

Session P8 **Thursday, May 24**
09:00 – 11:00 **Scala 2**

AUDIO EDUCATION

Chair: **Jan Berg**, Luleå University of Technology, Piteå, Sweden

09:00

P8-1 Does Spectral Flatness Affect the Difficulty of the Peak Frequency Identification Task in Technical Ear Training?—*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 common language and units used 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. The authors had investigated the correlation between the students' subjective ratings on the task difficulty and the physical measures calculated from the sound materials used in the training. However, the objective measure of the difficulty is still not known. Authors created the training materials with different spectral envelope but having the same music content and tested them in the ear training sessions.
Convention Paper 9945

09:30

P8-2 A Case Study of Cultural Influences on Mixing Practices—*Amandine Pras,*[1] *Brecht De Man,*[2,3] *Joshua D. Reiss*[3]
[1] University of Lethbridge, Lethbridge, Alberta, Canada
[2] Birmingham City University, Birmingham, UK
[3] Queen Mary University of London, London, UK

While sound mixers of popular music may share common principles across cultures, different engineers produce different mixes, and different listeners judge a mix differently. We designed a mixed-methods approach to examine this highly multidimensional problem in both style and perceived quality. Five student sound engineers from the Paris Conservatoire mixed the multitrack source of two pop songs and fully documented their mixing process. The resulting mixes were then used as stimuli for a blind, multi-stimulus listening test in a high-quality listening room that 13 students and 1 faculty member commented on and rated in terms of preference. Our outcomes highlight cultural and generational mixing specificities and offer a better understanding of the artistic side of the practice.
Convention Paper 9946

10:00

P8-3 Film Sound, Immersion and Learning: Field Study on 3D Surround Sound to Improve Foreign Language Learning—*Francisco Cuadrado,*[1] *Isabel López-Cobo,*[1] *Tania Mateos,*[2] *Beatriz Valverde*[1]
[1] Universidad Loyola Andalucía, Sevilla, AE, Spain
[2] University of Seville, Seville, Spain

This study focuses on the possibilities of film sound to improve the learning process, according to the immersion level elicited by 3D sound listening compared to stereo sound, and to the relation between emotion and learning. three-hundred-thirty students of English as a foreign language from Primary and High School Education watched an audiovisual sequence with one of two conditions: stereo and 3D surround mix. Learning evaluation (listening comprehension tests) and emotional impact measurement (electrodermal response, self-perception emotion test, and voice recording of the participants' reactions) have been the used instruments. The results show that students who watched the sequence listening to the 3D surround sound mix obtained better listening comprehension results than those that listened to the stereo sound mix.
Convention Paper 9947

10:30

P8-4 "Touch the Sound": Tangible Interface for Children Music Learning and Sound Experimentation— *Francisco Cuadrado,*[1] *Isabel López-Cobo,*[1] *Ana Tajadura-Jiménez,*[1,2,3] *David Varona*[1]
[1] Universidad Loyola Andalucía, Sevilla, AE, Spain
[2] Universidad Carlos III de Madrid, Madrid, Spain
[3] University College London, London, UK

"Touch the sound" is a music learning and sound experimenting system for children, composed by a technological tool (a tablet based interface that uses a series of physical and tangible elements for music and sound interaction) and a learning APP based on the IMLS (intelligent Music Learning System) project. In this paper we present and discuss the main pedagogical motivations for this tool: to create a tool based on accessible technology and to develop a learning tool that enhances the contact with the musical language through direct experimentation and that takes into count children's understanding of symbols. The design process of the whole system is also described. As presented in the outcomes section, the application possibilities of "Touch the Sound" go beyond the learning of music itself and open new paths for learning different contents based on the immersion generated by sound and on the emotional impact that sound has on the listener.
Convention Paper 9948

Workshop 8 **Thursday, May 24**
09:00 – 10:30 **Lobby**

RECORDING TECHNIQUE FOR 3D AUDIO—CAN CONVENIENCE AND AESTHETIC COEXIST?

Chair: **Toru Kamekawa**, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

Panelists: *Hyunkook Lee,* University of Huddersfield, Huddersfield, UK
Helmut Wittek, SCHOEPS GmbH, Karlsruhe, Germany

Several formats with height channels have been proposed, such as NHK's 22.2 channel system and the Auro 3D system, which are included in ITU-R BS.2159 standardized in 2012. Recording techniques for these 3D audio formats often include omnidirectional or cardioid microphones settings at certain distances (so-called spaced distribution array). Since these methods require much work for the setting, many methods have been proposed that expand the conventional stereo one-point microphones settings such as ORTF-3D, OCT-3D, Triple M/S Array. Recently Ambisonics is also the focus of the industry attention. In this workshop we compare the differences in impression due to the differences in settings of recording technique of these 3D audio techniques through actual recording examples and discusses how we can accomplish both convenience and aesthetics.

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

Workshop 9 **Thursday, May 24**
09:00 – 10:30 **Scala 3**

RECORDING, MIXING, AND MASTERING FOR DIFFERENT IMMERSIVE AUDIO FORMATS

Chair: **Joshua D. Reiss**, Queen Mary University of London, London, UK

Panelists: *Stefan Bilbao*, University of Edinburgh, Edinburgh, UK
Davide Rocchesso, University of Palermo, Palermo, Italy
Stefania Serafin, Aalborg University,

Copenhagen, Denmark
Vesa Välimäki, Aalto University, Espoo, Finland

Sound synthesis covers not just analog and MIDI synthesizers but the full range of algorithmic approaches to sound generation. Similarly, procedural audio is characterized by the philosophy of "sound as process, not samples," Procedural audio has been enthusiastically embraced by the games industry, and sound synthesis in all its forms may revolutionize sound design for animation, games, VR, augmented reality, and across the creative industries. This Workshop gives an overview of recent developments in the field. It brings together leading researchers to explain the key concepts and discuss new approaches and technologies. It will be relevant to practitioners, enthusiasts, and researchers wishing to gain a deeper understanding of this rapidly changing field.

This session is presented in association with the AES Technical Committee on Audio for Games and AES Technical Committee on Audio for Cinema

Tutorial 10 **Thursday, May 24**
09:30 – 10:30 **Scala 1**

FROM SEEING TO HEARING: SOUND DESIGN AND SPATIALIZATION FOR VISUALLY IMPAIRED FILM AUDIENCES

Presenters: **Gavin Kearney**, University of York, York, UK
Mariana Lopez, University of York, York, UK

This tutorial presents the concepts, processes, and results linked to the Enhancing Audio Description project (funded by AHRC, UK), which seeks to provide accessible audio-visual experiences to visually impaired audiences using sound design techniques and spatialization. Film grammars have been developed throughout film history, but such languages have matured with sighted audiences in mind and assuming that seeing is more important than hearing. We will challenge such assumptions by demonstrating how sound effects, first person narration as well as breaking the rules of sound mixing, can allow us to create accessible versions of films that are true to the filmmaker's conception. We will also discuss how the guidelines developed have been applied to the higher education context to train filmmakers on the importance of sound..

This session is presented in association with the AES Technical Committee on Audio for Cinema and AES Technical Committee on Spatial Audio

Session P9 **Thursday, May 24**
10:00 – 11:30 **Poster Area**

POSTERS: PERCEPTION

10:00

P9-1 Sound Localization of Beamforming-Controlled Reflected Sound from Ceiling in Presence of Direct Sound— *Hiroo Sakamoto, Yoichi Haneda*, University of Electro-Communications, Chofu-shi, Tokyo, Japan

Sound localization from the ceiling is important for 3D surround-sound systems, such as Dolby Atmos, 22.2 channel and higher-order ambisonics. Such systems are difficult to set up in a typical home environment owing to loudspeakers that must be placed on the ceiling. This problem can be solved by using the reflected sound from the ceiling through loudspeaker-array beamforming, such as in commercial sound bars. In the beamforming method, the listener always hears the jammer sound (side lobe) before he/she hears the main sound from the ceiling. To date, the relationship between the time and level differences in the direct and reflected sounds for sound image localization at the reflected position on the ceiling is not clear. This paper investigates this relationship through listening experiments. We also confirmed the localization at the ceiling by using a 32-element spherical loudspeaker array.
Convention Paper 9949

10:00

P9-2 Evaluation of Player-Controlled Flute Timbre by Flute Players and Non-Flute Players—*Mayu Kasahara, Atsushi Marui, Toru Kamekawa*, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

In order to investigate how flute players and non-flute players differ in the perception of the instrument, two listening experiments were carried out. The flute sounds were recorded to have changes in five levels of harmonic overtones energy levels played by three flute players. Through a listening experiment of attribute rating on "brightness," the flute players were found to evaluate the stimuli "brighter" as the harmonic overtones energy decreased while the non-flute players evaluated inversely. Through the second listening experiment of pairwise global dissimilarity rating among the stimuli, two dimensions corresponding to the harmonic overtones energy levels and to the noise levels were found. The experience of the flute performance did not seem to affect the result. These results indicate that the experience of the flute performance seemed to affect the result only when evaluating the stimuli using the word "brightness."
Convention Paper 9950

10:00

P9-3 Comparing the Effect of Audio Coding Artifacts on Objective Quality Measures and on Subjective Ratings—
Matteo Torcoli,[1] *Sascha Dick*[2]
[1] Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2] International Audio Laboratories Erlanged, a joint institution of Universität Erlangen-Nürnberg and Fraunhofer IIS, Erlangen, Germany

A recent work presented the subjective ratings from an extensive perceptual quality evaluation of audio signals, where isolated coding artifact types of varying strength were introduced. We use these ratings as perceptual reference for studying the performance of 11 well-known tools for objective audio quality evaluation: PEAQ, PEMO-Q, ViSQOLAudio, HAAQI, PESQ, POLQA, fwSNRseg, dLLR, LKR, BSSEval, and PEASS. Some tools achieve high correlation with subjective data for specific artifact types (Pearson's r > 0: 90, Kendall's t > 0: 70), corroborating their value during the development of a specific algorithm. Still, the performance of each tool varies depending on the artifact type and no tool reliably assesses artifacts from parametric audio coding. Nowadays, perceptual evaluation remains irreplaceable, especially when comparing different coding schemes introducing different artifacts.
Convention Paper 9951

10:00

P9-4 On the Accuracy and Consistency of Sound Localization

at Various Azimuth and Elevation Angles—*Maksims Mironovs, Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

This study examined the sound localization of a broadband pink noise burst at various azimuth and elevation angles in a critical listening room. A total of 33 source positions were tested, ranging from 0° to 180° azimuth and -30° to 90° elevation angles with 30° intervals. Results indicated that sound source elevation was localized inaccurately with a large data spread; however, it was improved on the off-center planes. It was observed that elevation localization accuracy was worse on the back compared to the front. Back to front confusion was observed for the sources raised to 60° elevation angle. Proposed listening response method showed consistent localization result and is therefore considered to be useful for future studies in 3D sound localization.
Convention Paper 9952

10:00

P9-5 Examination of the Factors Influencing Binaural Rendering on Headphones with the Use of Directivity Patterns—*Bartlomiej Mróz*, Gdansk University of Technology, Gdansk, Poland

This paper presents a study on the influence of the directional sound sources with the use of the directivity patterns. This contribution also includes a comparison to the work done by Wendt et al. where several directivity pattern designs used to gradually control the auditory source distance in a room were showed. While the tests of Wendt et al. were done by auralizing source and room using a loudspeaker ring in an anechoic chamber, this study aims at investigating whether the effect performs similarly in binaural auralization over headphone playback. The study includes not only the auditory source distance but also tries to discover the influence of auralized room characteristics, source-to-receiver distance, and signal on auditory externalization.
Convention Paper 9953

Technical Comittee Meeting
Thursday, May 24, 10:00 – 11:00 Room Brera 1

Technical Committee Meeting on Audio for Telecommunications

Professional Sound Expo PSE01
Thursday, May 24, 10:00 – 10:45 PSE Stage

THE IMPORTANCE OF THERMAL SIMULATION IN LOUDSPEAKER DESIGN: CASE STUDY FOR A COMPRESSION DRIVER

Presenters: **Marco Baratelli**, Faital S.p.A., Milan, Italy
Grazia Spatafora, Faital S.p.A.,Milan, Italy

Faital S.p.A. makes use of FEM simulation in order to predict heating phenomena and minimize the potential permanent damage to the coil for a wide variety of products. All heat transfer mechanisms are important: conduction, convection, and radiation. Compression drivers are high efficiency electroacoustic transducers dedicated to high frequencies. They are designed to produce high SPL levels, usually coupled to a horn or waveguide. Due to their spectrum of operation these devices experience incredibly small membrane excursions, hence they can be considered approximately static. In other words, compression drivers do not take advantage of the cooling effects that usually a woofer coil experiences while moving, as air is strongly pushed and pulled in and out of the air gap. For this reason these devices can be sensitive to severe heating, which in extreme cases can lead to permanent damage of the coil by combustion of the insulating enamel.

The outcomes of the presented case study were compared with an equivalent measurement. Results show a satisfactory match between the model and the measurements, with a worst case error of 5%.

Tutorial 11 Thursday, May 24
10:30 – 12:00 Arena 3 & 4

TOTAL TIMBRE: TOOLS AND TECHNIQUES FOR TWEAKING AND TRANSFORMING TONE

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

Recordists shape timbre through the coordinated use of several different signal processors. While equalization is a great starting point, the greatest tonal flexibility comes from strategic use of additional timbre-modifying signal processors: compression, delay, reverb, distortion, and pitch shift. This tutorial defines the timbre-driving possibilities of the full set of studio effects, connecting key FX parameters to their relevant timbral properties, with audio examples that reveal the results. This multi-effect approach to timbre enables you to extract more from the effects you already use, and empowers you to get the exact tones you want.

Standards Comittee Meeting
Thursday, May 24, 10:30 – 12:00 Room Castello 1

SC-02-08 Working Group on Audio-File Transfer and Exchange

Tutorial 12 Thursday, May 24
10:45 – 12:15 Scala 1

ACOUSTIC ENHANCEMENT SYSTEMS

Presenter: **Ben Kok**, BEN KOK - acoustic consulting, Uden, The Netherlands

Acoustic enhancement systems can be considered the ultimate in electro acoustics, as these systems influence the perceived natural acoustics of a given space. Due its complex nature-and the interaction with subjective perception-these systems often are considered mystic or even a “black art.” This tutorial will focus on how subjective and objective room acoustic qualities and parameters can be influenced with use of acoustic enhancement and how this relates to the requirements and design criteria for such a system.

In the presentation it is assumed that the attendees already have a basic knowledge of auditoria acoustics and acoustic enhancement systems, or at least have read the feature on acou stic enhancement by Francis Rumsey in JAES, Vol 26, No.6, 2014 June.

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

Tutorial 13 Thursday, May 24
10:45 – 12:15 Scala 3

PERCEPTUAL AND PHYSICAL EVALUATION OF GUITAR LOUDSPEAKERS

Presenter: **Wolfgang Klippel**, Klippel GmbH, Dresden, Germany

Loudspeaker and headphones generate distortion in the reproduced sound that can be assessed by objective measurements based on a physical or perceptual model or just by a subjective evaluation performed in systematic listening tests. This tutorial gives an overview on the various techniques and discusses the way how measurement and listening can be combined by auralization techniques to give a more reliable and comprehensive picture on the quality of the sound reproduction. The separation of signal distortion in speech signals allows to assess signal distortion which are for most stimuli below the audibility threshold. Further analysis of the separated distortion signals gives clues to identify the physical root cause of loudspeaker defects which is very crucial for fixing design problems.

Workshop 10 **Thursday, May 24**
10:45 – 12:15 **Lobby**

NEW SURROUND AND IMMERSIVE RECORDINGS

Presenters: **Jim Anderson**, Anderson Audio NY, New York, NY, USA; Clive Davis Institute of Recorded Music, New York University, New York, NY, USA
Ulrike Schwarz, Anderson Audio NY, New York, NY, USAl

Jim Anderson and Ulrike Schwarz have spent the past year recording and mixing music in high resolution and in immersive formats from New York to Norway to Havana. This presentation will highlight that work.

This session is presented in association with the AES Technical Committee on Recording Technology and Practices and AES Technical Committee on Spatial Audio and AES Technical Committee on High Resolution Audio

Professional Sound Expo **PSE02**
Thursday, May 24, 11:00 — 11:45 **PSE Stage**

THE ADVANTAGES OF ELECTRONIC BEAM STEERING AND ITS RELATION WITH FIR FILTERS

Presenter: **Daniele Mochi**, K-array, Florence, Italy

The ability to digitally adjust the dispersion of a line array element not only ensures the same listening experience to all audience members but, at the same time, limits the noise pollution in the areas where the sound pressure must be kept at a minimum. This presentation discusses advantages of Electronic Beam Steering (EBS) using FIR filters by providing extremely high frequency resolution.

Technical Comittee Meeting
Thursday, May 24, 11:00 – 12:00 **Room Brera 1**

Technical Committee Meeting on Broadcast and On-Line Delivery

Session P10 **Thursday, May 24**
11:15 – 12:45 **Scala 2**

AUDIO CODING, ANALYSIS, AND SYNTHESIS

Chair: **Jürgen Herre**, International Audio Laboratories Erlanged, a joint institution of Universität Erlangen-Nürnberg and Fraunhofer IIS, Erlangen, Germany

11:15

P10-1 Bandwidth Extension Method Based on Generative Adversarial Nets for Audio Compression—*Qingbo Huang, Xihong Wu, Tianshu Qu*, Peking University, Beijing, China

The compression ratio of core-encoder can be improved significantly by reducing the bandwidth of the audio signal, resulting in the poor listening perception. This paper proposes a bandwidth extension method based on generative adversarial nets (GAN) for extending the bandwidth of an audio signal, to create a more natural sound. The method uses GAN as a generative model to fit the distribution of the MDCT coefficients of the audio signals in the high-frequency components. Through minimax two-player gaming, more natural high-frequency information can be estimated. On this basis, a codec system is built up. To evaluate the proposed bandwidth extension system the MUSHRA experiments were carried on and the results show that there is comparable performance with HE-AAC.
Convention Paper 9954

11:45

P10-2 Device-Specific Distortion Observed in Portable Devices Available for Recording Device Identification—*Akira Nishimura*, Tokyo University Information Sciences, Chiba-shi, Japan

This study addresses device-specific distortion observed in recorded audio, to identify a built-in system-on-a-chip (SoC) in a portable device. A swept sinusoidal signal is emitted from a loudspeaker and is recorded by the portable device used in this study. The three types of distortion observed by spectral analysis of the recorded signals are the folded components at frequencies symmetrical across 4 kHz and 8 kHz of the signal component, non-harmonic and non-subharmonic distortion components whose frequencies are 4 kHz below and multiples of 4 kHz above the signal frequency, and mixed non-subharmonics and folded components in the low-frequency region. They are also observed using the correlation matrix on temporal amplitude variations among frequencies derived from the recorded speech signals.
Convention Paper 9955

12:15

P10-3 Physically Derived Synthesis Model of an Edge Tone—*Rod Selfridge, Joshua D. Reiss, Eldad J. Avital*, Queen Mary University of London, London, UK

The edge tone is the sound generated when a planar jet of air from a nozzle comes into contact with a wedge and a number of physical conditions are met. Fluid dynamics equations were used to synthesize authentic edge tones without the need for complex computation. A real-time physically derived synthesis model was designed using the jet airspeed and nozzle exit-to-wedge geometry. We compare different theoretical equations used to predict the tone frequency. A decision tree derived from machine learning based on previously published experimental results was used to predict the correct mode of operation. Results showed an accurate implementation for mode selection and highlighted areas where operation follows or deviates from previously published data.
Convention Paper 9956

Session P11 **Thursday, May 24**
11:45 – 13:15 **Arena 2**

POSTERS: MEASUREMENT

11:45

P11-1 Presence Detection by Measuring the Change of Total Sound Absorption—*Michal Luczynski*, Wroclaw University of Science and Technology, Wroclaw, Poland

The author of this paper analyzed potential possibilities of human presence detection inside a room based on determining the change in total sound absorption. The change between sound absorption with and without human leads to determine person presence. Limitations of this method were examined: in case of human presence, the change of reverberation time must be greater than measurement uncertainty. Important parameters are a volume of a room, total sound absorption, a frequency characteristic of the measurement signal, signal to noise ratio, etc. These parameters are criteria for assessing the reliability of the detection. Different types of measurement signals and systems were considered and tested. The aim was a maximum simplification of measurement method while keeping the minimum measurement uncertainty.
Convention Paper 9957

11:45

P11-2 The Evolution of Chirp-Based Measurement Techniques—*Mark Martin, Jayant Datta, Xinhui Zhou*, Audio Precision, Beaverton, OR, USA

Logarithmic chirp signals, also known as exponentially swept sines, have been used to determine the behavior of audio systems for more than two decades. This is done using a nonlinear deconvolution process that evaluates the direct and harmonic responses of a system. Despite the long history of this technique, improvements continue to be discovered and questions remain. Relatively subtle features of these signals are important for making accurate measurements. This paper describes how these signals have been used in measuring audio systems, describes the current state of the art, and clarifies the theoretical foundations and limitations of the technique.
Convention Paper 9958

11:45

P11-3 A Novel Measurement Procedure for Wiener/Hammerstein Classification of Nonlinear Audio Systems—*Andrea Primavera,*[1] *Michele Gasparini,*[1] *Stefania Cecchi,*[1] *Wataru Hariya,*[2] *Shogo Murai,*[2] *Koji Oishi,*[2] *Francesco Piazza*[1]
[1] Universitá Politecnica della Marche, Ancona, Italy
[2] KORG Inc. ,Tokyo, Japan

Non linear systems identification is a widespread topic and many techniques have been developed over the years in order to identify or synthesize black box models. Among the others, Wiener and Hammerstein structures are two of the more common nonlinear models, and identification techniques based on them are widespread in the literature. The choice of one structure over the other needs some a-priori knowledge. In this paper a novel method to determine if a system has a Wiener or Hammerstein nature is introduced. The method is based on the comparison of frequency responses in linear and nonlinear working region. Some simulated and real test results are reported in order to confirm the validity of the proposed approach.
Convention Paper 9959

11:45

P11-4 Moving Microphone Measurements for Room Response in Cinema—*Paul Peace,*[1,4] *Shawn Nageli,*[2,4] *Charles Sprinkle*[3,4]
[1] Community Professional Loudspeakers, Chester, PA, USA
[2] NXC Systems, West Jordan, UT, USA
[3] Kali Audio, Simi Valley, CA, USA
[4] Harman International, Stamford, CT, USA

Comparison of static multi-microphone measurements are made to moving microphone measurements for the purposes of determining room response and for application of appropriate equalization. Several benefits of moving microphone measurements are shown in this paper including consistency of measurements and the ability to measure interaction of correlated signal from multiple sources without the obfuscation of combing artifacts. It is also shown that moving microphone measurement is consistent with multiple microphone spatial average of sufficient resolution. This paper focuses on the moving microphone technique in the cinema application and experiments conducted were done in a medium sized cinema auditorium.
Convention Paper 9960

11:45

P11-5 Balloon Explosion, Wood-Plank, Revolver Shot, or Traditional Loudspeaker Large-Band Excitation: Which Is Better for Microphone Measurement?—*Balazs Kakonyi, Riyas Abdul Jaleel, Antonin Novak*, Université du Mans, Le Mans, France

The work presented in this paper investigates different sound sources used for microphone measurement in an anechoic room. The microphone under test is measured using three different physical impulse-like sources: balloon explosions, sounds created using wood-planks, and revolver shots; and the results are compared with a measurement in which the sound is generated by a loudspeaker excited with a large band swept-sine signal. The frequency response functions of three different commercial microphones are measured and compared with the data provided by the manufacturer.
Convention Paper 9961

Student Event/Career Development
SC2 STUDENT RECORDING CRITIQUES
Thursday, May 24, 12:00 – 13:00
Galleria (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 of their studies can sign up to participate. Students should sign up at the student (SDA) booth immediately on arrival at the convention, and deliver stereo 44.1 Khz, 24 bit AIFF or WAVE files to the SDA booth at that time. 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). These events are generously supported by Genelec.

Professional Sound Expo **PSE15**
Thursday, May 24, 12:00 – 12:45 **PSE Stage**

PROGRESSIVE DIRECTIVITY ARRAY: TECHNOLOGY OVERVIEW AND PERFORMANCE ADVANTAGES FOR SOUND REINFORCEMENT SYSTEMS

Presenter: **Moreno Zampieri**, Bose Professional

The concept and advantages of Progressive Directivity Array (PDA) was introduced at the ISEAT 2015 convention. The primary advantage of the PDA compared to conventional approaches such as line-array or point-source speakers is that the intended tonality can be achieved regardless of the room conditions (shape and material) with consistency throughout the target audience area not just for a single listening point. The study clearly indicates that precise coverage control to match the audience area with continuity of the sound source without gaps between modules are two crucial elements to achieve the primary objective. In this session adaptation of the PDA concept to the real product as a modular approach is described with performance of the product followed by application of the products to real world examples.

Technical Comittee Meeting
Thursday, May 24, 12:00 – 13:00 **Room Brera 1**

Technical Committee Meeting on Loudspeakers and Headphones

Workshop 11 **Thursday, May 24**
12:15 – 13:45 **Scala 3**

DEEP LEARNING FOR AUDIO APPLICATIONS

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

Panelists: *Qiuqiang Kong*, University of Surrey, Guildford, Surrey, UK
Stylianos Ioannis Mimilakis, Fraunhofer Institute for Digital Media Technology (IDMT), Ilmenau, Germany
Christian Uhle, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany; International Audio Laboratories Erlangen, Erlangen, Germany*l*

Deep learning is currently the most active research field. It has led to impressive improvements over former state-of-the-art methods in, for example, image classification, speech recognition, and text translation. This workshop introduces the concepts of deep learning and shows how it is applied to various problems of the audio engineering community, e.g., acoustic event detection, signal enhancement, source separation, and classification.

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

Student Development/Career Development
SC3 CAREERS IN THE PROFESSIONAL AUDIO INDUSTRY—FIRST STEPS
Thursday, May 24 12:15 – 13:15 **Arena 3 & 4**

Presenter: **Richard Wear**, Interfacio Ltd., Thames Ditton, Surrey, UK

Starting out on your career in audio? Not sure where to start? Looking for some advice on possible career paths?

Richard has over a decade of experience in recruitment but also comes from the audio industry himself. Having originally trained as an engineer he has worked in key sales, marketing, and prod-uct development roles for leading audio manufacturers, dealing with distributors and customers worldwide, and gaining a unique insight into the challenges and requirements of professional audio brand businesses. This seminar will be a fairly open format discussion and will cover marketing yourself to potential employers, how to format your CV/cover letter, and how to break into this exciting industry if you're just starting out in your career.

Tutorial 14 **Thursday, May 24**
12:30 – 13:15 **Lobby**

MUSIC IS THE UNIVERSAL LANGUAGE

Presenter: **Fei Yu**, Dream Studios, China; US

Music supervisor and music producer Fei Yu works on Chinese video games and movies for the last 8 years. During the work collaborating with Western composers, she has her unique perspective on this new form of international collaboration.

This event will discuss the critically acclaimed score to NetEase Games' "Revelation," the popular Chinese fantasy MMORPG that will soon be released internationally. As one of the first Chinese games to be scored by a westerner for the emerging Chinese video game market, we will show, by using musical examples, a way to work on this Chinese project. We will also talk about the recording techniques for projects like the movie Born In China that uses Chinese instruments—how to balance the sound, how to communicate with composer, etc.

Audio Applications Forum **Thursday, May 24**
12:30 – 13:30 **Arena 1**

HOW LINE ARRAY TECHNOLOGY HAS INSPIRED A NEW APPROACH TO MICROPHONES

Presenter: **Daniele Mochi**, K-array, Florence, Italy

Line array speakers have recently become extremely popular in the professional audio industry. With a high directivity on the vertical plan, a line array's sound beam can be focused solely on the audience area, ensuring greater intelligibility in highly reverberant environments and enabling uniform coverage of distances much greater than those attainable with point sources. The theory behind the line array configuration is well known and acoustic prediction software is able to simulate their behavior with high accuracy. But what happens when we apply this theory to line array microphones? What are the advantages of these kinds of microphones and what are their applications? This presentation offers an in-depth, easy-to-understand overview of line array technology applied to microphones, with practical examples and demonstrations:

Workshop 12 **Thursday, May 24**
13:00 – 14:00 **Scala 2**

AES67 PRACTICAL—A GUIDE ON HOW TO SETUP AES67 NETWORKS

Presenter: **Andreas Hildebrand**, ALC NetworX GmbH, Munich, Germany

Since the publication of the AES67 Standard on "High performance Streaming Audio-over-IP Interoperability" in September

2013, a myriad of AES67-compatible devices have been brought to market. Meanwhile, quite a few applications with AES67-compatible devices have been projected and put into operation. This session will give insight and provide tips on system planning, device and network configuration, management, and monitoring.

Professional Sound Expo **PSE04**
Thursday, May 24, 13:00 – 13:45 **PSE Stage**

FPGA EFFECTS, MICROPHONE AND PREAMP MODELING

Antelope Audio is a professional audio equipment manufacturer with over 20 years of experience in digital and analogue audio technologies. The company is designing advanced yet affordable audio interfaces; the industry's finest master clocks and premium FPGA FX. Among its latest endeavors find a collection of vintage mic emulations, interfaces with discrete preamps and modelling microphones.

Technical Comittee Meeting
Thursday, May 24, 13:00 – 14:00 **Room Brera 1**

Technical Committee Meeting on Hearing and Hearing Loss Prevention

Workshop 13 **Thursday, May 24**
13:15– 14:30 **Arena 3 & 4**

OBJECT-BASED AUDIO BROADCASTING: PRACTICAL ASPECTS

Chair: **Matthieu Parmentier**, francetélévisions, Paris, France

Panelists: *Dominique Brulhart*, Merging Technologies, Puidoux, Switzerland
Rupert Brun, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
Christophe Chabanne, Dolby
Michael Kratschmer, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
Scott Norcross, Dolby

This workshop will offer a state of the art of object-based audio for live broadcasting and post-production. The discussion will underline the role of the Audio Definition Model as an open format for production exchange.

Session P12 **Thursday, May 24**
13:30 – 15:30 **Scala 4**

SPATIAL AUDIO—PART 2

Chair: **Stefania Cecchi**, Universitá Politecnica della Marche, Ancona, Italy

13:30

P12-1 Binaural Room Impulse Responses Interpolation for Multimedia Real-Time Applications—*Victor Garcia-Gomez, Jose J. Lopez*, Universidad Politecnica de Valencia, Valencia, Spain

In this paper a novel method for the interpolation of Binaural Room Impulse Responses (BRIR) is presented. The algorithm is based on decomposition in time and frequency of the BRIRs combined with an elaborated peak searching and matching algorithm for the early reflections followed by interpolation. The algorithm has been tested with real room data, carrying out a perceptual subjective test. It outperforms, both in quality and in computational cost, the state-of-the-art algorithms.
Convention Paper 9962

14:00

P12-2 Evaluation of Binaural Renderers: Localization —*Gregory Reardon,*[1] *Andrea Genovese,*[1] *Gabriel Zalles,*[1] *Patrick Flanagan,*[2] *Agnieszka Roginska*[1]
[1] New York University, New York, NY, USA
[2] THX Ltd., San Francisco, CA, USA

Binauralrendererscanbeusedtoreproducespatialaudioover headphones. A number of different renderers have recently become commercially available for use in creating immersive audio content. High-quality spatial audio can be used to significantly enhance experiences in a number of different media applications, such as virtual, mixed and augmented reality, computer games, and music and movie. A large multi-phase experiment evaluating six commercial binaural renderers was performed. This paper presents the methodology, evaluation criteria, and main findings of the horizontal-plane source localization experiment carried out with these renderers. Significant differences between renderers' regional localization accuracy were found. Consistent with previous research, subjects tended to localize better in the front and back of the head than at the sides. Differences between renderer performance at the side regions heavily contributed to their overall regional localization accuracy.
Convention Paper 9963
Paper presented by Andrea Genovese]

14:30

P12-3 Characteristics of Vertical Sound Image with Two Parametric Loudspeakers—*Shigeaki Aoki, Kazuhiro Shimizu, Kouki Itou*, Kanazawa Institute of Technology, Nonoichi, Japan

A parametric loudspeaker utilizes nonlinearity of a medium and is known as a super-directivity loudspeaker. So far, the applications have been limited monaural reproduction sound system. We had discussed characteristics of stereo reproduction with two parametric loudspeakers. In this paper the sound localization in the vertical direction using the upper and lower parametric loudspeakers was confirmed by listening tests. The level difference between the upper and lower parametric loudspeakers were varied as a parameter. The direction of sound localization in the vertical plane was able to be controlled. We obtained interesting characteristics of the left-right sound localization in the horizontal plane. The simple geometrical acoustic model was introduced and analyzed. The analysis led to explain the measured characteristics.
Convention Paper 9964

15:00

P12-4 Virtual Hemispherical Amplitude Panning (VHAP): A Method for 3D Panning without Elevated Loudspeakers—*Hyunkook Lee, Dale Johnson, Maksims Mironovs*, University of
Huddersfield, Huddersfield, West Yorkshire, UK

This paper proposes "virtual hemispherical amplitude

panning (VHAP)," which is an efficient 3D panning method exploiting the phantom image elevation effect. Research found that a phantom center image produced by two laterally placed loudspeakers would be localized above the listener. Based on this principle, VHAP attempts to position a phantom image over a virtual upper-hemisphere using just four ear-level loudspeakers placed at the listener's left side, right side, front center, and back center. A constant-power amplitude panning law is applied among the four loudspeakers. A listening test was conducted to evaluate the localization performance of VHAP. Results indicate that the proposed method can enable one to locate a phantom image at various spherical coordinates in the upper hemisphere with some limitations in accuracy and resolution.
Convention Paper 9965

Workshop 14 **Thursday, May 24**
13:30 – 14:30 **Lobby**

PRODUCTION OF DANCE MUSIC IN 3D

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

Today's synthesizers produce almost exclusively mono and stereo sounds. Anyone who wants to implement productions for 3D audio must therefore compose corresponding sounds manually. The creation of multichannel sounds requires solid knowledge of psychoacoustics. Impressive 3D pad sounds can be constructed using very similar sounding stereo signals. And a meaningful assignment of sounds in the 3D space is subject to a number of rules. During the presentation, the most important basics of psychoacoustics, which are responsible for an impressive sound of 3D sounds, will be illustrated. In addition, methods how to create corresponding 3D sounds for a dance track, so that they can be played as 3D sounds directly from a keyboard, will be described. Various sound and video examples will be shown.

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

Student Event/Career Development
SC4 SAUL WALKER STUDENT DESIGN EXHIBITION
Thursday, May 24, 13:30 – 15:30 **Arena 2**

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.

Standards Comittee Meeting
Thursday, May 24, 13:30 – 15:00 **Room Castello 1**

SC-04-03 Working Group on Loudspeaker Modeling and Measurement

Tutorial 15 **Thursday, May 24**
14:00 – 15:30 **Scala 3**

LISTENING TESTS—UNDERSTANDING THE BASIC CONCEPTS

Presenter: **Jan Berg**, Luleå University of Technology, Piteå, Sweden

Listening tests are important tools for audio professionals as they assist our understanding of audio quality. There are numerous examples of tests, either formally recommended and widely used or specially devised for a single occasion. In order to understand listening tests and related methods, and also to potentially design and fully benefit from their results, some basic knowledge is required. This tutorial aims to address audio professionals without prior knowledge of listening test design and evaluation. The fundamentals of what to ask for, how to do it, whom to engage as listeners, what sort of results that may be expected, and similar issues will be covered..

This session is presented in association with the AES Technical Committee on Perception and Subjective Evaluation of Audio Signals

Professional Sound Expo **PSE05**
Thursday, May 24, 14:00 – 14:45 **PSE Stage**

THE NETWORKED STUDIO: A DREAM OR REALITY?

Presenter: **Jan Lykke**, NTP

The presentation will discuss the use of networked audio technologies such as Dante and AES67 in music and audio post-production studios. It will look at pros and cons of networked technology, what benefits and challenges you may run into, and illustrated with real-life examples. The topic of latency will also be covered.

Technical Comittee Meeting
Thursday, May 24, 14:00 – 15:00 **Room Brera 1**

Technical Committee Meeting on Fiber Optics for Audio

Session EB2 **Thursday, May 24**
14:15 – 16:00 **Scala 2**

APPLICATIONS & AUDIO EDUCATION

Chair: **Nyssim Lefford**, Luleå, Luleå, Sweden

14:15

EB2-1 New Packet Routing for 5G to Replace TCP/IP—*John Grant*, Nine Tiles, Cambridge, UK

While most of the attention has been focused on new radio and getting more bits over the wireless interface, operators also need 5G to have new packet routing technology that will make better use of those bits and support new services including low-latency live media. The new technology is being developed in ETSI ISG NGP, which the author chairs, and is expected to be standardized by 2020. This paper outlines the likely main features of the new technology, which is partly developed from AES47 and AES51, and discusses how they will make it more appropriate than IP for audio networking.
Engineering Brief 414

14:30

EB2-2 Miniaturized Noise Generation System—A Simulation of a Simulation—*Jan Banas, Przemek Maziewski,* *Sebastian Rosenkiewicz*, Intel Technology Poland, Gdansk, Poland

In the speech recognition industry, there is an everlasting need for evaluation of products in environments imitating real use cases. A widespread solution is to build a setup compliant with ETSI EG 202 396-1 standard, which defines a unified artificial laboratory environment to simulate real use scenarios of products to be tested. For space and cost reduction, a method is being developed to miniaturize the standard setup and simulate its behavior in a soundproof enclosure. In order to achieve high fidelity a number of spectral and temporal qualities of sound are measured in a laboratory and replicated in a box. The performance is evaluated using metrics specific to speech recognition.
Engineering Brief 415

14:45

EB2-3 FXive: A Web Platform for Procedural Sound Synthesis—*Parham Bahadoran*,[1,2] *Adan Benito*,[1,2] *Thomas Vassallo*,[1] *Joshua D. Reiss*[1]
[1] Queen Mary University London, London, UK
[2] FXive.com, London, UK

FXive is a real-time sound effect synthesis framework in the browser. The system is comprised of a library of synthesis models, audio effects, post-processing tools, temporal, and spatial placement functionality for the user to create the scene from scratch. The real-time nature allows the user to manipulate multiple parameters to shape the sound at the point of creation. Semantic descriptors are mapped to low level parameters in order to provide an intuitive means of user manipulation. Post-processing features allow for the auditory, temporal, and spatial manipulation of these individual sound effects.
Engineering Brief 416

15:00

EB2-4 Auto-EQ: Can Algorithms Replace a Sound Engineer?—*Daniil Sinev*,[1,2] *Guillaume Rossi-Ferrari*[1]
[1] ARKAMYS, Paris, France
[2] Le Mans University, Le Mans, France

This brief's aim is to present a work in progress on an automatic equalization algorithm. The algorithm's particular design, based on a parametric equalizer rather than inverse filtering, presents certain advantages as well as certain challenges. Being conceived and developed together with sound engineers, it is meant to mimic human decisions in filter choices. This necessitates a careful analysis of a sound engineer's workflow and a search for algorithmic solutions that correspond to decisions based on listening, experience and personal preference.
Engineering Brief 417

15:15

EB2-5 Challenging Changes for Live NGA Immersive Audio Production—*Peter Poers*, Junger Audio GmbH, Berlin, Germany

The world of broadcast audio is on the verge of a major revolution. Numerous "3D Immersive" formats are developing and will find their way into the mainstream of broadcast production and distribution in the near future. Along with Immersive Audio another category comes into game—Object Based Audio (OBA). All together it describes the Next Generation Audio formats—NGA. What does this mean and what challenge we need to fight with here? OBA will give the end user the option to personalize their experience by selecting personalized audio mixes. In object based audio, an "object" is essentially an audio stream with accompanying descriptive metadata. One of the major challenges for the production side of the industry will be to start OBA production. This means completely rethinking how we perform the final mix because, with OBA, it will be performed at home by the viewer rather than by a mixer in a post-production facility. What does this all mean for the broadcaster? A complete re-build of existing facilities and a total re-think about the audio processing equipment required for outside broadcast vehicles? Well, if we get it right, there will be some changes to overall workflow and hardware. And working with metadata for live streams and in files will become a major challenge. There will be new technical tools and new standards that will help to reach this new level of requirements. Some ideas and facts will be presented.
Engineering Brief 418

15:30

EB2-6 Mutebook.me—Interactive Online Tools for Teaching Music Technology—*Thilo Schaller*,[1] *Jan Burle*[2]
[1] Buffalo State University College, Buffalo, NY, USA
[2] University of Calgary, Calgary, AB, Canada

Mutebook is a project that aims to help audio arts students comprehend scientific concepts related to music technology, such as basic mathematics, fundamentals of acoustics, and digital audio theory. By using online interactive course material with visual and aural feedback, students of various audio arts disciplines can intuitively explore and understand relevant scientific concepts. Mutebook was in part funded by the Innovative Instruction Technology Grant (IITG) of the State University of New York (SUNY), and its first phase—i.e., the creation of an initial collection of interactive lecture notes with integrated applets—will be completed in May 2018.
Engineering Brief 419

15:45

EB02-7 CAE Support to Woofer Installation in a Car—*Andrzej Pietrzyk*, Volvo Car Corporation, Torslanda, Sweden

CAE support to the development of the audio system from an automotive OEM perspective is discussed, on the example of the development of the details of installation of a door woofer. For this application a loudspeaker model driven with electric voltage has to be integrated in the vibro-acoustic simulation. An example of an implementation of such a model is discussed and the accuracy of simulations is presented for the case of a woofer in free hanging door. It is further discussed in car models with different trim levels, from plane metal body with trimmed closures to fully trimmed body. Finally, the results obtained in a car model, with different variants of installation details of the woofer are presented and discussed.
Engineering Brief 420

Workshop 15 **Thursday, May 24**
14:30 – 16:00 **Arena 3 & 4**

STYLING YOUR LIVE AND RECORDED CLASSICAL, JAZZ, AND ACOUSTIC ENSEMBLE SOUND

Chair: **Ian Corbett**, Kansas City Kansas Community

College, Kansas City, KS, USA; off-beat-open-hats recording & sound reinforcement

Panelists: *Margaret Luthar*, Chicago Mastering Service, Chicago, IL, USA
Magdalena Piotrowska, Gdansk University of Technology, Gdansk, Poland
Kyle Snyder, Ohio University, Athens, OH, USA

Venue, ensemble and performers, musical material, engineer's/producer's vision, equipment available, purpose, and logistics. All of these things influence the choices a live recording or live sound engineer makes when planning and recording or reinforcing a live concert or event. Join our panel of live event recording and sound engineers as they discuss how and why they chose the techniques they used for varied classical, jazz, and other acoustic music situations and play some of the results for you. Material and situations presented will range from entry level non-professional ensemble events to professional productions.

Workshop 16 **Thursday, May 24**
14:45 – 15:45 **Lobby**

ANTON - UNIVERSE IN 3D: AMBISONICS IN ELECTRONIC MUSIC PRODUCTION

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

We would like to present collaboration between a sound engineer involved in spatial audio perception and ambisonic processing (Pawel Malecki) and a composer, arranger, and musician (Szymon Aleksander Piotrowski). Szymon's music, electronic project "ANTON," relates to interstellar travel and man's desolate journey through the universe. Spacious and three-dimensional thinking during composing inspired Pawel and Szymon to introduce these music concepts in 3D sound using ambisonics. This session includes playback in multichannel system, description of tools used, implementation and mixing process, technical, creative and aesthetic choices during the design process. Part of project "ANTON" (in stereo) is available at: https://www.youtube.com/watch?v=Ae8pwQQFFwY&feature=youtu.be

Workshop 17 **Thursday, May 24**
15:00 – 16:30 **Scala 1**

POWER AMPLIFICATION FOR HIGH RESOLUTION AUDIO

Presenter **John Dawson**, Jade Electronics Ltd., Cambridge, UK

Driving modern passive loudspeakers to realistic levels without audible distortion can be surprisingly demanding—more so when the source is high resolution audio. And, although the peak power levels are much lower, achieving full dynamic range with headphones can also present tricky design problems. This tutorial reviews the various types of linear amplifier—particularly classes A, AB, B, D, G and H - and examines some of the trade offs involved in designing and implementing some recent commercial designs for both speakers and headphones. A number of the author's measurements of real products—not all of them good—will be shown.

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

Professional Sound Expo **PSE06**
Thursday, May 24, 15:00 – 15:45 **PSE Stage**

EVENTIDE H9000 MULTICHANNEL FX PLATFORM

Presenter: **Patrick Flores**

In this presentation, we will be talking about Eventide's rich history of 47 years of audio innovations and introduce the H9000—our next generation Harmonizer. Designed for high-end recording, mastering, and post-production studios, the H9000 leverages the latest advances in technology and features the processing power of 16 DSP engines powered by four quad-core ARM processors. Topics covered will include processing up to 64 channels of audio, D.A.W. integration with the emote app, creating FX chains, and network capability with optional Dante, MADI and more.

Standards Comittee Meeting
Thursday, May 24, 15:00 – 16:00 **Room Castello 1**

SC-02-01 Working Group on Digital Audio Measurement Techniques

Technical Comittee Meeting
Thursday, May 24, 15:00 – 16:00 **Room Brera 1**

Technical Committee Meeting on Semantic Audio Analysis

Workshop 18 **Thursday, May 24**
15:30 – 17:00 **Scala 3**

RECRUITING AND TRAINING PARTICIPANTS FOR LISTENING TESTS

Chair: **Jon Francombe**, BBC Research and Development, Salford, UK

Panelists: *Tore Stegenborg-Andersen*, DELTA SenseLab, Hørsholm, Denmark
Todd Welti, Harman International Inc., Northridge, CA, USA

Listening tests can produce accurate and reliable results when conducted and analyzed correctly. The participants in such tests are an important contributing factor to the quality of results that are obtained. However, recruiting, training, and maintaining a committed panel of expert listeners is challenging, and reporting in scientific publications often lacks detail.

In this workshop, industrial and academic experts will share best practices from their extensive experience of convening listener panels. How can listeners be recruited, assessed, and trained? How should details of the listeners be presented when reporting results? How can diversity in panel membership be ensured? What about ethical issues and data security?

The workshop will feature a number of short presentations and plenty of time for questions and suggestions from the audience.

Session P13 **Thursday, May 24**
16:00 – 17:30 **Arena 2**

POSTERS: MODELING

16:00

P13-1 **Nonlinear Real-Time Emulation of a Tube Amplifier with a Long Short Time Memory Neural-Network—** *Thomas Schmitz, Jean-Jacques Embrechts*, University of Liege, Liege, Belgium

Numerous audio systems for musicians are expensive and bulky. Therefore, it could be advantageous to model them and to replace them by computer emulation. Their nonlinear behavior requires the use of complex models. We propose to take advantage of the progress made in the field of machine learning to build a new model for such nonlinear audio devices (such as the tube amplifier). This paper specially focuses on the real-time constraints of the model. Modifying the structure of the Long Short Term Memory neural-network has led to a model 10 times faster while keeping a very good accuracy. Indeed, the root mean square error between the signal coming from the tube amplifier and the output of the neural network is around 2%.
Convention Paper 9966

16:00

P13-2 Audio Control Room Optimization Employing BEM (Boundary Element Method)—*Robert Hersberger,*[1,2] *Gabriel Hauser,*[1] *Dirk Noy,*[1,2] *John Storyk*[3]
[1] Walters Storyk Design Group, Basel, Switzerland
[2] Fachschule für Akustik
[3] Walters-Storyk Design Group, Highland, NY, USA

The Boundary Element Method (BEM) is a state-of-the art tool in many engineering and science disciplines. In acoustics, the usage of BEM is increasing, especially for low frequency analysis, since the computational effort for small to medium geometries and long wavelengths is comparatively small. While BEM is well known to give reliable results for correctly programmed room shapes, the paper at hand demonstrates that the BEM model can also respond accurately to inserted absorptive materials, and hence the method is useful for virtually prototyping the efficiency of proposed acoustical modifications ahead of actual construction.
Convention Paper 9967

16:00

P13-3 A Machine Learning Approach to Detecting Sound-Source Elevation in Adverse Environments—*Hugh O'Dwyer, Enda Bates, Francis M. Boland*, Trinity College Dublin, Dublin, Ireland

Recent studies have shown that Deep neural Networks (DNNs) are capable of detecting sound source azimuth direction in adverse environments to a high level of accuracy. This paper expands on these findings by presenting research that explores the use of DNNs in determining sound source elevation. A simple machine-hearing system is presented that is capable of predicting source elevation to a relatively high degree of accuracy in both anechoic and reverberant environments. Speech signals spatialized across the front hemifield of the head are used to train a feedforward neural network. The effectiveness of Gammatone Filter Energies (GFEs) and the Cross-Correlation Function (CCF) in estimating elevation is investigated as well as binaural cues such as Interaural Time Difference (ITD) and Interaural Level Difference (ILD). Using a combination of these cues, it was found that elevation to within 10 degrees could be predicted with an accuracy upward of 80%.
Convention Paper 9968

16:00

P13-4 Design of an Acoustic System with Variable Parameters—*Karolina Prawda*, AGH University of Science and Technology, Kraków, Poland

The number of multifunctional halls with need for acoustic adaptation aimed at many different demands of the spaces is constantly growing. The present paper shows the design of an acoustic system with variable characteristics and adjustable sound absorption coefficient that may be used in such spaces.
Convention Paper 9969

16:00

P13-5 High Frequency Modelling of a Car Audio System—*Aleksandra Pyzik, Andrzej Pietrzyk*, Volvo Car Corporation, Torslanda, Sweden

Geometrical Acoustics is a widely used room acoustic modelling method. Since GA neglects the wave phenomena and is strictly applicable for short wavelengths relative to model and surface sizes, the application for the automotive industry is still subject of research. The paper studies the feasibility of using GA for high frequency simulations of a car sound system. The GA models of a vehicle at three production stages were created based on FE models. An impedance gun was used for in-situ measurements of the properties of the car interior materials. The directivity of the midrange and tweeter speakers was measured in anechoic conditions. In subsequent simulations, various GA software settings were tested. Simulation results were verified with measurements in the car.
Convention Paper 9970

Professional Sound Expo **PSE07**
Thursday, May 24, 16:00 – 16:45 **PSE Stage**

THE ITALIAN WAY—ITALIAN RECORDING STUDIOS

Presenters: **Marco Borsatti, Sabino Cannone, and Luca Pilla**, Audiofader.com, Milano, IT

Discover how two of the very best Italian sound engineer work for international production, staying in Italy. Marco Borsatti (marcoborsatti.com) and Sabino Cannone (morevox.com) will explain how they adapt recording, mixing and mastering for international audience in respect of Italian sound, with audio example and plug-ins explanation. We will analyze also how technology has changed their workflow, their sound and their recording studios. Presenter: Luca Pilla, editor of Audiofader.com. Workshop will be handling only in Italian language.

Standards Comittee Meeting
Thursday, May 24, 16:00 – 17:30 **Room Castello 1**

SC-07-01 Working Group on Audio Metadata

Session P14 **Thursday, May 24**
16:15 – 17:45 **Scala 4**

PERCEPTION—PART 1

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

16:15

P14-1 The Standard Deviation of the Amplitude Spectrum as a Predictor of the Perception of the "Power" of Distorted Guitar Timbre—*Koji Tsumoto, Atsushi Marui, Toru Kamekawa*, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

The perception of the wildness and heaviness of distorted guitar timbre can be compiled as one attribute associated with the "power" according to our previous study. The amount of distortion is one of the predictors of the perception of the "power." Although the predictor of the "power" other than the amount of distortion is yet to be known, the existence of a predictor for the subtle differences of the "power" can be assumed from the engineering perspective. Specifically, the amount of even and odd harmonics are altered by the symmetrical or asymmetrical placements of the diodes in the distortion effect pedal. We investigated how these changes affect the perception of "power." The spectral centroids of the stimuli were equalized to eliminate the influence of the perception of "brightness" over "power." A pairwise comparison was conducted for the stimuli of three different amounts of distortion and three types of diode placements. Regression analysis indicated that the Standard Deviation of the Amplitude Spectrum seemed to be an appropriate predictor of the perception of "power."
Convention Paper 9971

16:45

P14-2 Categorization of Isolated Sounds on a Background—Neutral—Foreground Scale—*William Coleman,*[1] *Charlie Cullen,*[2] *Ming Yan*[3]
[1] Dublin Institute of Technology, Dublin, Ireland
[2] University of the West of Scotland, UK
[3] DTS Licensing Inc., UK

Recent technological advances have driven changes in how media is consumed in home, automotive, and mobile contexts. Multichannel audio home cinema systems are not ubiquitous but have become more prevalent. The consumption of broadcast and gaming content on smartphone and tablet technology via telecommunications networks is also more common. This has created new possibilities and consequently poses new challenges for audio content delivery such as how media can be optimized for multiple contexts while minimizing file size. For example, a stereo audio file may be adequate for consumption in a mobile context using headphones, but it is limited to stereo presentation in the context of a surround-sound home entertainment system. Another factor is the variability of telecommunications network bandwidths. Object-based audio may offer a solution to this problem by providing audio content on an object level with metadata that controls how the media is delivered depending on the consumption paradigm. In this context, insight into the relative importance of different sounds in the auditory scene will be useful in forming content delivery strategies. This paper presents the results of a listening test investigating categorization of isolated sounds on a Background (BG) — Neutral (N) — Foreground (FG) scale. A continuum of importance was observed among the sounds tested and this shows promise for application in object-based audio delivery.
Convention Paper 9972

17:15

P14-3 The Relevance of Auditory Adaptation Effects for the Listening Experience in Virtual Acoustic Environments—*Florian Klein, Stephan Werner*, Technische Universität Ilmenau, Ilmenau, Germany

Virtual acoustic environments can provide a plausible reproduction of real acoustic scenes. Since the perceived quality is based on expectations and previous sound exposure, a reliable measure of the listening experience is difficult. Listeners are able to learn how to interpret spatial cues and room reflections for certain tasks. To discuss the relevance of auditory adaptation effects, this paper summarizes a series of listening experiments that show adaptation processes with effect on localization accuracy, externalization, and the ability of a listener to identify the own position in a virtual acoustic environment.
Convention Paper 9973

Student Event/Career Development
SC5 RECORDING COMPETITION—PART 1
Thursday, May 24, 16:00 – 18:00 Lobby

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.

Workshop 19 Thursday, May 24
16:15 – 17:45 Arena 3 & 4

MICROPHONES—CAN YOU HEAR THE SPECS? A MASTER CLASS

Chair: **Helmut Wittek**, SCHOEPS GmbH, Karlsruhe, Germany

Panelists: *Eddy B. Brixen*, EBB-consult, Smørum, Denmark; DPA Microphones
Kelly Kay, Josephson Engineering, Santa Cruz, CA, USA
Hans Riekehof, SCHOEPS Mikrofone GmbH, Karlsruhe, Germany
Martin Schneider, Georg Neumann GmbH, Berlin, Germany

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.

Session P15 Thursday, May 24
16:30 – 18:00 Scala 2

MEASUREMENTS

Chair: **John Mourjopoulos**, University of Patras, Patras, Greece

16:30

P15-1 Comparison of Effectiveness of Acoustic Enhancement Systems—Comparison of In-Line, Regenerative, and Hybrid-Regenerative Enhancement Methods—*Takayuki Watanabe,*[1] *Dai Hashimoto,*[1] *Hideo Miyazaki,*[1] *Ron Bakker*[2]
[1] Yamaha Corp., Hamamatsu, Shizuoka, Japan
[2] Yamaha Commercial Audio Systems Europe, Rellingen, Germany

Acoustic enhancement systems have become popular in recent years and are broadly used in various kinds of facilities because of their acoustic naturalness and system stability. Today, demand for acoustic enhancement systems exists not only for multi-purpose halls but also for highly absorptive spaces, especially lecture halls and theaters. To investigate an effective enhancement system for highly absorptive spaces, we compared several enhancement methods that are commonly applied in a small auditorium. This paper summarizes the features and the acoustical characteristics of systems configured according to each of the considered enhancement methods in the small auditorium.
Convention Paper 9974

17:00

P15-2 Comparison of Methods for Estimating the Propagation Delay of Acoustic Signals in an Audience Service for Live Events—*Marcel Nophut, Robert Hupke, Stephan Preihs, Jürgen Peissig*, Leibniz Universität Hannover, Hannover, Germany

Our novel audience service for live events uses supplementary audio content presented through transparent headphones to enhance the traditional audio playback of a PA loudspeaker system. The service requires to estimate the propagation delay of sound waves from the PA loudspeakers to the listener in order to individually delay the supplementary audio content and temporally align it with the PA playback. This paper compares two different correlation-based methods regarding their computational complexity and their performance in estimating the above mentioned time delay using realistic recordings of music and speech samples. Additional measures, that make the estimation more robust, were developed and are also presented. Typical issues like tonal components, room reflections, crosstalk, and a large number of correlation lags are addressed.
Convention Paper 9975

17:30

P15-3 Experimental Results on Active Road Noise Cancellation in Car Interior—*Carlo Tripodi, Alessandro Costalunga, Lorenzo Ebri, Marco Vizzaccaro, Luca Cattani, Emanuele Ugolotti, Tiziano Nili*, Ask Industries S.p.A., Montecavolo di Q.Castella (RE), Italy

We discuss the implementation and the performance of an active road noise control system. We review the design of a system based on the Least Mean Square (LMS) adaptive algorithm, suitable for the wideband road noise reduction inside a car cabin. As the system is based on a feedforward control approach, we discuss the method for the selection of the sensors providing the best noise reference signals. We then give a computational complexity analysis of the overall system and discuss the system implementation into a prototype hardware for a mid-size sedan. Performance are then evaluated on different road noise scenarios in real driving situations.
Convention Paper 9976
[Paper presented by Caro Tripodi]

Tutorial 16 **Thursday, May 24**
16:45 – 17:45 **Scala 1**

AUDIO LOCALIZATION METHOD FOR VR APPLICATION

Presenter: **Joo Won Park**, Columbia University, New York, NY, USA

Audio localization is a crucial component in the Virtual Reality (VR) projects as it contributes to a more realistic VR experience to the users. In this event a method to implement localized audio that is synced with user's head movement is discussed. The goal is to process an audio signal real-time to represent three-dimensional soundscape. This tutorial introduces a mathematical concept, acoustic models, and audio processing that can be applied for general VR audio development. It also provides a detailed overview of an Oculus Rift- MAX/MSP demo that conceptualizes the techniques.

Tutorial 17 **Thursday, May 24**
17:00 – 18:00 **Scala 3**

USE OF DELAY-FREE IIR FILTERS IN MUSICAL SOUND SYNTHESIS AND AUDIO EFFECTS PROCESSING

Presenter: **Federico Fontana**, University of Udine, Udine, Italy

The delay-free loop problem appears when an audio electronic system, typically an analog processor, is transformed in the digital domain by means of a filter network preserving the structural connections among nonlinear components. If such connections include delay-free loopbacks then there is no explicit procedure allowing for the computation of the corresponding digital filter output.

The tutorial will show how a delay-free IIR filter network is designed, realized, and finally computed; and why they have led to successful real-time digital versions of the Dolby B, the Moog and EMS VCS3 voltage-controlled filter, as well as nonlinear oscillators and RLC networks, magnitude-complementary parametric equalizers, and finite-difference time-domain scheme-based models of membranes characterized by low wave dispersion.

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

Professional Sound Expo **PSE03**
Thursday, May 24, 17:00 – 17:45 **PSE Stage**

THE EFFECT OF ACOUSTIC CENTER IN SUBWOOFER

Presenter: **Paolo Martignon**, Contralto Audio srl, Parma (PR), Italy

This presentation is the outcome of an investigation begun in 2017 which sees the collaboration between Contralto Audio and Merlijn Van Veen. It was already presented briefly in AES143 in New York and is proposed again in this context where more room for details and discussion is allowed. As explained by J.Vanderkooy in his paper "The Low-Frequency Acoustic Centre: Measurement, Theory and Application," the acoustic center of a direct radiating subwoofer unit is placed ahead respect to the driver membrane, at a distance depending on driver and cabinet dimensions. This has effects on acoustic simulations and it deserves some attention to avoid errors. Measurements are shown that confirm acoustic center position theoretical calculation and a discussion is made about its effect on the definition of models for accurate simulations.

Special Event
SE03 THE RICHARD C. HEYSER MEMORIAL LECTURE
Thursday, May 24, 18:30 – 20:00 Arena 3 & 4

Lecturer: **Malcolm Hawksford**, Emeritus Professor, Essex University, UK

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 Malcolm Hawksford, Emeritus Professor and recent AES Gold Medal Winner. The title of his talk is “Understanding High Quality Audio—A Personal Journey.”

Audio engineering is bounded by theoretical principles that ultimately establish what can be achieved in practical terms. The mapping of theory into a practical realization is inevitably moderated by system-specific approximations and increasing entropy as information travels through an audio system. While the theory may be deceptively simple, the means of implementation can embrace diverse concepts and ingenious design techniques. My career as an academic and researcher mirrors this diversity, which has witnessed remarkable transformations as the industry has adapted from analogue to digital (and back again!). Working with highly talented students, my research projects have been principally motivated by an underlying ambition to extend the horizons of sound reproduction. Alongside personal reflections about factors critical to high quality sound reproduction, my lecture will concentrate on important circuit and system factors, both analogue and digital. A new means of down-sample rate conversion will be demonstrated too.

Malcolm O. J. Hawksford was educated at Aston University in Birmingham, UK, where he received his B.Sc. with first-class honours in 1968, Ph.D. in 1972, and D.Sc. in 2008. He is now Emeritus Professor within the School of Computing Science and Electronic Engineering at Essex University, Colchester, UK. Early research embraced delta- and sigma–delta modulation (SDM) applied to color TV coding that under the award of a BBC Research Scholarship lead to a method of luminance and chrominance multiplexing exploiting digital time-compression, a forerunner of MAC/DMAC. Principal interests include audio engineering, electronic circuit design, error correction in amplifiers, and signal processing focussing on loudspeakers, SDM, PWM linearization, spatial audio and telepresence. Malcolm is recipient of the AES Publications Award for the best contribution by an author of any age for *JAES*, volumes 45 and 46 and holds the AES Silver Medal awarded 2006 for major contributions to engineering research in the advancement of audio reproduction. Also, in 2014 he was awarded the Peter Barnett Memorial Award by the Institute of Acoustics, UK, for his contributions in the field of electroacoustics extending over four decades. He is a chartered engineer and fellow of the AES, IET, and IOA. Malcolm has been formative chair of the AES Technical Committee on High-Resolution Audio and was a founding member of Acoustic Renaissance for Audio (ARA).

Student Event/Career Development
SC6 AES STUDENT PARTY
Thursday, May 24, 20:00 – 24:00 tbd

The AES Student Party is open to any 144th 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 full details. It will be hosted at a venue to be announced at the Opening and Student Delegate Assembly Meeting—Part 1.

Session P16 Friday, May 25
08:45 – 10:45 Scala 4

SPATIAL AUDIO—PART 3

Chair: **Ville Pulkki**, University of Aalto, Espoo, Finland

08:45

P16-1 Surround with Depth on First-Order Beam-Controlling Loudspeakers—*Thomas Deppisch*,[1,2] *Nils Meyer-Kahlen*,[1,2] *Franz Zotter*,[2] *Matthias Frank*[2]
[1] University of Technology, Graz, Austria
[2] University of Music and Performaing Arts Graz, Graz, Austria

Surround systems are typically based on fixed-directivity loudspeakers pointing towards the listener. Laitinen et al. showed for a variable-directivity loudspeaker that directivity control can be used to influence the distance impression of the reproduced sound. As we have shown in a listening experiment, using beam-controlling loudspeakers, stable auditory events at directions additional to the loudspeaker positions can be created by exciting specific wall reflections. We use these two effects to enable distance control and increase the number of effective surround directions in two different surround setups. We present IIR filter design derived from a physical model, which achieves low frequency beam-control for our novel cube-shaped 4-channel loudspeakers.
Convention Paper 9977

09:15

P16-2 A Method to Reproduce a Directional Sound Source by Using a Circular Array of Focused Sources in Front of a Linear Loudspeaker Array—*Kimitaka Tsutsumi*,[1,2] *Yoichi Haneda*,[2] Ken'ichi Noguchi,[1] *Hideaki Takada*[1]
[1] NTT Service Evolution Laboratoties, Yokosuka-shi, Kanagawa, Japan
[2] University of Electro-Communications, Chofu-shi, Tokyo, Japan

We propose a method to create a directional sound source in front of a linear loudspeaker array. The method creates a virtual circular loudspeaker array comprising multiple focused sources to reproduce directivity patterns. In the proposed method, the driving functions for the secondary sources are defined as a cascade combination of two driving functions: The first one for directivity control derived by an analytical conversion of circular harmonic modes, and the second one for creating focused sources. The proposed driving functions can deal with directivity rotation by changing the position of focused sources, thereby avoiding recalculations of driving functions. Using computer simulation, we obtained accuracy and algorithmic complexity comparable or better than those of a conventional method.
Convention Paper 9978

09:45

P16-3 Subjective Evaluation of Multichannel Upmix Method Based on Interchannel Coherent Component Extraction—*Yuta Hashimoto, Yasuto Goto, Akio Ando*, University of Toyama, Toyama, Japan

We developed a method that extracted the interchannel coherent component from multichannel audio signal. In this paper two subjective evaluation experiments for testing the upmix performance of our method are shown. In the first experiment, stereo signals were upmixed to 4-channel signals in which channels were set at ±30° and ±60°. The subjective evaluation with MUSHRA method showed that our method was superior to the conventional methods. In the second experiment, signals of 4 channels located at ±30° and ±110° were upmixed to 8-channel signals in which four of the channels were set at the upper layer. The subjective evaluation showed that there were no significant differences between the upmixed 8-channel sound and the original 4-channel sound in terms of spatial impression.
Convention Paper 9979

10:15

P16-4 Comparison between Different Microphone Arrays for 3D-Audio—*Lucca Riitano, Jorge Enrique Medina Victoria*, University of Applied Sciences Darmstadt, Darmstadt, Germany

The growing need of 3D recordings for film, virtual reality, and games started the development and research on different microphone arrays for 3D-Audio such as the ORTF-3D, MS-3D, and a wide range of experimental and particular setups. Comparison between the different microphone arrays has been rather the exception. For this paper three different arrays are placed together to record a piece of music. Based on a listening test, the advantages and disadvantages between the three different microphone arrays are compared and discussed in order to find the most suitable array for music recording in 3D.
Convention Paper 9980

Technical Comittee Meeting
Friday May 25, 09:00 – 10:00 Room Brera 1

Technical Committee Meeting on Automotive Audio

Workshop 20 Friday, May 25
09:15 – 10:15 Scala 1

STEREOPHONIC TECHNIQUES FOR VR AND 360º CONTENT

Presenters: **Hannes Dieterle**, SCHOEPS Mikrofone GmbH, Karlsruhe, Germany
Kacper Sagnowski, SCHOEPS Mikrofone GmbH, Karlsruhe, Germany
Helmut Wittek, SCHOEPS GmbH, Karlsruhe, Germany

In head-tracked binaural audio, the two-channel output is produced by a real-time convolution of a limited number of sources with HRTFs. There are elaborate systems that measure the HRTFs of each individual source. More general solutions define a grid of sources on a sphere. An arbitrary source is then mapped to this grid of sources based on the (higher-order) Ambisonics principle. With this method-utilized by many state-of-the-art binauralizers-it is possible to binauralize stereophonic virtual loudspeakers without performance problems. The workshop will give examples of stereophonic techniques for VR production and will show a workflow suggestion using practical examples. The advantages of using conventional stereo microphone arrays instead of first-order Ambisonics microphones will be shown.

Session P17 Friday, May 25
09:30 – 11:00 Poster Area

POSTERS: ANALYSIS / SYNTHESIS

09:30

P17-1 A Preliminary Study of Sounds Emitted by Honey Bees in a Beehive—*Stefania Cecchi, Alessandro Terenzi, Simone Orcioni, Paola Riolo, Sara Ruschioni, Nunzio Isidoro*, Universita Politecnica delle Marche, Ancona, Italy

Honey bees (*Apis mellifera L.*) are well known insects that have positive effects on a human being's life. They are so important that the honey bee colonies decline of the last years has produced an increasing interest for their safeguard. In this context, the proposed work aims at studying and developing an innovative system capable of monitoring the beehive's condition exploiting the sound emitted by the beehives in combination with measurable parameters such as temperature, humidity, CO2, hive weight, and weather conditions. In this paper preliminary results will be reported describing the developed platform and the first results obtained in a real scenario.
Convention Paper 9981

09:30

P17-2 Analysis of Reports and Crackling Sounds with Associated Magnetic Field Disturbances Recorded during a Geomagnetic Storm on March 7, 2012, in Southern Finland—*Unto K. Laine*, Aalto University, Aalto, Finland

Audio- and magnetic field signals were recorded during a geomagnetic storm on March 7, 2012, on open fields at Karkkila, approximately 70 km north of Helsinki by using a Zoom H4n recorder. Almost 90 distinct sound events like short claps, loud reports, or even a crackling sound were recorded. The paper describes the methods used and the results obtained in the audio- and magnetic field signal analysis. Relationship between the instances of the sound events and the geomagnetic activity is described. It is shown that the spectral properties of the crackling sound and the reports are similar. The challenges in finding connections between individual sounds and the corresponding magnetic field fluctuations are demonstrated and discussed.
Convention Paper 9982

09:30

P17-3 Harmonics and Intermodulation Distortion Analysis of the Even-Order Nonlinearity Controlled Effector Pedal—*Masaki Inui, Kanako Takemoto, Toshihiko Hamasaki*, Hiroshima Institute of Technology, Hiroshima, Japan

It is well known that components of an electric guitar system contain even-harmonics properties inherently, such as vacuum tubes or Fuzz pedal feedback circuits. Furthermore, there are several popular pedals in the "Overdrive" category and the difference in timbers attributes to even-harmonics. However, it is difficult to identify the correlation between transfer characteristics and auditory nuances, which seems to involve psychoacoustic factors such as masking effect. In this study we developed a novel pedal that can continuously control the strength of even-harmonics and odd-harmonics, and clarified the high-order distortion

effects from spectrum analysis of intermodulation distortion taking Loudness K-weighted full scale (LKFS) correction. This effector pedal must be a powerful tool for perceptual distortion analysis by using real instrument signal where masking effect occurs.
Convention Paper 9983

09:30

P17-4 Bandwidth Extension with Auditory Filters and Block-Adaptive Analysis—*Sunil G. Bharitkar,*[1] *Timothy Mauer,*[1] *Charles Oppenheimer,*[1] *Teresa Wells,*[1] *David Berfanger*[2]
[1] HP Labs., Inc., San Francisco, CA, USA
[2] HP, Inc., Vancouver, WA, USA

Bandwidth limits incurred by an audio signal due to low-excursion and narrow-bandwidth speakers reduces the perception of bass. A method to overcome this is to synthesize the harmonics of the fundamental frequency using side-chain processing. Depending on the input signal, intermodulation distortion could be introduced resulting in artifacts. A recent approach selects relevant portions of the low-frequency signal for reproduction using perceptually motivated filters, resulting in cleaner bass reproduction as confirmed through listening tests. However, one of the limitations is the need for large-duration frames or blocks (e.g., 5296 samples/block at 48 kHz) to obtain adequate frequency resolution at low-frequencies. In this paper we present an alternative approach that scales well in performance with respect to smaller block-sizes using 1/6-octave filterbank and power analysis.
Convention Paper 9984

09:30

P17-5 Evaluating Similarity of Temporal Amplitude Envelopes of Violin Sounds—*Magdalena Dziecielska, Krzysztof Martyn, Ewa Lukasik*, Poznan University of Technology, Poznan, Poland

The paper presents a method for evaluating similarity of temporal amplitude envelope of violin sounds. The experienced violinmakers are able to separate good sounding from bad sounding violins just by regarding the envelope of individual sounds. The contours of two-sided envelopes of individual open-string sounds are considered as images. Four uncorrelated visual descriptors are used to form a feature vector characterizing image shape. Individual distance measures are selected for each feature. The similar objects are grouped using k-means method. Violin sounds from AMATI database have been used in experiments.
Convention Paper 9985

Session EB3 — Friday, May 25
09:30 11:15 — Scala 2

SIGNAL PROCESSING / AUDIO EFFECTS & INSTRUMENTATION / MEASUREMENTS / FORENSICS

Chair: **Aki Mäkivirta**, Genelec Oy, Iisalmi, Finland

09:30

EB3-1 SysId—A System Identification Measurement Package—*Jont Allen*, University of Illinois, Urbana, IL, USA

SysID is a computer program that was developed c1980 at Bell Labs, for measurement of linear and nonlinear systems. At that time it was ported to the IBM-PC, and was sold by the Ariel Corp., Highland Park, NJ. Today it is a Matlab/Octave script that works with special hardware manufactured by Mimosa Acoustics of Champaign IL. SysID can measure the complex frequency response and impulse response of any linear system, such as loudspeakers, earphones, and rooms. When coupled with Matlab/Octave, it may be used as an audio-band network analyzer providing a pole-zero analysis of measured complex impedance, either electrical or acoustical. By using synchronous analysis, SysID complements spectrum analyzer functions, and in many cases, can extend the function of a two-channel spectrum analyzer, quickly resulting in highly accurate magnitude and phase results, limited only by the bit-accuracy of the codecs. It can more accurately characterize harmonic distortion and inter-modulation distortion, group delay, phase, impedance, and many other important system features, in near real time, along with a pole-zero analysis. In this presentation I will describe the theory behind SysID, and give a demonstration on the hand-held portable system. There is a long history behind such system. For example a number-theory method called MLS is believed by many to be superior, however this is a topic that needs clarification. MLS uses binary sequences, and therefore cannot directly measure TDH+N, or inter-modulation distortion. These issues are easily overcome by using int-32 sequences having powers of 2 sequence lengths. SysID has been used to measure auditoriums, conference rooms, loudspeaker impulse responses, cochlear potentials, ear canal impedance and reflectance, and many other two-port measurements such as a detailed loudspeaker analysis. The system is used in ECE-403 by students to analyze loud speaker characteristics, and was used to characterize and then model hearing aid receivers [1, 2].
Engineering Brief 421

09:45

EB3-2 Capacitor Distortion in High-Order Active Filters—*Douglas Self*, The Signal Transfer Company, London, UK

Some non-electrolytic capacitor types such as polyester generate distortion when they have a significant signal voltage across them. This can be avoided by using polypropylene or polystyrene types but they are larger and more expensive. I have previously shown that in 2nd-order Sallen & Key filters, both lowpass and highpass, only one of the two capacitors has to be of the expensive sort to obtain the same freedom of distortion achieved with two. The Geffe configuration allows 3rd and 4th-order filters to be realized economically in one stage, and it is shown that here too only one linear capacitor is required in both lowpass and highpass cases, saving a lot of money.
Engineering Brief 422

10:00

EB3-3 Graphical Development Design for a Heterogeneous DSP Core Architecture—*Miguel Chavez*, Analog Devices, Wilmington, MA, USA

For over 15 years, Analog Devices has continued improving its graphical programming environment to support several audio specific and general purpose digital signal processors (DSPs). As of 2016, all supported processors have either contained single or dual DSP cores whereas both of them have had the same architecture. With the need to have a heterogeneous DSP architecture the team

had the challenge to program both cores within the same environment. This paper describes challenges, trade-offs, and design decisions made when programming a new heterogeneous DSP core architecture.
Engineering Brief 423

10:15

EB3-4 Room Acoustic Measurements with Logarithmic Sine Sweeps on Android Phones—Lorenzo Rizzi, Giulio Scotti, *Gabriele Ghelfi*, Suono e Vita - Acoustic Engineering, Lecco, Italy

An application for room acoustics measurements has been developed for Android devices: its novelty represents the use of the logarithmic sine sweep method, which is better than typical direct methods used so far. The article describes the main points of the design phase and stresses the instrument on-site testing results in common use rooms. The testing of this Android instrument gives insights on small-room acoustics and on acoustical parameters measurement quality.
Engineering Brief 424

10:30

EB3-5 DirPat—Database and Viewer of 2D/3D Directivity Patterns of Sound Sources and Receivers—*Manuel Brandner, Matthias Frank, Daniel Rudrich*, University of Music and Performing Arts Graz, Graz, Austria

A measurement repository (DirPat) has been set up to archive all 3D and 2D directivity patterns measured at the Institute of Electronic Music and Acoustics, University of Music and Performing Arts in Graz. Directivity measurements have been made of various loudspeakers, microphones, and also of human speakers/singers for specific phonemes. The repository holds time domain impulse responses for each direction of the radiating or incident sound path. The data can be visualized with the provided 2D and 3D visualization scripts programmed in MATLAB. The repository is used for ongoing scientific research in the field of directivity evaluation of sources or receivers regarding localization, auditory perception, and room acoustic modeling.
Engineering Brief 425

10:45

EB3-6 The Anatomy, Physiology, and Diagnostics of Smart Audio Devices—*Xinhui Zhou, Mark Martin, Jayant Datta, Vijay*
Badawadagi, Audio Precision, Beaverton, OR, USA

Smart audio devices are becoming ubiquitous and their popularity has been skyrocketing. By current standards, a smart audio device is voice-controlled through interaction with an Internet-based intelligent virtual assistant and usually provides access to remote repositories of music or information. This paper focuses on the smart speakers (e.g., Amazon Echo, Google Home, etc.)—the most popular smart audio device. Though they are usually composed of relatively simple audio components, these devices incorporate very sophisticated audio signal processing, a plethora of audio pathways, and functional audio subsystems—posing significant challenge in testing. This paper explores the audio subsystems and pathways found on this type of device and suggests ways to test and validate their functionality and performance.
Engineering Brief 426

11:00

EB3-7 Proposed AES Test Standard for Specifying Continuous and Short-Term Maximum Power and SPL for Electronic
and Electro-Acoustic Systems: Part 1—*D. B. (Don) Keele, Jr.,*[1] *Steven Hutt,*[2] *Marshall Kay,*[3] *Hugh Sarvis*[4]
[1] DBK Associates and Labs, Bloomington, IN, USA
[2] Equity Sound Investments, Bloomington, IN, USA
[3] Keysight Technologies, Apex, NC, USA
[4] Presonus Audio Electronics-Worx Audio Technologies, Baton Rouge, LA, USA

This paper describes a proposed test method that allows both the continuous and short-term maximum peak output and SPL for electronic and electroacoustic systems to be measured. Systems such as power amplifiers, loudspeakers, and sound reinforcement/cinema systems can be measured and specified over the complete audible range. The test is divided into two parts that individually assesses: (1) the system's broadband continuous maximum output using various steady-state low-crest-factor test signals and (2) the system's short-term narrow-band maximum output using high-crest-factor test signals. The combination of both tests completely specifies the system's maximum output on a continuous and short-term narrow-band basis. A future Part 2 paper will go into detail concerning the tests and will illustrate with measured test results.
Engineering Brief 427

Workshop 21 **Friday, May 25**
09:30 – 11:00 **Scala 3**

BEYOND AES67, A BETTER AND UNIFIED EXPERIENCE OF THE AUDIO NETWORK

Presenters: **Jeff Berryman**, Bosch Communications, Claymont, Delaware, USA
Greg Shay, The Telos Alliance, Cleveland, OH, USA
Nicolas Sturmel, Merging Technologies, Puidoux, Switzerland

With the recent publication of the ST2110 standard, and the industry wide adoption of AES67, multi-vendor audio networks become a reality. But AES67 and ST2110 only consider transport on the network, and this is only a small portion of the many features needed to fully integrate a standard based system. This workshop will address the issue of control and monitoring of the audio network by presenting different open initiatives providing a level of unification on the network. The standardized control protocol AES70 (and especially its connection management part), the open interface specification and API for connection management NMOS and some other network oriented tools will be presented by world class experts on the matter before leaving room for a Q&A session.

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

Workshop 22 **Friday, May 24**
09:30 – 10:30 **Lobby**

10 YEARS PLOUD IN EUROPE —- THE PAST, PRESENT AND FUTURE

Presenters: **Florian Camerer**, ORF, Vienna, Austria
Eelco Grimm, HKU University of the Arts, Utrecht, Netherlands; Grimm Audio, Eindhoven, The Netherlands

Matthieu Parmentier, francetélévisions, Paris, France

In 2008, the European loudness group PLOUD was founded within the EBU, the European Broadcasting Union. A dedicated tribe of audio enthusiasts set out to conquer peak normalization and its detrimental effects on basically everything, but specifically on audio quality and listener satisfaction. Much has been been achieved in the 10 years since its gestation—but there remain a few loose ends; this was to be expected during such a fundamental change of the way audio is treated.

Florian Camerer, Eelco Grimm, and Matthieu Parmentier are all core members of PLOUD and share different aspects of the work done, being done and to be done:

Florian in his role as the chairman of PLOUD is giving a brief overview of the major changes that have already happened and are now well established. Loudness normalization is now the norm in TV, but other areas in broadcasting are still lagging behind, most notably Radio. An outlook will be given on the most burning issues in this area under the light of the recent work of AES on Recommended Practices for Streaming.

Eelco will offer a brief update about the work on Loudness in cinema. Topics that will be touched are the relationship between overall electric loudness and the dialog level, some data about acoustic playback levels of movies, and the master fader level setting at film festivals, and efforts to come to a new standard for trailers and commercials.

Matthieu will present a view on the activities of France Televisions, also targeting the tricky terrain of loudness in an object-based production environment

Student Event/Career Development
SC7 CLASSICAL MUSIC RECORDING EDUCATION PANEL DISCUSSION: CONTEMPORARY PRODUCTION PRACTICES AND TRAINING ENGINEERS FOR TODAY AND THE FUTURE
Friday, May 25 09:30 – 11:00 Arena 3 & 4

Chair: **Nyssim Lefford**, Luleå University of Technology, Luleå, Sweden

Panelists: *David Gleeson*, Royal Academy of Music, London, UK
Theresa Leonard, Freelance Music Producer / Audio Educator, Victoria, BC, Canada
Thilo Schaller, SUNY Buffalo State, Buffalo, NY, USA
Denis Vautrin, Conservatoire de Paris, Paris, France; Herisson.TV
Mark Willsher, Pin3hot Ltd., London, UK

To engage (new) audiences, orchestras and opera houses around the world are experimenting with new and often simultaneous modes of distribution, emerging formats, and new technologies-for example, VR, 3D audio, and “live from…” simulcasts. Consequently, classical productions today often need to be captured in a manner that can be repurposed in a multitude of formats. This panel will consider how workflows and aesthetics are changing to embrace new possibilities for classical music. We will review what audio engineering programs are doing to cover both established re-cording techniques and modern approaches; and also discuss how educators can anticipate future eventualities. What do audio engineering students need to know now to innovate in classical music recording tomorrow?

Professional Sound Expo PSE08
Friday, May 25, 10:00 – 10:45 PSE Stage

GENERATION 3 - E-COUSTIC SYSTEMS ELECTRONIC ARCHITECTURE

Presenter: **Steve Barbar**, E-Coustic Systems, Belmont, MA, USA

E-coustic Systems Electronic Architecture™ is the product of over 30 years research in the fields of digital audio, neuroscience, and acoustics—refined by more than 20+ years of successful installations in a wide variety of venues and applications. E-coustic Systems incorporates patented signal processing techniques that overcome the fundamental physics problem of using microphones and loudspeakers together—acoustic feedback. This, combined with our third generation acoustics processing, delivers astonishing acoustic realism that was heretofore unattainable. Generation 3 hardware provides a range of systems that are optimally scaled for different venue sizes and performance requirements.

Technical Comittee Meeting
Friday May 25, 10:00 – 11:00 Room Brera 1

Technical Committee Meeting on Acoustics and Sound Reinforcement

Tutorial 18 Friday, May 25
10:30 – 12:00 Scala 1

THE SOUND WILL TAKE A LEAD IN VR

Presenters: **Daniel Deboy**, DELTA Soundworks, Germany
Ana Monte, DELTA Soundworks, Germany
Michal Sokolowski
Paulina Szczepaniak

This session focuses on details of the work on two games in VR. This innovative platform introduces not only new possibilities and perspectives but also new challenges and questions.

Game ““Slavic Legacy VR” - First and the largest VR game based on mysterious Slavic mythology and culture. Through virtual reality, the player immerses himself into a dark story in Eastern Europe. From the very beginning, he has to sneak and hide from a deadly threat and its consequences. The unique fairytale atmosphere and uncommon in VR, award-winning graphics, is a mixture of realism and artistic vision of the creators. All of the tools, elements of the environment and encountered mythical beasts were created based on ethnographic research so that they faithfully reflect the life and beliefs of the then people. The game also confronts the problem of motion sickness with a completely new way of moving and designing the locations. These and many other features make Slavic Legacy a valued, one-of-a-kind and immersive production for the educational and entertainment market. Authors will present a gameplay with binaural playback and discuss technical, perception, and esthetic aspects of production process and choices. Panelists will also bring up challenges of sounds implementation in Unreal Engine.

Game “The Stanford Virtual Heart” Pediatric cardiologists at Lucile Packard Children’s Hospital Stanford are using immersive virtual reality technology to explain complex congenital heart defects, which are some of the most difficult medical conditions to teach and understand. The Stanford Virtual Heart experience helps families understand their child’s heart conditions by employing a new kind of interactive visualization that goes far beyond diagrams, plastic models and hand-drawn sketches. For medical trainees, it provides an immersive and engaging new way to learn about the two dozen most common and complex congenital heart anomalies by allowing them to inspect and manipulate the affected heart, walk around inside it to see how the blood is flowing, and watch how a particular defect interferes with the heart’s normal function. The panelists will give an insight about the challenges for the sound design and how it was integrated in Unity.

Standards Comittee Meeting
Friday May 25, 10:30 – 12:00 Room Castello 1

SC-05-05 Working Group on Grounding and EMC Practices

Session P18 Friday, May 25
11:00 – 13:00 Scala 4

PERCEPTION—PART 2

Chair: **Franz Zotter**, University of Music and Performing Arts Graz, Graz, Austria

11:00

P18-1 The Effect of the Rise Time and Frequency Character of the Sound Source Signal on the Sense of the Early LEV—*Toru Kamekawa, Atsushi Marui*, Tokyo University of the Arts, Tokyo, Japan

The effect of the rise-time of the sound source signal on the sense of the early LEV was investigated. Authors conducted the Scheffe's pairwise comparison method using seven kinds of bandpass noise that has eight kinds of rise-time convolved with impulse responses and played back from seven loudspeakers. From the result, the octave-band noises up to 1 kHz, the early LEV is felt strongly when the rise-time is 40 to 60 ms and the evaluation due to the difference in the rise-time varies in the case of the high frequency band. Additionally the early LEV in the mid-low range is related to the early reflections in 80 ms based on the first wavefront law, and there was almost no relationship between IACC.
Convention Paper 9986

11:30

P18-2 Plausibility of an Interactive Approaching Motion towards a Virtual Sound Source Based on Simplified BRIR Sets—*Annika Neidhardt, Alby Ignatious-Tommy, Anson Davis Pereppadan*, Technical University Ilmenau, Ilmenau, Germany

In this paper the interactive approaching motion towards a virtual loudspeaker created with dynamic binaural synthesis is subject to research. A realization based on a given set of measured binaural room impulse responses (BRIRs) was rated as plausible by all participants in a previous experiment. In this study the same BRIR-data is systematically simplified to investigate the consequences for the perception. This is of interest in the context of position-dynamic reproduction, related interpolation and extrapolation approaches as well as attempts of parameterization. The potential of inaudible data simplification is highly related to the human sensitivity to position-dependent changes in room acoustics. The results suggest a high potential for simplification, while some kinds of BRIR-impairment clearly affect the plausibility.
Convention Paper 9987

12:00

P18-3 The Impact of Trajectories of Head and Source Movements on Perceived Externalization of a Frontal Sound Source—*Song Li, Jiaxiang E, Roman Schlieper, Jürgen Peissig*, Leibniz Universität Hannover, Hannover, Germany

Two listening experiments were performed to investigate the influence of different trajectories of head and source movements on perceived externalization of a frontal sound source. In the first listening test, virtual moving sound sources with seven various trajectories were presented over headphones, while subjects' heads remained stationary. In the second test, subjects were asked to rotate their heads on three predefined trajectories coupled with real-time binaural rendering, while the simulated virtual sound source was kept stationary. After each presentation, subjects should rate the degree of perceived externalization. Results suggested that large head and source movements can improve perceived externalization, except source movements in the front/back direction. In addition, small source or head movements do not have the influence on externalization.
Convention Paper 9988

12:30

P18-4 Evaluation of Binaural Renderers: Externalization, Front/Back and Up/Down Confusions—*Gregory Reardon,[1] Gabriel Zalles,[1] Andrea Genovese,[1] Patrick Flanagan,[2] Agnieszka Roginska[1]*
[1] New York University, New York, NY, USA
[2] THX Ltd. - San Francisco, CA, USA

Binaural renderers can be used to reproduce dynamic spatial audio over headphones and deliver immersive audio content. Six commercially available binaural renderers with different rendering methodologies were evaluated in a multi-phase subjective study. This paper presents and discusses the testing methodology, evaluation criteria, and main findings of the externalization, front/back discrimination and up/down discrimination tasks that are part of the first phase. Statistical analysis over a large number of subjects revealed that the choice of renderer has a significant effect on all three dependent measures. Further, ratings of perceived externalization for the renderers were found to be content-specific, while renderer reversal rates were much more robust to different stimuli.
Convention Paper 9989

Workshop 23 Friday, May 25
11:00 – 12:00 Scala 3

AES67 IN REAL WORLD APPLICATIONS

Presenters: **Claudio Becker-Foss**, DirectOut GmbH, Mittweida, Germany
Andreas Hildebrand, ALC NetworX GmbH, Munich, Germany
Nicolas Sturmel, Merging Technologies, Puidoux, Switzerland

Since the publication of AES67 and its recent adoption by the ST 2110 media over IP standard, we see a growing number of installations and studio built around this standard. This workshop and Q&A session will focus on practical aspects of using AES67 and ST 2110-30 in the field: what are the benefits of using AES67? What are the difficulties? What should I do / not do, when configuring my network? Panelists experienced in real-world applications will provide valuable insights and share opinions on the importance of AES67 and its role as part of ST 2110 for the wider broadcast market.

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

Workshop 24 Friday, May 25
11:00 – 12:30 Lobby

NEW SURROUND AND IMMERSIVE RECORDINGS:

LISTENING SESSION

Presenters: **Jim Anderson**, Anderson Audio NY, New York, NY, USA; Clive Davis Institute of Recorded Music, New York University, New York, NY, USA
Ulrike Schwarz, Anderson Audio NY, New York, NY, USA

Multi-Grammy winner producer and engineer in surround productions Jim Anderson and Ulrike Schwarz have spent the past year recording and mixing music in high resolution and in immersive formats from venues in New York to Norway to Havana. Their recordings have been made in various 3D recording formats and feature solo piano, big band, jazz trio and quartet, and orchestral performances. Mixing has taken place at Skywalker Sound and mastering has been by Bob Ludwig and Darcy Proper. Recordings will highlight performances by Jane Ira Bloom, Gonzalo Rubalcaba, the Jazz Ambassadors, and Norway's Stavanger Symphony Orchestra.

Student Event/Career Development
SC8 EDUCATION AND CAREER/JOB FAIR
Friday, May 25, 11:00 – 13:00 Arena 2

The combined AES 144th 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 141st 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.

Professional Sound Expo PSE09
Friday, May 25, 11:00 – 11:45 PSE Stage

THE LONGER, THE BETTER: K-ARRAY'S PURE ARRAY TECHNOLOGY EXPLAINED

Presenter: **Daniele Mochi**, K-array, Florence, Italy

Ever since K-array launched its first slim column speaker in 2007, it has maintained its position of leader of sleek line arrays. With every subsequent product release, the company has perfected its line arrays composed of closely-spaced, full-range sound sources by incorporating its Pure Array Technology. The scope of this presentation is to explain this innovative technology in detail and provide examples of its benefits.

Technical Comitee Meeting
Friday May 25, 11:00 – 12:00 Room Brera 1

Technical Committee Meeting on Microphones and Applications

Special Event
SE4 BACK TO THE FUTURE, A TECHNOLOGY PROJECT REVIEW: OUTDOOR SOUND REINFORCEMENT OF SYMPHONY AND OPERA FOR EXTREMELY LARGE AUDIENCES
Friday, May 25, 11:15 – 12:15 Arena 3 & 4

Presenter: **David Scheirman**, Bose Professional, Carlsbad, CA USA

Serving audiences up to 200,000 persons for outdoor operatic and symphonic programs, a fully-distributed portable sound reinforcement system was deployed in public parks throughout New York City including Central Park. Supporting onstage sound from a tensile-membrane acoustical shell designed in a 36m (118 ft) wide and 21m (68 ft) high pyramid-like structure for the US $3.385M outdoor music pavilion, 24 battery-powered, digitally-delayed portable loudspeaker towers received wireless transmissions from a centrally-located mixing position to cover asymmetrical audience areas. A 24m (78 ft) wide stage area was served with only seven microphones. From 1991-95, the author led a team to deploy and operate this futuristic outdoor sound reinforcement system, pioneering innovative and advanced technologies now taken for granted. System design attributes and operating principles are detailed.

Workshop 25 Friday, May 25
12:00 – 13:30 Scala 3

THE CHALLENGE OF LOUDSPEAKER INTEGRATION IN AUTOMOTIVE AUDIO APPLICATIONS

Presenter **Alfred Svobodnik**, MVOID Group

The integration of loudspeakers into a vehicle, both in terms of acoustical and mechanical aspects, is a highly challenging topic. Parasitic vibrations of panels excited by the loudspeaker can lead to significant deterioration of the sound quality. Acoustical short-circuits and Helmholtz resonances are another potential source for performance degradation. This workshop aims to discuss the major challenges of automotive loudspeaker packaging and possible solutions to improve the acoustical quality.

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

Profesional Sound Expo PSE10
Friday, May 25, 12:00 – 12:45 PSE Stage

SETTING UP A CONTROL ROOM WITH GIK ACOUSTICS

Presenter: **Lukas Rimbach**, GIK Acoustics

Lukas Rimbach discusses the basics of room acoustics, how to set up and treat a control room - and the new Room Acoustics Visualizer from GiK, an augmented reality 3D rendering app for smartphones.

Technical Comitee Meeting
Friday May 25, 12:00 – 13:00 Room Brera 1

Technical Committee Meeting on High Resolution Audio

Workshop 26 Friday, May 25
12:15 – 13:30 Scala 1

THE EBU ADM RENDERER

Co-chairs: **Chris Pike**, BBC R&D, Salford, Greater Manchester, UK; University of York, Heslington, York, UK
Michael Weitnauer, IRT, Munich, Germany

Panelists: *Nicolas Epain*, b<>com
Thomas Nixon, BBC R&D, Salford, UK

This workshop will introduce the use of the Audio Definition Model to store and distribute Object-Based Audio master files to transport channel, object and scene-based contents.

The discussion will underline the role of OBA renderer within the post-production workflow and introduce EAR, the EBU ADM Renderer recently published as an open-source software.

Audio Applications Forum **Friday, May 24**
12:30 – 13:30 **Arena 1**

MAINS TO ACOUSTIC EFFICIENCY

Presenter: **Claudio Lastrucci**, Powersoft S.p.a., Scandicci (FI), Italy

A case study related to Power requirements to feed a subwoofer cabinet at high levels is pursued. Standardized signals and program audio signals are applied and results are directly measured on real devices. Measurements on true power input, true power output, and overall efficiency in the amplification chain has been addressed including different amplification topologies. As a result, surprising high overall mains input to acoustic output chain efficiency is evidenced in the specific, usable, passband of the speaker.

Tutorial 19 **Friday, May 25**
12:45 – 13:45 **Lobby**

**SUPPORTING THE STORY WITH SOUND—
AUDIO FOR FILM AND ANIMATION**

Presenters: **Kris Górski**, AudioPlanet, Koleczkowo, Poland
Kyle P. Snyder, Ohio University, School of Media Arts & Studies, Athens, OH, USA

Storytelling is a primary goal of film and there's no better way to ruin the story than with bad sound. This session will focus on current workflows, best practices, and techniques for film and animation by breaking down recent projects from the panelists.

Kyle Snyder will present audio from the "Media in Medicine" documentary *The Veterans' Project,* a film that seeks to bring civilian physicians and the American public closer to the individual voices of veterans of all ages, as they return to civilian life and seek to live long productive lives after they have served. *The Veterans' Project* is a feature-length documentary that weaves together the testimonies of injured or ill combat veterans who navigate the complexities of military, VA, and civilian medical systems in seeking treatment and reintegration into the civilian world.

Rodzina Treflików presented by Kris Górski is a classic, stop motion animation series for children. It uses original solutions for the mouth animation and generally speaking tries to combine the old and the new techniques of today to make an appealing proposition to the youngest audience. It is meant primarily for airing on children network TV but it is also shown in selected movie theaters. At the present moment we are in the process of finishing the 4th season. The movie's soundtrack is produced in 5.1 and stereo.

Student Event/Career Development
SC9 STUDENT RECORDING CRITIQUES
Friday, May 25, 13:00 – 14:00 **Galleria (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 of their studies can sign up to partici-pate. Students should sign up at the student (SDA) booth immediately on arrival at the convention, and deliver stereo 44.1 Khz, 24 bit AIFF or WAVE files to the SDA booth at that time. 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). These events are generously supported by Genelec.

Professional Sound Expo **PSE11**
Friday, May 25, 13:00 – 13:45 **PSE Stage**

AUDIO INTERFACES

Event sponsored by Prism Sound

Standards Comittee Meeting
Friday, 25 May, 13:30 – 15:30 **Room Castello 1**

SC-02-02 Working Group on Digital Input/Output Interfacing

Session P19 **Friday, May 25**
13:15 – 14:45 **Arena 2**

POSTERS: AUDIO PROCESSING / AUDIO EDUCATION

13:15

P19-1 Combining Fully Convolutional and Recurrent Neural Networks for Single Channel Audio Source Separation—*Emad M. Grais, Mark D. Plumbley*, University of Surrey, Guildford, Surrey, UK

Combining different models is a common strategy to build a good audio source separation system. In this work we combine two powerful deep neural networks for audio single channel source separation (SCSS). Namely, we combine fully convolutional neural networks (FCNs) and recurrent neural networks, specifically, bidirectional long short-term memory recurrent neural networks (BLSTMs). FCNs are good at extracting useful features from the audio data and BLSTMs are good at modeling the temporal structure of the audio signals. Our experimental results show that combining FCNs and BLSTMs achieves better separation performance than using each model individually.
Convention Paper 9990

13:15

P19-2 A Group Delay-Based Method for Signal Decorrelation—*Elliot K. Canfield-Dafilou, Jonathan S. Abel*, Stanford University, Stanford, CA, USA

By breaking up the phase coherence of a signal broadcast from multiple loudspeakers, it is possible to control the perceived spatial extent and location of a sound source. This so-called signal decorrelation process is commonly achieved using a set of linear filters and

finds applications in audio upmixing, spatialization, and auralization. Allpass filters make ideal decorrelation filters since they have unit magnitude spectra and therefore can be perceptually transparent. Here, we present a method for designing allpass decorrelation filters by specifying group delay trajectories in a way that allows for control of the amount of correlation as a function of frequency. This design is efficiently implemented as a cascade of biquad allpass filters. We present statistical and perceptual methods for evaluating the amount of decorrelation and audible distortion.
Convention Paper 9991

13:15

P19-3 Designing Quasi-Linear Phase IIR Filters for Audio Crossover Systems by Using Swarm Intelligence—*Ferdinando Foresi,*[1] *Paolo Vecchiotti,*[1] *Diego Zallocco,*[2] *Stefano Squartini*[1]
[1] Universita Politecnica delle Marche, Ancona, Italy
[2] Elettromedia s.r.l., Potenza Piena, Italy

In sound reproduction systems the audio crossover plays a fundamental role. Nowadays, digital crossover based on IIR filters are commonly employed, of which non-linear phase is a relevant topic. For this reason, solutions aiming to IIR filters approximating a linear phase behavior have been recently proposed. One of the latest exploits Fractional Derivative theory and uses Evolutionary Algorithms to explore the solution space in order to perform the IIR filter design: the IIR filter phase error is minimized to achieve a quasi-linear phase response. Nonetheless, this approach is not suitable for a crossover design, since the single filter transition band behavior is not predictable. This shoved the authors to propose a modified design technique including suitable constraints, as the amplitude response cut-off frequency, in the ad-hoc Particle Swarm Optimization algorithm exploring the space of IIR filter solutions. Simulations show that not only more performing filters can be obtained but also fully flat response crossovers achieved.
Convention Paper 9992

13:15

P19-4 Graduate Attributes in Music Technology: Embedding Design Thinking in a Studio Design Course—*Malachy Ronan, Donagh O'Shea*, Limerick Institute of Technology, Limerick, Ireland

Student acquisition of graduate attributes is an increasingly important consideration for educational institutes, yet embedding these attributes in the curriculum is often challenging. This paper recounts the process of embedding design thinking in a studio design course. The process is adapted to suit music technology students and delivered through weekly interactive workshops. Student adaptation to design thinking is assessed against the characteristics of experienced designers to identify issues and derive heuristics for future iterations of the course.
Convention Paper 9993

13:15

P19-5 Dynamic Range Controller Ear Training: Analysis of Audio Engineering Student Training Data—*Denis Martin, George Massenburg, Richard King*, McGill University, Montreal, Quebec, Canada; The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada

Eight students from the McGill University Graduate Program in Sound Recording participated in a dynamic range controller ear training program over a period of 14 weeks. Analysis of the training data shows a significant improvement in the percentage of correct responses over time. This result agrees with previous findings by other researchers and demonstrates a positive effect of this technical ear training program.
Convention Paper 9994

Session P20 **Friday, May 25**
13:15 – 15:45 **Scala 4**

AUDIO PROCESSING AND EFFECTS—PART 1

Chair: **Filippo Maria Fazi***,* University of Southampton, Southampton, UK

13:15

P20-1 Deep Neural Networks for Cross-Modal Estimations of Acoustic Reverberation Characteristics from Two-Dimensional Images—*Homare Kon, Hideki Koike*, Tokyo Institute of Technology, Meguro-ku, Tokyo, Japan

In augmented reality (AR) applications, reproduction of acoustic reverberation is essential for creating an immersive audio experience. The audio component of an AR experience should simulate the acoustics of the environment that users are experiencing. Earlier, sound engineers could program all the reverberation parameters in advance for a scene or if the audience was in a fixed position. However, adjusting the reverberation parameters using conventional methods is difficult because all such parameters cannot be programmed for AR applications. Considering that skilled acoustic engineers can estimate reverberation parameters from an image of a room, we trained a deep neural network (DNN) to estimate reverberation parameters from two-dimensional images. The results suggest a DNN can estimate the acoustic reverberation parameters from one image.
Convention Paper 9995

13:45

P20-2 Deep Learning for Timbre Modification and Transfer: An Evaluation Study—*Leonardo Gabrielli,*[1] *Carmine Emanuel Cella,*[2] *Fabio Vesperini,*[1] *Diego Droghini,*[1] *Emanuele Principi,*[1] *Stefano Squartini*[1]
[1] Universitá Politecnica delle Marche, Ancona, Italy
[2] IRCAM, Paris, France

In the past years, several hybridization techniques have been proposed to synthesize novel audio content owing its properties from two audio sources. These algorithms, however, usually provide no feature learning, leaving the user, often intentionally, exploring parameters by trial-and-error. The introduction of machine learning algorithms in the music processing field calls for an investigation to seek for possible exploitation of their properties such as the ability to learn semantically meaningful features. In this first work we adopt a Neural Network Autoencoder architecture, and we enhance it to exploit temporal dependencies. In our experiments the architecture was able to modify the original timbre, resembling what it learned during the training phase, while preserving the pitch envelope from the input.
Convention Paper 9996

14:15

P20-3 Feature Selection for Dynamic Range Compressor Parameter Estimation—*Di Sheng, György Fazekas*, Queen Mary University of London, London, UK

Casual users of audio effects may lack practical experience or knowledge of their low-level signal processing parameters. An intelligent control tool that allows using sound examples to control effects would strongly benefit these users. In a previous work we proposed a control method for the dynamic range compressor (DRC) using a random forest regression model. It maps audio features extracted from a reference sound to DRC parameter values, such that the processed signal resembles the reference. The key to good performance in this system is the relevance and effectiveness of audio features. This paper focusses on a thorough exposition and assessment of the features, as well as the comparison of different strategies to find the optimal feature set for DRC parameter estimation, using automatic feature selection methods. This enables us to draw conclusions about which features are relevant to core DRC parameters. Our results show that conventional time and frequency domain features well known from the literature are sufficient to estimate the DRC's threshold and ratio parameters, while more specialized features are needed for attack and release time, which induce more subtle changes to the signal.
Convention Paper 9997

14:45

P20-4 Effect of Delay Equalization on Loudspeaker Responses—*Aki Mäkivirta,*[1] *Juho Liski,*[2] *Vesa Välimäki*[2]
[1] Genelec Oy, Iisalmi, Finland
[2] Aalto University, Espoo, Finland

The impulse response of a generalized two-way loudspeaker is modeled and is delay equalized using digital filters. The dominant features of a loudspeaker are low and high corner roll-off characteristics and the behavior at the crossover points. The proposed model characterizes also the main effects of the mass-compliance resonant system. The impulse response, its logarithm and spectrogram, and the magnitude and group delay responses are visualized and compared with those measured from a two-way loudspeaker. The model explains the typical group-delay variations and magnitude-response deviations from a flat response in the passband. The group-delay equalization of the loudspeaker is demonstrated in two different methods. The first method, the time-alignment of the tweeter and woofer elements using a bulk delay, is shown to cause ripple in the magnitude response. The second method, which flattens the group delay of the speaker model in the whole audio range, leads to pre-ringing in the impulse response.
Convention Paper 9998

15:15

P20-5 An Allpass Chirp for Constant Signal-to-Noise Ratio Impulse Response Measurement—*Elliot K. Canfield-Dafilou, Jonathan S. Abel*, Stanford University, Stanford, CA, USA

A method for designing an allpass chirp for impulse response measurement that ensures a constant signal-to-noise ratio (SNR) in the measurement is presented. By using the background noise and measurement system's frequency responses, a measurement signal can be designed by specifying the group delay trajectory. This signal will have a small crest factor and will be optimally short such that the measured impulse response will have a desired and constant SNR.
Convention Paper 10014

Tutorial 20 Friday, May 25
13:30 – 15:00 Arena 3 & 4

BEFORE THE STUDIO: THE ART OF PREPRODUCTION

Presenters: **Bill Crabtree**, MTSU, Murfreesboro, TN, USA
Wes Maebe, RAK Studios/Sonic Cuisine, London, UK
Barry Marshall, The New England Institute of Art, Boston, MA, USA
Mandy Parnell, Black Saloon Studios, London, UK
Marek Walaszek, Addicted to Music Studio, Warsaw, Poland

Preproduction is as essential to the success of music productions as it is to the success of any film or television show. So why do so many young producers skimp-on or even skip preproduction? This panel will take a look at the preparation of both the act and the act's material for the recording process. We will focus on lyric and melody as well as musical arrangements and instrumentation.

Tutorial 21 Friday, May 25
14:00 – 15:00 Lobby

KRAFTWERK AND BOOKA SHADE—THE CHALLENGE TO CREATE ELECTRO POP MUSIC IN IMMERSIVE / 3D AUDIO

Presenter: **Tom Ammermann**, New Audio Technology GmbH, Hamburg, Germany

Music has not a cinematic approach where spaceships are flying around the listeners. Nonetheless, music can become a fantastic spatial listening adventure in Immersive / 3D. How this sounds will be shown with the new Kraftwerk (Grammy nominated) and Booka Shade Blu-ray releases this year. Production philosophies, strategies, and workflows to create Immersive / 3D in current workflows and DAWs will be shown and explained.

Tutorial 22 Friday, May 25
14:00 – 15:00 Scala 1

A PRACTICAL USE OF SPATIAL AUDIO FOR STORYTELLING, AND DIFFERENT DELIVERY PLATFORMS

Presenter: **Axel Drioli**, 3D Sound Designer and Producer, London, UK / Lille, France

An overview of Spatial Audio's characteristics, strengths, and weaknesses for storytelling. A case study will be analyzed to better understand what can be done at every stage of the spatial audio production process, from recording to decoding to different delivery platforms. This presentation is tailored for an audience of content creators and engineers, who may not know enough (or not at all) the Spatial Audio capabilities for storytelling, using demos and examples to explain and clarify concepts.

Professional Sound Expo PSE12
Friday, May 25, 14:00 – 14:45 PSE Stage

PRO TOOLS IN THE CLOUD

Presenter: **Dave Tyler**, AVID

Event sponsored by AVID

Technical Comittee Meeting
Friday May 25, 14:00 – 15:00 **Room Brera 1**

Technical Committee Meeting on Perception and Subjective Evaluation of Audio Signals

Session EB4 **Friday, May 25**
14:15 – 15:45 **Scala 2**

SPATIAL AUDIO

Chair: **Frank Schultz**, University of Music and Performing Arts Graz, Graz, Austria

14:15

EB4-1 Ambilibrium—A User-Friendly Ambisonics Encoder/Decoder-Matrix Designer Tool—*Michael Romanov*, University of Music and Performing Arts Graz, Graz, Austria

The name "Ambilibrium" is composed of the terms "Ambi" (prefix)—Latin: around and "aequilibrium"—Latin: balance. This engineering brief discusses the implementation of a tool that creates the spatial balance within any spherical microphone array to any loudspeaker array in the form of custom Ambisonics encoder / decoder matrices packed in a user friendly interface. A technique for automatic loudspeaker position estimation using similar approach to GPS and the optimization of the AllRAD are also discussed in this engineering brief.
Engineering Brief 428

14:30

EB4-2 Distant Speech Beamforming Improving Multi-User ASR—*Adam Kupryjanow,*[1] *Raghavendra R. R.,*[2] *Przemek Maziewski,*[1] *Lukasz Kurylo*[1]
[1] Intel Technology Poland, Gdansk, Poland
[2] Intel Technology, Devarabeesanahalli, KA, India

In this paper an algorithm that improves side speaker attenuation for super directive beamformer like MVDR (Minimum Variance Distortionless Response) is presented. This technique can be utilized in a scenario where there are multiple people in a room intending to interact with an ASR- (automatic speech recognition) enabled device, e.g., smart speaker. The experiments show that the proposed solution gives a reduction of WER (word error rate) up to 23.93% calculated for command uttered by one user when a second user was treated as the side speaker.
Engineering Brief 429

14:45

EB4-3 An Efficient Method for Producing Binaural Mixes of Classical Music from a Primary Stereo Mix—*Tom Parnell, Chris Pike*, BBC R&D, Salford, Greater Manchester, UK

Radio audiences in the UK are increasingly listening using headphones, and binaural mixes are likely to offer more natural and immersive classical musical experiences than stereo broadcasts. However, the stereo mix is currently a priority for broadcasters, and producers have limited resources to create an additional, binaural mix. This engineering brief describes the semi-automated workflow used to produce binaural mixes of performances from the BBC Proms. Spatial audio mixes were created by repositioning the individual microphone signals from the stereo broadcast in three dimensions, and adding ambient signals captured using a 3D microphone array. A commercial mixing application was used for spatial panning and binaural rendering, and the resulting binaural audio was streamed live online. Comments on the production workflow were collected from the music balancers, and audience responses were surveyed.
Engineering Brief 430

15:00

EB4-4 The Anaglyph Binaural Audio Engine—*David Poirier-Quinot, Brian F. G. Katz*, Sorbonne Université, CNRS, Paris, France

Anaglyph is part of an ongoing research effort into the perceptual and technical capabilities of binaural rendering. The Anaglyph binaural audio engine is a VST audio plugin for binaural spatialization integrating the results of over a decade of spatial hearing research. Anaglyph has been designed as an audio plugin to both support ongoing research efforts as well as to make accessible the fruits of this research to audio engineers through traditional existing DAW environments. Among its features, Anaglyph includes a personalizable morphological ITD model, near-field ILD and HRTF parallax corrections, a Localization Enhancer, an Externalization Booster, and SOFA HRIR file support. The basic architecture and implementation of each audio-related component is presented here.
Engineering Brief 431

15:15

[Engineering Brief 432 withdrawn]

15:30

EB4-6 HOBA-VR: HRTF On Demand for Binaural Audio in Immersive Virtual Reality Environments—*Michele Geronazzo,*[1] *Jari Kleimola,*[2] *Erik Sikström,*[3] *Amalia de Götzen,*[1] *Stefania Serafin,*[1] *Federico Avanzini*[4]
[1] Aalborg University, Copenhagen, Denmark
[2] Hefio Ltd., Espoo, Finland
[3] Virsabi ApS, Copenhagen, Denmark
[4] University of Milan, Milan, Italy

One of the main challenges of spatial audio rendering in headphones is the personalization of the so-called head-related transfer functions (HRTFs). HRTFs capture the listener's acoustic effects supporting immersive and realistic virtual reality (VR) contexts. This e-brief presents the HOBA-VR framework that provides a full-body VR experience with personalized HRTFs that were individually selected on demand based on anthropometric data (pinnae shapes). The proposed WAVH transfer format allows a flexible management of this customization process. A screening test aiming to evaluate user localization performance with selected HRTFs for a non-visible spatialized audio source is also provided. Accordingly, is might be possible to create a user profile that contains also individual non-acoustic factors such as localizability, satisfaction, and confidence.
Engineering Brief 433

Tutorial 23 **Friday, May 25**
15:00 – 16:30 **Scala 3**

OPTIMIZING TRANSDUCER DESIGN FOR SYSTEMS WITH ADAPTIVE NONLINEAR CONTROL

Presenters: **Gregor Höhne**, Klippel GmbH, Dresden, Germany
Marco Raimondi, STMicroelectronics SRL, Cornaredo (MI), Italy

Modern loudspeakers increasingly incorporate digital signal processing, amplification and the transducer itself in one unit. The utilized algorithms not only comprise linear filters but adaptively control the system, deploying measured states and complex models. Systems with adaptive nonlinear control can be used to equalize, stabilize, linearize, and actively protect the transducer. Thus, more and more demands can be taken care of by digital signal processing than by pure transducer design, opening new degrees of freedom for the latter. The tutorial focuses on how this freedom can be utilized to design smaller and more efficient loudspeaker systems. Examples are given for how existing transducer designs can be altered to increase their efficiency and which new challenges arise when driving speaker design to its limits.

Workshop 27 **Friday, May 25**
15:00 – 16:30 **Arena 3 & 4**

ARTIFICIAL INTELLIGENCE IN YOUR AUDIO

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

Panelists: *Jonathan Bailey*, iZotope
Joshua D. Reiss, Queen Mary University of London, London, UK

AI has been part of the listener's experience for many years . . . . now it is influencing the development of tools made for music production and music creation. We will look at how it is being used, the promise it holds in developing new tools and the challenges it presents for music engineers and producers.

This session is presented in association with the AES Technical Committee on Recording Technology and Practices

Student Event/Career Development
SC10 RECORDING COMPETITION—PART 2
Friday, May 25, 15:00 – 18:00 **Lobby**

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 afternoon. 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.

Professional Sound Expo **PSE13**
Friday, May 25, 15:00 – 15:45 **PSE Stage**

REFERENCE MONITORING AND TRANSLATION

Presenter: **Aki Mäkivirta**, Genelec Oy, Iisalmi, Finland

Once you have overcome the circles of confusion in pro monitoring, you can create content that translates well between rooms and between in-room and headphones. Topics covered: In-room spectral calibration, level calibration, monitor dispersion, loudness-based production and monitoring, prevention of listener fatigue.

Technical Comittee Meeting
Friday May 25, 15:00 – 16:00 **Room Brera 1**

Technical Committee Meeting on Network Audio Systems

Tutorial 24 **Friday, May 25**
15:15 – 16:15 **Scala 1**

360 & VR—CREATE OZO AUDIO AND HEAD LOCKED SPATIAL MUSIC IN COMMON DAWS

Presenter: **Tom Ammermann**, New Audio Technology GmbH, Hamburg, Germany

VR and 360 are mainly headphone applications. Of course, binaural headphone virtualization, well known since decades, becomes important as never before. Fortunately current binaural virtualization technologies offer a fantastic new experience far away from creepy low quality 'surround simulations' in the past. The presentation will show how to create content for NOKIA's new 360 audio format OZO Audio and how to create head locked binaural music in common DAWs.

Standards Comittee Meeting
Friday May 25, 15:30 – 17:00 **Room Castello 1**

AESSC Plenary

Technical Comittee Meeting
Friday May 25, 16:00 – 17:00 **Room Brera 1**

Technical Committee Meeting on Audio Forensics

Tutorial 25 **Friday, May 25**
16:15 – 17:45 **Scala 3**

LIVE SOUND SUBWOOFER SYSTEM OPTIMIZATION

Presenter: **Adam J. Hill**, University of Derby, Derby, Derbyshire, UK

There is little reason this day in age to accept undesirable low-frequency sound coverage in live sound reinforcement. The theories behind subwoofer system optimization are well-known within academia and various branches of industry, although this knowled ge isn't always fully-translated into practical terms for end-users. This tutorial provides a comprehensive overview of how to achieve desirable low-frequency sound coverage including: subwoofer polar response control, array and cluster configuration, signal routing/processing options, performance stage effects, source decorrelation, acoustic barriers and perceptual considerations. The tutorial is suitable for practitioners, academics and students, alike, providing practical approaches to low-frequency sound control and highlighting recent technological advancements.

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

Session P21 **Friday, May 25**
16:30 – 18:00 **Arena 2**

POSTERS: AUDIO CODING AND QUALITY

16:30

P21-1 Quantization with Signal Adding Noise Shaping Using Long Range Look-Ahead Optimization—*Akihiko Yoneya*, Nagoya Institute of Technology, Nagoya, Japan

A re-quantization approach for digital audio signals using noise shaping by extra signal addition is studied. The approach has been proposed by the author but its properties have not been studied well. In this paper, the feature and performance of the approach is investigated. As a result, the noise shaping performance is a little better than the conventional one and perceptual evaluation is superior in terms of the fineness of the sound source image especially when the optimization horizon used in the additional signal calculation is wide. Since a wide horizon requires a lot of computation, a pruning scheme of the optimization is proposed to reduce the calculation time and the amount of computation is evaluated experimentally.
Convention Paper 9999

16:30

P21-2 A Comparison of Clarity in MQA Encoded Files vs. Their Unprocessed State as Performed by Three Groups—Expert Listeners, Musicians, and Casual Listeners—*Mariane Generale, Richard King, Denis Martin*, McGill University, Montreal, QC, Canada; CIRMMT, Montreal, QC, Canada

This paper aims to examine perceived clarity in MQA encoded audio files compared to their unprocessed state (96-kHz 24-bit). Utilizing a methodology initially proposed by the authors in a previous paper, this study aims to investigate any reported differences in clarity for three musical sources of varying genres. A double-blind test is conducted using three groups—expert listeners, musicians, and casual listeners—in a controlled environment using high-quality loudspeakers and headphones. The researchers were interested in comparing the responses of the three target groups and whether playback systems had any significant effect on listeners' perception. Data shows that listeners were not able to significantly discriminate between MQA encoded files and the unprocessed original due to several interaction effects.
Convention Paper 10000

16:30

P21-3 A Subjective Evaluation of High Bitrate Coding of Music—*Kristine Grivcova, Chris Pike, Thomas Nixon*, BBC Research & Development, Salford, UK

The demand to deliver high quality audio has led broadcasters to consider lossless delivery. However the difference in quality over formats used in existing services is not clear. A subjective listening test was carried out to assess the perceived difference in quality between AAC-LC at 320 kbps and an uncompressed reference, using the method of ITU-R BS.1116. Twelve audio samples were used in the test, which included orchestral, jazz, vocal music, and speech. A total of 18 participants with critical listening experience took part in the experiment. The results showed no perceptible difference between AAC-LC at 320 kbps and the reference.
Convention Paper 10001

16:30

P21-4 Subjective Evaluation of a Spatialization Feature for Hearing Aids by Normal-Hearing and Hearing-Impaired Subjects—*Gilles Courtois,[1] Hervé Lissek,[1] Philippe Estoppey,[2] Yves Oesch,[3] Xavier Gigandet[3]*
[1] Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
[2] Acoustique Riponne, Lausanne, Switzerland
[3] Phonak Communciations AG, Murten, Switzerland

Remote microphone systems significantly improve speech intelligibly performance offered by hearing aids. The voice of the speaker(s) is captured close to the mouth by a microphone, then wirelessly sent to the hearing aids. However, the sound is rendered in a diotic way, which bypasses the spatial cues for localizing and identifying the speaker. The authors had formerly proposed a feature that localizes and spatializes the voice. The current study investigates the perception of that feature by normal-hearing and hearing-impaired subjects with and without remote microphone system experience. Comparing the diotic and binaural reproductions, subjects rated their preference over various audiovisual stimuli. The results show that experienced subjects mostly preferred the processing achieved by the feature, contrary to the other subjects.
Convention Paper 10002

16:30

P21-5 Virtual Reality for Subjective Assessment of Sound Quality in Cars—*Angelo Farina,[1] Daniel Pinardi,[1] Marco Binelli,[1] Michele Ebri,[1]* Lorenzo Ebri[2]
[1] Università di Parma, Parma, Italy
[2] Ask Industries S.p.A., Monte San Vito, Italy

Binaural recording and playback has been used for decades in the automotive industry for performing subjective assessment of sound quality in cars, avoiding expensive and difficult tests on the road. Despite the success of this technology, several drawbacks are inherent in this approach. The playback on headphones does not have thebenefit of head-tracking, so the localization is poor. The HRTFs embedded in the binaural rendering are those of the dummy head employed for recording the sound inside the car, and finally there is no visual feedback, so the listener gets a mismatch between visual and aural stimulations. The new Virtual Reality approach solves all these problems. The research focuses on obtaining a 360° panoramic video of the interior of vehicle, accompanied by audio processed in High Order Ambisonics format, ready for being rendered on a stereoscopic VR visor. It is also possible to superimpose onto the video a real-time color map of noise levels, with iso-level curves and calibrated SPL values. Finally, both sound level color map and spatial audio can be filtered by the coherence with one or multiple reference signals, making it possible to listen and localize very precisely noise sources and excluding all the others. These results have been acquired employing a massive spherical microphone array, a 360° panoramic video recording system, and accelerometers or microphones for the reference signals.
Convention Paper 10003

16:30

P21-6 Quality Evaluation of Sound Broadcasted Via DAB+ System Based on a Single Frequency Network—*Stefan Brachmanski, Maurycy Kin*, Wroclaw University of Technology, Wroclaw, Poland

This paper presents the results of quality assessment of speech and music signals transmitted via the DAB+ system. The musical signals have been evaluated in both

overall quality and some particular attributes. The subjective research was provided with the use of ACR procedure according to the ITU recommendation and the results have been presented as the MOS values for various bit rates. The speech signals were additionally examined with PESQ method. The results have shown that the assumed quality of 4 MOS, for this kind of broadcasting could be achieved at 48 kbit/s. This fact was confirmed by both subjective and objective research. The differences of results obtained for overall sound quality and particular sound attributes are discussed.
Convention Paper 10004

16:30

P21-7 An Investigation into Spatial Attributes of 360° Microphone Techniques for Virtual Reality—*Connor Millns, Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

Listening tests were conducted to evaluate perceived spatial attributes of two types of 360° microphone techniques for virtual reality (First Order Ambisonics (FOA) and the Equal Segment Microphone Array (ESMA)). Also a binaural dummy head was included as a baseline for VR audio. The four attributes tested were: source shift/ensemble spread, source/ensemble distance, environmental width, and environmental depth. The stimuli used in these tests included single and multi-source sounds consisting of both human voice and instruments. The results indicate that listeners can distinguish differences in three of the four spatial attributes. The binaural head was rated the highest for each attribute and FOA was rated the least except for in environmental depth.
Convention Paper 10005

16:30

P21-8 Statistical Tests with MUSHRA Data—*Catarina Mendonça,*[1] *Symeon Delikaris-Manias*[2]
[1] Aalto University, Espoo, Finland
[2] Aalto University, Helsinki, Finland

This work raises concerns regarding the statistical analysis of data obtained with the MUSHRA method. There is a widespread tendency to prefer the ANOVA test, which is supported by the recommendation. This work analyzes four assumptions underlying the ANOVA tests: interval scale, normality, equal variances, and independence. Data were collected from one experiment and one questionnaire. It is found that MUSHRA data tend to violate all of the above assumptions. The consequences of each violation are debated. The violation of multiple assumptions is of concern. The violation of independence of observations leads to the most serious concern. In light of these findings, it is concluded that ANOVA tests have a high likelihood of resulting in type 1 error (false positives) with MUSHRA data and should therefore never be used with this type of data. The paper finishes with a section devoted to statistical recommendations. It is recommended that when using the MUSHRA method, the Wilcoxon or Friedman tests be used. Alternatively, statistical tests based on resampling methods are also appropriate.
Convention Paper 10006

16:30

P21-9 Investigation of Audio Tampering in Broadcast Content—*Nikolaos Vryzas, Anastasia Katsaounidou, Rigas Kotsakis, Charalampos A. Dimoulas, George Kalliris*, Aristotle University of Thessaloniki, Thessaloniki, Greece

Audio content forgery detection in broadcasting is crucial to prevent misinformation spreading. Tools for the authentication of audio files can be proven very useful, and several techniques have been proposed. In the current paper a database for evaluation of such techniques is introduced. A script was created for automatic generation of tampered audio files, given a number of original source files that contain recorded speech, while they have been encoded in different audio formats (Mp3, AAC, AMR, FLAC) and bitrates and finally they were used to generate the tampered audio files. The database was subjectively evaluated by experts in terms of samples changing audibility. The effect of tampering on several audio features was tested, in order to propose semi-automatic methods for discrimination between the original and tampered files. The database and the scripts are publically accessible so that researchers can use the pre-generated files or use the script to create datasets oriented to their research interests.
Convention Paper 10007

Session P22 **Friday, May 25**
16:00 – 18:00 **Scala 4**

PERCEPTION—PART 3

Chair: **Jürgen Peissig**, Leibniz Universität Hannover, Hannover, Germany

16:00

P22-1 Audibility of Loudspeaker Group-Delay Characteristics—*Juho Liski,*[1] *Aki Mäkivirta,*[2] *Vesa Välimäki*[1]
[1] Aalto University, Espoo, Finland
[2] Genelec Oy, Iisalmi, Finland

Loudspeaker impulse responses were studied using a paired-comparison listening test to learn about the audibility of the loudspeaker group-delay characteristics. Several modeled and six measured loudspeakers were included in this study. The impulse responses and their time-reversed versions were used in order to maximize the change in the temporal structure and group delay without affecting the magnitude spectrum, and the subjects were asked whether they could hear a difference. Additionally, the same impulse responses were compared after convolving them with a pink impulse, defined in this paper, which causes a low-frequency emphasis. The results give an idea of how much the group delay of a loudspeaker system can vary so that it is unlikely to cause audible effects in sound reproduction. Our results suggest that when the group delay in the frequency range from 300 Hz to 1 kHz is below 1.0 ms, it is inaudible. With low-frequency emphasis the group delay variations can be heard more easily.
Convention Paper 10008

16:30

P22-2 The Influence of Hearing and Sight on the Perceptual Evaluation of Home Speaker Systems —*Hans-Joachim Maempel, Michael Horn*, Federal Institute for Music Research, Berlin, Germany

Home speaker systems are not only functional but also

aesthetic objects with both acoustic and optical proper-ties. We investigated the perceptual evaluation of four home speaker systems under acoustic, optical, and opto-acoustic conditions (factor *Domain*). By varying the speakers' acoustic and optical properties under the opto-acoustic condition in a mutually independent manner (factors *Acoustic loudspeaker, Optical loudspeaker*), we also investigated their proportional influence on perception. To this end, 40 non-expert participants rated 10 auditory, 2 visual, and 4 audiovisual features. The acoustic stimuli were generated by means of data-based dynamic binaural synthesis. Noticeably, participants did not realize that the speakers were acoustically simulated. Results indicated that only the mean ratings of two auditory and one audiovisual feature were significantly influenced by the factor *Domain*. There were speaker-dependent effects on three further auditory features. Small crossmodal effects from *Optical loudspeaker* on six auditory features were observed. Remarkably, the audiovisual features, particularly *monetary value*, were dominated by the optical properties instead of the acoustic. This is due to a low acoustic and a high optical variance of the speakers. Results give reason to the hypothesis that the optical properties imply an overall quality that in turn may influence the rating of auditory features.
Convention Paper 10009

17:00

P22-3 A VR-Based Mobile Platform for Training to Non-Individualized Binaural 3D Audio—*Chungeun Kim, Mark Steadman, Jean-Hugues Lestand, Dan F. M. Goodman, Lorenzo Picinali*, Imperial College London, London, UK

Delivery of immersive 3D audio with arbitrarily-positioned sound sources over headphones often requires processing of individual source signals through a set of Head-Related Transfer Functions (HRTFs), the direction-dependent filters that describe the propagation of sound in an anechoic environment from the source to the listener's ears. The individual morphological differences and the impracticality of HRTF measurement make it difficult to deliver completely individualized 3D audio in this manner, and instead lead to the use of previously-measured non-individual sets of HRTFs. In this study a VR-based mobile sound localization training prototype system is introduced that uses HRTF sets for audio. It consists of a mobile phone as a head-mounted device, a hand-held Bluetooth controller, and a network-enabled PC with a USB audio interface and a pair of headphones. The virtual environment was developed on the mobile phone such that the user can listen-to/navigate-in an acoustically neutral scene and locate invisible target sound sources presented at random directions using non-individualized HRTFs in repetitive sessions. Various training paradigms can be designed with this system, with performance-related feedback provided according to the user's localization accuracy, including visual indication of the target location, and some aspects of a typical first-person shooting game, such as enemies, scoring, and level advancement. An experiment was conducted using this system in which 11 subjects went through multiple training sessions, using non-individualized HRTF sets. The localization performance evaluations showed reduction of overall localization angle error over repeated training sessions, reflecting lower front-back confusion rates.
Convention Paper 10010

17:30

P22-4 Speech-To-Screen: Spatial Separation of Dialogue from Noise towards Improved Speech Intelligibility for the Small Screen—*Philippa Demonte, Yan Tang, Richard J. Hughes, Trevor Cox, Bruno Fazenda, Ben Shirley*, University of Salford, Salford, Greater Manchester, UK

Can externalizing dialogue when in the presence of stereo background noise improve speech intelligibility? This has been investigated for audio over headphones using head-tracking in order to explore potential future developments for small-screen devices. A quantitative listening experiment tasked participants with identifying target words in spoken sentences played in the presence of background noise via headphones. Sixteen different combinations of three independent variables were tested: speech and noise locations (internalized/externalized), video (on/off), and masking noise (stationary/fluctuating noise). The results revealed that the best improvements to speech intelligibility were generated by both the video-on condition and externalizing speech at the screen while retaining masking noise in the stereo mix.
Convention Paper 10011

Professional Sound Expo **PSE14**
Friday, May 25, 16:00 – 16:45 **PSE Stage**

MAINS TO ACOUSTIC EFFICIENCY

Presenter: **Claudio Lastrucci**, Powersoft S.p.a., Scandicci (FI), Italy

A case study related to Power requirements to feed a subwoofer cabinet at high levels is pursued. Standardized signals and program audio signals are applied and results are directly measured on real devices.

Measurements on true power input, true power output and overall efficiency in the amplification chain has been addressed including different amplification topologies. As a result, surprising high overall mains input to acoustic output chain efficiency is evidenced in the specific, usable, passband of the speaker.

Session EB5 **Friday, May 25**
16:30 – 18:00 **Arena 2**

POSTERS 2

16:30

EB5-1 Musical Polyphony Estimation—*Saarish Kareer, Sattwik Basu*, University of Rochester, Rochester, NY, USA

Knowing the number of sources present in a mixture is useful for many computer audition problems such as polyphonic music transcription, source separation, and speech enhancement. Most existing algorithms for these applications require the user to provide this number thereby limiting the possibility of complete automatization. In this paper we explore a few probabilistic and machine learning approaches for an autonomous source number estimation. We then propose an implementation of a multi-class classification method using convolutional neural networks for musical polyphony estimation. In addition, we use these results to improve the performance of an instrument classifier based on the same dataset. Our final classification results for both the networks, prove that this method is a promising starting point for further advancements in unsupervised source counting and separation algorithms for music and speech.
Engineering Brief 434

16:30

EB5-2 Variability of Speech to Reverberation Modulation Energy Ratio—*Przemek Maziewski, Adam Kupryjanow*, Intel Technology Poland, Gdansk, Poland

This paper illustrates variability of the speech to reverberation modulation energy ratio (SRMR). The presented experiments indicate SRMR inconsistencies per user, per utterance, and per microphone position. Additionally, the results show that the normalization available in the reference SRMR implementation does not limit the variability to an acceptable range. Further the paper presents a study of SRMR and reverberation time (RT) correlation. Experiments suggest that a precise relation between SRMR and RT can only be obtained for a specific utterance coming from a known user.
Engineering Brief 435
This e-Brief was not presented

16:30

EB5-3 Introducing a Dataset of Guitar Sounds for Electric Guitars Model Recognition—*Renato Profeta, Gerald Schuller*, Ilmenau University of Technology, IImenau, Germany;

This engineering brief introduces a dataset of electric guitar sounds. The main goal of the dataset is to provide a set of electric guitar recordings that can be used for research in identification and/or classification of different electric guitar types. The dataset, at its current stage, consists of recordings from 30 guitars of different manufacturers and types, with around 3500 music events. All audio files are acquired in one-channel, 16-bit waveform audio file format with a sampling rate of 44100 Hz and are accompanied by parameter annotations in xml format. The dataset is planned to include recordings of over 50 guitars and will be released in uncompressed wav file format under Creative Commons License.
Engineering Brief 436

16:30

EB5-4 Development of the 4-pi Sampling Reverberator, VSVerb—Preliminary Experiments—*Masataka Nakahara,*[1] *Akira Omoto,*[1,2] *Yasuhiko Nagatomo*[3]
[1] ONFUTURE Ltd., Tokyo, Japan
[2] Kyushu University, Fukuoka, Japan
[3] Evixar Inc., Tokyo, Japan

The strategy for capturing and restoring acoustic properties of a 4-pi sound field is one of the important issues for immersive-sound productions. Especially for post-production work, they are required to have a lot of flexibility. While some strategies have been already proposed such as Ambisonics, WFS, and BoSC, they require much effort and have less flexibility. Therefore, the authors developed a 4-pi sampling reverberator, VSVerb, which restores 4-pi reverberation by using captured virtual sound sources on site. Moreover, it enables to adjust various acoustic parameters at the post-production stage. In order to verify feasibility of the proposed method, some preliminary experiments were conducted. As a result, effectiveness of the VSVerb is confirmed, and some future subjects are also found.
Engineering Brief 437

16:30

EB5-5 Practical Evaluation of Sweet Spot in Current Noise Reproduction Systems—*Piotr Klinke, Roksana Kostyk, Jan Banas, Przemek Maziewski, Dominik Stanczak*, Intel Technology Poland, Gdansk, Poland

Quality assessment for speech recognition systems is a complex problem involving many factors but it all starts and ends with a proper synthesis of user scenarios in a controlled acoustic environment. In this paper industry leading background noise reproduction methods are presented. Implementations of two ETSI standards for noise reproduction, the ES 202 396-1 and TS 103 224 are measured and compared to determine the vulnerability for sound pressure level and frequency response deviation around the setup center.
Engineering Brief 438

16:30

EB5-6 Open Hardware Mobile Wireless Serial Audio Transmission Unit for Acoustical Ecological Momentary Assessment Using Bluetooth RFCOMM—*Sven Franz, Holger Groenewold, Inga Holube, Jörg Bitzer*, Jade Hochschule Oldenburg, Oldenburg, Germany

Acoustical Ecological Momentary Assessment (EMA) is a necessary step towards understanding noise exposure in everyday life and can also help identify obstacles in the speech understanding of hearing impaired people. Smartphones would represent a desirable tool for this task, but no simple solution has been proposed yet. Stereo audio transmission between mobile devices via Bluetooth using the Advanced Audio Distribution Profile (A2DP) is a common technique. However, due to software restrictions in Android, this profile is limited to act as a source and is prohibited from receiving a stereo audio stream. In this contribution we present a solution to transmitting uncompressed stereo audio data via Radio Frequency Communication (RFCOMM) enabling an Android smartphone to act as a receiver. Although, in contrast to A2DP, this solution is limited to stereo, 16 kHz and 16 bit, the resulting audio quality is sufficient for speech signals and acoustical measurements.
Engineering Brief 439

16:30

EB5-7 Sound Masking on iOS Devices. Masking Everyday Noises and Tinnitus—*Lorenzo Rizzi, Nadir Bertolasi, Gabriele Ghelfi*, Suono e Vita - Acoustic Engineering, Lecco, Italy

The research started studying the mobile devices abilities and limits to perform an environmental noise analysis using the internal microphone recording. Perceptual masking algorithms have been implemented to specifically modify natural sounds for the analyzed noise. Then sound pleasantness algorithms have been applied to obtain better masking sounds. This set-up is being extended to the masking of tinnitus: the tinnitus frequency selection is being proposed to the user through sine-sweep listening.
Engineering Brief 440

16:30

EB5-8 Wearable Mobile Bluetooth Device for Stereo Audio Transmission to a Modified Android Smartphone—*Holger Groenewold, Sven Franz, Inga Holube, Jörg Bitzer*, Jade Hochschule Oldenburg, Oldenburg, Germany

For acoustical long-term measurements based on ecolog-

ical momentary assessment (EMA), we needed a solution for sending A2DP audio data to an Android smartphone. Due to Android's software limitations, the common smartphones cannot act as the necessary sink for stereo audio. Our innovation is a wearable Bluetooth device containing a wireless transmission unit with two microphones connected to a Nexus 5 smartphone. The Nexus 5 platform was chosen, since Google performed tests featuring an Android-powered car audio system, which enables a stereo Bluetooth sink. The proposed open hardware and software solution can be used for several audio application areas, where non-intrusive long-term observations are necessary, e.g., noise exposure dosimeters.
Engineering Brief 441

16:30

EB5-9 Theoretical and Experimental Study of an Acoustic Monopole Source—*Pierluigi A. Argenta, Francesco Martellotta, Leonardo Soria*, Politecnico di Bari, Bari, Italy

As it is well known, the acoustic monopole source plays a fundamental role in the fields of theoretical and experimental acoustics. In this work a sound source that well approximates the dynamic response of a theoretical monopole is designed and the parameters affecting its operational behavior are highlighted and optimized in terms of sound power and range of reproducible frequencies. We develop a hybrid lumped-parameter-finite-element model for describing the source operation. The model is first experimentally tested to validate its predictive effectiveness. Then, we perform a non-dimensional parametric optimization of the source directionality, obtaining relationships with which general design guidelines are identified, to minimize the far field directionality and, thus, achieve a quite omnidirectional behavior.
Engineering Brief 442

16:30

EB5-10 Development and Validation of a Full Range Acoustic Impedance Tube—*Roman Schlieper, Song Li, Jürgen Peissig*, Leibniz Universität Hannover, Hannover, Germany

The knowledge about the physical properties of materials is of high importance for research and development in acoustics. A standardized method for the determination of acoustic impedances is the impedance measuring tube based on the transfer functions method according to ISO 10534-2. This engineering brief presents the development of an impedance measuring tube with an internal diameter of 8 mm for acoustical impedance measurements in the range of 60 Hz to 20 kHz. The impedance tube was validated by comparison of the measurement results to analytical results of the rigid termination, the open-ended tube, and the empty sample holder.
Engineering Brief 443

16:30

EB5-11 Directivity and Electro-Acoustic Measurements of the IKO—*Frank Schultz, Markus Zaunschirm, Franz Zotter*, University of Music and Performing Arts, Graz, Austria

The icosahedral loudspeaker (IKO) as a compact spherical array is capable of 3rd order Ambisonics (TOA) beamforming, and it is used as a musical and technical instrument. To develop and verify beamforming with its 20 loudspeakers flush-mounted into the faces of the regular icosahedron, electroacoustic properties must be measured. We offer a collection of measurement data of IEM's IKO1, IKO2, and IKO3 along with analysis tools to inspect these properties. Multiple-input-multiple-output (MIMO) data comprises: (i) laser vibrometry measurements of the 20x20 transfer functions from driving voltages to loudspeaker velocities, (ii) 20x16 finite impulse responses (FIR) of the TOA decoding filters, and (iii) 648x20 directional impulse responses from driving voltages to radiated sound pressure. With the open data sets, open source code, and resulting directivity patterns, we intend to support reproducible research about beamforming with spherical loudspeaker arrays.
Engineering Brief 444

Session EB6 Friday, May 25
16:15 – 17:30 Scala 2

TRANSDUCERS & PSYCHOACOUSTICS

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

16:15

EB6-1 Woofer Performance Variance Due to Components and Assembly Process—*Maria Costanza Bellini, Angelo Farina*, Università di Parma, Parma, Italy

This paper presents an experimental study of the main causes of scrap during the production of a woofer loudspeaker. After analyzing the most critical components of a transducer, samples with reference and modified components have been built and characterized in terms of frequency-response and linear distortion curves and electrical, mechanical, acoustical parameters. In addition, a second set of samples has been built using reference components but varying the assembly process parameters; these samples also have been characterized as the previous ones. Measurements have been performed both in an anechoic chamber, along a production line, and inside a car. By the analysis of acquired data, the authors have individuated the most influential components and assembly parameters in terms of required performance.
Engineering Brief 446
[This e-Brief is presented by Angelo Farina]

16:30

EB6-2 Design and Measurement of a First-Order, Horizontally Beam-Controlling Loudspeaker Cube—*Nils Meyer-Kahlen,[1,2] Franz Zotter,[2] Katharina Pollack[1,2]*
[1] University of Technology, Graz, Austria
[2] University of Music and Performaing Arts Graz, Graz, Austria

This paper describes a loudspeaker cube with four transducers on its horizontal facets, designed to enable sound radiation with adjustable first-order beam control. Design and equipping of the cubical loudspeaker is presented along with two open data sets containing multiple-input-multiple-output impulse responses (MIMO-IRs) of our measurements. The first one contains 648x 4 MIMO-IRs from input voltages to a grid of microphones at a fixed distance. The second set contains 4x 4 MIMO-IRs from input voltages to loudspeaker cone velocities, and it characterizes the active and passive transducer coupling through the enclosure that we aim to equalize/decouple. Based on these measurements we present a simple FIR filter design required for beam control of which we discuss

operation range and limitations.
Engineering Brief 447

16:45

EB6-3 Ambisonics Directional Room Impulse Response as a New Convention of the Spatially Oriented Format for Acoustics—*Andrés Pérez-López,*[1,2] *Julien De Muynke*[1]
[1] Eurecat ,Barcelona, Spain
[2] Pompeu Fabra University, Barcelona, Spain

Room Impulse Response (RIR) measurements are one of the most common ways to capture acoustic characteristics of a given space. When performed with microphone arrays, the RIRs inherently contain directional information. Due to the growing interest in Ambisonics and audio for Virtual Reality, new spherical microphone arrays recently hit the market. Accordingly, several databases of Directional RIRs (DRIRs) measured with such arrays, referred to as Ambisonics DRIRs, have been publicly released. However, there is no format consensus among databases. With the aim of improving interoperability, we propose an exchange format for Ambisonics DRIRs, as a new Spatially Oriented Format for Acoustics (SOFA) convention. As a use-case, some existing databases have been converted and released following our proposal.
Engineering Brief 448

17:00

EB6-4 Fidelity of Low Frequency Reproduction in Cars in a Sound Field Control Context—*Hans Lahti,*[1] *Anders Löfgren,*[2] *Adrian Bahne*[3]
[1] Harman, Gothenburg, Sweden
[2] Volvo Cars, Torlsanda, Sweden
[3] Dirac Research AB - München, Germany

Overall sound quality of factory-delivered automotive sound systems has reached a very high standard. Particularly branded high-end systems comprise great components and are well-tuned. The low frequency reproduction in automotive sound systems is, however, typically flawed. The most prominent flaw consists of resonant bass reproduction and undesirable spectral decay characteristics with strong ringing in wide frequency bands. To overcome this challenge, we adapt an algorithm allowing for simultaneous equalization of multiple channels, assuring full exploitation of the acoustic degrees of freedom inherent to a multichannel system. The sound field can be spatially controlled, yielding a uniform and tight reproduction of low frequencies in all regions of interest throughout the car compartment, with controlled and improved spectral decay characteristics.
Engineering Brief 449

17:15

EB6-5 A Distributed Audio System for Automotive Applications—*Johannes Boehm, Dirk Olszewski, Zafar Baig Mirza, Philipp Rathmann, Antonio Prados-Vilchez, Vitalie Botan, Juergen Binder, Klaus Rodemer*, paragon AG, Delbrück, Germany

With a trend to higher levels of drivetrain electrification and autonomous driving, the technology to increase audio performance is becoming a more significant factor of request. Instead of centralizing related signal processing in a single powerful hardware platform, distributing it in a more intelligent way can lead to several advantages such as optimized cabling, reduced weight, improved system scalability, performance, and costs. The distributed audio system proposed in this work is connected to an automobile head unit that serves as human machine interface and media source. Portions of the data acquisition, signal processing, and amplification are placed within distributed processing nodes. We present a realization with 34 loudspeakers and 16 microphones featuring seat-individual 3D audio rendition, in-car communication, and further innovative use cases.
Engineering Brief 450

Workshop 28 **Friday, May 25**
16:30 – 18:00 **Scala 1**

MIXING VR BEING IN VR

Presenters: **Daniel Deboy**, DELTA Soundworks, Germany
Christian Sander, Dear Reality GmbH, Germany

Mixing Virtual Reality (VR) content such as 360° Film can be a frustrating job. Current workflows either show an equirectangular projection of the 360° film or just a small portion of the full view on regular 2D Displays, next to the common DAW environment. A representation of the audio objects as an overlay may help to map the object to the correct visual location, but is not a replacement of viewing the film with a head mounted display (HMD). We present a new workflow that enables the engineer to mix object based audio directly in VR without leaving the HMD.

Special Event
SE5 AUDIO AESTHETICS—DESIGNED FOR THE FUTURE
Friday, May 25, 16:30 – 18:00 **Arena 3& 4**

Presenters: **Jacob Mathew**
Andrea Pivetta

With Milan as the home for the Audio Engineering Society's 144th Convention, this special event is dedicated to the aesthetics of audio design and in particular the innovative world of loudspeaker design aesthetics. Focusing on the work of a selection of inspirational audio designers, the session will introduce a brief history of audio design; design principles; design processes; technologies; and the materials that are being used to develop these icons of audio.

Special Event
SE6 BANQUET
Friday, May 25, 20:00 – ???
Grand Hotel Et De Milan
Via Alessandro Manzoni, 29, Milano

Come and enjoy an evening in Milan with your colleagues.

For this year's banquet join us in the heart of Milan at the famous Grand Hotel Et De Milan (www.grandhoteletdemilan.it/en) an 18th-century mansion best known as the Milanese house to Giuseppe Verdi, where the composer wrote the "Othello" and "The Falstaff." Conveniently located on Via Manzoni, facing fashionable Via Montenapoleone, the Grand Hotel et de Milan is only a few minutes walk to La Scala Opera House, The Duomo Cathedral, and the Galleria Vittorio Emanuele.

After the aperitif served in the large Chimney Hall and in the adjacent Gerry's bar, we will move to Puccini lounge for the dinner. Decorated with a prestigious collection of paintings by Maria Luigia d'Austria, Puccini lounge represents a classic example of the nineteenth century liberty architecture. The menu includes a 3-course meal with traditional Milanese dishes, a selection of wines, coffee, and friandises.

The Grand Hotel Et De Milan is in the very center of Milan. From NH Hotel it is easily reachable by Metro. The transfer takes about 45 minutes: take Metro Line 2 - green - (direction to "Cologno Nord" or "Gessate") for 12 stops, change to Metro Line 3 - yellow -

in "Centrale FS" (direction to "Comasina" or "San Donato") and get off after 3 stops at "Montenapoleone", right in front of the Grand Hotel.

Note that you need the extra-urban ticket (2,5 Euros) to go from NH to the city center.

Tickets will be available at the registration area. Spacing is limited.

Session P23 **Saturday, May 26**
09:00 – 11:30 **Scala 4**

AUDIO PROCESSING AND EFFECTS—PART 2

Chair: **Balázs Bank**, Budapest University of Technology and Economics, Budapest, Hungary

09:00

P23-1 Stage Compression in Transaural Audio—*Filippo Maria Fazi, Eric Hamdan*, University of Southampton, Southampton, UK

The reproduction of binaural audio with loudspeakers, also referred to as transaural audio, is affected by a number of artifacts. This work focuses on the effect of reproduction error on low frequency Interaural Time Difference (ITD). Transaural systems do not provide perfect cross-talk cancellation between the left and right ear signals, especially at low frequencies. It is shown that increase in cross-talk leads to a perceived source azimuth angle that is smaller than intended. The authors show that in ideal theoretical conditions the angular error calculated from the interaural phase difference indicates stage compression for frequencies for which high cross-talk occurs. This trend is shown in the resultant ITD calculated from Interaural Cross Correlation (IACC), examined in one-third octave bands.
Convention Paper 10012

09:30

P23-2 Multi-Track Crosstalk Reduction Using Spectral Subtraction—*Fabian Seipel,*[1] *Alexander Lerch*[2]
[1] Technical University of Berlin, Berlin, Germany
[2] Georgia Institute of Technology, Atlanta, GA, USA

While many music-related blind source separation methods focus on mono or stereo material, the detection and reduction of crosstalk in multi-track recordings is less researched. Crosstalk or "bleed" of one recorded channel in another is a very common phenomenon in specific genres such as jazz and classical, where all instrumentalists are recorded simultaneously. We present an efficient algorithm that estimates the crosstalk amount in the spectral domain and applies spectral subtraction to remove it. Randomly generated artificial mixtures from various anechoic orchestral source material were employed to develop and evaluate the algorithm, which scores an average SIR-Gain result of 15.14 dB on various datasets with different amounts of simulated crosstalk.
Convention Paper 10013

10:00

P23-4 Wave Digital Modeling of the Diode-Based Ring Modulator—*Alberto Bernardini,*[1] *Kurt James Werner,*[2] *Paolo Maffezzoni,*[1] *Augusto Sarti*[1]
[1] Politecnico di Milano, Milan, Italy
[2] Queen's University Belfast, Belfast, UK

The ring modulator is a strongly nonlinear circuit common in audio gear, especially as part of electronic musical instruments. In this paper an accurate model based onWave Digital (WD) principles is developed for implementing the ring modulator as a digital audio effect. The reference circuit is constituted of four diodes and two multi-winding transformers. The proposed WD implementation is based on the Scattering Iterative Method (SIM), recently developed for the static analysis of large nonlinear photovoltaic arrays. In this paper SIM is shown to be suitable for implementing also audio circuits for Virtual Analog applications, such as the ring modulator, since it is stable, robust and comparable to or more efficient than state-of-the-art strategies in terms of computational cost.
Convention Paper 10015

10:30

P23-5 Improving the Frequency Response Magnitude and Phase of Analogue-Matched Digital Filters—*John Flynn,*[1,2] *Joshua D. Reiss*[2]
[1] Balance Mastering, London, UK
[2] Queen Mary University of London, London, UK

Current closed-form IIR methods for approximating an analogue prototype filter in the discrete-domain do not match frequency response phase. The frequency sampling method can match phase, but requires extremely long filter lengths (and corresponding latency) to perform well at low frequencies. We propose a method for discretizing an analogue prototype that does not succumb to these issues. Contrary to the IIR methods, it accurately approximates the phase, as well as the magnitude response. The proposed method exhibits good low frequency resolution using much smaller filter lengths than design by frequency sampling.
Convention Paper 10016

11:00

P23-6 Optimization of Personal Audio Systems for Intelligibility Contrast—*Daniel Wallace, Jordan Cheer*, University of Southampton, Southampton, Hampshire, UK

Personal audio systems are designed to deliver spatially separated regions of audio to individual listeners. This paper demonstrates a method of personal audio system design that provides a level of contrast in the perceived speech intelligibility between bright and dark audio zones. Limitations in array directivity which would lead to a loss of privacy are overcome by reproducing a synthetic masking signal in the dark zone. This signal is optimized to provide effective masking whilst remaining subjectively pleasant to listeners. Results of this optimization from a simulated personal audio system are presented.
Convention Paper 10017

Tutorial 26 **Saturday, May 26**
09:00 – 10:30 **Scala 1**

CANCELED

Student Event/Career Development
SC11 CREATING AUDIO PLUGINS WITH MATLAB
Saturday, May 26 09:00 – 10:30
Arena 3 & 4

Presenter: **Gabriele Bunkheila**, MathWorks, Cambridge, UK

The first AES MATLAB Plugin Student Competition is now open for submissions – The entry deadline is August 15, 2018, with show-case and awards scheduled for the upcoming 145th Con-vention of the AES in New York City.

This optional tutorial covers the technical foundations needed to enter the competition. After attend-ing, you will be able to build a simple VST plugin using MATLAB. You will learn about structur-ing code for real-time efficiency, defining clear interfaces to tune parameters via interactive con-trols, 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. A video recording will be available online after the Convention.

For more information about the competition, please visit "http://www.aes.org/students/awards/mpsc/"

Tutorial 27 **Saturday, May 26**
09:15 – 10:15 **Scala 2**

HEARING THE PAST: USING ACOUSTIC MEASUREMENT TECHNIQUES AND COMPUTER MODELS TO STUDY HERITAGE SITES

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

The relationship between acoustics and theater performances throughout history has been an area of extensive research. However, studies on pre-seventeenth century acoustics have focused either on Greek and Roman or Elizabethan theater. The study of medieval acoustics has been centered on churches both as worship spaces and as sites of liturgical drama, leaving aside the spaces used for secular drama performances. This tutorial explores how a combination of acoustic measurement techniques adapted to challenging outdoor spaces and the use of a multiplicity of computer models to tackle the unknowns regarding heritage sites can help enhance our knowledge on how our ancestors 'heard' drama performances in connection to the acoustics of the performance spaces and the surrounding soundscapes.

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

Session P24 **Saturday, May 26**
09:30 – 11:00 **Poster Area**

POSTERS: SPATIAL AUDIO

09:30

P24-1 Acoustic and Subjective Evaluation of 22.2- and 2-Channel Reproduced Sound Fields in Three Studios—*Madhu Ashok,*[1] *Richard King,*[2] *Toru Kamekawa,*[3] *Sungyoung Kim*[4]
[1] University of Rochester, Rochester, NY, USA
[2] McGill University, Montreal, Quebec, Canada; The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada
[3] Tokyo University of the Arts, Adachi-ku, Tokyo, Japan
[4] Rochester Institute of Technology, Rochester, NY, USA

Three studios of similar outer-shell dimensions, with varying acoustic treatments and absorptivity, were evaluated via both recorded and simulated binaural stimuli for 22.2- and 2-channel playback. A series of analysis, including acoustic modelling in CATT-Acoustic and subjective evaluation, was conducted to test whether the 22.2-channel playback preserved common perceptual impressions regardless of room-dependent physical characteristics. Results from multidimensional scaling (MDS) indicated that listeners used one perceptual dimension for differentiating between reproduction format, and others for physical room characteristics. Clarity and early decay time measured in the three studios illustrated a similar pattern when scaled from 2- to 22.2-channel reproduced sound fields. Subjective evaluation revealed a tendency to preserve inherent perceptual characteristics of 22.2-channel playback in spite of different playback conditions.
Convention Paper 10018

09:30

P24-2 Audio Source Localization as an Input to Virtual Reality Environments—*Agneya A. Kerure, Jason Freeman*, Georgia Institute of Technology, Atlanta, GA, USA

This paper details an effort towards incorporating audio source localization as an input to virtual reality systems, focusing primarily on games. The goal of this research is to find a novel method to use localized live audio as an input for level generation or creation of elements and objects in a virtual reality environment. The paper discusses the current state of audio-based games and virtual reality, and details the design requirements of a system consisting of a circular microphone array that can be used to localize the input audio. The paper also briefly discusses signal processing techniques used for audio information retrieval and introduces a prototype of an asymmetric virtual reality first-person shooter game as a proof-of-concept of the potential of audio source localization for augmenting the immersive nature of virtual reality.
Convention Paper 10019

09:30

P24-3 High Order Ambisonics Encoding Method Using Differential Microphone Array—*Shan Gao, Xihong Wu, Tianshu Qu*, Peking University, Beijing, China

High order Ambisonics (HOA) is a flexible way to represent and analyze the sound field. In the process of the spherical Fourier transform, the microphone array need be uniformly scattered on the surface of a sphere, which limits the application of the theory. In this paper we introduce a HOA encoding method using the differential microphone arrays (DMAs). We obtain the particular beam patterns of different orders of spherical functions by the weighted sum of time-delayed outputs from closely-spaced differential microphone array. Then HOA coefficients are estimated by projecting the signals to the beam patterns. The coefficients calculated by the DMA are compared to the results derived from the theoretical spherical harmonics, which proves the effectiveness of our method.
Convention Paper 10020

09:30

P24-4 Development of a 64-Channel Spherical Microphone Array and a 122-Channel Loudspeaker Array System for 3D Sound Field Capturing and Reproduction Technology Research—*Shoken Kaneko,*[1] *Tsukasa Suenaga,*[1] *Hitoshi Akiyama,*[1] *Yoshiro Miyake,*[1] *Satoshi Tominaga,*[1] *Futoshi Shirakihara,*[1] *Hiraku Okumura*[1,2]
[1] Yamaha Corporation, Iwata-shi, Japan
[2] Kyoto University, Kyoto, Japan

In this paper we present our recent activities on building facilities to drive research and development on 3D sound field capturing and reproduction. We developed a 64-channel spherical microphone array, the ViReal Mic, and a

122-channel loudspeaker array system, the ViReal Dome. The ViReal Mic is a microphone array whose microphone capsules are mounted on a rigid sphere, with the positions determined by the spherical Fibonacci spiral. The ViReal Dome is a loudspeaker array system consisting of 122 active coaxial loudspeakers. We present the details of the developed systems and discuss directions of future research.
Convention Paper 10021

09:30

P24-5 A Recording Technique for 6 Degrees of Freedom VR— *Enda Bates, Hugh O'Dwyer, Karl-Philipp Flachsbarth, Francis M. Boland*, Trinity College Dublin, Dublin, Ireland

This paper presents a new multichannel microphone technique and reproduction system intended to support six degrees of freedom of listener movement. The technique is based on a modified form of the equal segment microphone array (ESMA) concept and utilizes four Ambisonic (B-format) microphones in a near-coincident arrangement with a 50cm spacing. Upon playback, these Ambisonic microphones are transformed into virtual microphones with different polar patterns that change based on the listener's position within the reproduction area. The results of an objective analysis and an informal subjective listening test indicate some inconsistencies in the on and off-axis response, but suggest that the technique can potentially support six degrees of freedom in a recorded audio scene using a compact microphone array that is well suited to Virtual Reality (VR) and particularly Free View Point (FVV) applications.
Convention Paper 10022

09:30

P24-6 On the Use of Bottleneck Features of CNN Auto-Encoder for Personalized HRTFs—*Geon Woo Lee,*[1] *Jung Min Moon,*[1] *Chan Jun Chun,*[2] *Hong Kook Kim*[1]
[1] Gwangju Institute of Science and Technology (GIST), Gwangju. Korea
[2] Korea Institute of Civil Engineering and Building Technology (KICT), Goyang, Korea

The most effective way of providing immersive sound effects is to use head-related transfer functions (HRTFs). HRTFs are defined by the path from a given sound source to the listener's ears. However, sound propagation by HRTFs differs slightly between people because the head, body, and ears differ for each person. Recently, a method for estimating HRTFs using a neural network has been developed, where anthropometric pinna measurements and head-related impulse responses (HRIRs) are used as the input and output layer of the neural network. However, it is inefficient to accurately measure such anthropometric data. This paper proposes a feature extraction method for the ear image instead of measuring anthropometric pinna measurements directly. The proposed method utilizes the bottleneck features of a convolutional neural network (CNN) auto-encoder from the edge detected ear image. The proposed feature extraction method using the CNN-based auto-encoder will be incorporated into the HRTF estimation approach.
Convention Paper 10023

Workshop 29 — Saturday, May 26
09:30 – 11:00 — Lobby

SPATIAL AUDIO MICROPHONES

Chair: **Helmut Wittek**, SCHOEPS GmbH, Karlsruhe, Germany

Panelists: *Gary Elko*, mh acoustics
Johannes Kares, Sennheiser
Hyunkook Lee, University of Huddersfield, Huddersfield, UK
Oliver Thiergart, International Audio Laboratories Erlangen, Erlangen, Germany
Tomasz Zernicki, Zylia sp. z o.o., Poznan, Poland

Multichannel loudspeaker setups as well as Virtual Reality applications enable Spatial Sound to be reproduced with large resolution. However, on the recording side it is more complicated to gather a large spatial resolution. Various concepts exist in theory and practice for microphone arrays. In this workshop the different concepts are presented by corresponding experts and differences, applications as well as pros and cons are discussed. The different array solutions include coincident and spaced Ambisonics arrays as well as Stereophonic multi-microphone arrays.

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

Professional Sound Expo — PSE16
Saturday, May 26, 10:00 – 10:45 — PSE Stage

AUDIO POST-PRODUCTION PRIMER—OUT OF TIME AND OVER BUDGET

Presenter: **Glenn Lorbecki**, Glenn Sound Inc., Seattle, WA, USA

A workshop geared for engineers seeking to understand the modern audio post-production workflow in a world where everyone is competing to do the work faster, better, AND cheaper.

Tutorial 28 — Saturday, May 26
10:30 – 12:00 — Scala 2

INTELLIGENT ACOUSTIC INTERFACES FOR HIGH-DEFINITION 3D AUDIO

Presenter: **Danilo Comminiello**, Sapienza University of Rome, Rome, Italy

This tutorial aims at introducing a new paradigm for interpreting 3D audio in acoustic environments. Intelligent acoustic interfaces involve both sensors for data acquisition and signal processing methods with the aim of providing a high-quality 3D audio experience to a user. The tutorial explores the motivations for 3D intelligent interfaces. In that sense, the recent standardization of the MPEG-H has provided an incredible boost. Then, how to design 3D acoustic interfaces is analyzed, involving ambisonics and small arrays. Moreover, the main methodologies are introduced for the processing of recorded 3D audio signals in order to "provide intelligence" to interfaces. Here, we leverage the properties of signal processing in the quaternion domain. Finally, some examples of 3D audio applications are shown involving intelligent acoustic interfaces.

Workshop 30 — Saturday, May 26
10:30 – 12:00 — Scala 1

INVOLVING PERFORMERS' CREATIVITY IN INNOVATIVE AUDIO RESEARCH

Co-chairs: **Jan Berg**, Luleå University of Technology, Piteå, Sweden

Amandine Pras, Paris Conservatoire (CNSMDP), Paris, France; University of Lethbridge, Lethbridge, Alberta, Canada

Panelists: *Valentin Bauer*, Paris Conservatoire (CNSMDP), Paris, France
Dimitri Soudoplatoff, Paris, France
Petter Sundkvist, Luleå University of Technology, Piteå, Sweden

With a multidisciplinary panel of performers and researchers, this workshop highlights the benefits of examining the impact of sound engineers' manipulations of acoustics and audio technology on musical performance and creative decisions. Based on previous studio and stage experiments that involved orchestra conductors and instrumentalists, we discuss novel research approaches that incorporate Audio in the emergent fields of Performance Science and Creative Cognition to study how performers are affected by sound, e.g. when monitoring in binaural to improvise in studio and when performing in venues with different acoustic conditions. Whereas performers' feedback can inform the development of innovative audio technology, our studies can make musicians more aware of sound engineering possibilities and hall acoustics' impact, also providing students with a broader perspective on artistic practices.

This workshop will be addressed to students, researchers, and technology developers who want to discuss novel approaches to study the interconnections between technology/acoustics and musical performance/creativity.

Special Event
SE7 IMMERSIVE AUDIO FOR REALITY OR VIRTUAL REALITY—DOES IT MAKE A DIFFERENCE?
Saturday, May 26, 10:30 – 12:00 Arena 3 & 4

Chair: **Daniel Deboy**, DELTA Soundworks, Germany

Panelists: *Tom Ammerman,* New Audio Technology
Feli Andriessens, Ton un Meister, Germany
Ana Monte, DELTA Soundworks, Germany
Tom Parnell, BBC, UK
Martin Rieger, VRTONUNG, Germany
Agnieszka Roginska, New York University, NY, USA
Christian Sander, Dear Reality, Germany

In this roundtable discussion we invite experts in the field of immersive audio production for music, cinema, and 360° film content to discuss similarities and differences regarding formats, workflows, and aesthetics. Channel and object based immersive audio systems as well as headphone virtualization using binaural rendering both are trending topics in audio production. With the rise of new applications like virtual reality, 3D audio in cinema, and home entertainment, compatibility for cross-platform distribution is more and more requested, yet engineers often are only familiar with either playback on loudspeaker arrays or binaural rendering. In this event we will try to settle a common ground for all production types and look at the different requirements, both from a technical and aesthetic point of view.

Student Event/Career Development
SC12 STUDENT RECORDING CRITIQUES
Saturday, May 26, 11:00 – 12:00 Galleria (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 of their studies can sign up to partici-pate. Students should sign up at the student (SDA) booth immediately on arrival at the convention, and deliver stereo 44.1 Khz, 24 bit AIFF or WAVE files to the SDA booth at that time. 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 feed-back as part of the competition process). These events are generously supported by Genelec.

Professional Sound Expo PSE17
Saturday, May 26, 11:00 – 11:45 PSE Stage

SETTING UP A CONTROL ROOM WITH GIK ACOUSTICS

Presenter: **Lukas Rimbach**, GIK Acoustics

Lukas Rimbach discusses the basics of room acoustics, how to set up and treat a control room - and the new Room Acoustics Visualizer from GiK, an augmented reality 3D rendering app for smartphones.

Workshop 31 Saturday, May 26
11:15 – 12:15 Lobby

INTRODUCING THE IEM PLUG-IN SUITE

Presenter: **Daniel Rudrich**, Institute of Electronic Music and Acoustics Graz, Graz, Austria

The "IEM Plug-in Suite" is a free and open-source audio plug-in suite created by staff and students at IEM. It features 3D higher-order Ambisonic plug-ins up to 7th order including encoders, decoders, surround visualizer, dynamic processors, and delay and reverb effects, which also allow to auralize directional sources. Moreover, the plug-in suite includes a tool to design Ambisonic decoders for arbitrary loudspeaker layouts employing the AllRAD and imaginary-loudspeaker downmix approach. All implementation is research-driven, and the suite is meant to continuously grow by recent developments, e.g., improvements in binaural Ambisonic rendering. The workshop presents the plug-ins and practically demonstrates the basic applications and workflow in Reaper.

Professional Sound Expo PSE18
Saturday, May 26, 12:00 – 12:45 PSE Stage

THE NETWORKED STUDIO, A DREAM OR REALITY?

Presenter: **Jan Lykke**, NTP

The presentation will discuss the use of networked audio technologies such as Dante and AES67 in music and audio post-production studios. It will look at pros and cons of networked technology, what benefits and challenges you may run into, and illustrated with real-life examples. The topic of latency will also be covered.

Student Event/Career Development
SC13 STUDENT DELEGATE ASSEMBLY MEETING—PART 2
Saturday, May 26, 12:15 – 14:00
Arena 3 & 4

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..

Tutorial 29 **Saturday, May 26**
12:30 – 13:15 **Scala 1**

ICOSAHEDRAL LOUDSPEAKER LISTENING SESSION (IKO)

Presenters: **Frank Schultz**, University of Music and Performing Arts Graz, Graz, Austria
Franz Zotter, IEM, University of Music and Performing Arts, Graz, Austria

In this listening session we show examples of how to employ focused sound beams, played by the powerful 20-channel spherical loudspeaker array IKO. Just as with sound bars but more powerfully, its beams enables utilizing the available wall reflections as passive surround system, but there is more to it. In our artistic research project OSIL, we explored artistic and psychoacoustical subtleties of the IKO. You are going to hear that sound positioning by beamforming in a room is not restricted to distinct wall reflections but supports a whole range of intermediate and room traversing positions. Artistically, the control of beam of the IKO allows to compose sculptural effects. We will explain and listen to some of the basic effects, before listening to advanced musical compositions.

Workshop 32 **Saturday, May 26**
12:30 – 13:15 **Lobby**

COMPARISON OF STEREOPHONIC, ASSEMBLED STEREO AND FIRST-ORDER AMBISONIC RECORDINGS FOR VR AND 360° CONTENT

Presenters: **Hannes Dieterle**, SCHOEPS Mikrofone GmbH, Karlsruhe, Germany
Kacper Sagnowski, SCHOEPS Mikrofone GmbH, Karlsruhe, Germany

For the research concerning multichannel microphone techniques for VR or 360° videos, several examples have been recorded using either stereophonic approaches (3D microphone arrays, assembled stereophonic ambiences) or first-order Ambisonic recordings. The differences in sound that those techniques produce will be shown and discussed in this listening session.

Audio Applications Forum **Saturday, May 24**
12:30 – 13:30 **Arena 1**

SINUSOID OR TRANSIENT? NUMBERS OR MUSIC? ALL THE QUESTIONS THAT LEAD TO ARTHUR

Presenter: **Stephan Schertler**, Schertler SA, Mendrisio, Switzerland

• Sinusoidal and transient signals
• How an amplifier reacts
• How the NFB behaves
• How we can interpret non-linear distortions
• What the specifications of a device tell us
• Electrodynamic contact microphones: mechanical and electrical operation.

Session P25 **Saturday, May 26**
13:00 – 15:00 **Scala 4**

AUDIO APPLICATIONS

Chair: **Annika Neidhardt**, Technical University of Ilmenau, Ilmenau, Germany

13:00

P25-1 Films Unseen: Approaching Audio Description Alternatives to Enhance Perception, Immersion, and Imagery of Audio-Visual Mediums for Blind and Partially Sighted Audiences: Science Fiction—*Cesar Portillo*, SAE Institute London, London, UK

"Films Unseen" is a research conducted to analyze the nature of audio description and the soundtrack features of Science Fiction films/content. The paper explores the distinctive immersive, sound spatialization and sound design features that could allow blind and partially sighted audiences to perceive and conduct an accurate interpretation of the optical elements, presented within visually complex audio visual mediums, such as the film used as a case study called *How to Be Human* (Centofanti, 2017). Correspondingly, the results collected from 15 experienced audio description users demonstrated the efficiency of SFX, immersive audio and binaural recording techniques to stimulate the perception of visual performances by appealing to auditory senses, evoking a more meaningful and understandable experience to visually impaired audiences in correlation with experimental approaches to sound design and audio description.
Convention Paper 10024

13:30

P25-2 "It's about Time!" A Study on the Perceived Effects of Manipulating Time in Overdub Recordings—*Tore Teigland, Pal Erik Jensen, Claus Sohn Andersen*, Westerdals Oslo School of Arts, Oslo, Norway

In this study we made three separate recordings using both close, near, and room microphones. These recordings were then the subject for a listening test constructed to study a variety of perceived effects due to manipulating time in overdub recordings. While the use of time alignment to decrease comb filtering has been widely studied, there has been little work on investigating other perceived effects. Time alignment has become more and more common, but as this paper concludes, it should not be used without concern. The findings will shed light on a range of important factors affected by manipulating time between microphones in overdub recordings, while also concluding on which of, and when, the investigated techniques are normally preferred or not.
Convention Paper 10025

14:00

P25-3 Musicians' Binaural Headphone Monitoring for Studio Recording—*Valentin Bauer,*[1] *Hervé Déjardin,*[2] *Amandine Pras*[1,3]
[2] Radio France, Paris, France
[3] University of Lethbridge, Lethbridge, Alberta, Canada

This study uses binaural technology for headphone monitoring in world music, jazz, and free improvisation recording sessions. We first conducted an online survey with 12 musicians to identify the challenges they face when performing in studio with wearable monitoring devices. Then, to investigate musicians' perceived differences between binaural and stereo monitoring, we carried out three comparative tests followed by semi-directed focus groups. The survey analysis highlighted the main challenges of coping with an unusual performance situation and a lack of realism and sound quality of the auditory scene. Tests showed

that binaural monitoring improved the perceived sound quality and realism, musicians' comfort and pleasure, and encouraged better musical performances and more creativity in the studio.
Convention Paper 10026

14:30

P25-4 Estimation of Object-Based Reverberation Using an Ad-Hoc Microphone Arrangement for Live Performance—*Luca Remaggi,[1] Philip Jackson,[1] Philip Coleman,[1] Tom Parnell[2]*
[1] University of Surrey, Guildford, Surrey, UK
[2] BBC Research & Development, Salford, UK

We present a novel pipeline to estimate reverberant spatial audio object (RSAO) parameters given room impulse responses (RIRs) recorded by ad-hoc microphone arrangements. The proposed pipeline performs three tasks: direct-to-reverberant-ratio (DRR) estimation; microphone localization; RSAO parametrization. RIRs recorded at Bridgewater Hall by microphones arranged for a BBC Philharmonic Orchestra performance were parametrized. Objective measures of the rendered RSAO reverberation characteristics were evaluated and compared with reverberation recorded by a Soundfield microphone. Alongside informal listening tests, the results confirmed that the rendered RSAO gave a plausible reproduction of the hall, comparable to the measured response. The objectification of the reverb from in-situ RIR measurements unlocks customization and personalization of the experience for different audio systems, user preferences, and playback environments.
Convention Paper 10028

Professional Sound Expo — PSE19
Saturday, May 26, 13:00 – 13:45 — PSE Stage

DIGITAL CLOCKING: WHAT IS CLOCKING AND WHY IS IT IMPORTANT?

Presenter: **Marcel James**, Antelope Audio

Marcel James will be lecturing on clocking and why is it important for digital audio.

Tutorial 30 — Saturday, May 26
13:30 – 14:15 — Scala 1

SURROUND WITH DEPTH LISTENING SESSION ON IEM LOUDSPEAKER CUBES

Presenters: **Thomas Deppisch**, University of Technology, Graz, Austria; University of Music and Performing Arts Graz, Graz, Austria
Matthias Frank, University of Music and Performing Arts Graz, Graz, Austria
Nils Meyer-Kahlen, University of Technology, Graz, Austria; University of Music and Performaing Arts Graz, Graz, Austria
Franz Zotter, IEM, University of Music and Performing Arts, Graz, Austria

This listening event shows effects and music on a new quadraphonic surround-with-depth playback system, whose control, measurement, and design are described in our paper engineering brief at the convention. Our playback system consists of four loudspeaker cubes, each of which using four loudspeakers driven by a first-order beam control. We show how to involve reflections involved to create a stable sound image for a big audience, and how direct and first-order-reflected sound is used to get smooth surround panning (close sounds), while a set of second-order reflections permit panning to distant sounds. We will explain and play back the effects we programmed before listening to a mix of a surround-with-depth mixed song together.

Workshop 33 — Saturday, May 26
13:30 – 14:30 — Lobby

THE 5TH ELEMENT—HOW A SCI-FI CLASSIC SOUNDS WITH A NEW 3D AUDIO MIX

Presenter: **Tom Ammermann**, New Audio Technology GmbH, Hamburg, Germany

The 5th Element it's certainly a milestone in Sci-Fi film history. Recently it was completely overworked doing a completely new film scan in 4k and mixing the whole audio elements again in Dolby Atmos and Headphone Surround 3D. This version was released in Germany as UHD Blu-ray and offers a fantastic new adventure of this great production form Luc Besson.

Professional Sound Expo — PSE20
Saturday, May 26, 14:00 – 14:45 — PSE Stage

IMMERSIVE AUDIO WORKFLOWS

Presenter: **Dave Tyler**, AVID

Immersive audio workflows are evolving rapidly both for cinema and VR. Dave Tyler will explore the latest developments in immersive audio and explain how Pro Tools can facilitate these new audio workflows.

Professional Sound Expo — PSE21
Saturday, May 26, 15:00 — 15:45 — PSE Stage

REFERENCE MONITORING AND TRANSLATION

Presenter: **Aki Mäkivirta**, Genelec Oy, Iisalmi, Finland

Once you have overcome the circles of confusion in pro monitoring, you can create content that translates well between rooms and between in-room and headphones. Topics covered: In-room spectral calibration, level calibration, monitor dispersion, loudness-based production and monitoring, prevention of listener fatigue.