AES 132nd Convention Program

April 26 – 29, 2012

Budapest Congress and World Trade Center

At recent AES conventions, authors have had the option of submitting complete 4- to 10-page manuscripts for peer-review by subject-matter experts. The following two papers have been recognized as co-winners of the AES 132nd Convention Peer-Reviewed Paper Award.

Linear Mixing Models for Active Listening of Music Productions in Realistic Studio Conditions— *Nicolas Sturmel,[1] Antoine Liutkus,[2] Jonathan Pinel,[3] Laurent Girin,[3] Sylvain Marchand,[4] Gaël Richard,[2] Roland Badeau,[2] Laurent Daudet[1]*
[1]Université Paris Diderot, Paris, France
[2]Telecom ParisTech, Paris, France
[3]GIPSA-Lab, Grenoble INP, Grenoble, France
[4]Université de Bretagne Occidentale, Brest, France

Convention Paper 8594

To be presented on Thursday, April 26 in Session P4 —Sound Reinforcement and Studio Technologies

and

MPEG Unified Speech and Audio Coding—The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types— *Max Neuendorf,[1] Markus Multrus,[1] Nikolaus Rettelbach,[1] Guillaume Fuchs,[1] Julien Robilliard,[1] Jérémie Lecomte,[1] Stephan Wilde,[1] Stefan Bayer,[1] Sascha Disch,[1] Christian Helmrich,[1] Roch Lefebvre,[2] Philippe Gournay,[2] Bruno Bessette,[2] Jimmy Lapierre,[2] Kristofer Kjörling,[3] Heiko Purnhagen,[3] Lars Villemoes,[3] Werner Oomen,[4] Erik Schuijers,[4] Kei Kikuiri,[5] Toru Chinen,[6] Takeshi Norimatsu,[7] Chong Kok Seng,[7] Eumi Oh,[8] Miyoung Kim,[8] Schuyler Quackenbush,[9] Bernhard Grill[1]*
[1]Fraunhofer IIS, Erlangen, Germany
[2]Université de Sherbrooke, Sherbrooke, Quebec, Canada
[3]Dolby Sweden AB, Stockholm, Sweden
[4]Philips Research Laboratories, Eindhoven, The Netherlands
[5]NTT DOCOMO, INC., Yokosuka, Kanagawa, Japan
[6]Sony Corporation, Shinagawa, Tokyo, Japan
[7]Panasonic Corporation
[8]Samsung Electronics, Suwon, Korea
[9]Audio Research Labs, Scotch Plains, NJ, USA

Convention Paper 8654

To be presented on Saturday, April 28 in Session P16 —High Resolution and Low Bit Rate

* * * *

The AES has launched a new 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 132nd 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 132nd AES Convention Student Paper Award is:

Robustness of a Mixed-Order Ambisonics Microphone Array for Sound Field Reproduction — *Marton Marschall,[1] Sylvain Favrot,[1] Jörg Buchholz[2]*
[1]Technical University of Denmark, Lyngby, Denmark
[2]National Acoustic Laboratories, Chatswood, Australia
Convention Paper 8645

To be presented on Saturday, April 28 in Session P15 —Spatial Audio: Part 1

* * * * *

Thurs., April 26 09:00 Bartók Meeting Room

Technical Committee Meeting on Acoustics and Sound Reinforcement

Student/Career Development Event
STUDENT DELEGATE ASSEMBLY MEETING PLACE
Thursday – Sunday 9:00 – 18:30
Foyer

Come visit the SDA Booth to find out about AES student events at chapters around the world. This is also where you will see postings about finalists in the recording competition and be able to purchase tickets to the Student Social.

Session P1 **Thursday, April 26**
09:30 – 11:00 **Room Lehar**

APPLICATIONS IN AUDIO

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

09:30

P1-1 Efficient Binaural Audio Rendering Using Independent Early and Diffuse Paths—
Fritz Menzer, MN Signal Processing, Schwerzenbach, Switzerland

A multi-source binaural audio rendering structure is proposed that efficiently implements plausible binaural reverberation including early reflections and late reverberation. The structure contains delay lines and a feedback delay network that operate independently, modeling early reflections and diffuse reverberation, respectively. Computationally efficient heuristics are presented for the implementation of an HRTF set and for the diffuse reverberation A real-time implementation on a mobile device is presented.
Convention Paper 8584

10:00

P1-2 The Hand Clap as an Impulse Source for Measuring Room Acoustics—*Prem Seetharaman, Stephen P. Tarzia*, Northwestern University, Evanston, IL, USA

We test the suitability of hand clap recordings for measuring several acoustic features of musical performance and recording rooms. Our goal is to make acoustic measurement possible for amateur musicians and hobbyists through the use of a smartphone or web app. Hand claps are an attractive acoustic stimulus because they can be produced easily and without special equipment. Hand claps lack the high energy and consistency of other impulse sources, such as pistol shots, but we introduce some signal processing steps that mitigate these problems to produce reliable acoustical measurements. Our signal processing tool chain is fully-automated, which allows both amateurs and technicians to perform measurements in just a few seconds. Using our technique, measuring a room's reverberation times and frequency response is as easy as starting a smartphone app and clapping several times.
Convention Paper 8585

10:30

P1-3 Subjective Sound Quality Evaluation of a Codec for Digital Wireless Transmission—
Matthias Frank,[1] *Alois Sontacchi*,[1] *Thomas Lindenbauer,*[2] *Martin Opitz*[2]
[1]University of Music and Performing Arts Graz, Graz, Austria
[2]AKG Acoustics GmbH, Vienna, Austria

This paper presents a subjective evaluation of a proprietary sub-band ADPCM (Adaptive Differential Pulse Code Modulation) codec for digital wireless transmission. The evaluation is carried out with 40 expert listeners and is divided into several experimental stages. First, the audibility threshold for codec artifacts is determined for each frequency sub-band, separately. In the next stage, different configurations are ranked on a scale of subjective sound quality ratings, with the resolutions varied across all bands. Finally, selected configurations corresponding to different quality ratings are compared to signals of analog wireless transmission in a multi-dimensional test. This test reveals the characteristic artifacts for each transmission method. Overall, digital transmission can achieve a better sound quality than analog transmission.
Convention Paper 8586

Session P2 **Thursday, April 26**
09:30 – 11:00 **Room Liszt**

EMERGING AND INNOVATIVE AUDIO

Chair: **Francis Rumsey**

09:30

P2-1 Virtual Microphones: Using Ultrasonic Sound to Receive Audio Waves—*Tobias Merkel*,[1] *Hans-G. Lühmann*,[2] *Tom Ritter*[1]
[1]Beuth Hochschule für Technik, Berlin, Germany
[2]Lütronik Elektroakustik GmbH, Berlin, Germany

A highly focused ultrasound beam was sent through the room. At a distance of several meters the ultrasonic wave was received again with an ultrasonic microphone. The wave field of a common audio source was overlaid with the ultrasonic beam. It was found that the phase shift of the received ultrasonic signal obtains the audio information of the overlaid field. Since the ultrasonic beam itself acts as sound receiver, there is no technical device like membranes necessary at direct vicinity of sound reception. Because this kind of sound receiver is not visible or touchable we call it "Virtual Microphone."
Convention Paper 8587

10:00

P2-2 Implementation and Evaluation of Autonomous Multi-Track Fader Control—
Stuart Mansbridge, Saoirse Finn, Joshua D. Reiss, Queen Mary, University of London, London, UK

A new approach to the autonomous control of faders for multitrack audio mixing is presented. The algorithm is designed to generate an automatic sound mix from an arbitrary number of monaural or stereo audio tracks of any sample rate and to be suitable for both live and postproduction use. Mixing levels are determined by the use of the EBU R-128 loudness measure, with a cross-adaptive process to bring each track to a time-varying average. An hysteresis loudness gate and selective smoothing prevents the adjustment of intentional dynamics in the music. Real-time and off-line software implementations have been created. Subjective evaluation is provided in the form of listening tests, where the

method is compared against the results of a human mix and a previous automatic fader implementation.
Convention Paper 8588

10:30

P2-3 A Voice Classification System for Younger Children with Applications to Content Navigation—*Christopher Lowis, Christopher Pike, Yves Raimond*, BBC R&D, London, UK

A speech classification system is proposed that has applications for accessibility of content for younger children. To allow a young child to access online content (where typical interfaces such as search engines or hierarchical navigation would be inappropriate) we propose a voice classification system trained to recognize a range of sounds and vocabulary typical of younger children. As an example we designed a system for classifying animal noises. Acoustic features are extracted from a corpus of animal noises made by a class of young children. A Support Vector Machine is trained to classify the sounds into 1 of 12 corresponding animals. We investigated the precision and recall of the classifier for various classification parameters. We investigated an appropriate choice of features to extract from the audio and compared the performance when using mean Mel-frequency Cepstral Coefficients (MFCC), a single-Gaussian model fitted to the MFCCs, as well as a range of temporal features. To investigate the real-world applicability of the system we paid particular attention to the difference between training a generic classifier from a collected corpus of examples and one trained to a particular voice.
Convention Paper 8589

Thurs., April 26 10:00 Bartók Meeting Room

Technical Committee Meeting on Semantic Audio Analysis

Workshop 1 Thursday, April 26
10:30 – 12:30 Room Bartók

CANCELLED

Session P3 Thursday, April 26
11:00 – 12:30 Room Lehar

MUSIC AND MODELING

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

11:00

P3-1 Physical Model of the Slide Guitar: An Approach Based on Contact Forces—*Gianpaolo Evangelista*, Linköping University, Campus Norrköping, Sweden

In this paper we approach the synthesis of the slide guitar, which is a particular play mode of the guitar where continuous tuning of the tones is achieved by sliding a metal or glass piece, the bottleneck, along the strings on the guitar neck side. The bottleneck constitutes a unilateral constraint for the string vibration. Dynamics is subject to friction, scraping, textured displacement, and collisions. The presented model is physically inspired and is based on a dynamic model of friction, together with a geometrical model of the textured displacements and a model for collisions of the string with the bottleneck. These models are suitable for implementation in a digital waveguide computational scheme for the 3-D vibration of the string, where continuous pitch bending is achieved by all-pass filters to approximate fractional delays. Friction is captured by nonlinear state-space systems in the slide junction and textured displacements by signal injection at a variable point in the waveguide.
Convention Paper 8590

11:30

P3-2 Measuring Spectral Directivity of an Electric Guitar Amplifier—*Agnieszka Roginska*,[1] *Alex U. Case*,[2] *Andrew Madden*,[1] *Jim Anderson*[1]
[1]New York University, New York, NY, USA
[2]University of Massachusetts Lowell, Lowell, MA, USA

The recorded timbre of an electric guitar amplifier is highly dependent on the position of the microphone. Small changes in the location of the microphone can yield significant spectral differences, particularly at positions very close to the amp. This paper presents densely measured radiation pattern characteristics of an electric guitar amplifier on a 3-D grid in front, beside, behind, and above the amplifier in a hemi-anechoic space. We use this data to analyze the change in spectral differences between the numerous points on the measurement grid. Differences between acoustically measured and estimated frequency responses (predicted, using interpolation) are used to study the change in the acoustic field in order to gain insight and an understanding of the spectral directivity sensitivity factor of the electric guitar amplifier.
Convention Paper 8592

12:00

P3-3 Magnitude-Priority Filter Design for Audio Applications—*Balázs Bank*, Budapest University of Technology and Economics, Budapest, Hungary

In audio, often specialized filter design methods are used that take into account the logarithmic frequency resolution of hearing. A notable side-effect of these quasi-logarithmic frequency design methods is a high-frequency attenuation for non-minimum-phase targets due to the frequency-dependent windowing effect of the filter design. This paper presents two approaches for the correction of this high-frequency attenuation, based either on the iterative update of the magnitude or the iterative update of the phase of the target specification. As a result, the filter follows both magnitude and phase in those frequency regions where it can, while where this is not possible, it focuses on the magnitude. Thus, the new method combines the advantages of traditional complex and magnitude-only filter designs. The algorithms are demonstrated by parallel filter designs, but since the method does not make ➥

any assumption on the filter design algorithm used in the iteration, it is equally applicable to other techniques.
Convention Paper 8591

Session P4 Thursday, April 26
11:00 – 12:30 Room Liszt

SOUND REINFORCEMENT AND STUDIO TECHNOLOGIES

Chair: **Diemer de Vries**, Amsterdam, The Netherlands

11:00

P4-1 Full Room Equalization at Low Frequencies with Asymmetric Loudspeaker Arrangements—*Balázs Bank*, Budapest University of Technology and Economics, Budapest, Hungary

For rectangular rooms with symmetric loudspeaker arrangements, full room equalization can be achieved at low frequencies, as demonstrated by previous research. The method is based on generating a plane wave that propagates along the room. However, often the room is not rectangular, and/or a symmetric loudspeaker setup cannot be assured, leading to a deteriorated equalization performance. In addition, the performance of the method drops significantly above a cutoff frequency where a plane wave cannot be generated. These problems are addressed by the proposed method by prescribing only the magnitude in the control points, while the phase is determined by an iterative optimization process starting from the plane wave solution. A true "magnitude-only" variant of the method is also presented. Comparison is given to the plane-wave based methods by introducing asymmetries to the loudspeaker setup in a simulated environment, showing that the new methods result in smaller average magnitude deviations compared to the previous plane-wave based approach.
Convention Paper 8593

11:30

P4-2 Linear Mixing Models for Active Listening of Music Productions in Realistic Studio Conditions—*Nicolas Sturmel*,[1] *Antoine Liutkus*,[2] *Jonathan Pinel,*[3] *Laurent Girin*,[3] *Sylvain Marchand*,[4] *Gaël Richard,*[2] *Roland Badeau*,[2] *Laurent Daudet*[1]
[1]Université Paris Diderot, Paris, France
[2]Telecom ParisTech, Paris, France
[3]GIPSA-Lab, Grenoble INP, Grenoble, France
[4]Université de Bretagne Occidentale, Brest, France

The mixing/demixing of audio signals as addressed in the signal processing literature (the "source separation" problem) and the music production in studio remain quite separated worlds. Scientific audio scene analysis rather focuses on "natural" mixtures and most often uses linear (convolutive) models of point sources placed in the same acoustic space. In contrast, the sound engineer can mix musical signals of very different nature and belonging to different acoustic spaces, and exploits many audio effects including nonlinear processes. In the present paper we discuss these differences within the strongly emerging framework of active music listening, which is precisely at the crossroads of these two worlds: it consists in giving to the listener the ability to manipulate the different musical sources while listening to a musical piece. We propose a model that allows the description of a general studio mixing process as a linear stationary process of "generalized source image signals" considered as individual tracks. Such a model can be used to allow the recovery of the isolated tracks while preserving the professional sound quality of the mixture. A simple addition of these recovered tracks enables the end-user to recover the full-quality stereo mix, while these tracks can also be used for, e.g., basic remix / karaoke/soloing and re-orchestration applications.
Convention Paper 8594

12:00

P4-3 Capturing Height: The Addition of Z Microphones to Stereo and Surround Microphone Arrays—*Paul Geluso*, New York University, New York, NY, USA

As surround systems with height channels become more commonplace, new microphone techniques to capture sound in 3-D are needed. In order for the height channels to be effective, they must contain sonic information that is compatible with the 5.1 surround channels in order to improve the listener's sense of three-dimensional sound imaging, space, and immersion. Complex height information can be captured by pairing horizontally oriented microphones with vertically oriented bi-directional microphones. In this paper the author presents a rationale, methodology, and preliminary evaluation of a microphone technique based on this concept.
Convention Paper 8595

Workshop 2 Thursday, April 26
11:00 – 12:30 Room Brahms

GAME AUDIO

Chair: **Michael Kelly**, DTS, London, UK

Panelists: *Zoltan Nyakacska*, Nemesys Games
Zach Zebrowski, Nemesys Games

This workshop starts with a general introduction to the games industry and goes on to discuss the specific implementation of audio in games.

The workshop, presented by Nemesys Games, looks at the pros and cons of selecting different audio engines for game development in the context of their current title, *Ignite*. The talk shows how the whole audio system for the game is built and discusses the challenging task of implementing car engine sounds; outlining the technical and creative problems and how they are solved.

The workshop concludes by comparing the nonlinear nature of game-sound development to the production of a more linear piece such as a short movie or trailer.

This event is aimed at a general audience who may not be familiar with game development but presents an appropriate level of detail for those who want to dig deeper.

Thurs., April 26 11:00 Bartók Meeting Room

Technical Committee Meeting on Audio Recording and Mastering Systems

Thurs., April 26 12:00 Bartók Meeting Room

Technical Committee Meeting on Automotive Audio

Special Event
AWARDS PRESENTATION
AND KEYNOTE ADDRESS
Thursday, April 26, 13:00 – 14:00
Room Bartók

Opening Remarks:
- Executive Director Bob Moses
- President Jan A. Pedersen
- Convention Chair Janos Gyori

Program:
- AES Awards Presentation
- Introduction of Keynote Speaker by Convention Chair Janos Gyori
- Keynote Address by John Buckman

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.

Keynote Speaker
This year's Keynote Speaker is **John Buckman**, founder of Berkeley, CA-based Magnatune.com (subscription music for consumers), a record label known for its eclectic artist roster, its successful application of Creative Commons licensing, and its patently artist-friendly business practices. Buckman, whose egalitarian business practices are distinguished by non-exclusive agreements with musicians, sharing profits equally with them, and allowing them to retain full rights to their own music, has pioneered a philosophy known as "fair trade music." Since founding Magnatune, Buckman has signed more than 250 recording artists across multiple genres including classical, electronica, world, alt rock, jazz, and hard rock. Buckman is also the founder/owner of a string of aditional successful small businesses working with music: MoodMixes.com (background music for restaurants), ToneGnome.com (audio engineering services over the Internet), and is about to launch iLicenseMusic.com. He also runs BookMooch.com (over a million books swapped per year), founded and sold Lyris Inc. (email newsletter software), and is chairman of the Electronic Frontier Foundation, a US-based Internet civil liberties organization. Profiled by *Inc. Magazine, The Economist*, and other major publications, Buckman is the co-author of an article in *SysAdmin Magazine* entitled "Which OS Is Fastest for High-Performance Network Applications?" and an article in *Linux Journal* entitled "Magnatune, an Open Music Experiment." The title of his keynote address is "Small and Beautiful: Models for Successful Independent Music Businesses."

Buckman will explain how to find profitable niches and how to build a small business where the work can be done by you and your staff—regardless of how many hands you have.

Thurs., April 26 09:00 Bartók Meeting Room

Technical Committee Meeting on Acoustics and Sound Reinforcement

Thurs., April 26 13:00 Room Fortuna

Standards Committee Meeting on Digital Input/Output Interfacing

Thurs., April 26 14:00 Room Fortuna

Standards Committee Meeting on Acoustics and Sound Source Modeling

Session P5 **Thursday, April 26**
14:30 – 18:30 **Room Lehar**

RECORDING AND PRODUCTION

Chair: **Balazs Bank**

14:30

P5-1 Automated Horizontal Orchestration Based on Multichannel Musical Recordings—*Maximos Kaliakatsos-Papakostas*,[1] *Andreas Floros*,[2] *Michael N. Vrahatis*[1]
[1]University of Patras, Patras, Greece
[2]Ionian University, Corfu, Greece

Orchestration of computer-aided music composition aims to approximate musical expression using vertical instrument sound combinations, i.e., through finding appropriate sets of instruments to replicate synthesized sound samples. In this paper we focus on horizontal orchestration replication, i.e., the potential of replicating the instantaneous intensity variation of a number of instruments that comprise an existing, target music recording. A method that efficiently performs horizontal orchestration replication is provided, based on the calculation of the instrumental Intensity Variation Curves. It is shown that this approach achieves perceptually accurate automated orchestration replication when combined with automated music generation algorithms.
Convention Paper 8596

15:00

P5-2 The Effect of Scattering on Sound Field Control with a Circular Double-Layer Array of Loudspeakers —*Ji-ho Chang, Finn Jacobsen*, Technical University of Denmark, Lyngby, Denmark

A recent study has shown that a circular double-layer array of loudspeakers makes it possible to achieve a sound field control that can generate a controlled field inside the array and reduce sound waves propagating outside the array. This is useful if it is desirable not to disturb people outside the array or to prevent the effect of reflections from the room. The study assumed free field condition, however in practice a listener will be located inside the array. The listener scatters sound waves, which propagate outward. Consequently, the scattering effect can be expected to degrade the performance of the system. This paper computationally examines the scattering effect based on the simple assumption that the listener's head is a rigid sphere. In addition, methods to solve the problem are discussed.
Convention Paper 8597

15:30

P5-3 The Equidome, a Personal Spatial Reproduction Array—*James L. Barbour*, Swinburne University of Technology, Melbourne, Australia

Research into spatial sound requires the use of a loudspeaker array capable of reproducing sounds around and above a listening position. Different arrays have been developed that are often limited in their accuracy due to room size, the number of speakers employed, and their positions. They are also expensive to develop, which restricts access for many researchers. A unique array, The Equidome, has been developed with speakers around and above the listening position and it uses audio software capable of locating a sound anywhere over the array, and moving a sound throughout the array. The Equidome is a personal research facility capable of 3-D spatial composition with mono, stereo, multichannel, or ambisonic source files, and reverberation software is also utilized to create immersive sound fields throughout the array.
Convention Paper 8598
[Paper was not presented but is available for purchase]

16:00

P5-4 A Clipping Detector for Layout-Independent Multichannel Audio Production—*Giulio Cengarle, Toni Mateos*, Fundació Barcelona Media, Barcelona, Spain

In layout-independent audio production, content is produced independently from the number of channels and their location, so that it can be played-back in different multichannel setups. In such contexts, sound is monitored through a playback system that might differ from the potentially many exhibition layouts. Signals combine to the outputs of each playback system in a different way and may produce clipping in some loudspeakers. A method is presented for detecting and quantitatively estimating clipping in the output stage of such systems, based on a suitable definition of a worst-case loudspeaker layout, and associated audio scene rotation and decoding. Practical examples are provided to validate the algorithm.
Convention Paper 8599

16:30

P5-5 Evaluation of Spatial Impression Comparing Surround with Height Channels for 3-D Imagery—*Toru Kamekawa,*[1] *Atsushi Marui,*[1] *Toshihiko Date,*[2] *Aiko Kawanaka,*[2] *Masaaki Enatsu*[3]
[1]Tokyo University of the Arts, Tokyo, Japan
[2]AVC Networks Company, Panasonic Corporation, Osaka, Japan
[3]marimoRECORDS Inc., Tokyo, Japan

Three-dimensional (3-D) imagery is now widely spreading as one of the next visual formats for Blu-ray or other future media. Since more audio channels are available with future media, the authors aim to find the suitable sound format for 3-D imagery. Semantic Differential method using 24 attributes such as "presence," "naturalness," and "preference" was carried out comparing combinations of 3-D and 2-D imagery with 2-channel stereo, 5-channel surround, and 7-channel surround sound (5-channel surround plus 2 height channels). Three factors (spatial factor, preference factor, and quality factor) were extracted from the results of Factor Analysis. Combination of the 3-D imagery with 7-channel surround gives higher scores at all those three factors.
Convention Paper 8600

17:00

P5-6 Microphone Array Design for Localization with Elevation Cues—*Michael Williams*, Sounds of Scotland, Le Perreux sur Marne, France

Analysis of the HRTF characteristics with respect to both azimuth and elevation localization cues would seem to suggest that, while inter-aural time difference and inter-aural level difference information give strong azimuth localization cues, we only have spectral variations in the vertical plane to generate localization cues with respect to elevation. Of course positioning of a second layer of loudspeakers above the horizontal reference plane will already introduce listener spectral differences relative to the horizontal plane reproduction, but the microphone array that feeds this second layer must not generate information that will be in conflict with the localization cues already generated in the horizontal plane of the first layer or main array. In the event of time difference and level difference information being generated between microphones in both layers, localization characteristics must be considered as projected onto the main horizontal plane of localization information.
Convention Paper 8601

17:30

P5-7 General Integral Equation for the Width-Height Reproduction of a Focused Source Inside—*Jung-Woo Choi, Yang-Hann Kim*, Korea Advanced Institute of Science and Technology, Deajeon, Korea

A general integral formula to reproduce sound field from a virtual sound source located inside secondary sources is proposed. The proposed formula extends the theory on the focused sound source of Wave Field Synthesis for the reproduction of three-dimensional sound field. To resolve the non-existence problem involved with the reproduction of a source inside, an alternative sound field satisfying homogeneous wave equation is derived. The Kirchhoff-Helmholtz integral is reformulated in such a way that the alternative field is reproduced in terms of the secondary sources distributed on a surface. Then the general equation is reduced to simpler forms using various approximations, such as single layer formula reproducing the sound field only by monopole sound sources.
Convention Paper 8602

18:00

P5-8 Evaluation of a New Active Acoustics System in Performances of Five String Quartets—

Doyuen Ko, Wieslaw Woszczyk, Sonh Hui Chon, McGill University, Montreal, Quebec, Canada

An innovative electro-acoustic enhancement system, based on measured high-resolution impulse responses, was developed at the Virtual Acoustics Technology (VAT) lab of McGill University and was permanently installed in MultiMedia Room, a large rectangular scoring stage. Standard acoustic measures confirmed that the system was able to effectively improve room acoustic conditions in both spectral and spatial parameters. Subjective evaluation was conducted with twenty professional musicians from five string quartets on three different acoustic conditions. Spatial impression, stage support, and tonal quality were found to be three most dominant perceptual dimensions, while “naturalism of reverberation” was the most salient attribute affecting musicians’ preferences. Results showed a strong preference for enhanced acoustics over natural acoustics of the space.
Convention Paper 8603

Session P6 **Thursday, April 26**
14:30 – 17:00 **Room Liszt**

MULTIMODAL APPS AND BROADCAST

Chair: **Bozena Kostek**, Gdansk University of Technology, Gdansk, Poland

14:30

P6-1 Immersive Audiovisual Environment with 3-D Audio Playback—*Javier Gómez Bolaños, Ville Pulkki*, Aalto University, Espoo, Finland

The design of an immersive audiovisual environment for researching the aspects of the perception of spatial sound in the presence of a surrounding moving visual image is presented. The system consists in a visual screen with wide field-of-view based in acoustically transparent screens that span 226° in the horizontal plane and 57° in the vertical plane. In addition, a 3-D multichannel sound reproduction system with 29 active loudspeakers is installed. The total system is optimized for audio playback, and measured acoustical system responses are presented. The system is equipped with a tracking system based on infrared cameras, which enable head-tracking for head phone listening and also interaction based on gestures. This audiovisual system aims to be a tool for researching spatial audio, crossmodal interaction and psychoacoustics, auralization, and gaming.
Convention Paper 8604

15:00

P6-2 Sensitive Audio Data Encryption for Multimodal Surveillance Systems—*Janusz Cichowski, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

Novel algorithms for data processing in audiovisual surveillance systems were developed allowing for a better personal data protection. The solution merging the image and audio encryption for privacy-sensitive data protection employing audio stream is described. The main objectives of this research study including motivation and the state of the art are presented with a comprehensive explanation of audio stream relation to the surveillance. The invertible encryption methodology for privacy preserving using audio container is applied. The experiments are described and obtained results are reported including prospects for future improvements.
Convention Paper 8605

15:30

P6-3 Loudness Normalization of Wide-Dynamic Range Broadcast Material—*Scott Norcross, Michel Lavoie*, Communications Research Centre, Ottawa, Ontario, Canada

Various techniques are being used by broadcasters to normalize the loudness levels of their programs. For long-form content, EBU R128 recommends that the full program be measured using the algorithm described in ITU-R BS.1770-2. ATSC A/85 recommends instead that only the “Anchor Element” of long-form content need be measured. For narrow dynamic range material, the differences between the two measures are not large, but there can be large differences between the two approaches when the material has a wide dynamic range. This paper compares these two measurement approaches and explores their subjective consequences.
Convention Paper 8606

16:00

P6-4 Creating Mood Dictionary Associated with Music—*Magdalena Plewa, Bozena Kostek*, Gdansk University of Technology, Gdansk, Poland

The paper presents an attempt to create a dictionary of words related to mood associated with music. Two parts of a listening test were designed and carried out with a group of students, most of them users of social music online services. The audience task was to propose adjectives well-describing music tracks. These words were given in Polish and then compared to their English equivalents. The obtained results show that terms associated with music are language-specific and in addition there is a need to use multi-label mood description.
Convention Paper 8607

16:30

P6-5 Redundancy Optimization for Networked Audio Systems—*Damian Kowalski, Piotr Z. Kozlowski*, Wroclaw University of Technology, Wroclaw, Lower Silesia, Poland

Networked audio systems can be simply defined as a connection of IT and professional audio. Nowadays, we can use protocols developed by IT specialists to ensure system recovery without human intervention. There is a possibility to improve the recovery time of the system after failure by optimizing the protocols responsible for network redundancy. The paper is a summary of research completed at Wroclaw University of Technology on May 2011. It contains guidelines ➥

on how to optimize the network redundancy in order to achieve the best results.
Convention Paper 8608
[Paper presented by Piotr Kozlowski]

Workshop 3 **Thursday, April 26**
14:30 – 16:30 **Room Brahms**

AUDIO HARDWARE IN SMARTPHONES

Chair: **Antti Kelloniemi**, Nokia Corporation, Espoo, Finland

Panelists: *Juha Backman*, Nokia Corporation, Espoo, Finland
Mika Hanski, GoerTeck, Finalnd
Jörg Rehder, Knowles Electronics, Denmark
Friedrich Reining, Knowles Electronics GmbH, Vienna, Austria
Anders Weckström, GoerTek, Finland
Erick Wiederholtz, Knowles Electronics GmbH, Vienna, Austria

Comparison between novel smartphones and any other professional or consumer audio equipment reveals that phones actually provide highly sophisticated audio functions for their size and price. High quality noise reduction and echo cancellation for telephony and multichannel audio capture with wide frequency response and high dynamic range are expected, as well as loud and clear sound output for alarm tones and multimedia playback. The quality expectations keep rising, while devices should be kept small and affordable.

To start this workshop, the audio solutions in a smartphone are presented on an introductory level. After that, the panel, who are representatives of major audio component manufacturers, discuss the current audio performance requirements and present examples of state-of-the-art in audio component technology.

Student/Career Development Event
OPENING AND STUDENT DELEGATE ASSEMBLY MEETING – PART 1
Thursday, April 26, 14:30 – 16:00
Room Bartók

Chair: **Magdalena Plewa**

Vice Chair: **Philip Waldenberger**

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 European and International Regions, announce the finalists in the four new recording competition categories, and announce any upcoming 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.

All students and educators are invited to participate in this meeting. Election results and Recording Competition Awards will be given at the Student Delegate Assembly Meeting–Part 2 on Sunday, April 29.

Thurs., April 26 **15:00** **Bartók Meeting Room**

Technical Committee Meeting on Coding of Audio Signals

Thurs., April 26 **16:00** **Bartók Meeting Room**

Technical Committee Meeting on Loudspeakers and Headphones

Tutorial 1 **Thursday, April 26**
16:30 – 18:30 **Room Bartók**

MIX KITCHEN: MIXING TECHNIQUES OF ELECTRONIC MUSIC

Presenter: **Marek Walaszek**, Addicted to Music Media Group, Gorchow, Poland

This tutorial will be based on already-done mixes that will show techniques of modern mixing in electronic and semi-electronic modern music. Marek Walaszek will show a few mixes and talk through the approach to the mixes, effects used, and mix philosophy. He will talk about main production errors, how to prevent them, and how to cure them in the mixing stage.

This tutorial will be supported by a power point presentation and an extensive Q&A.

Session P7 **Thursday, April 26**
17:00 – 18:00 **Room Liszt**

PERCEPTION

Chair: **Bozena Kostek**, Gdansk University of Technology, Gdansk, Poland

17:00

P7-1 **Detection of Two Subwoofers: Effect of Broad-Band-Channel Level and Crossover Frequency**—*Jussi Rämö, Sakari Bergen, Julian Parker, Veli-Matti Yli-Kätkä, Ville Pulkki*, Aalto University, Espoo, Finland

The use of multiple subwoofers can be advantageous compared to a setup consisting of a single subwoofer due to the cancellation of room modes. We investigate the effect of subwoofer crossover frequency and program material on the perceived localization of bass frequencies using single or dual subwoofers, via a listening test. Test results show that dual subwoofer setups are harder to detect than single subwoofer setups and also exhibit the well-known relationship between crossover frequency and difficulty of localization.
Convention Paper 8609
[Paper presented by Ville Pulkki]

17:30

P7-2 **Pitch, Timbre, Source Separation, and the Myths of Loudspeaker Imaging**—*David Griesinger*, David Griesinger Acoustics, Cambridge, MA, USA

Standard models for both timbre detection and sound localization do not account for our acuity of localization in reverberant environments or when there are several simultaneous sound sources. They also do not account for our near instant ability to determine whether a sound is near or far. This paper presents data on how both semantic content and localization information is encoded in the harmonics of complex

tones and the method by which the brain separates this data from multiple sources and from noise and reverberation. Much of the information in these harmonics is lost when a sound field is recorded and reproduced, leading to a sound image which may be plausible but is not remotely as clear as the original sound field.
Convention Paper 8610

Workshop 4 **Thursday, April 26**
17:00 – 18:00 **Room Brahms**

THE REPLAY OF HISTORICAL MAGNETIC TAPE—MORE THAN PRESSING THE PLAY BUTTON

Chair: **Nadja Wallaszkovits**, Phonogrammarchiv, Austrian Academy of Sciences, Vienna, Austria

Panelists: *Dietrich Schüller*, Phonogrammarchiv, Austrian Academy of Sciences, Vienna, Austria
TBA

Since analog magnetic tape technology is no longer a part of the audio production process, specific knowledge is endangered to decline. This workshop discusses the various problems occurring in the transfer process of historical magnetic audio tapes. Starting with a definition of historical tape brands and an overview of early magnetic tape developments, the practical handling of critical tapes is outlined: based on an analysis of the physical and chemical preservation status of the individual tape, the choice and adjustment of the replay equipment and parameters are discussed. Problems of carrier handling and physical as well as chemical restoration are outlined, as well as possible signal enhancement on the playback process only. The workshop focuses on the handling and reproduction of original tapes from the early ages of magnetic recording, stored under irregular conditions.

Thurs., April 26 **17:00** **Bartók Meeting Room**

Technical Committee Meeting on Audio for Games

Thurs., April 26 **17:00** **Room Fortuna**

Standards Committee Meeting on Audio-File Transfer and Exchange

Special Event
MIXER PARTY
Thursday, April 26, 18:30 – 19:30
Room Bartók

A Mixer Party will be held on Thursday evening to enable Convention attendees to meet in a social atmosphere to catch up with friends and colleagues from the world of audio. There will be a cash bar.

Session P8 **Friday, April 27**
9:00 – 11:00 **Room Lehar**

LISTENING TESTS: PART 1

Chair: **Jan Plogsties**, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

09:00

P8-1 Comparison of Localization Performance of Blind and Sighted Subjects on a Virtual Audio Display and in Real-life Environments—*György Wersényi, József Répás*, Széchenyi István University, Györ, Hungary

Localization performance of blind subjects was measured in a virtual audio environment using non-individualized but customized HRTFs. Results were compared with former results of sighted users using the same measurement setup. Furthermore, orientation and navigation tasks in a real-life outdoor environment were performed in order to compare localization ability of sighted and visually impaired including "walking straight" tasks with and without acoustic feedback and test runs using the white cane as an acoustic tool during navigation.
Convention Paper 8611

09:30

P8-2 HELM: High Efficiency Loudness Model for Broadcast Content—*Alessandro Travaglini,*[1] *Andrea Alemanno,*[2] *Aurelio Uncini*[2]
[1]FOX International Channels Italy, Rome, Italy
[2]University of Rome "La Sapienza," Rome, Italy

In this paper we propose a new algorithm for measuring the loudness levels of broadcast content. It is called the High Efficiency Loudness Model (HELM) and it aims to provide robust measurement of programs of any genre, style, and format, including stereo and multichannel audio 5.1 surround sound. HELM was designed taking into account the typical conditions of the home listening environment, and it is therefore particularly good at meeting the needs of broadcast content users. While providing a very efficient assessment of typical generic programs, it also successfully approaches some issues that arise when assessing unusual content such as programs heavily based on bass frequencies, wide loudness range programs, and multichannel programs as opposed to stereo ones. This paper details the structure of HELM, including its channel-specific frequency weighting and recursive gating implementation. Finally, we present the results of a mean opinion score (MOS) subjective test that demonstrates the effectiveness of the proposed method.
Convention Paper 8612
[Paper presented by Andrea Alemanno]

10:00

P8-3 Defining the Listening Comfort Zone in Broadcasting through the Analysis of the Maximum Loudness Levels—*Alessandro Travaglini,*[1] *Andrea Alemanno,*[2] *Fabrizio Lantini*[3]
[1]FOX International Channels Italy, Rome, Italy
[2]University of Rome "La Sapienza," Rome, Italy
[3]Electric Light Studio, Rome (RM), Italy

Over the last few years, the broadcasting industry has finally approached the loudness issue by standardizing its measurement and recommending target loudness levels with which all programs are required to comply. If the recommendations are applied and all programs are normalized at the target level, viewers ought to experience consistent perceived loudness levels throughout transmissions. However, due to the

inner loudness modulation of the programs themselves, this is not always the case. In fact, even if the overall program loudness levels perfectly match the required target level, excessive loudness modulations can still generate annoyance to viewers if the foreground sound levels exceed the so-called "comfort zone." The fact is that we still have no clear data on which metering can provide visual/numeric feedback on the perception of "hearing annoyance." This paper investigates this issue and aims to provide objective evidence of which parameters would better represent this phenomenon. In particular, we describe an extensive subjective test performed for both the typical Stereo TV and the 5.1 home theater set reproductions and analyze its results in order to verify whether the Maximum Momentary Loudness Level, the Maximum Short Loudness Level, and Loudness Range (LRA) values described in EBU R128 can provide robust and reliable numeric references to generate a comfortable listening experience for viewers. Furthermore, we perform a similar analysis for the loudness descriptors of the algorithm HELM and finally indicate the values of those parameters that show the most consistent and reliable figures.
Convention Paper 8613
[Paper presented by Andrea Alemanno]

10:30

P8-4 The Relative Importance of Speech and Non-Speech Components for Preferred Listening Levels—*Ian Dash,*[1] *Miles Mossman,*[2] *Densil Cabrera*[2]
[1]Consultant, Marrickville, NSW, Australia
[2]The University of Sydney, Sydney, NSW, Australia

In a prior paper the authors reported on a listening test that attempted to establish the relative importance of speech and non-speech components of a mixed soundtrack when matching loudness to reference audio items. That paper concluded that listeners match loudness by overall content rather than by the loudness of the speech or non-speech components. This paper reports on a follow-up listening test that attempts to establish the relative importance of speech and non speech components in setting preferred listening level without any external reference. The results indicate that while speech levels are set more consistently than non-speech levels, listeners tend to set the overall levels more consistently than either of these components.
Convention Paper 8614

Session P9 — Friday, April 27
09:00 – 12:00 — Room Liszt

ANALYSIS AND SYNTHESIS OF SOUND: PART 1

Chair: **Juha Backman,** Nokia Corporation, Espoo, Finland

09:00

P9-1 Emergency Voice/Stress-Level Combined Recognition for Intelligent House Applications—*Konstantinos Drosos,*[1] *Andreas Floros,*[1] *Kyriakos Agkavanakis,*[2] *Nicolas-Alexander Tatlas,*[3] *Nikolaos-Grigorios Kanellopoulos*[1]
[1]Ionian University, Corfu, Greece
[2]BlueDev Ltd., Patras, Greece
[3]Technological Educational Institute of Piraeus, Piraeus, Greece

Legacy technologies for word recognition can benefit from emerging affective voice retrieval, potentially leading to intelligent applications for smart houses enhanced with new features. In this paper we introduce the implementation of a system, capable to react to common spoken words, taking into account the estimated vocal stress level, thus allowing the realization of a prioritized, affective aural interaction path. Upon the successful word recognition and the corresponding stress level estimation, the system triggers particular affective-prioritized actions, defined within the application scope of an intelligent home environment. Application results show that the established affective interaction path significantly improves the ambient intelligence provided by an affective vocal sensor that can be easily integrated with any sensor-based home monitoring system.
Convention Paper 8615

09:30

P9-2 Loudness Range (LRA)—Design and Evaluation—*Esben Skovenborg,*
TC Electronic A/S, Risskov, Denmark

Loudness Range (LRA) is a measure of the variation of loudness on a macroscopic time-scale. Essentially, LRA is the difference in loudness level between the soft and loud parts of a program. In 2009 the algorithm for computing LRA was published by TC Electronic and was then included in the EBU R-128 recommendation for loudness normalization. This paper describes the design choices underlying the LRA algorithm. For each of its parameters the interval of optimal values is presented, supported by analyses of audio examples. Consequences of setting parameter values outside these intervals are also described. Although the LRA measure has already proven its usefulness in practice, this paper provides background knowledge that could support further refinement and standardization of the LRA measure.
Convention Paper 8616

10:00

P9-3 Statistical Properties of the Close-Microphone Responses—*Elias K. Kokkinis, Eleftheria Georganti, John Mourjopoulos,*
University of Patras, Patras, Greece

The close-microphone technique is widely used in modern sound engineering practice. It is mainly used to minimize the effect of leakage and room acoustics on the received signal. In this paper the properties of the close-microphone response are investigated from a signal processing point of view, through the respective frequency domain statistical moments. Room impulse response measurements were made in various rooms and source-microphone distances, and statistical moments were calculated over fre-

quency and distance. It is shown that the statistical properties of the impulse responses remain relatively constant for short source-microphone distances and this in turn provides a consistent sound, which was validated through a subjective evaluation test.
Convention Paper 8617

10:30

P9-4 Reproduction of Proximity Virtual Sources Using a Line Array of Loudspeakers— *Jung-Min Lee, Jung-Woo Choi, Dong-Soo Kang, Yang-Hann Kim*, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea

For reproducing a desired sound field from virtual sources, Wave Field Synthesis (WFS) assumes that virtual sources are positioned at far-field from the loudspeaker array. This far-field assumption inevitably produces reproduction errors when the virtual source is adjacent to the array. In this paper we propose a method that can render the sound field from virtual sources positioned near the loudspeaker array. The driving functions of loudspeakers are derived for the planar array geometry, and then the surface integral is reduced to a line integral by utilizing different approximations from WFS. In addition, a modified equation for the discrete loudspeaker distribution is presented. Numerical simulations show that the proposed method can reduce the reproduction error to a practically acceptable level.
Convention Paper 8618

11:00

P9-5 On the Statistics of Binaural Room Transfer Functions—*Eleftheria Georganti*,[1] *Tobias May,*[2] *Steven van de Par*,[2] *John Mourjopoulos*[1]
[1]University of Patras, Patras, Greece
[2]University of Oldenburg, Oldenburg, Germany

The well-known property of the spectral standard deviation of Room Transfer Functions (RTFs), that is, its convergence to 5.57 dB, is extended to reverberant Binaural Room Transfer Functions (BRTFs). The BRTFs are related to the anechoic Head Related Transfer Functions (HRTFs) and the corresponding RTFs. Consequently, the statistical properties of the RTFs and HRTFs can be systematically related to the statistical properties of the BRTFs. In this paper the standard deviation of BRTFs measured in different types of rooms, for various source/receiver distances and azimuth angles is computed. The derived values are compared to the ones obtained from the single channel RTFs measured at the same positions. Their relationship to the 5.57 dB value is discussed.
Convention Paper 8619

11:30

P9-6 On Acoustic Detection of Vocal Modes— *Eddy B. Brixen*,[1] *Cathrine Sadolin,*[2] *Henrik Kjelin*[2]
[1]EBB-consult, Smorum, Denmark
[2]Complete Vocal Institute, Copenhagen, Denmark

According to the Complete Vocal Technique four vocal modes are defined: Neutral, Curbing, Overdrive, and Edge. These modes are valid for both the singing voice and the speaking voice. The modes are clearly identified both from listening and from visual laryngograph inspection of the vocal cords and the surrounding area of the vocal tract. For many reasons it would be preferred to apply a simple acoustic analysis to identify the modes. This paper looks at the characteristics of the voice modes from an acoustical perspective based on voice samples from four male and two female subjects. The paper describes frequency domain criteria for the discrimination of the various modes.
Convention Paper 8620
[Paper not presented but is available for purchase]

Workshop 5 **Friday, April 27**
09:00 – 10:30 **Room Bartók**

USE CASES FOR HIGH PERFORMANCE AUDIO-OVER-IP IN BROADCAST FACILITIES

Chair: **Stefan Ledergerber**, Lawo Group, Zurich, Switzerland

Panelists: *Axel Holzinger*, ALC Networx, Munich, Germany
Lars Jonsson, Swedish Radio, Stockholm, Sweden
Sonja Langhans, IRT, Munich, Germany
Greg Shay, Telos Systems Inc., Cleveland, OH, USA

In today's broadcasting houses the topic of audio networking is heavily discussed. In the overall discussion the audio network is very often seen as a solution for everything. But what are the real use cases today for an audio-over-IP solution within facilities? The panelists will each make a 15 minute presentation about their views, followed by an open discussion with the audience about the hot topics of the migration from established wires into network land.

Student/Career Development Event
EDUCATION FORUM
Friday, April 27, 09:00 – 10:30
Room Brahms

Moderators: **Ezequiel Morfi**, Chair, AES SDA, North and Latin American Regions
Colin Pfund, Vice Chair AES SDA, North and Latin American Regions
Magdalena Plewa, Chair SDA, Europe and International Regions
Philip Waldenberger, Vice Chair, AES SDA, Europe and International Regions

Student Roundtable

Come share your unique experience as a student of audio. Bring your thoughts and perspectives to an open discussion to be moderated by the AES Student Delegate Assembly Officers who want to encourage this dialogue. How are you learning about audio? What is unique about your program and its facilities? How do co-curricular activities like the ones sponsored by AES and other organizations contribute to your experience?

Explore strategies for making connections with the professional world and discuss the curriculums and ➡

philosophies of your programs. Students, faculty, alumni, industry professionals, and anyone interested in commenting on the state of audio education are welcome to participate.

Fri., April 27 09:00 Bartók Meeting Room

Technical Committee Meeting on Human Factors in Audio Systems

Fri., April 27 09:00 Room Fortuna

Standards Committee Meeting on Microphone Measurement and Characterization

Fri., April 27 10:00 Bartók Meeting Room

Technical Committee Meeting on Spatial Audio

Session P10 Friday, April 27
11:00 – 12:00 Room Lehar

EDUCATION

Chair: **Jan Plogsties**, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

11:00

P10-1 Audio DSP from Scratch for Students of Computing—*Ewa Lukasik*, Poznan University of Technology, Poznan, Poland

The paper presents a set of dedicated programming applications for teaching digital audio basics to the novices with a background in computing. These applications have been used in the Institute of Computing Science Poznan University of Technology for over 15 years. Each application is devoted to an individual subject, is self explanatory, illustrates the given topic, and includes several simple problems to be solved by students. The topics include: the idea of signal energy and RMS, signals orthogonality, signals approximations, simple sinusoidal synthesis, spectrum, sampling theorem, upsampling and downsampling, convolution, idea of digital filters and interpretation of frequency characteristics, FIR and IIR filters—their zeroes and poles, windowing, DFT and its time and frequency resolution, cepstral analysis, and others. This set of programs proved useful as a starting point to more advanced audio projects in the domain of MIR, speech and speaker recognition, speech synthesis, audio coding, audio transmission, and others.
Convention Paper 8621

11:30

P10-2 A Comparison of Audio Frameworks for Teaching, Research, and Development—
Martin Robinson,[1] *Jamie Bullock*[2]
[1]University of the West of England, Bristol, UK
[2]Birmingham Conservatoire, Birmingham, UK

This paper compares a range of audio frameworks for the support of teaching, research, and the development of audio applications. The authors employ a range of metrics with which to compare the frameworks including: licensing terms, portability across different architectures, audio data-type support, efficiency of processing code, expressiveness, usability, and community activity. Conclusions are drawn that none of these frameworks score highly in all of these domains. This suggests that while there are already a large number of such frameworks there remain areas to be addressed. The authors suggest that this might be through the development of existing systems or the development of new frameworks to meet these needs.
Convention Paper 8622

Workshop 6 Friday, April 27
11:00 – 13:00 Room Brahms

OPEN QUESTIONS IN SPATIAL AUDIO

Chair: **Karlheinz Brandenburg**

Panelists: *David Griesinger*
Frank Melchior
Stefan Werner

For more than 60 years spatial audio systems are in use. Just in the last 20 years the number of loudspeakers and transmission channels used have been increased, and with Walkman and MP3, player listening via headphone has become popular, too. Both scenarios provide much better audio quality than the old stereo systems, but very often strange, unexpected effects are limiting the perceived quality. This workshop will list aspects where the knowledge in spatial audio quality is not perfect. Examples are influence of room reflections, mixing of the room acoustical properties of recording and reproduction rooms, influence of room on binaural headphone reproduction (sic!), audio-visual coherence in large rooms, and issues about perception of elevated sound sources.

Tutorial 2 Friday, April 27
11:00 – 13:00 Room Bartók

EAR TRAINING FOR MASTERING ENGINEERS

Presenters: **Andrés Mayo**, Andres Mayo Mastering, Buenos Aires, Argentina
Darcy Proper, Wisseloord Studios, Hilversum, The Netherlands

This tutorial includes a comprehensive set of tools to improve ear training, focused on what the Mastering Engineer needs to do his job. Dynamic EQ, de-essing, de-woofing, and many other techniques will be shown, aiming to help audience to recognize different ranges of frequency.

Fri., April 27 11:00 Bartók Meeting Room

Technical Committee Meeting on Audio for Telecommunications

Fri., April 27 11:00 Room Fortuna

Standards Committee Meeting on Audio Connectors

Fri., April 27 12:00 Bartók Meeting Room

Technical Committee Meeting on Signal Processing

Workshop 7 **Friday, April 27**
12:15 – 13:45 **Room Liszt**

DISTRIBUTED MUSIC PANEL

Chair: **Alex Carôt**, Hochschule Anhalt, Köthen/Anhalt, Germany

Panelists: *Alvaro Barbosa*, UCP Porto, Porto, Portugal
Nathan Brock, University of California at San Diego, San Diego, CA, USA
Karl Steinberg, Digitalmusician, Hamburg
Andrea Szigetvári, Ferenc Liszt Academy of Music, Budapest, Hungary
David Willyard, Musicianlink, San Jose, CA, USA

This workshop covers the emerging field of music intended for performance over networks, including both advanced research networks and the public Internet. The discussion will cover experimentally determined limits for delay between locations, and methods for performance both above and below these thresholds. Representatives of several different dedicated systems and projects will be present, each with many years' experience in the networked music scene. Soundjack and the Jamlink hardware interface will be discussed as examples of realistic interaction over networks, while JackTrip and the Digital Musician Link will show methods for overcoming long latencies on long-distance network links. There will also be a discussion of artistic strategies for overcoming excessive delay in performance.

Session P11 **Friday, April 27**
12:30 – 14:00 **Foyer**

POSTERS: QUALITY EVALUATION AND SPATIAL AUDIO

12:30

P11-1 Influence of Resolution of Head Tracking in Synthesis of Binaural Audio—*Mikko-Ville Laitinen, Tapani Pihlajamäki, Stefan Lösler, Ville Pulkki*, Aalto University, Espoo, Finland

The use of head tracking in binaural synthesis of spatial sound increases the quality of reproduction. The required quality of a head-tracking system for this purpose is the topic of interest in this paper. A listening test was performed to evaluate the effect of four common sources of error in head-tracking systems. The listeners rated the naturalness of binaural reproduction in different head-tracking conditions. According to the test, the main requirement for the head-tracking system is to obtain an accurate and unrestricted azimuth angle. Furthermore, lower update rate of the tracking system affects the quality. The results showed prominent dependence on program material, and individual differences were also notable.
Convention Paper 8623

12:30

P11-2 Objective and Subjective Tests of Consumer-Class Audio Devices—*Marek Pluta, Pawel Malecki*, AGH University of Science and Technology, Krakow, Poland

The paper presents a new stand for subjective listening tests in the Department of Mechanics and Vibroacoustics of Krakow University of Science and Technology, as well as a procedure and results of preliminary tests that utilize it. The most important difference, compared to the previously used stand, is exclusion of an analog mixing console, which reduces the length of electroacoustic channel, and substituting active studio monitors with 12 pairs of closed headphones in order to provide a larger group of listeners with the same listening conditions. The tests study popular consumer audio reproduction devices, therein standalone and motherboard integrated sound cards, as well as portable players. The research included measurements and comparison of objective parameters, as well as ABX method listening test, with a studio-class audio interface as a reference device. A group of listeners of various professional profiles, including students of the Academy of Music in Krakow and Acoustic Engineering of Krakow University of Science and Technology, took part in these listening tests. Results are compared to the data obtained using the previous test stand.
Convention Paper 8624

12:30

P11-3 Subjective Evaluations of Perspective Control Microphone Array (PCMA)—*Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

Perspective Control Microphone Array (PCMA) is a technique that allows one to flexibly render spatial audio images depending on the desired virtual listening position in a reproduced sound field. Two subjective listening experiments have been conducted to evaluate the effectiveness of PCMA on the controls over the perceived source/ensemble distance and width attributes. The first experiment verified a hypothesis suggesting that perceived width would decrease as source-listener distance increased, using anechoic trumpet and conga sources convolved with binaural impulse responses of a concert hall. It was shown that the distance and width linearly changed at doubled distances and were negatively correlated. The second experiment tested three reference virtual array configurations of PCMA on the same attributes using the same sources. The results agreed with the perceptual patterns observed in the concert hall situation in that there was a linear decrease in perceived width at an increased perceived distance. The main effect of PCMA configuration was found to be statistically significant. These results seem to strongly validate the effectiveness of PCMA for postproduction and user-interactive applications.
Convention Paper 8625

12:30

P11-4 Objective Profiling of Perceived Punch and Clarity in Produced Music—*Steven Fenton, Jonathan Wakefield*, University Of Huddersfield, Huddersfield, UK

This paper describes and evaluates an objective measurement that profiles a complex musical signal over time in terms of identification of dynamic content and overall perceived quality.

The authors have previously identified a potential correlation between inter-band dynamics and the subjective quality of produced music excerpts. This paper describes the previously presented Inter-Band Relationship (IBR) descriptor and extends this work by conducting more thorough testing of its relationship to perceived punch and clarity over varying time resolutions. A degree of correlation is observed between subjective test scores and the objective IBR descriptor suggesting it could be used as an additional model output variable (MOV) to describe punch and clarity with a piece of music. Limitations have been identified in the measure however and further consideration is required with regard to the choice of threshold adopted based on the range of dynamics detected within the musical extract and the possible inclusion of a gate as utilized in some loudness algorithms.
Convention Paper 8626

12:30

P11-5 A Hybrid Method Combining Synthesis of a Sound Field and Control of Acoustic Contrast—*Martin Bo Møller,*[1] *Martin Olsen*,[1] *Finn Jacobsen*[2]
[1]Bang & Olufsen a/s, Struer, Denmark
[2]Technical University of Denmark, Lyngby, Denmark

Spatially confined regions with different sound field characteristics, in the following referred to as sound zones, may be desired in some situations. Recently, various sound field control methods for generating separate sound zones have been proposed in the literature. The different algorithms introduce different levels of control over the physical characteristics of the resulting sound fields. This paper introduces a hybrid of two existing methods employed for obtaining sound zones: "Energy Difference Maximization" for control of the sound field energy distribution and "Pressure Matching," which contributes with synthesis of a desired sound field. The hybrid method introduces a tradeoff between acoustic contrast between two sound zones and the degree to which the phase is controlled in the optimized sound fields.
Convention Paper 8627

12:30

P11-6 Validation of Room Plane Wave Decomposition as a Tool for Spatial Ecogram Analysis of Rooms—*Ana Torres*,[1] *Jose J. Lopez*,[2] *Basilio Pueo*[3]
[1]Universidad de Castilla La Mancha, Ciudad Real, Castile-La Mancha, Spain
[2]Technical University of Valencia, Valencia, Spain
[3]Universidad de Alicante, Alicante, Spain

The classical point based analysis of sound attributes in room acoustics are not usually enough for analyzing the acoustic complexity of a room. Some spatial attributes that provide more information are been employed. Circular arrays of microphones and the subsequently plane wave decomposition have been proposed in the literature and employed successfully by the authors and others. In this paper we validate our previous work comparing the resulting measured echograms with the simulation of the room acoustics using a basic room acoustics modeling software, ROOMSIM, freely available. The resulting echograms of both methods are compared identifying strengths and weaknesses of the measurement method based on circular arrays and proposing some ideas for a future more in deep analysis.
Convention Paper 8628

Fri., April 27 13:00 Bartók Meeting Room

Technical Committee Meeting on High Resolution Audio

Fri., April 27 13:00 Room Fortuna

Standards Committee Meeting on EMC

Session P12 Friday, April 27
14:00 – 16:30 Room Lehar

AMPS AND MEASUREMENT

Chair: **Mark Sandler**, Queen Mary University of London, London, UK

14:00

P12-1 Current-Driven Switch-Mode Audio Power Amplifiers—*Arnold Knott, Niels Christian Buhl, Michael A. E. Andersen*, Technical University of Denmark, Lyngby, Denmark

The conversion of electrical energy into sound waves by electromechanical transducers is proportional to the current through the coil of the transducer. However virtually all audio power amplifiers provide a controlled voltage through the interface to the transducer. This paper presents a switch-mode audio power amplifier not only providing controlled current but also being supplied by current. This results in an output filter size reduction by a factor of 6. The implemented prototype shows decent audio performance with $THD + N$ below 0.1%.
Convention Paper 8629

14:30

P12-2 Debugging of Class-D Audio Power Amplifiers—*Lasse Crone, Jeppe Arnsdorf Pedersen, Jakob Døllner Mønster, Arnold Knott*, Technical University of Denmark, Lyngby, Denmark

Determining and optimizing the performance of a Class-D audio power amplifier can be very difficult without knowledge of the use of audio performance measuring equipment and of how the various noise and distortion sources influence the audio performance. This paper gives an introduction on how to measure the performance of the amplifier and how to find the noise and distortion sources and suggests ways to remove them. Throughout the paper measurements of a test amplifier are presented along with the relevant theory.
Convention Paper 8630

15:00

P12-3 Investigation of Crosstalk in Self Oscillating Switch Mode Audio Power Amplifier— *Thomas Haagen Birch, Rasmus Ploug, Niels Elkjær Iversen, Arnold Knott*, Technical University of Denmark, Lyngby, Denmark

Self oscillating switch mode power amplifiers are known to be susceptible to interchannel disturbances also known as crosstalk. This phenomenon has a significant impact on the performance of an amplifier of this type. The goal of this paper is to investigate the presence and origins of crosstalk in a two-channel self oscillating switch mode power amplifier (class D). A step-by-step reduction of elements in an amplifier built for this task is used for methodically determining the actual presence and origins of crosstalk. The investigation shows that the crosstalk is caused by couplings in the self oscillating pulse width modulation circuits, but also that the output filter has a major impact on the level of crosstalk.
Convention Paper 8631

15:30

P12-4 Measuring Mixing Time in Non-Sabinian Rooms: How Scattering Influences Small Room Responses —*Lorenzo Rizzi, Gabriele Ghelfi*, Suono e Vita – Acoustic Engineering, Lecco, Italy

The goal of this work is to optimize a DSP tool for extrapolating from room impulse response information regarding the way in which the transition between early reflections and late reverberation occurs. Two different methods for measuring this transition (usually referred as mixing time, tmix) have been found in literature, both based on statistical properties of acoustic spaces. Appropriate changes have been implemented and the algorithms have been tested on I.R. measured in eight different environments. Particular attention is given to non-Sabinian environments such as small-rooms for music. It has been also measured a relationship between sound diffusion and tmix, showing how the presence of scattering elements contributes to lower tmix altering the statistical properties of I.R.
Convention Paper 8632

16:00

P12-5 Separation of High Order Impulse Responses in Methods Based on the Exponential Swept-Sine—*Stephan Tassart*,[1] *Aneline Grand*[1,2]
[1]ST-Ericsson, Paris, France
[2]Now with Arkamys, Paris, France

Many real analog systems (e.g., electroacoustic loudspeaker, audio amplifiers, filters, etc.) exhibit weakly nonlinear features when driven by large amplitude signals. A large scale of such electromechanical devices are well modeled by the cascade of Hammerstein models. The exponential swept-sine is a natural excitation vector in order to identify the structural elements from those models. This paper extends the original swept-sine principle to the case of band-limited test vectors, suggests an intermodulation law for the generation of band-limited test vectors, and shows that a long-duration swept-sine can be replaced by a series of slightly phase-shifted short-duration swept-sines. High order impulse responses are separable even in case of temporal overlap with a linear combination of the measurements. The method is demonstrated on examples.
Convention Paper 8633

Session P13 — Friday, April 27
14:00 – 16:30 — Room Liszt

ANALYSIS AND SYNTHESIS OF SOUND: PART 2; CONTENT MANAGEMENT

Chair: **Michael Kelly**, DTS Inc., London, UK

14:00

P13-1 Overview of Feature Selection for Automatic Speech Recognition—*Branislav Gerazov, Zoran Ivanovski*, Faculty of Electrical Engineering and Information Technologies, Skopje, Macedonia

The selection of features to be used for the task of Automatic Speech Recognition (ASR) is critical to the overall performance of the ASR system. Throughout the history of development of ASR systems, a variety of features have been proposed and used, with greater or lesser success. Still, the research for new features, as well as modifications to the traditional ones, continues. Newly proposed features as well as traditional feature optimization focus on adding robustness to ASR systems, which is of great importance for applications involving noisy environments. The paper seeks to give a general overview of the various features that have been used in ASR systems, giving details to an extent granted by the space available.
Convention Paper 8634

14:30

P13-2 Evaluating the Influence of Source Separation Methods in Robust Automatic Speech Recognition with a Specific Cocktail-Party Training—*Amparo Marti*,[1] *Maximo Cobos*,[2] *Jose J. Lopez*[1]
[1]Universitat Politècnica de València, Valencia, Spain
[2]Universitat de València, Burjassot, Valencia, Spain

Automatic Speech Recognition (ASR) allows a computer to identify the words that a person speaks into a microphone and convert it to written text. One of the most challenging situations for ASR is the cocktail party environment. Although source separation methods have already been investigated to deal with this problem, the separation process is not perfect and the resulting artifacts pose an additional problem to ASR performance in case of using separation methods based on time-frequency masks. Recently, the authors proposed a specific training method to deal with simultaneous speech situations in practical ASR systems. In this paper we study how the speech recognition perfor- ➥

mance is affected by selecting different combinations of separation algorithms both at the training and test stages of the ASR system under different acoustic conditions. The results show that, while different separation methods produce different types of artifacts, the overall performance of the method is always increased when using any cocktail-party training.
Convention Paper 8635

15:00

P13-3 Automatic Regular Voice, Raised Voice, and Scream Recognition Employing Fuzzy Logic **—** *Kuba Lopatka, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

A method of automatic recognition of regular voice, raised voice, and scream used in an audio surveillance system is presented. The algorithm for detection of voice activity in a noisy environment is discussed. Signal features used for sound classification, based on energy, spectral shape, and tonality are introduced. Sound feature vectors are processed by a fuzzy classifier. The method is employed in an audio surveillance system working in real-time both in an indoor and outdoor environment. Achieved results of classifying real signals are presented and discussed.
Convention Paper 8636

15:30

P13-4 Enhanced Chroma Feature Extraction from HE-AAC Encoder—*Marco Fink*,[1] *Arijit Biswas*,[2] *Walter Kellermann*[1]
[1]University of Erlangen-Nuremberg, Erlangen, Germany
[2]Dolby Germany GmbH, Nuremberg, Germany

A perceptually enhanced chroma feature extraction during the HE-AAC audio encoding process is proposed. Extraction of chroma features from the MDCT-domain spectra of the encoder and its further enhancement utilizing the perceptual model of the encoder is investigated. The main advantage of such a scheme is a reduced computational complexity when both chroma feature extraction and encoding is desired. Specifically, the system is designed to produce reliable chroma features irrespective of the block switching decision of the encoder. Three methods are discussed to circumvent the poor frequency resolution during short blocks. All proposed enhancements are evaluated systematically within a well-known state-of-the-art chord recognition framework.
Convention Paper 8637

16:00

P13-5 Hum Removal Filters: Overview and Analysis **—***Matthias Brandt, Jörg Bitzer*, Jade University of Applied Sciences, Oldenburg, Germany

In this contribution we analyze different methods for removing sinusoidal disturbances from audio recordings. In order to protect the desired signal, high frequency selectivity of the used filters is necessary. However, due to the time-bandwidth uncertainty principle, high frequency selectivity brings about long impulse responses. This can result in audibly resonating filters, causing artifacts in the output signal. Thus, the choice of the optimal algorithm is a compromise between frequency selectivity and acceptable time domain behavior. In this context, different filter structures and algorithms have different characteristics. To investigate their influence on the hum disturbance and the desired signal, we have evaluated three methods using objective error measures to illustrate advantages and drawbacks of the individual approaches.
Convention Paper 8638

Workshop 8 **Friday, April 27**
14:00 – 15:30 **Room Brahms**

HIGHLY DIRECTIONAL MICROPHONES FOR SOUND RECORDING

Chair: **Helmut Wittek**, Schoeps Mikrofone GmbH, Karlsruhe, Germany

Panelists: *Christof Faller*, Illusonic LLC
Michael Millitzer, Microtech Gefell

Times are changing quickly with regard to highly directional microphones both in theory and in practice. Most systems aim at optimal speech transmission and intelligibility in communication applications. However, there are also new approaches that sound engineers can consider using for certain recording purposes as well (e.g., the Eigenmike and the SuperCMIT). Along with improved communication microphone solutions, they will compete with conventional methods such as short and long interference tubes as well as parabolic mirrors.

So which method fits which application best? Can these new microphone types be used for applications such as music recording, location sound, nature recording? Each has different advantages with regard to frequency response, directivity, flexibility, price, robustness, predictability of results, and practicability. The workshop will compare them on the basis of technical data, practical experiences and sample recordings

Workshop 9 **Friday, April 27**
14:00 – 16:00 **Room Bartók**

ACTIVE ACOUSTICS TODAY: CHALLENGES, SOLUTIONS, APPLICATIONS

Chair: **Wieslaw Woszczyk**, McGill University, Montreal, Quebec, Canada

Panelists *Leo de Klerk*, Bloomline Acoustics b.v., ‘s-Gravendeel, The Netherlands
Doyuen Ko, McGill University, Montreal, Quebec, Canada
Karoly Molnar, Meyer Sound Labs, Inc., Berkeley, CA, USA

Active acoustics uses digital signal processing and electroacoustics to deliver adjustable correction and enhancement of the response of acoustic enclosure used for music and speech communication. Using variable virtual architecture, correction and adaptation of existing acoustics can be done quickly and efficiently for each specific application, something not possible using passive variable acoustics. Recent advances in active acoustics allow us to create superb stage and auditorium acoustics for performing musicians and listening audiences, as well as for recordings. The panelists will discuss about the state of the art in active acoustics, the challenges and solutions encountered in the field, and brainstorm about future directions.

Fri., April 27 14:00 Bartók Meeting Room

Technical Committee Meeting on Fiber Optics for Audio

Fri., April 27 14:00 Room Fortuna

Standards Committee Meeting on Loudspeaker Modeling and Measurement

Fri., April 27 15:00 Bartók Meeting Room

Technical Committee Meeting on Network Audio Systems

Special Event
FUTURE DIRECTIONS IN MULTICHANNEL
Friday, April 27, 16:00 – 18:00
Room Brahms

There will be two separate presentations of new multichannel systems.

Recording Techniques for Auro 3D /3D Audio
16:00 – 17:00

Chair: **Gregor Zielinsky**, Sennheiser Electronic, Wedemark, Germany

Panelists: *TBA*

Auro3D is a rather recent additional proposal for reproduction of musical recordings including height information. Its comparatively simple, pragmatic approach makes it all the more interesting for adaption to film and home theater loudspeaker setups. Benefits of Auro3D especially for music recordings will be discussed.

For the Tonmeister then, the question arises how to produce 3-D recordings, either founding the mix mostly on spot microphones, or on new main microphone placement techniques, including height microphones. A vast set of different setups were seen at the recent ICSA 2011 Conference on Spatial Audio, in Detmold. Several microphone setups and their sonic differences will be described. The use of microphones, combinations and adjustable polar patterns will be discussed.

As in stereo, with main microphones a further question arises on if and how to delay compensate additional spot microphones, relative to the main microphones. Also, how need the main microphones need to be delayed in themselves.

Including height loudspeakers produces enormous differences in timbre and spatial perception. How can they be described, and explained? Finally, comments will be presented regarding the differences between single-source and multi-source (phantom source) based localization in 3-D fields.

Mercedes-Benz Signature Sound—
A mediaHYPERIUM Production Mixed and Mastered at Skywalker Sound
17:00 – 18:00

Chair: Herbert Waltl, mediaHYPERIUM, Los Angeles, CA, USA

Panelists: *Stefan Bock*, msm-studios, Munich, Germany
Mario Fresner, Daimler AG, Sindelfingen, Germany
Leslie Ann Jones, Skywalker Sound, Marin County, California, USA
Norbert Niemczyk, Daimler AG, Sindelfingen, Germany
Wieslaw Woszczyk, McGill University, CIRMA, Montreal, QC, Canada

A “first-of-its-kind” surround sound project: music mixed and mastered for the new Mercedes SL Roadster model. To create the ultimate listening experience the producer and engineers tailored discrete surround mixes to the car’s audio system, its specific acoustic environment, and the position of driver and passenger.

The “hands-on” approach of the Signature Sound production concept aims to reflect the true artistic intention of the musical pieces in an immersive listening experience.

Daimler AG has underwritten the production and collaborated on the exclusive, limited music DVD with the title “Mercedes-Benz Signature Sound.”

We will discuss the concept, how it was done and the importance of an artistic approach in surround sound productions.

Fri., April 27 16:00 Bartók Meeting Room

Technical Committee Meeting on Studio Practices and Production

Fri., April 27 16:00 Room Fortuna

Standards Committee Meeting on Metadata for Audio

Session P14 Friday, April 27
16:30 – 18:00 Room Lehar

LISTENING TESTS: PART 2

Chair: **Thomas Sporer**, Fraunhofer Institute for Integrated Circuits IIS, Ilmenau, Germany

16:30

P14-1 Determining the Threshold of Acceptability for an Interfering Audio Program—
Jon Francombe,[1] *Russell Mason,*[1] *Martin Dewhirst*,[1] *Søren Bech*[2]
[1]University of Surrey, Guildford, Surrey, UK
[2]Bang & Olufsen, Struer, Denmark

An experiment was performed in order to establish the threshold of acceptability for an interfering audio program on a target audio program, varying the following physical parameters: target program, interferer program, interferer location, interferer spectrum, and road noise level. Factors were varied in three levels in a Box-Behnken fractional factorial design. The experiment was performed in three scenarios: information gathering, entertainment, and reading/working. Nine listeners performed a method of adjustment task to determine the threshold values. Produced thresholds were similar in the information and entertainment scenarios, however there were significant differences between subjects, and factor levels also had a significant effect: interferer program was the most important factor across the three scenarios, while interferer location was the least important.
Convention Paper 8639

17:00

P14-2 Signal Processing Framework for Virtual Headphone Listening Tests in a Noisy Environment—*Jussi Rämö, Vesa Välimäki*, Aalto Univesity, Espoo, Finland

A signal processing framework is introduced to enable parallel evaluation of headphones in a virtual listening test. It is otherwise impractical to conduct a blind comparison of several headphones. The ambient noise isolation capability of headphones has become an important design feature, since the mobile usage of earphones takes place in noisy listening environments. Therefore, the proposed signal processing framework allows a noise signal to be fed through a filter simulating the ambient sound isolation at the same time when music is played. This enables the simultaneous evaluation of the timbre and background noise characteristics, which together define the total headphone listening experience. Methods to design FIR filters for compensating the reference headphone response and for simulating the frequency response and isolation curve of the headphones to be tested are presented. Furthermore, a real-time test environment implemented using Matlab and Playrec is described.
Convention Paper 8640

17:30

P14-3 Perceptual Evaluation of Stochastic-Event-Based Percussive Timbres for Use in Statistical Sonification—*William Martens, Mark McKinnon-Bassett, Densil A. Cabrera*, University of Sydney, Sydney, NSW, Australia

The results of statistical data analysis have typically been presented using visual displays, but the sonification of data for auditory display, particularly using sound varying along realistic timbral dimensions, can offer an attractive alternative means for rendering such results. It was hypothesized that stochastic-event-based percussive timbres could be useful in communicating the details of statistical data, and so a preliminary study was designed to investigate the use of these timbres in data sonification. This study examined the ability of listeners to estimate the variation in a physical parameter for stimuli selected from a set of recorded percussive events. Specifically, two experiments were executed to determine whether listeners are generally able to estimate the number of small pellets that were present inside a container based upon the sounds that were made when the container was shaken a few times. The experimental results showed that a 6-dB trial-to-trial variation in reproduction level had no significant effect on obtained estimates, whereas variation in spectral energy distribution did significantly affect estimates of the number of small pellets in the shaken container. While the capacity to discriminate between sonification system outputs has been established, investigation of system effectiveness in applications remains to be done.
Convention Paper 8642

Student/Career Development Event
RECORDING COMPETITION—PART 1
Friday, April 27, 16:30 – 18:30
Room Bartók

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. This event presents stereo and surround recordings in these categories:

• Traditional Non-Multitrack Acoustic Recording – 16:30–17:30

• Sound for Visual Media – 17:30–18:30

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 Sunday 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 if your project isn't one of the finalists, and it's a great chance to meet other students and faculty members.

Session EB1 Friday, April 27
17:00 – 18:30 Room Liszt

ENGINEERING BRIEFS—PAPERS

Chair: **Josh Reiss**, Queen Mary University of London, London, UK

17:00

EB1-1 What Is the Worth of Pre-DSP Traditional" Production Values in the Contemporary Manipulation-Oriented Context?—*Justin Paterson*, London College of Music, University of West London, London, UK

Moorefield's "illusion of reality" was focused on the perception of authenticity in recordings, something that often continues to concern music producers. Editing techniques have increasingly been applied to move beyond this mimetic reproduction and "re-perform" musical elements, and furthermore, DSP commonly offers such extensive manipulation possibilities that all identifiable components of authenticity might be masked, even subverted, offering "virtual timbres" and revised sonic meaning. Here, editing and processing are considered along with their aesthetic and technical implications, placed in a historical perspective and augmented through the synthesis of contributions from a number of professional producers. Several perspectives are presented, and their tensions considered. The fluxive nature of authenticity will be further revealed in the trajectory toward the "reality of illusion."
Engineering Brief 41

17:15

EB1-2 Subjective Differences in Digital Audio Workstation Math—*Brett Leonard, Padraig Buttner-Schnirer*, McGill University, Montreal, Quebec, Canada, Centre for Interdisciplinary

Research in Music, Media & Technology, Montreal, Quebec, Canada

The subject of sound quality often arises when discussing the merits of various digital audio workstations (DAWs). While many engineers argue that one DAW "sounds better" than another, very little quantified data exists on the subject. In order to test these claims, multiple DAWs are fed the same multi-track digitized audio from a single converter. This audio is then processed by lowering all faders in each DAW by a fixed, arbitrary amount, generating identical mixes, save the internal math performed through the gain change and the internal summing of each DAW. The resulting mixes are then tested for discriminability by trained listeners through the use of ABX testing. While there were differences between mixes, most listeners struggled to discriminate between DAWs.
Engineering Brief 42

17:30

EB1-3 Virtual Microphones: Optimum Real-Time Demodulation of Phase Modulated Ultrasound—*Tom Ritter*,[1] *Tobias Merkel*,[1] *Hans-G. Lühmann*[2]
[1]Beuth University of Applied Sciences Berlin, Berlin, Germany
[2]Lütronic Elektroakustik GmbH, Berlin, Germany

The project "Virtual Microphones" aims to replace conventional audio microphones by a system consisting of ultrasonic sensors. The fundamental physical principle utilized is that ultrasonic waves are phase-modulated when superimposed by audio waves. Due to the physical relationships the phase changes of the ultrasonic carrier are very small. The precise demodulation of the audio signal was in the focus of the investigation among other measures, like providing a very stable carrier frequency or the focusing of the ultrasound beam on the receivers. Several methods of demodulation for phase-modulated signals are compared. The comparatively low resulting signal-to-noise ratio after the demodulation required further signal processing in form of noise reduction to be investigated and its performance accessed.
Engineering Brief 43

17:45

EB1-4 The Audio Plugin Generator: Rapid Prototyping of Audio DSP Algorithms—*Robert Cerny, Fritz Menzer*, dlab GmbH, Winterthur, Switzerland

This paper introduces the Audio Plugin Generator (APG), a rapid prototyping tool that allows transforming Simulink models into VST plugins with little effort. VST is a well-known audio plugin interface, supported by a variety of host programs, to process audio signals with low latency. On the other hand, Simulink enables DSP engineers to use clean, human-readable descriptions of their algorithms. The generated plugin supports parameter tuning as well as data logging in real-time. Furthermore, signals can be analyzed inside the plugin with signal scopes. The paper explains the APG workflow and how it bridges the gap between the Simulink Coder and the VST standard. A light version of APG can be obtained free of charge from the APG website: http://audioplugingenerator.com/
Engineering Brief 44

18:00

EB1-5 MS Mastering of Stereo Microphone Signals—*Martin Schneider*, Georg Neumann GmbH, Berlin, Germany

MS (mid-side) technique is a classic tool for recording and mastering based on coincident microphone setups. For these, localization theory tells us that on reproduction via loudspeakers, the interaural level differences (ILDs) will yield stable and fully mono-compatible results. Yet, when a mastering engineer obtains a mix, he will rarely be told how the specific sources were recorded. Was it as mono signals, level panned to a certain reproduction angle, or with coincident, spaced, or mixed stereo microphone setups. While the first two techniques only involve ILDs, spaced microphone techniques also imply interaural time differences (ITDs). MS matrixing techniques can be applied to change the apparent width of a full stereo mix in postproduction. The engineer should be aware that he then cross-feeds signals between the channels. The signal paths from a signal sound source, via the microphones, the matrixing, and the loudspeakers, to finally reach the ears, can now be seen as eight, instead of four, for coincident techniques. An increased sensation of spaciousness or envelopment may arise, but one should be aware of the physics behind the process. Some of the arising comb filters are shown here, for spaced and ORTF setups.
Engineering Brief 45

18:15

EB1-6 A New Visual Paradigm to Surround Sound Mixing—*Thierry Dilger*, www.sonabilis.com, Paris, France

Today's tools for surround mixing are like add ons to old style linear software. Their paradigm for surround mixing are almost the same: a sound file (or a track) is routed to a surround panner, with the help of the timeline and automation, sounds are moving in space. But you cannot prepare several sound movements at once, you cannot have a whole timeline overview of the spatialization process, you cannot compare easily two surround mixes, you can nearly not work on non-traditional set-ups, you cannot create generative surround sounds. The model presented during this presentation is based on a color coding system (RGB) working on three axes which are: time sounds are playing (x), the speakers number and position (y), and the intensity of each sound (z). I discuss its pros (it gives answers to all the flaws previously announced) and cons compared to traditional tools and about its field of application.
Engineering Brief 46

EB1-7 Sounds Not Signals: A Perceptual Audio Format—*Michael J. Terrell, Andrew J. R. Simpson, Mark Sandler,* Queen Mary University of London, London, UK

This Engineering Brief presents a brief overview of a novel audio transmission and storage format. The main feature of the format is the use of absolute sound pressure within the format. A secondary feature is the use of perceptual features to describe the associated sound content. Error reporting and correction systems are briefly described and a case study used to demonstrate the principles involved.
Engineering Brief 52

Special Event
OPEN HOUSE OF THE TECHNICAL COUNCIL AND THE RICHARD C. HEYSER MEMORIAL LECTURE
Friday, April 27, 19:00 – 20:00
Room Bartók

Lecturer: **Graham Blyth**

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

The Richard C. Heyser distinguished lecturer for the 132nd AES Convention is Graham Blyrh. Blyth was born in 1948, began playing the piano aged 4 and received his early musical training as a Junior Exhibitioner at Trinity College of Music in London, England. Subsequently at Bristol University, where he read Electrical Engineering, he founded the University Music Society, conducting their Chamber Orchestra and Choir. He holds diplomas in Organ Performance from the Royal College of Organists, The Royal College of Music and Trinity College of Music.

In 1973 he founded Soundcraft with Phil Dudderidge, and has been Technical Director from the beginning. Soundcraft has since grown to be one of the most significant designers and manufacturers of audio mixing consoles in the world. In 1988, the company was bought by Harman, whose professional audio companies now include JBL, DBX, Lexicon, AKG, and Studer.

In the late 1980s he renewed his musical studies with Sulemita Aronowsky for piano and Robert Munns for organ. He gives numerous concerts each year, principally as organist and pianist, but also as a conductor and harpsichord player. He made his international debut with an organ recital at St. Thomas Church, New York in 1993 and since then has given concerts on many of the finest organs in Europe, including the Madeleine and St. Etienne du Mont in Paris, and the Liebfrauen Dom in Munich, and in North America, including Grace Cathedral, San Francisco and St. Ignatius Loyola in New York.

He has lived in Wantage, Oxfordshire, since 1984 where he is currently Artistic Director of Wantage Chamber Concerts and Director of the Wantage Summer Festival. In 1995 he built the Challow Park Recital Hall, an 80 seat auditorium with completely variable acoustics, using the Lexicon LARES system, designed by his Harman colleague David Griesinger. This allows for performance and recording of music ranging from String Quartets through to Organ Recitals.

Today he divides his time between audio engineering and organ design activities. In 2003 he founded the Veritas Organ Company to address the top end of the digital classical organ market, specialising in creating new organs by adding digital voices to existing pipes. He is a Fellow of the Royal Society of Arts and the Audio Engineering Society.

The title of his speech is “In Pursuite of Elegant Siplicity: Life, Luck, and Learning in Music and Audio.”

Blyth will talk about how he became a design engineer in the audio industry, the development of the mixing console from a personal perspective during his 41 years in the business, and, in particular, his approach to microphone preamp design with illustrated examples. He will also talk about the importance of the analog engineer in a mostly digital world and about the technical and musical challenges in designing high-quality digital classical organs, with audio examples. Blyth is well known to attendees at AES conventions for his popular organ recitals.

Blyth’s presentation will be followed by a reception hosted by the AES Technical Council.

Session P15 **Saturday, April 28**
09:00 – 11:30 **Room Lehar**

SPATIAL AUDIO: PART 1

Chair: **Christof Faller**, Illusonic LLC

09:00

P15-1 Analysis on Error Caused by Multi-Scattering of Multiple Sound Sources in HRTF Measurement—*Guangzheng Yu, Bosun Xie, Zewei Chen, Yu Liu*, South China University of Technology, Guangzhou, China

A model consisting of two pulsating spherical sound sources and a rigid-spherical head is proposed to evaluate the error caused by multi-scattering of multiple sound sources in HRTF measurement. The results indicate that the ipsilateral error below 20 kHz caused by multi-scattering is within ± 1 dB when the radius of sound sources does not exceed 0.025 m, the source distance to head center is not less than 0.5 m, and the angular interval between the two adjacent sources is not less than 25 degrees. This accuracy basically satisfies the requirement of ipsilateral HRTF measurements. For improving the accuracy in contralateral HRTF measurement, some sound absorption treatments on the source surface are necessary.
Convention Paper 8643

09:30

P15-2 Personalization of Headphone Spatialization Based on the Relative Localization Error in an Auditory Gaming Interface—*Aki Härmä, Ralph van Dinther, Thomas Svedström, Munhum Park, Jeroen Koppens*, Philips Research Europe, Eindhoven, The Netherlands

In binaural sound reproduction applications using head-related transfer functions (HRTFs) it is beneficial that the properties of the HRTFs correspond to the personal characteristics of the real HRTFs of the user. In this paper we propose a method to choose HRTFs using a relative lo-

calization test. This allows us to make the selection of the best HRTFs using a simple auditory interface. It is possible to design the HRTF personalization interface in a consumer device as an auditory game where the task of the user is to place sound objects in relation to each other. Two different interfaces are compared in a listening test. The results of the tests reported in the current paper are mixed and do not give a conclusive picture on the performance of the proposed system, however, they do give interesting insights about the properties of binaural listening.
Convention Paper 8644

10:00

P15-3 Robustness of a Mixed-Order Ambisonics Microphone Array for Sound Field Reproduction—*Marton Marschall,*[1] *Sylvain Favrot*,[1] *Jörg Buchholz*[2]
[1]Technical University of Denmark, Lyngby, Denmark
[2]National Acoustic Laboratories, Chatswood, Australia

Spherical microphone arrays can be used to capture and reproduce the spatial characteristics of acoustic scenes. A mixed-order Ambisonics (MOA) approach was recently proposed to improve the horizontal spatial resolution of microphone arrays with a given number of transducers. In this paper the performance and robustness of an MOA array to variations in microphone characteristics as well as self-noise was investigated. Two array processing strategies were evaluated. Results showed that the expected performance benefits of MOA are achieved at high frequencies, and that robustness to various errors was similar to that of HOA arrays with both strategies. The approach based on minimizing the error of the reproduced spherical harmonic functions showed better performance at high frequencies for the MOA layout.
Convention Paper 8645

10:30

P15-4 An Algorithm for Efficiently Synthesizing Multiple Near-Field Virtual Sources in Dynamic Virtual Auditory Display—*Bosun Xie, Chengyun Zhang*, South China University of Technology, Guangzhou, China

An algorithm for efficiently synthesizing multiple near-field virtual sources in dynamic virtual auditory display (VAD) is proposed. Applying the method of principal component analysis, a set of measured near-field head-related impulse responses (HRIRs) for KEMAR manikin at various source directions and distances are decomposed into a weighted sum of 15 time-domain basic functions along with a mean time-domain function, in which the time-independent weights represent the location dependence of HRIRs. Accordingly, multiple virtual sources synthesis at various locations is implemented by a common bank of 16 filters representing the time -domain basis functions and the mean time-domain function, in which the adjustable gains of the filters for each input stimulus (as well as an overall gain and delay for each input stimulus) control intended source locations. The computational cost of the proposed algorithm is reduced compared with that of conventional ones. Psychoacoustic experiments via a dynamic VAD with head-tracking validate the performance of the proposed algorithm.
Convention Paper 8646

11:00

P15-5 Scalable Coding of Three-Dimensional Multichannel Sound—Design of Conversion Matrix and Modeling of Unmasking Phenomenon—*Akio Ando*, NHK Science and Technical Research Laboratories, Setagaya, Tokyo, Japan, and Tokyo Institute of Technology, Meguro, Tokyo, Japan

We propose two methods for coding and transmitting three-dimensional multichannel sound signals: scalable coding and transmission, and modeling of the quantization error. The first method converts *N*-channel sound signals into *M*-channel basic signals and (*N-M*)-channel additional signals using a matrix operation. The matrix is trained by simulated annealing to minimize its condition number and the energy of additional signals. The unmasking artifact may occur when the *N*-channel signals are restored from the decoded signals using the inverse matrix. The second method estimates the quantization error signals by the polynomial approximation of the decoded signals. Experimental results showed that the combination of both methods could realize a 1.2 Mbps scalable transmission of 22-channel sounds without a notable sound degradation.
Convention Paper 8647

Session P16 Saturday, April 28
09:00 – 12:30 Room Liszt

HIGH RESOLUTION AND LOW BIT RATE

Chair: **Vesa Välimäki**, Aalto Univesity, Espoo, Finland

09:00

P16-1 Time Domain Performance of Decimation Filter Architectures for High Resolution Sigma-Delta Analog to Digital Conversion—*Yonghao Wang,*[1,2] *Joshua D. Reiss*[1]
[1]Queen Mary University of London, London, UK
[2]Birmingham City University, Birmingham, UK

We present the results of a comparison of different decimation architectures for high resolution sigma delta analog to digital conversion in terms of passband, transition band performance, simulated signal-to-noise ratio, and computational cost. In particular, we focus on the comparison of time domain group delay response of different filter architectures including classic multistage FIR, cascaded integrator-comb (CIC) with FIR compensation filters, particularly multistage polyphase IIR filter, cascaded halfband minimum phase FIR filter, and multistage minimum phase FIR filter designs. The analysis shows that the multistage minimum phase FIR filter and multistage polyphase IIR filter are most promising for

low group delay audio applications.
Convention Paper 8648

09:30

P16-2 A Delta-Sigma Modulator Using Dual NTF for 1-Bit Digital Switching Amplifier—*Jungmin Choi, Jaeyong Cho, Haekwang Park*, Samsung Electronics, Suwon, Korea

In this paper a fifth-order single-loop single-bit delta-sigma modulator (DSM), which is constructed by cascade-of-integrator, feed-back (CIFB) form for a 1-bit digital audio switching amplifier is proposed. High order DSM can achieve high signal-to-noise ratio (SNR), but it has probability that the oscillation occur. To achieve high SNR and improve the stability of the modulator for a large input range, we propose the DSM which is composed of dual noise transfer function (NTF). The one is high SNR mode that maximizes SNR of DSM and the other is stable mode that enhances stability of DSM. The proposed architecture is simulated in the register transfer level (RTL) and implemented in the FPGA board.
Convention Paper 8649

10:00

P16-3 9 Years HE AAC—Technical Challenges Using an Open Standard in Real-World Applications—*Martin Wolters*,[1] *Gregory McGarry*,[2] *Andreas Schneider,*[1] *Robin Thesing*[1]
[1]Dolby Germany GmbH, Nuremberg, Germany
[2]Dolby Australia Pty Ltd., Sydney, Australia

The technical work on creating the MPEG HE AAC standard was finished nine years ago. Since then the format has become very popular in specific markets and devices such as PCs and mobile phones mainly due to its high-compression efficiency. However, creating a reliable eco-system based on this open standard remains a technical challenge. In this paper results of several compatibility tests, which were conducted over the last two years on both mobile phones and broadcast receivers, are presented. The problems encountered and recommended solutions are described.
Convention Paper 8650

10:30

P16-4 Subjective Tests on Audio Mix Dedicated to MP3 Coding—*Szymon Piotrowski*,[1] *Magdalena Plewa*[2]
[1]AGH University of Science and Technology, Krakow, Poland
[2]Gdansk University of Technology, Gdansk, Poland

Over the past years the Internet has become very popular as a means of distributing audio. MP3 coded audio is present in the Internet, in a bus, in broadcast. Sound engineers agree that there can often be a lack of control over the downstream processing that is applied to final material. The aim of the presented work is to compare audio mixes dedicated to CD format and converted to MP3 format with MP3-dedicated productions to and evaluate them.
Convention Paper 8651

11:00

P16-5 New Enhancements for Improved Image Quality and Channel Separation in the Immersive Sound Field Rendition (ISR) Parametric Multichannel Audio Coding System—*Hari Om Aggrawal*,[1] *Deepen Sinha*[2]
[1]ATC Labs, Noida, India
[2]ATC Labs, NJ, USA

Consumer audio applications such as satellite broadcasts, HDTV, multichannel audio streaming, gaming, and playback systems are highlighting newer challenges in low-bit-rate parametric multichannel audio coding. This paper describes the continuation of our research related to the Immersive Sound field Rendition (ISR) parametric multichannel encoding system. We focus on the recent enhancements for the surround and center channel generation components of the ISR system. The emphasis being on improving the fidelity and quality of reconstructed 5/5.1-channel audio so that it achieves a level of transparency desirable for high end applications. Furthermore, it is being attempted to improve the robustness of the coding scheme to various difficult signals and listening environments by reducing inter-channel leakage to a minimum. We describe challenging case, various algorithmic improvements to the ISR algorithm to address these and also discuss the subjective impact of these algorithmic improvements.
Convention Paper 8652

11:30

P16-6 Novel Decimation-Whitening Filter in Spectral Band Replication—*Han-Wen Hsu, Chi-Min Liu*, National Chiao Tung University, Hsinchu, Taiwan

MPEG-4 high-efficiency advanced audio coding (HE-AAC) has adopted spectral band replication (SBR) to efficiently compress high-frequency parts of the audio. In SBR, the linear prediction is applied to low-frequency subbands to clip the tonal components and smooth the associated spectrum for replicating to high-frequency bands. Such a process is referred to as the whitening filtering. In SBR, to avoid the alias artifact from spectral adjustment, a complex filterbank instead of real filterbank is adopted. For the QMF subbands, this paper analyzes that the linear prediction defined in SBR standard results in the predictive biases. A new whitening filter, called the decimation-whitening filter, is proposed to eliminate the predictive bias and provide advantages in terms of noise-to-signal ratio measure, frequency resolution, energy leakage, and computational complexity for SBR.
Convention Paper 8653

12:00

P16-7 MPEG Unified Speech and Audio Coding—The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types—*Max Neuendorf,*[1] *Markus Multrus,*[1] *Nikolaus Rettelbach,*[1] *Guillaume Fuchs,*[1] *Julien Robilliard,*[1] *Jérémie Lecomte,*[1] *Stephan Wilde,*[1] *Stefan Bayer,*[1] *Sascha Disch,*[1] *Christian Helmrich,*[1] *Roch Lefebvre,*[2] *Philippe Gournay,*[2] *Bruno Bessette,*[2] *Jimmy Lapierre,*[2] *Kristofer Kjörling,*[3] *Heiko Purnhagen,*[3] *Lars Villemoes,*[3]

Werner Oomen,[4] *Erik Schuijers,*[4] *Kei Kikuiri,*[5] *Toru Chinen,*[6] *Takeshi Norimatsu,*[7] *Chong Kok Seng,*[7] *Eumi Oh,*[8] *Miyoung Kim,*[8] *Schuyler Quackenbush,*[9] *Bernhard Grill*[1]
[1]Fraunhofer IIS, Erlangen, Germany
[2]Université de Sherbrooke, Sherbrooke, Quebec, Canada
[3]Dolby Sweden AB, Stockholm, Sweden
[4]Philips Research Laboratories, Eindhoven, The Netherlands
[5]NTT DOCOMO, INC., Yokosuka, Kanagawa, Japan
[6]Sony Corporation, Shinagawa, Tokyo, Japan
[7]Panasonic Corporation
[8]Samsung Electronics, Suwon, Korea
[9]Audio Research Labs, Scotch Plains, NJ, USA

In early 2012 the ISO/IEC JTC1/SC29/WG11 (MPEG) finalized the new MPEG-D Unified Speech and Audio Coding standard. The new codec brings together the previously separated worlds of general audio coding and speech coding. It does so by integrating elements from audio coding and speech coding into a unified system. The present publication outlines all aspects of this standardization effort, starting with the history and motivation of the MPEG work item, describing all technical features of the final system, and further discussing listening test results and performance numbers that show the advantages of the new system over current state-of-the-art codecs.
Convention Paper 8654

Workshop 10 **Saturday, April 28**
09:00 – 10:00 **Room Bartók**

SCREEN-LESS NAVIGATION FOR HIGH-RESOLUTION AUDIO ON BLU-RAY DISC

Chair: **Stefan Bock**, msm-studios GmbH, Munich, Germany

Panelists: *Jim Anderson*, New York University, New York, NY, USA
Tomi Pietilä, Producer/Engineer, Tampere, Finland
Darcy Proper, Wisseloord Studios, Hilversum, The Netherlands
David Walstra, Consultant AV Entertainment Industry

High-resolution audio, presented as uncompressed LPCM, has been waiting for a suitable transport format for some time. The Blu-ray Disc (BD) format offers such a transport and supports the necessary linear and lossless codecs as part of its basic specification. While many BD players can be found in home theater and games environments, there are some issues that need to be addressed before they can be introduced into a hi-fi environment that does not have a screen to present visual menus for audio stream setup and track selection. This recommended method specifies a structure for authoring a BD ROM to enable playback in screen-less consumer systems, and to provide simple track selection from the remote control.

Tutorial 3 **Saturday, April 28**
09:00 – 10:30 **Room Brahms**

CANCELLED

Sat., April 28 **09:00** **Bartók Meeting Room**

Technical Committee Meeting on Electro Magnetic Compatibility

Sat., April 28 **10:00** **Bartók Meeting Room**

Technical Committee Meeting on Audio Forensics

Workshop 11 **Saturday, April 28**
10:30 – 12:00 **Room Brahms**

MUSHRA RELOADED

Chair: **Judith Liebetrau**, Technical University Ilmenau, Ilmenau, Germany

Panelists: *Poppy Crum*, Dolby Laboratories, San Francisco, CA, USA
Frederik Nagel, Fraunhofer Institute for Integrated Circuits IIS/International Audio Laboratories, Erlangen, Germany
Thomas Sporer, Fraunhofer Institute for Integrated Circuits IIS, Ilmenau, Germany

Since its finalization in 2001 Recommendation ITU-R BS.1534, nicknamed MUSHRA, has become very popular. MUSHRA is an acronym for “multi stimulus with hidden reference and anchors.” Many researchers created variants of the original design and called it still MUSHRA, and very often test reports are missing important information about test parameters and statistical analysis. Very often the mandatory anchor is not used.

ITU-R WP 6C is currently working on a revision of MUSHRA. This revision will specify in more detail how to perform the test and the statistical analysis, will add some of the most popular variants as official alternatives, and will offer new, more powerful anchors.

This workshop will present current discussions and results from the spring 2012 meeting of ITU-R.

Tutorial 4 **Saturday, April 28**
10:15 – 12:15 **Room Bartók**

LOUDNESS LEVELING AND NORMALIZATION—BASICS, MISUNDERSTANDINGS, DANGERS, AND SOLUTIONS

Presenter: **Florian Camerer**, ORF, Vienna, Austria, Chairman EBU-group PLOUD

The switch from peak to loudness normalization is in full swing not only in Europe, but worldwide. There is an international measurement algorithm (ITU-R BS.1770) as well as a few recommendations, all based on that algorithm. This is arguably the biggest change in the audio leveling world in decades. This tutorial will cover the basics of loudness work and look into some details of where potential pitfalls lie, where there might be or are misunderstandings or misconceptions, and give a brief overview of the current status of loudness implementations in Europe and the world. Examples will be played to illustrate the presentation.

Student/Career Development
EDUCATION AND CAREER/JOB FAIR
Saturday, April 28, 11:00 – 12:30
Foyer

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 to participate.

The **Career Fair** will feature several companies from the exhibit floor. All attendees of the Convention, students and professionals alike, are welcome to come talk with representatives from the companies and find out more about job and internship opportunities in the audio industry. Bring your resume! Admission is free and open to all Convention attendees.

Sat., April 28 11:00 Bartók Meeting Room

Technical Committee Meeting on Microphones and Applications

Session P17 Saturday, April 28
11:30 – 13:00 Room Lehar

AUDIO EFFECTS

Chair: **Christof Faller**, Illusonic LLC

11:30

P17-1 Amplitude Manipulation for Perceived Movement in Depth—*Sonia Wilkie, Tony Stockman, Joshua D. Reiss*, Queen Mary University of London, London, UK

The presentation of objects moving in depth toward the viewer (looming) is a technique used in film (particularly those in 3-D) to assist in drawing the viewer into the created world. The sounds that accompany these looming objects can affect the extent to which a viewer can perceptually immerse within the multidimensional world and interact with moving objects. However the extent to which sound parameters should be manipulated remains unclear. For example, amplitude, spectral components, reverb and spatialization can all be altered, but the degree of their alteration and the resulting perception generated, need greater investigation. This paper presents the results from an investigation into one of the sound parameters used as an audio cue in looming scenes by the film industry, namely amplitude, reporting the degree and slope of their manipulation.
Convention Paper 8655

12:00

P17-2 Virtual 5.1 Channel Reproduction of Stereo Sound for Mobile Devices—*Kangeun Lee, Changyong Son, Dohyung Kim, Shihwa Lee*, Samsung Advanced Institute of Technology, Suwon, Korea

With rapid development in mobile devices, consumer demand for a premium sound experience is growing. In this paper a method for 5.1 channel upmixing is introduced. First, the primary and ambience components are separated and the primary signal is decomposed into the standard 5.1 channel direction. Since all separated sources should appear in the upmixed output, we designed the new masking scheme for panning coefficient. In order to reproduce the 5.1 channel sound field in the earphone or headset, the decomposed multichannel signal is virtually rendered by means of HRTF. The proposed method was compared with conventional upmixing methods and demonstrated better spatiality image and panning effects with very low complexity requirements, which allows easy implementation on a wide variety of platforms.
Convention Paper 8656

12:30

P17-3 From Short- to Long-Range Signal Tunneling—*Alexander Carôt*,[1] *Horst Aichmann*[2]
[1]Hochschule Anhalt, Köthen/Anhalt, Germany
[2]Agilent Technologies, Kronberg, Germany

Research results in superluminal pulse transmission indicate propagations with velocities faster than the speed of light. This has so far been considered a short-range effect to be applied within distances of several centimeters. Based on these results and the corresponding theory of quantum tunneling the authors revisit such experiments with distances of several meters. They show that long-range superluminal signal transmission is possible and that it is only restricted by the actual signal bandwidth.
Convention Paper 8657

Sat., April 28 12:00 Bartók Meeting Room

Technical Committee Meeting on Hearing and Hearing Loss Prevention

Workshop 12 Saturday, April 28
12:30 – 14:30 Room Brahms

SPATIAL SOUND REPRODUCTION WITH HEIGHT: WHY, WHERE, HOW, AND WHEN?

Co-Chairs: **Frank Melchior**, BBC R&D, Salford, UK
Florian Völk, Technical University Munich, Munich, Germany

Panelists: *Jan-Mark Batke*, Technicolor, Research and Innovation, Hannover, Germany
Kimio Hamasaki, NHK Science and Technical Research Laboratories, Tokyo, Japan
Michael Kelly, DTS, Inc., London, UK
Stephan Mauer, IOSONO GmbH, Erfurt, Germany
Norbert Niemczyk, Daimler AG, Sindelfingen, Germany
Jan Plogsties, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
Nicolas Tsingos, Dolby Laboratories, Inc., San Francisco, CA, USA

At the present time, expanding spatial audio reproduction to the vertical plane seems to be the next logical step from an engineering point of view. The consequence is that new transmission formats, reproduction and production techniques, and processing methods are required. Before starting the discussion from a technical point of view, we propose to determine applications and environments for such new reproduction systems. The aim of this workshop is to ask and discuss the six W's of information gathering in terms of spatial audio with height: Who is it about? What will happen? Where will it take place? When will it take place? Why will it happen? How will it happen? The workshop will bring together experts of different approaches and use cases where spatial audio reproduction with height is considered being important now and in the future.

Tutorial 5 **Saturday, April 28**
12:30 – 14:30 **Room Bartók**

REALITY IS NOT A RECORDING/ A RECORDING IS NOT REALITY

Presenter: **Jim Anderson**, New York University, New York, NY, USA

The former *New York Times* film critic, Vincent Canby, wrote "all of us have different thresholds at which we suspend disbelief, and then gladly follow fictions to conclusions that we find logical." Any recording is a "fiction," a falsity, even in its most pure form. It is the responsibility, if not the duty, of the recording engineer and producer, to create a universe so compelling and transparent that the listener isn't aware of any manipulation. Using basic recording techniques, and standard manipulation of audio, a recording is made, giving the listener an experience that is not merely logical but better than reality. How does this occur? What techniques can be applied? How does an engineer create a convincing loudspeaker illusion that a listener will perceive as a plausible reality? Recordings will be played.

Sat., April 28 **12:30** **Room Fortuna**

Standards Committee Meeting on Digital Audio Measurement Techniques

Session P18 **Saturday, April 28**
13:00 – 14:30 **Foyer**

POSTERS: EDUCATION AND HUMAN FACTORS; APPLICATIONS

13:00

P18-1 **Optimizing Teaching Room Acoustics: A Comparison of Minor Structural Modifications to Dereverberation Based on Smoothed Responses —***Panagiotis D. Hatziantoniou,*[1] *Nicolas-Alexander Tatlas*[2]
[1]University of Patras, Patras, Greece
[2]Institute of Piraeus, Aigaleo-Athens, Greece

In this work a comparison between traditional acoustical treatment such as building material substitution and digital room acoustics dereverberation is presented for teaching rooms. Measured responses for a number of listening positions for two rooms are shown as well as relevant parameters namely T30, EDT, C-50, and STI. Corresponding values are calculated by employing a simulation model in order to verify its accuracy; minor changes are then introduced to the model aiming to improve speech intelligibility and a new set of parameters is obtained. Finally, dereverberation achieved from inversion of appropriately modified measured room responses based on Complex Smoothing is employed and the acoustical parameters from the filtered impulse responses are derived.
Convention Paper 8658

13:00

P18-2 **Designing an Audio Engineer's User Interface for Microphone Arrays—***Stefan Weigand,*[1,2] *Thomas Görne,*[2] *Johann-Markus Batke*[1]
[1]Technicolor, Research and Innovation, Hannover, Germany
[2]University of Applied Science (HAW), Hamburg, Germany

Microphone arrays are rarely used in artistic recordings, despite the benefits they offer. We think this is due to a lack of user interfaces considering audio engineers' needs, enabling them to address the features in an easy and suited way. This paper contributes to solving this problem by outlining guidelines for such interfaces. A graphical user interface (GUI) for microphone arrays employing Higher Order Ambisonics (HOA), incorporating the audio engineers' aims and expectations, has been developed by analyzing their common activities. The presented solution offers three operation modes, covering the most frequent tasks in (professional) audio productions, thus making it more likely to engage audio engineers in using microphone arrays in the future.
Convention Paper 8659

13:00

P18-3 **User Interface Evaluation for Discrete Sound Placement in Film and TV Post-Production—** *Braham Hughes, Jonathan Wakefield*, University of Huddersfield, Huddersfield, UK

This paper describes initial experiments to evaluate the effectiveness of different 3-D input interfaces combined with visual feedback for discrete sound placement in film and TV post-production. The experiments required the user to control the 3-D position of a sound object for a moving target object within a video clip on screen using a range of physical interfaces with and without visual feedback. Inclusion of visual feedback had a statistically significant impact on the accuracy of the tracking of the target object. The Wii remote controller appeared to perform the best in the tests and in the user preference ranking. The traditional desk-based input method performed worst in all tests and the user preference ranking.
Convention Paper 8660

13:00

P18-4 **KnuckleTap—Exploring the Possibilities of Audio Input in a Mobile Rhythmic Notepad Application—***Julian Rubisch,*[1] *Michael Jaksche*[2]
[1]breakingwav.es, Vienna, Austria
[2]University of Applied Sciences, St. Poelten, Austria

Apart from some significant contributions in the scientific community and a few notable product innovations, audio input for musical interaction purposes on mobile devices has until now been widely neglected as a control parameter. This situation does not correspond to the fact that musical ideas are often vocalized and developed by musicians using their voice or other sounding ➥

objects, nor does it take advantage of contemporary mobile devices' most rigid and reliable sensor—the microphone. As an example case exploiting these possibilities, we conceived a notepad application to track rhythmic ideas by recording taps on a surface with a smartphone's built-in microphone, refined by subsequent detection of onsets, clustering, instantaneous rearranging of the detected events, and export capabilities.
Convention Paper 8661

13:00

P18-5 An Optical System to Track Azimuth Head Rotations for Use in Binaural Listening Tests of Automotive Audio Systems—*Anthony Price*, Bang and Olufsen a/s, Struer, Denmark, presently at University of Surrey, Guildford, Surrey, UK

Binaural technology is used to capture elements of an automotive audio system and reproduce them over headphones. This requires the tracking of azimuth head rotations of listening test participants in order to assist source localization. The parameters and faults of the currently implemented system are discussed and a new method of tracking azimuth head rotations is described. The system is tested, implemented, and found to have an error within 0.26 degrees. The potential for its further development, and the development of the field, is discussed.
Convention Paper 8662

13:00

P18-6 Investigation of Salient Audio-Features for Pattern-Based Semantic Content Analysis of Radio Productions—*Rigas Kotsakis, George Kalliris, Charalampos Dimoulas*, Aristotle
University of Thessaloniki, Thessaloniki, Greece

The paper focuses on the investigation of salient audio features for pattern-based semantic analysis of radio programs. Most "news and music" radio programs have many structure similarities with respect to the appearance of different content types. Speech and music are continuously interchanged and overlapped, whereas specific speakers and voice patterns are more important to recognize. Recent research showed that various taxonomies and hierarchical classification schemes can be effectively deployed in combination with supervised and unsupervised training for semantic audio content analysis. Undoubtedly, audio feature extraction and selection is very important for the success of the finally trained expert system. The current paper employs feature ranking algorithms, investigating audio features saliency in various classification taxonomies of radio production content.
Convention Paper 8663

13:00

P18-7 Listeners Who Have Low Hearing Thresholds Do Not Perform Better in Difficult Listening Tasks—*Piotr Kleczkowski, Marek Pluta, Paulina Macura, Elzbieta Paczkowska*, AGH University of Science and Technology, Krakow, Poland

The relationship between measures of hearing acuity and performance in listening tasks for normally hearing subjects has not found a solid evidence. In this work six one-parameter measures of hearing acuity, based on audiograms, were used to investigate whether a relationship between those measures and listeners' performance existed. The quantifiable results of several listening tests were analyzed, using speech and non-speech stimuli. The results showed no correlation between hearing acuity and performance thus demonstrating that hearing acuity should not be a critical factor in the choice of listeners.
Convention Paper 8641
[Paper presented by Paulina Macura]

Sat., April 28 13:00 Bartók Meeting Room

Technical Committee Meeting on Transmission and Broadcasting

Session P19 Saturday, April 28
14:30 – 17:30 Room Lehar

SPATIAL AUDIO: PART 2

Chair: **Antti Kelloniemi**, Nokia Corporation, Espoo, Finland

14:30

P19-1 Sound Field Reproduction Method in Spatio-Temporal Frequency Domain Considering Directivity of Loudspeakers—*Shoichi Koyama, Ken'ichi Furuya, Yusuke Hiwasaki, Yoichi Haneda*, NTT Cyber Space Laboratories, NTT Corporation, Musashino-shi, Tokyo, Japan

A method for transforming received signals of a microphone array into driving signals of a loudspeaker array for sound field reproduction is needed to achieve real-time sound field transmission systems from the far-end to the near-end. We recently proposed a transform method using planar or linear microphone and loudspeaker arrays in the spatio-temporal frequency domain, which is more efficient than conventional methods based on a least squares algorithm. In this method, the directivity of loudspeakers in the array is assumed to be omnidirectional to derive the transform filter. However the directivity of common loudspeakers is not always omnidirectional, especially at high frequencies. We therefore propose a transform method that takes into consideration the directivity of loudspeakers in the array, which is derived using analytical and numerical approaches. Numerical simulation results indicated that the accurately reproduced region of the proposed method was larger than that of the method with an omnidirectional assumption.
Convention Paper 8664

15:00

P19-2 Practical Applications of Chameleon Subwoofer Arrays—*Adam J. Hill, Malcolm O. J. Hawksford*, University of Essex, Colchester, Essex, UK

Spatiotemporal variations of the low-frequency

response in a closed-space are predominantly caused by room-modes. Chameleon subwoofer arrays (CSA) were developed to minimize this variance over a listening area using multiple independently-controllable source components and calibrated with one-time measurements. Although CSAs are ideally implemented using hybrid (multiple source component) subwoofers, they can alternatively be realized using conventional subwoofers. This capability is exploited in this work where various CSA configurations are tested using commercially-available subwoofers in a small-sized listening room. Spectral and temporal evaluation is performed using tone-burst and maximum length sequence (MLS) measurements. The systems are implemented with practicality in mind, keeping the number of subwoofers and calibration measurements to a minimum while maintaining correction benefits.
Convention Paper 8665

15:30

P19-3 Localization in Binaural Reproduction with Insert Headphones—*Marko Hiipakka, Marko Takanen, Symeon Delikaris-Manias, Archontis Politis, Ville Pulkki*, Aalto University, Espoo, Finland

Circumaural headphones are commonly used in binaural reproduction and it is well known that individual equalization of the headphones improves the quality of the reproduction. The suitability of insert headphones to binaural reproduction has not been studied partly due the lack of a commonly accepted individual equalization method for insert headphones. Recently, a method to estimate the frequency response evoked by insert headphones has been presented. In this paper the localization accuracy of test subjects is evaluated in binaural listening with insert headphones and high-quality circumaural headphones. The results show that the accuracy with inserts is similar to that with circumaural headphones when the recently proposed method is applied for equalization, which motivates their use in binaural reproduction.
Convention Paper 8666

16:00

P19-4 A Comparative Evaluation between Numerical Techniques for Implementing the Acoustic Diffusion Equation Model—*Juan M. Navarro,*[1] *Juan E. Noriega*,[1] *Jose Escolano*,[2] *Jose J. Lopez*[3]
[1]San Antonio's Catholic University of Murcia, Guadalupe, Spain
[2]University of Jaén, Linares, Spain
[3]Universidad Politécnica de Valencia, Valencia, Spain

The acoustic diffusion equation model is an energy-based model that is being successfully applied in room acoustics for predicting the late part of the decay, in the past few years. Early researches usually used a finite element method to solve the diffusion equation model. Recently, an alternative implementation, using finite difference methods has been proposed. A comparison between both numerical techniques could be helpful to clarify the pros and cons of each method. In this paper this evaluation is made by several simulations in a cubic shaped room. Both prediction accuracy and computational performance are compared using different absorption distributions. It is suggested that the finite difference implementation is less computationally intensive than the finite element method. Moreover, the obtained values in the simulations are accurate, at least as well as other geometrical models.
Convention Paper 8667

16:30

P19-5 A Bayesian Framework for Sound Source Localization—*José Escolano*,[1] *Maximo Cobos*,[2] *Jose M. Pérez-Lorenzo*,[1] *José J. López*,[3] *Ning Xiang*[4]
[1]University of Jaén, Linares, Spain
[2]University of Valencia, Burjassot, Valencia, Spain
[3]Universidad Politécnica de Valencia, Valencia, Spain
[4]Rensselaer Polytechnic Institute, Troy, NY, USA

The localization of sound sources, and particularly speech, has a numerous number of applications to the industry. This has motivated a continuous effort in developing robust direction-of-arrival detection algorithms. Time difference of arrival-based methods, and particularly, generalized cross-correlation approaches have been widely investigated in acoustic signal processing. Once a probability function distribution is obtained, indicating those directions of arrival with highest probability, the vast majority of methods have to assume a certain number of sound sources in order to process the information conveniently. In this paper a model selection based on a Bayesian framework is proposed in order to determine, in an unsupervised way, how many sound sources are estimated. Real measurements using two microphones are used to corroborate the proposed model.
Convention Paper 8668

17:00

P19-6 A Comparison of Modal versus Delay-and-Sum Beamforming in the Context of Data-Based Binaural Synthesis—*Sascha Spors, Hagen Wierstorf, Matthias Geier*, Deutsche Telekom Laboratories, Technische Universität Berlin, Berlin, Germany

Several approaches to data-based binaural synthesis have been published that capture a sound field by means of a spherical microphone array. The captured sound field is decomposed into plane waves that are then auralized using head-related transfer functions (HRTFs). The decomposition into plane waves is often based upon modal beamforming techniques that represent the captured sound field with respect to surface spherical harmonics. An efficient and numerically stable approximation to modal beamforming is the delay-and-sum technique. This paper compares these two beamforming techniques in the context of data-based binaural synthesis. Their frequency- and time-domain properties are investigated, as well as the perceptual properties of the resulting binaural synthesis according to a binaural model.
Convention Paper 8669

Workshop 13 **Saturday, April 28**
14:30 – 16:30 **Room Brahms**

DEPLOYING THE LOUDNESS CONCEPT IN EUROPE—HOW, WHEN AND WHERE

Chair: **Florian Camerer**, ORF, Vienna, Austria

Panelists: *Alfio DiFazio*, tpc Switzerland
Matthieu Parmentier, France Television
Askan Siegfried, NDR
Alessandro Travaglini, Fox International Channels Italy
Richard Van Everdingen, Dutch Loudness Committee

European broadcasters are gearing up fast putting the concept of loudness normalization into practice on the basis of the EBUB recommendation R 128. The panelists are representatives of broadcasters who are on the forefront of this development and will shed a light on the process, the intricacies and solutions.

Tutorial 6 **Saturday, April 28**
14:30 – 16:30 **Room Liszt**

DRUM AND PERCUSSION PROGRAMMING

Presenter: **Justin Paterson**, London College of Music, University of West London, London, UK

Drum programming has been evolving at the heart of many studio productions for some 30 years. Over this period, technological opportunities for enhanced creativity have multiplied in numerous directions, from sequenced MIDI one-shots to sample loops, and from DAW cut and stretch techniques to deterministic beat-slicer plug-ins, etc. The palette of sounds available today ranges from ever more realistic to ever more synthetic/exotic. This tutorial will embrace all of these techniques and more—and include numerous live demonstrations. Although introducing all key concepts from scratch, its range and hybridization should provide inspiration even for experienced practitioners, leading up to the state-of-the-art. A number of genres will be covered from the pseudo-realism of jazz and funk, to the exploitation of synthetic textures in “intelligent dance music.”

Tutorial 7 **Saturday, April 28**
15:00 – 16:30 **Room Bartók**

HOW DOES IT SOUND NOW? THE EVOLUTION OF AUDIO

Moderator: **Gary Gottlieb**

Panelists: *Jim Anderson*
Andres Mayo
Ronald Prent
Darcy Proper

One day Chet Atkins was playing guitar when a woman approached him. She said, “That guitar sounds beautiful.” Chet immediately quit playing. He asked, “How does it sound now?” The quality of sound in Chet’s case clearly rested with the player, not the instrument, and the technical and aesthetic quality of our product lies with our engineers and producers, not solely the equipment. The dual significance of this question, “How does it sound now,” informed my research for the last three years and will inform our discussion, since it addresses both the engineer as the driver and the changes we have seen and heard as our methodology evolved through the decades. The book that resulted from this research, *How Does It Sound Now?* received the 2010 ARSC Award for Excellence in Historical Recorded Sound Research for the Best Research in General History of Recorded Sound. One of the most interesting facets of the research, comprised of interviews with top engineers and producers, was the way the conversation kept returning to the thread of quality. They loved to talk about how they strived for quality then, and still do. Let’s talk about how engineers and producers retain quality and create a product that conforms to their own high standards. This may lead to other conversations about musicians, consumers, and the differences and similarities between their standards and our own. How high should our standards be? How did it sound then? How does it sound now? How should it sound tomorrow?

Session EB2 **Saturday, April 28**
15:30 – 17:00 **Foyer**

ENGINEERING BRIEFS—POSTERS

15:30

EB2-1 An Exact Harmonic Computing Technique for Polynomial Nonlinearities—*Nay Oo*,[1] *Woon-Seng Gan*[2]
[1]Fraunhofer Institute for Digital Media Technology, IDMT, Oldenburg, Germany
[2]Nanyang Technology University, Singapore, Singapore

An exact harmonic computing technique for polynomial nonlinearities is developed from first principles. By applying this technique, when an input sinusoid with arbitrary amplitude, frequency and phase information and a polynomial nonlinearity are given, the exact computation of DC component and output harmonics’ amplitudes, frequencies, and phases is computationally possible without discrete Fourier transform (DFT). Two basic mathematical results such as power of cosine and harmonic addition theorem are utilized to develop this technique.
Engineering Brief 47

15:30

EB2-2 Suspension Creep Models for Miniature Loudspeakers—*Holger Hiebel***,** Knowles Electronics Austria GmbH, Vienna, Austria, Graz University of Technology, Graz, Austria

Different models for describing the suspension creep behavior of loudspeakers are available at the moment. They include the “LOG”-model from Knudsen, the model used in the Klippel Analyzer system (same as Knudsen, but neglecting the frequency dependent losses) and the 3-parameter model of Ritter and Agerkvist. In contrast to HiFi speakers the differences between the creep models can be seen more clearly in miniature loudspeakers using compound (multilayer) membranes. Benefits from choosing the right creep model are tabulated in terms of RMS fit errors and evaluation results for miniature loudspeaker samples are shown.
Engineering Brief 48

15:30

EB2-3 Smart Microphone Sensor System Platform— *Elias Kokkinis,*[1] *Konstantinos Drossos,*[2] *Nicolas-Alexander Tatlas,*[3] *Andreas Floros,*[2] *Alexandros Tsilfidis,*[1] *Kyriakos Agavanakis*[3]
[1]University of Patras, Patras, Greece
[2]Ionian University, Corfu, Greece
[3]BLUEdev Ltd., Patras, Greece

A platform for a flexible, smart microphone system using available hardware components is presented. Three subsystems are employed, specifically: (a) a set of digital MEMs microphones, with a one-bit serial output; (b) a preprocessing/digital-to-digital converter; and (c) a CPU/DSP-based embedded system with I2S connectivity. Basic preprocessing functions, such as noise gating and filtering can be performed in the preprocessing stage, while application-specific algorithms such as word spotting, beam-forming, and reverberation suppression can be handled by the embedded system. Widely used high-level operating systems are supported including drivers for a number of peripheral devices. Finally, an employment scenario for a wireless home automation speech activated front-end sensor system using the platform is analyzed.
Engineering Brief 49

15:30

EB2-4 Sound Field Synthesis Toolbox—*Hagen Wierstorf, Sascha Spors*, TU Berlin, Berlin, Germany

An open source toolbox for Sound Field Synthesis (SFS) is introduced. The toolbox is able to numerically simulate sound fields synthesized by SFS methods like Wave Field Synthesis or higher order Ambisonics. Various loudspeaker driving signals for the mentioned methods are provided for 2-, 2.5-, and 3-dimensional synthesis. The toolbox allows mono-frequent as well as broadband excitation signals. The latter allows generation of snapshots of the spatio-temporal impulse response of a chosen reproduction technique. The toolbox furthermore includes the computation of binaural room impulse responses (BRIR) for a given SFS setup. These can be used to simulate different sound field synthesis methods via binaural resynthesis. The toolbox is provided for Matlab/Octave and comes with an online documentation.
Engineering Brief 50

15:30

EB2-5 Listening to the Large Hadron Collider— *Daniel Deboy,*[1,2] *Ralph W. Aßmann,*[1] *Roderik Bruce,*[1] *Florian Burkart,*[1] *Marija Cauchi,*[1] *Clement Derrez,*[1] *Alessandro Masi,*[1] *Stefano Redaelli,*[1] *Belen Salvachua,*[1] *Gianluca Valentino,*[1] *Daniel Wollmann*[1]
[1]CERN, Geneva, Switzerland
[2]University of Music and Performing Arts, Graz, Austria

The Large Hadron Collider (LHC) at CERN is a high-energy particle accelerator in a 27 km long tunnel located in the underground of the Geneva area, Switzerland. Protons are accelerated to 99.9999991 percent of the speed of light before they collide with a total momentum of up to 14 TeV. It is the largest machine that human mankind has ever built. Over 10.000 sensors monitor the state of the LHC during operation. Recently, microphones have been added as an experimental setup. An acoustic monitoring system to detect and localize beam accidents is under current investigation. Such a system may reduce expensive downtime dramatically in an accident scenario. The acquired signals can also be used for other applications, e.g., sonification, media art installations, etc.
Engineering Brief 51

Sat., April 28 16:00 Bartók Meeting Room

Technical Committee Meeting on Sound for Digital Cinema & Television (Formative)

Tutorial 8 Saturday, April 28
16:30 – 18:00 Room Liszt

SEMANTIC WEB AND SEMANTIC AUDIO TECHNOLOGIES

Presenters: **Gyorgy Fazekas**, Queen Mary University of London, London, UK
Thomas Wilmering

The emerging Semantic Web provides a powerful framework for the expression and reuse of structured data. Recent efforts have brought this framework to bear on the field of Semantic Audio, as well as information management in audio applications. This tutorial will provide an introduction to Semantic Web concepts and how they can be used in the context of music-related studies. We will outline the use of the Resource Description Framework (RDF) and related ontology and query languages. Using practical examples, we will demonstrate the use of the Music and Studio Ontologies, and show how they facilitate interoperability between audio applications and linked data sets on the Web. We will explore how signal processing tools and results can be described as structured data and utilized in audio production.

Tutorial 9 Saturday, April 28
16:30 – 18:00 Room Brahms

LISTENING TESTS PART 2: STATISTICAL ANALYSIS

Presenters: **Poppy Crum**, Dolby Laboratories, San Francisco, CA, USA
Frederik Nagel, Fraunhofer Institute for Integrated Circuits IIS/International Audio Laboratories, Erlangen, Germany
Thomas Sporer, Fraunhofer Institute for Digital Media Technology, IDMT, Ilmenau, Germany

This tutorial aims at discussing current practices of listening test evaluation. Topics to be covered are post-screening of tests results enabling to reject data from unreliable subjects and analysis of the results. This will cover parametric and non-parametric statistics, and the decision when which of these is necessary. Special emphasis will be set on which conclusions from the analysis are adequate from a mathematical perspective. This tutorial aims at providing researchers with a better understanding of the results of listening tests.

Student/Career Development Event
RECORDING COMPETITION—PART 2
Saturday, April 28, 16:30 – 18:30
Room Bartók

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. This event presents stereo and surround recordings in these categories:

• Traditional Multitrack Recording – 16:30–17:30

• Non-Traditional Multitrack Recording – 17:30–18:30

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 Sunday 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 if your project isn't one of the finalists, and it's a great chance to meet other students and faculty members.

Sat., April 28 16:30 Room Fortuna

Standards Committee Meeting on Audio Applications of Networks

Special Event
ORGAN RECITAL BY GRAHAM BLYTH
Saturday, April 28, 20:00 – 21:30
Belvárosi plébánia templom (Inner City Parish Church)
1056 Budapest, Március 15. tér 2

Graham Blyth's traditional organ recital will be given on the new organ of the Inner-City Parish Church, Belvarosi Plebania Templom. The location at the Pest side of Elisabeth Bridge, across from Gellert Hill, is originally the gravesite of the martyr, Bishop Saint Gellért. Dating to the 12th century, this is the oldest—and, in terms of architectural history, the most extraordinary—ecclesiastical building in Pest. It's actually built on even older structures-an 11th-century Romanesque church commissioned by Hungary's first king, Saint Stephen, plus the remains of the Contra Aquincum (a third-century Roman fortress and tower), parts of which are visible. There is hardly any architectural style that cannot be found in some part or another, starting with a single Romanesque arch in its south tower. The single nave, with an equally high transept and unusually slender columns, still has its original Gothic chancel and some frescoes from the 14th-15th centuries. Two side chapels contain beautifully carved Renaissance altarpieces and tabernacles of red marble from the early 16th century. During Budapest's Turkish occupation the church served as a mosque—a mihrab, a Muslim prayer niche, is a reminder. During the 18th century the church was given two baroque towers and its present facade. In 1808 it was enriched with a rococo pulpit. From 1867 to 1875 Franz Liszt lived in a town house a few steps away, where he held regular "musical Sundays" at which Richard and Cosima Wagner were frequent guests. Liszt's own musical Sunday mornings often began in this church. He conducted many masses here, including the first Budapest performance of his Missa Choralis in 1872. The church contains the relics of Saint Gellért, the bishop who was initially buried here in 1046 after pagans pushed him off a hill across the river. And in 2006 the relics of the legendary 11th-century Hungarian king, Saint László, found their way here as well.

The program will include Prelude & Fugue in B minor by Bach, Chorale No.2 in B minor by Franck, and 1st Symphony by Vierne (selected movements).

Graham Blyth was born in 1948, began playing the piano aged 4 and received his early musical training as a Junior Exhibitioner at Trinity College of Music in London, England. Subsequently, at Bristol University, he took up conducting, performing Bach's St. Matthew Passion before he was 21. He holds diplomas in Organ Performance from the Royal College of Organists, The Royal College of Music and Trinity College of Music. In the late 1980s he renewed his studies with Sulemita Aronowsky for piano and Robert Munns for organ. He gives numerous concerts each year, principally as organist and pianist, but also as a conductor and harpsichord player. He made his international debut with an organ recital at St. Thomas Church, New York in 1993 and since then has played in San Francisco (Grace Cathedral), Los Angeles (Cathedral of Our Lady of Los Angeles), Amsterdam, Copenhagen, Munich (Liebfrauendom), Paris (Madeleine and St. Etienne du Mont) and Berlin. He has lived in Wantage, Oxfordshire, since 1984 where he is currently Artistic Director of Wantage Chamber Concerts and Director of the Wantage Festival of Arts.

He divides his time between being a designer of professional audio equipment (he is a co-founder and Technical Director of Soundcraft) and organ related activities. In 2006 he was elected a Fellow of the Royal Society of Arts in recognition of his work in product design relating to the performing arts.

Special Event
BANQUET
Saturday, April 28, 21:00 – 23:30
Academy Club
Hungarian Academy of Sciences
9 István Széchenyi square, Budapest

This year, the Banquet will be held at the Academy Club, part of the Hungarian Academy of Sciences. This elegant and exclusive building is a perfect setting for dining and mixing with friends and colleagues. One can also indulge in the breathtaking views of the Buda Castle and enjoy the stunning beauty of Budapest and the Danube illuminated in the evening while listening to intimate live music. An exciting choice of seasonal dishes combining the traditional Hungarian flavors with modern presentation will be offered, together with wines and other drinks.

Situated near to the church where the organ concert with be held, together these two events make for a perfect evening after a day at the convention.

60 Euros for AES members and nonmembers
Tickets will be available at the Special Events desk.

Session P20 Sunday, April 29
09:00 – 11:30 Room Lehar

TRANSDUCERS

Chair: **David Griesinger,** David Griesinger Acoustics, Cambridge, MA, USA

09:00

P20-1 Loudspeaker for Low Frequency Signal Driven by Four Piezoelectric Ultrasonic Motors—*Juro Ohga*,[1] *Ryousuke Suzuki,*[2] *Keita Ishikawa*,[2] *Hirokazu Negishi*,[3] *Ikuo Oohira*,[4] *Kazuaki Maeda*,[5] *Hajime Kubota*[2]

[1]Shibaura Institute of Technology/MIX Corporation, Kamakura, Japan
[2]Chiba Institute of Technology, Narashino, Japan
[3]MIX Corporation, Yokosuka, Japan
[4]I. Oohira and Associates, Yokohama, Japan
[5]TOA Corporation, Takarazuka, Japan

The authors have been developing a completely new direct-radiator loudspeaker construction that is driven by continuous revolution of piezoelectric ultrasonic motors. It converts continuous revolution of ultrasonic motors to reciprocal motion of a cone radiator. This loudspeaker shows almost flat phase frequency characteristics in low frequency region, because it includes no resonance in low frequency region. Therefore it is useful for radiation of the lowest frequency part of the audio signal. At this convention the authors are going to present a practical model of this loudspeaker driven by co-operation of four ultrasonic motors.
Convention Paper 8670

09:30

P20-2 Comprehensive Measurements of Head Influence on a Supercardioid Microphone—*Hannes Pomberger,*[1] *Franz Zotter,*[1] *Dominik Biba*[2]

[1]University of Music and Performing Arts Graz, Graz, Austria
[2]AKG Acoustics GmbH, Vienna, Austria

The directional pickup pattern of microphones is designed as to assist the audio engineer in avoiding acoustic feedbacks or interference from other sound sources. If the pattern deviates from the specified one, it is important for the audio engineer to know. This paper presents comprehensive measurements of a supercardioid microphone directivity under the influence of a dummy head. This dummy models the diffraction of a human talker or singer in front of the microphone. The discussed measurements collect information on a 15x15 degree grid in azimuth and elevation for different distances between the microphone and dummy head. Based on this data, we are able to discuss the influence of a singer's body on the free field directivity of the supercardioid microphone, its directivity index, and its front-to-back random ratio in detail.
Convention Paper 8671

10:00

P20-3 Simulation of a 4-Inch Compression Driver Using a Fully Coupled Vibroacoustic Finite Element Analysis including Viscous and Thermal Losses—*René Christensen, Ulrik Skov*, iCapture ApS, Gadstrup, Denmark

A 4-inch JBL compression driver is simulated using a finite element analysis, FEA. Compared to a conventional electrodynamic driver a compression driver has a phase plug with slits in front of the diaphragm. The slits are acoustically narrow and the diaphragm is separated from the phase plug only by a thin gap so an accurate model must include viscothermal effects to account for the losses associated with the narrow air gaps. Air domains, structural domains, and viscothermal domains are all fully coupled to ensure the proper continuity of their variables. Simulated results are compared to experimental measurements and it is demonstrated how the viscothermal effects dampen out acoustic and structural modes.
Convention Paper 8672

10:30

P20-4 Design of Vented Boxes Using Current Feedback Filters—*Juha Backman,*[1] *Tim Mellow*[2]

[1]Nokia Corporation, Espoo, Finland
[2]MWT Acoustics, Farnham, UK

A current feedback arrangement for a loudspeaker system that can be tuned to provide a pre-determined frequency-response shape over a fairly wide and continuous range of box volumes is discussed. A conventional high-pass filter only allows the system to be tuned to give a particular frequency-response shape if the box volume is correct. The traditional arrangement is either a flat response amplifier, a passive filter between the amplifier and loudspeaker, or an active filter before the amplifier. This paper discusses an alternative arrangement where current feedback and filter provide the desired amplifier output impedance and voltage transfer function characteristics. These interact directly with the complex load-impedance of the loudspeaker. Practical realizations of the current feedback implementations are presented.
Convention Paper 8673

11:00

P20-5 Midrange Coloration Caused by Resonant Scattering in Loudspeakers—*Juha Backman*, Nokia Corporation, Espoo, Finland

One of the significant sources of midrange coloration in loudspeakers is the resonant scattering of the exterior sound field from ports, recesses, or horns. This paper discusses a computationally efficient model for such scattering, based on waveguide models for the acoustical elements (ports, etc.), and mutual radiation impedance model for their coupling to the sound field generated by the drivers. This allows rapid evaluation of the effect of port placement, suitable for numerical optimization of loudspeaker enclosure layouts.
Convention Paper 8674

Workshop 14 **Sunday, April 29**
09:00 – 10:00 **Room Brahms**

SEMANTIC AUDIO ANALYSIS IN THE REAL WORLD

Chair: **Andrew Mason**, BBC Research & Development, London, UK

Panelists: *Andy Hill*, I Like Music, London, UK
Mark Sandler, Queen Mary University of London, London, UK
Nigel Smith, BBC Audio and Music Interactive, London, UK
Christian Uhle, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Semantic audio analysis presents many opportunities for ➥

research, and a huge amount of academic effort is expended. Naturally, a significant proportion of the effort is directed toward solving problems of interest to those doing the research. As a result there is a perception that many of the emergent applications are of limited commercial interest: a musicologist might be interested in minutiae of a composer's style, or in creating a play-list from a collection of 10,000 MP3 files. However, in the wider world there are other potential users: archivists with hundreds of thousands of hours of material to manage, and huge value to be realized from their archives; program makers searching for sound bites; and so on. This workshop aims to present some industry needs and to seek out new opportunities for developments in semantic audio analysis.

Tutorial 10 **Sunday, April 29**
09:00 – 10:30 **Room Liszt**

NOISE ON THE BRAIN—HEARING DAMAGE ON THE OTHER SIDE

Presenter: **Poppy Crum**, Dolby Laboratories, San Francisco, CA, USA

Did you know that drinking a glass of orange juice every day may actually protect your hearing? Most discussions of hearing damage focus on what happens to the cochlea and inner ear. While this understanding is crucial to predicting and avoiding trauma that can lead to hearing loss, both acoustic and chemical stimuli can also have significant effects on higher brain areas. In some cases, thresholds and audiograms can look completely normal but listeners may have great difficulty hearing a conversation in a noisy environment. This session will explore the latest research regarding the effects of acoustic and chemical trauma, and how this damage manifests throughout the auditory pathway as changes in hearing sensitivity, cognition, and the experience of tinnitus. We will also consider recent research in chemically preserving hearing and combating these conditions with supplements as common as Vitamin C!

Tutorial 11 **Sunday, April 29**
09:00 – 10:30 **Room Bartók**

DESIGN OF A DYNAMIC RANGE COMPRESSOR

Presenter: **Josh Reiss**, Queen Mary University of London, London, UK

Despite being one of the most widely used audio effects, dynamic range compression, remains poorly understood. And there is little formal knowledge and analysis of compressor design techniques. This tutorial will describe several approaches to digital dynamic range compressor design. It will demonstrate how to build a compressor from the ground up, and provide audio examples showcasing differences between designs. It will explain why the designs sound different, and provide distortion-based metrics to analyze their quality. It will also provide recommendations for high performance compressor design.

Sun., April 29 **09:00** **Room Fortuna**

Standards Committee Meeting: AESSC Plenary

Sun., April 29 **10:00** **Bartók Meeting Room**

Technical Committee Meeting on Perception and Subjective Evaluation of Audio Signals

Workshop 15 **Sunday, April 29**
10:30 – 12:30 **Room Brahms**

SYNCHRONIZATION IN A MULTICHANNEL DIGITAL AGE

Chair: **John Grant**

Panelists: *Steve Harris*
Stefan Heinzmann
Thomas Sporer, Fraunhofer Institute for Digital Media Technology, IDMT, Ilmenau, Germany

Once upon a time, not so long ago, in any system built up from separate items of digital audio equipment everything would be locked to house sync. Now, systems are increasingly built from equipment that is connected to an asynchronous network without a separate sync input; indeed, in many environments (such as in the home) there is no house sync available. And, increasingly, sound fields are produced by independent digitally-interfaced loudspeakers. This workshop examines requirements for synchronization in these scenarios and techniques for achieving it.

Workshop 16 **Sunday, April 29**
11:00 – 13:00 **Room Bartók**

LISTEN PROFESSIONALLY OR TRAIN YOUR EAR!

Chair: **Sungyoung Kim**, Yamaha Corporation, Hamamatsu, Japan

Panelists: *Kazuhiko Kawahara*, University of Kyusyu, Fukuoka, Japan
Doyuen Ko, McGill University, Montreal, Quebec, Canada
Søren Vase Legarth, Delta SenseLab, Hørsholm, Denmark
Atsushi Marui, Tokyo University of the Arts, Tokyo, Japan
Mark McKinnon-Bassett, University of Sydney, Sydney, NSW, Australia
Sean Olive, Harman International, Northridge, CA, USA

It has been generally accepted that critical listening ability is essential for audio engineers. Recent training programs provide multiple trainees with fast acquisition of such listening ability through a systematic curriculum optimized for the required task. Considering the interests and growth of ear training in the audio communities, it is timely and important to have a chance to share and discuss the opinions from the experts about necessary features and methods that assist trainees in acquiring the critical listening ability with efficiency, both for personal and group training. For this purpose, a workshop at the previous AES 131st convention invited panelists from all around the world who shared their in-depth experience, know-how, and insights in ear training. The current workshop is an extension of the previous one that aims to let workshop attendees experience and compare the characteristic functions of various ear training programs through hands-on demonstrations by the panelists. While the workshop locally aims to provide the attendees with a chance to experience theoretical and empirical matters of ear training programs around the world, it also globally aims to consider the importance of "listening" in the current video-oriented society.

Tutorial 12 **Sunday, April 29**
11:00 – 13:00 **Room Liszt**

LARGE ROOM ACOUSTICS: BASICS AND NEW DEVELOPMENTS

Presenter: **Diemer de Vries**, Amsterdam, The Netherlands

In the sound recording and reproduction chain, the audio engineer has to deal with the acoustics of the environment—usually an enclosed space, a "room"—he or she is working in. The physical properties of the acoustics of a room can be studied by advanced measurements or simulations, but what counts at the end is the perceptual quality of the acoustics. The challenge of the modern room acoustics researcher is to find an unambiguous relation between physics and perception. In the tutorial —which is focused on large rooms—it will be discussed how, long ago, only "reverberation time" was considered as the important parameter. Later, and until now, parameters were derived from impulse responses. Recent developments show that also these responses may have had their longest time as the scientific foundation of room acoustics.

Session P21 **Sunday, April 29**
12:00 – 13:30 **Foyer**

POSTERS: AUDIO EQUIPMENT AND INSTRUMENTATION

12:00

P21-1 Evaluation of Vibrating Sound Transducers with Glass Membrane Based on Measurements and Numerical Simulations—*György Wersényi*, Széchenyi István University, Györ, Hungary

In recent years manufacturers introduced so-called "invisible sound" solutions. In-wall, surface mount, or glass mount versions of different vibrating transducers are commercially available. The entire surface becomes a loudspeaker delivering sound, and the frequency response is said to be equivalent to conventional diaphragm speakers. Furthermore, the sound is omnidirectional at nearly all frequencies (60 Hz–15 kHz) while channel separation is maintained. This paper presents measurement results of the SolidDrive SD1g transducer mounted on different glass surfaces, including vibration measurements and acoustic parameters. Furthermore, based on a numerical FEM-model using COMSOL, comparison between measured and simulated results and estimation of transfer function and directional characteristics are presented.
Convention Paper 8675

12:00

P21-2 Voice Coil Inter-Turn Faults Modeling and Simulation—*German Ruiz, J. A. Ortega, J. Hernández*, UPC-Universitat Politecnica de Catalunya, Terrassa, Spain

The purpose of this paper is to present a new model to study inter-turn short circuit faults in a dynamic loudspeaker. The loudspeaker is modeled by using a classical voice coil parametric model attached to a mechanical piston, and the equations are modified to take into account the voice coil inter-turn faults. The loudspeaker model is global and can work in both normal and fault conditions due to a fictitious resistance in the winding circuit. Various simulation results have been presented indicating the fault instant in time and its corresponding effect in power spectral density. The model can serve as a step toward development of fault detection and diagnosis algorithm.
Convention Paper 8676

12:00

P21-3 Headphone Selection for Binaural Synthesis with Blocked Auditory Canal Recording—*Florian Völk*, AG Technische Akustik, MMK, Technische Universität München, Munich, Germany

Binaural synthesis aims at eliciting the reference scene hearing sensations by recreating the sound pressures at the eardrums, typically using headphones. If all transfer functions involved are approximated based on eardrum probe microphone or traditional artificial head measurements, the headphones have been shown not to influence the synthesis. It is also possible to achieve correct binaural synthesis with transfer functions measured at the entrances to the blocked auditory canals. Then, the headphones may influence the results. In this paper a blocked auditory canal headphone selection criterion (HPC) for binaural synthesis is proposed. Further, a procedure is derived, which allows to evaluate the (HPC) for specific circum-aural headphones based on four measurements using a specifically designed artificial head.
Convention Paper 8677

12:00

P21-4 A Low Latency Multichannel Audio Processing Evaluation Platform—*Yonghao Wang,*[1] *Xiangyu Zhu,*[2] *Qiang Fu*[3]
[1]Queen Mary University of London, London, UK
[2]Hebel University of Science and Technology, Shijiazhuang, China
[3]Shijiazhuang Mechanical Engineering College, Shijiazhuang, China

For live digital audio system with high-resolution multichannel functionalities, it is desirable to have accurate latency control and estimation over all of the stages of the digital audio processing chain. The evaluation system we designed supports 12 channel- 24-bit sigma delta based ADC/DAC, incorporating both a programmable FPGA and digital signal processor. It can be used for testing and evaluation of different ADC/DAC digital filter architectures, audio sample buffer subsystem design, interrupt and scheduling, high level audio processing algorithms, and other system factors that might cause the latency effects. It also can estimate the synchronization and delay of multiple channels.
Convention Paper 8678

Tutorial 13 **Sunday, April 29**
12:00 – 13:00 **Room Lehar**

BINAURAL AUDITORY MODELS

Presenter: **Ville Pulkki**, Aalto University, Helsinki, Finland

The working principles of brain mechanisms of binaural hearing have been debated during the last decade. In 1990s the common thinking was that human binaural decoding is based on delay lines and coincidence counters, as proposed by the Jeffress model. Later, some neurophysiological findings questioned the existence of such delay lines, and some evidence was found bolstering the count-comparison model proposed by Bekesy. In count-comparison model, the binaural differences are rate-coded between the left and brain right hemispheres. This tutorial will introduce the basic principles of most common binaural auditory models, and review some latest improvements in the models.

Student/Career Development Event
STUDENT DELEGATE ASSEMBLY MEETING —PART 2
Sunday, April 29, 13:00 – 14:30
Room 5

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. Plans for future student activities at local, regional, and international levels will be summarized.

Session P22 **Sunday, April 29**
14:00 – 16:30 **Room Lehar**

QUALITY EVALUATION

Chair: **György Wersényi**, Széchenyi István University, Györ, Hungary

14:00

P22-1 Evaluating Spatial Congruency of 3-D Audio and Video Objects—*Kristina Kunze, Judith Liebetrau, Thomas Korn*, Fraunhofer Institute for Digital Media Technology, IDMT, Ilmenau, Germany

In this paper we demonstrate the evaluation of spatial congruency of object based audio and 3-D video reproduction. With current developments in 3-D video representation we are able to introduce a depth dimension. Furthermore, audio reproduction systems like Wave Field Synthesis are able to reproduce the sound field of virtual sound sources at various positions in the room, also in front of or behind a video screen. When combing these technologies audio objects can be placed at the positions of 3-D video objects. Subjective evaluations are needed to investigate the quality of such combinations. In our experiment we displaced the audio and video objects with respect to certain angles and evaluated the noticeable displacement angle. Displacements of more than 5° are noticeable and become annoying above 10°.
Convention Paper 8679

14:30

P22-2 Some New Evidence that Teenagers and College Students May Prefer Accurate Sound Reproduction—*Sean Olive*, Harman International Industries Inc., Northridge, CA, USA

A group of 18 high school and 40 college students with different expertise in sound evaluation participated in two separate controlled listening tests that measured their preference choices between music reproduced in (1) MP3 (128 kbp/s) and lossless CD-quality file formats, and (2) music reproduced through four different consumer loudspeakers. As a group, the students preferred the CD-quality reproduction in 70% of the trials and preferred music reproduced through the most accurate, neutral loudspeaker. Critical listening experience was a significant factor in the listeners' performance and preferences. Together, these tests provide some new evidence that both teenagers and college students can discern and appreciate a better quality of reproduced sound when given the opportunity to directly compare it against lower quality options.
Convention Paper 8683

15:00

P22-3 Evaluation of Cultural Similarity in Playlist Generation—*Mariusz Kleæ*, Polish-Japanese Institute of Information Technology, Warsaw, Poland

Choosing appropriate songs that satisfy one's needs is often frustrating, tiresome, and ineffective due to the increasing number of music collections and their sizes. Successive songs should fit the situation, our mood, or at least have common features. Consequently, there is a need to develop solutions that would enrich our experience of listening to music. In this paper musical similarity has been studied at the cultural level in playlist generation process. Also, an author's program designed for testing different playlists will be described. It is used to perform an experiment examining quality of playlists created using a cultural similarity model.
Convention Paper 8681

15:30

P22-4 Toward an Unbiased Standard in Testing Laptop PC Audio Quality—*Ravi Kondapalli, Ben-Zhen Sung*, CCCRMA, Stanford University, Stanford, CA, USA

For a rapidly growing population, laptop PCs and tablet devices have become a primary means of watching and listening to electronic media. Nonetheless, form factor restrictions and placement of device functionality over audio quality have left a gap in the overall quality of the laptop listening experience. Multiple audio Post-processing solutions exist from audio DSP developers that aim to improve this. However, the quality of audio produced by pairing these post-processing solutions with PCs from a variety of manufacturers varies greatly. To date, there have been no blind and systematic comparisons

of audio quality resulting from unique PC/post-processing algorithm pairings. Here we present an unbiased methodological approach for evaluating such combinations, looking at audio quality for three different commercially available post-processing solutions as implemented on laptop PCs from two different manufacturers.
Convention Paper 8682

16:00

P22-5 User Evaluation on Loudness Harmonization on the Web—*Gerhard Spikofski, Peter Altendorf, Christian Hartmann*, Institut für Rundfunktechnik, Munich, Germany

The problem of annoying loudness jumps occurring between different TV and radio stations or programs became generally known over the last decades. Since the publication of the recommendations ITU-R BS.1770-1 (in 2006) and EBU-R 128 (in 2010), and the actual version of the ITU Recommendation BS.1770-2 (2011), more and more tools have become available that allow loudness harmonization in broadcasting. The European research project NoTube integrates loudness harmonization in its research concepts in order to investigate the applicability of these tools in web environments where the situation is even worse. The user evaluation of different loudness and loudness range adaptations presented in this paper was carried out online and thus considers real web conditions. In particular the interdependence between parameters of the listening environment and the loudness/loudness range adaptations are highlighted based on the results of nearly 100 participants.
Convention Paper 8680

Tutorial 14 Sunday, April 29
14:00 – 15:30 Room Liszt

SMALL ROOM ACOUSTICS

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

Acoustic basics of small rooms will be discussed. Specific issues related to the size of the room (room-modes) will be addressed. Absorption, reflection, diffraction, diffusion and how to use it, as well as specific aspects regarding low frequency treatment will be discussed.

Although this will not be a studio design class, specifics and differences of recording rooms and control rooms will be identified, including considerations for loudspeaker and microphone placement.

Tutorial 15 Sunday, April 29
15:00 – 16:30 Room Brahms

THE MAKING OF THE BEACH BOYS *SMILE* SESSIONS

Presenters: **Mark Linett**
Barry Marshall

Arguably the greatest “lost” album of all time, The Beach Boys *Smile* album sessions were finally released last November, in both a 2 CD and 5 CD version, along with vinyl and other digital configurations. The co-producer of the project, Mark Linett, will be interviewed by producer and educator Barry Marshall about the legendary 1966–67 sessions. The presentation will provide historical context on the role of producer Brian Wilson, as well as a focus on the technical and logistical challenges (like mixing and mastering from 45-year-old tapes at different configurations, speeds and sizes!) faced by co-producer and mix engineer Mark Linett in compiling this landmark release.