AES 130th Convention Program

May 13 – 16, 2011

Novotel London West, London, UK

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 130th 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 130th AES Convention Student Paper Award is:

Design of a Compact Cylindrical Loudspeaker Array for Spatial Sound Reproduction – *Mihailo Kolundzija, Christof Faller, Martin Vetterli*, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
Convention Paper 8336

To be presented on Friday, May 13 in Session P4 —Multichannel and Spatial Signal Processing, Part 1

* * * * *

Friday, May 13 09:00 Room Saint Julien
Technical Committee Meeting on Transmission and Broadcasting

Session P1 Friday, May 13 09:30 – 12:30
Room 4

SPEECH AND HEARING

Chair: **Bob Schulein**, Asius Technologies LLC, Longmont, CO, USA

09:30

P1-1 The Evolution of the Speech Transmission Index– *Herman J. M. Steeneken,*[1,2] *Sander J. van Wijngaarden,*[2] *Jan A. Verhave*[2]
[1]TNO Human Factors (retired), Soesterberg, The Netherlands
[2]Embedded Acoustics, Delft, The Netherlands

This year, the Speech Transmission Index celebrates its 40th anniversary. While the first measuring device built in the 1970s could barely fit inside a car, inexpensive pocket-size STI measuring solutions are now available to the world. Meanwhile, the STI method has continually evolved in order to deal with an increasing array of measuring challenges. This paper investigates how the STI kept up with these challenges and analyzes possible room for further improvement. Also, a roadmap for further development of the STI is proposed.
Convention Paper 8315

10:00

P1-2 Prosody Generation Module for Macedonian Text-to-Speech Synthesis– *Branislav Gerazov, Zoran Ivanovski*, Faculty of Electrical Engineering and Information Technologies, Skopje, Macedonia

The paper presents a fully functional prosody generation module developed for Macedonian text-to-speech (TTS) synthesis. The module is based on research of prosody generation modules in high-end TTS synthesis systems, previous prosody experiences in Macedonian TTS, as well as original research of prosody carried through by the authors. The paper starts with an overview of the basic tasks, problems, and solutions in prosody generation modules. Then it continues to give a detailed account of the workings of the developed module. The module first segments the input speech into intonation phrases and determines their intonation type. Next it generates durations for each of the units that will be used to synthesize the speech output. Then it determines the positions of the lexical stresses and modifies these units' durations. After determining the intonation phrase's pitch accent location, it generates an adequate pitch contour and calculates the pitch targets needed for unit modification. The synthesis module uses this data to generate prosody in the output speech. Generated prosody patterns in the output speech are of satisfactory quality for arbitrary text input. The presented results are of significant value for Macedonian TTS and can be used for other under-represented languages.
Convention Paper 8316

10:30

P1-3 The Influence of Transmission Channel on the Admissibility of Speech Sample for Forensic Speaker Identification– *Andrey Barinov*, Speech Technology Center Ltd., St. Petersburg, Russia

This paper is an extension and addition to papers previously published and presented during the AES 39th Conference and AES 129th Convention, regarding voice sample quality requirements and compensation for the influence of transmission channels for further forensic or automatic speaker identification. In this paper we provide the analysis of different types of transmission channels such as land line, GSM, radio, and VoIP. We analyze the important parameters of voice samples obtained from these channels and compare the influence of different channels on important speaker identification voice biometric features. At the end of the paper, there are some conclusions provided regarding the circumstances under which the particular recording can be accepted/rejected for forensic speaker identification.
Convention Paper 8317

11:00

P1-4 Optimizing the Acoustic and Intelligibility Performance of Assistive Audio Systems and Program Relay for the Hard of Hearing and Other Users– *Peter Mapp*, Peter Mapp Associates, Colchester, UK

Around 10% of the general population suffer from a noticeable degree of hearing loss and would benefit from some form of hearing assistance or deaf aid. DDA legislation and requirements mean that many more hearing assistive systems are being installed—yet there is continuing evidence to suggest that many of these systems fail to perform adequately and provide the benefit expected. This paper reports on the results of acoustic performance testing of a number of trial ALS systems. The use of STIPA, as a practical measure for assessing the potential intelligibility of ALS systems, is discussed. The ALS microphone type, distance, and angular location from the target acoustic source are shown to have a significant effect on the resultant potential intelligibility performance. The effects of typical ALS signal processing have been investigated and are shown to have a small but significant effect on the STIPA result. The requirements for a suitable acoustic test source to mimic a human talker are discussed as is the need to adequately assess the effects of both reverberation and noise. The findings of this paper are also relevant to the installation and testing of educational sound field systems as well as boardroom and conference room systems.
Convention Paper 8318

11:30

P1-5 Sound Reproduction within a Closed Ear Canal: Acoustical and Physiological Effects
– *Samuel Gido*,[1,2] *Robert B. Schulein*,[1] *Stephen Ambrose*[1]
[1]Asius Technologies LLC, Longmont, CO, USA
[2]University of Massachusetts, Amherst, MA, USA

When a sound producing device such as insert earphones or a hearing aid is sealed in the ear canal, the fact that only a tiny segment of the sound wave can exist in this small volume at any given instant, produces an oscillation of the static pressure in the ear canal. This effect can greatly boost the SPL in the ear canal, especially at low frequencies, a phenomena that we call Trapped Volume Insertion Gain (TVIG). In this study the TVIG has been found by numerical modeling as well as direct measurements using a Zwislocki coupler and the ear of a human subject, to be as much as 50 dB greater than sound pressures typically generated while listening to sounds in an open environment. Even at moderate listening volumes, the TVIG can increase the low frequency SPL in the ear canal to levels where they produce excursions of the tympanic membrane that are 100 to 1000 times greater than in normal open-ear hearing. Additionally, the high SPL at low frequencies in the trapped volume of the ear canal, can easily exceed the threshold necessary to trigger the Stapedius reflex, a stiffing response of the middle ear, which reduces its sensitivity and may lead to audio fatigue. The addition of a compliant membrane covered vent in the sound tube of an insert ear tip was found to reduce the TVIG by up to 20 dB, such that the Stapedius reflex would likely not be triggered.
Convention Paper 8319

12:00

P1-6 Can We Compare the Sound Quality of Noise Reduction between Hearing Aids? A Method to Level the Ground between Devices– *Rolph Houben, Inge Brons, Wouter A. Dreschler*, Academic Medical Center, Amsterdam, The Netherlands

This paper proposes the application of an equalization filter to remove unwanted differences in frequency response between recordings from different hearing aids. The filter makes it possible to compare the perceptual effects (such as user preference) of a specific signal processing feature (e.g., noise reduction) between different hearing aids, without the dominant influence of differences in their frequency responses. Both an objective quality measure (PESQ) and a listening experiment have shown that the filter was able to “level the ground” between the devices included. The potential application of the inverse filter is to use it on recordings from hearing aids so that we can directly compare the noise reduction between devices. This allows one to determine if users prefer a certain noise reduction over another, which could lead to improved rehabilitation of hearing impaired listeners.
Convention Paper 8320

Session P2 Friday, May 13 09:30 – 12:30
Room 1

LOUDSPEAKERS

Chair: **John VanderKooy**, University of Waterloo, Waterloo, Ontario, Canada & BW Group, Steyning, UK

09:30

P2-1 Vibrations in the Loudspeaker Enclosure Evaluated by Hot Wire Anemometry and Laser Interferometry—*Danijel Djurek*,[1] *Nazif Demoli*[2]
[1]AVAC (Alessandro Volta Applied Ceramics) Laboratory for Nonlinear Dynamics, Zlatar-Bistrica, Croatia
[2]Institute of Physics, Zagreb, Croatia

Vibration states of the loudspeaker enclosure were examined by the laser interferometry and hot wire anemometry. According to simple expressions, air pressure in the enclosure is independent on air density. This statement has been tested by the use of gases indicated by very different densities (air, He, SF6), and interferometric data show a strong dependence of vibration states of the enclosure walls on density, when driving frequency increases from 500 Hz up to 1 kHz. The main reason for a deviation resides in the imaginary part of the Morse impedance exerted by the vibrating gas within the enclosure.
Convention Paper 8321

10:00

P2-2 Losses and Coupling in Long Multi-Wire Loudspeaker Cables—*Xavier Meynial, Guennolé Gapihan*, Active Audio, Saint-Herblain, France

It is quite frequent in PA systems to use very long loudspeaker cables. These PA systems sometimes carry several audio signals in adjacent wires of the same cable, or several cables may be neighboring over very long distances such as several hundreds of meters. This paper deals with losses and coupling in these long cables. It is shown that even though losses and coupling may introduce significant crosstalk between wires and affect the amplifier loads, it is possible to use very long loudspeaker cables to connect line arrays in PA systems. Coupling between cables is also investigated, and strategies for reducing losses and coupling are presented.
Convention Paper 8322

10:30

P2-3 Subwoofers in Rooms: Modal Analysis for Loudspeaker Placement—*Juha Backman*, Nokia Corporation, Espoo, Finland and Aalto University, Espoo, Finland

Use of multiple subwoofers in a room is known to help in reducing the variation of response both as a function of frequency and a function of place, but simple geometry-based placement rules guarantee good results only in symmetrical cases. The paper discusses the use of experimental modal analysis and numerical optimization based on modal behavior to determine the optimal placement of single or multiple subwoofers in rooms with arbitrary geometry and surface properties.
Convention Paper 8323

11:00

P2-4 Losses in Loudspeaker Enclosures—*Claus Futtrup*, Scan Speak A/S, Videbæk, Denmark

In recent papers a lumped parameter model, which can simulate the impedance of conventional electro-dynamic transducers accurately, has been presented. The new model includes frequency-dependent damping, which questions traditional engineering practices in simulations of loudspeaker enclosures and, in particular, associated losses. In this paper the consequences of frequency-dependent damping are evaluated to aid the development of simulations and models of loudspeaker enclosures.
Convention Paper 8324

11:30

P2-5 Comparison of Measurement Methods for the Equalization of Loudspeaker Panels Based on Bending Wave Radiation—*Lars Hörchens, Diemer de Vries*, Delft University of Technology, Delft, The Netherlands

Spatially extended panel loudspeakers based on bending wave radiation, such as distributed mode loudspeakers or multi-actuator panels, exhibit a complex radiation pattern with frequency-dependent directivity characteristics. This paper seeks to determine the kind, position, and number of measurements needed to obtain an average radiation spectrum enabling efficient equalization. To this end, three approaches are compared: wave field extrapolation of the panel surface normal velocity, extrapolation of a near-field pressure measurement, and actual in-situ measurements at a number of random positions inside the listening space. The choice of the positions and the required number of measurements are discussed. Measurements taken on a multi-actuator panel are used to compare the different approaches and present numerical results.
Convention Paper 8325

12:00

P2-6 The Effect of Finite-Sized Baffles on Mobile Device Personal Audio—*Jordan Cheer*,[1] *Stephen J. Elliott*,[1] *Youngtae Kim*,[2] *Jung-Woo Choi*[2]
[1]University of Southampton, Southampton, UK
[2]Samsung Electronics Co. Ltd., Korea

To reduce the annoyance from the use of loudspeakers on mobile devices, previous work has investigated the use of acoustic contrast control to optimize the performance of small arrays of loudspeakers. These investigations have assumed that the baffle dimensions are negligible so that the loudspeakers are omnidirectional, which is reasonable at low frequencies; however, in practice the effect of a finite-sized baffle on the optimized performance is important at higher frequencies. This paper reports the results of using a finite-element model of a two-source array, positioned on a mobile phone sized baffle, to investigate the influence of the baffle on the predicted array performance. The baffle is shown to reduce the performance of the array at frequencies greater than around 1 kHz, but then the directivity of the individual drivers enhances performance at these higher frequencies.
Convention Paper 8326

Friday, May 13 10:00 Room Saint Julien
Technical Committee Meeting on Audio Recording and Mastering Systems

Workshop 1 Friday, May 13
10:30 – 12:30 Room 5

MUSIC AND SEMANTIC WEB

Co-Chairs: **David De Roure**, University of Oxford
Yves Raimond, BBC

Panelists: *David Bretherton*, Southampton University
Gregg Kellogg, Connected Media Experience
Alexandre Passant, Seevl
Evan Stein, Decibel

This workshop will provide an inter-disciplinary forum to explore and promote the combination of music and Semantic Web technologies—specifically, to develop an understanding of how Semantic Web technologies can contribute to the growth of music-related data on the Web, as well as applications of this data. Activities in this area are becoming well established, driven by projects in both research and industry, and they span personal applications together with all aspects of the creation, management, discovery, delivery and analysis of musical content. The ambitions of this inaugural workshop are to bring together this emerging and energetic community, to share information and practice, and especially to articulate the research agenda in Music and the Semantic Web. Through this we aim to facilitate innovation, inform Semantic Web research, and establish the basis and momentum for future events.

Session P3 Friday, May 13 11:00 – 12:30
Foyer

POSTERS: SOUND FIELD ANALYSIS

11:00

P3-1 Localization of Multiple Speech Sources Using Distributed Microphones—*Maximo Cobos, Amparo Marti, Jose J. Lopez*, Universidad Politecnica Valencia, Valencia, Spain

Source localization is an important task in many speech processing systems. There are many microphone array techniques intended to provide accurate source localization, but their performance is severely affected by noise and reverberation. The Steered-Response Power Phase Transform (SRP-PHAT) algorithm has been shown to perform very robustly in adverse acoustic environments; however, its computational cost can be an issue. Recently, the authors presented a modified version of the SRP-PHAT algorithm that improves its performance without adding a significant cost. However, the performance of the modified algorithm has only been analyzed in single source localization tasks. This paper explores further the possibilities of this localization method by considering multiple speech sources simultaneously active. Experiments considering different number of sources and acoustic environments are presented using simulations and real data.
Convention Paper 8327

11:00

P3-2 Detection of "Solo Intervals" in Multiple Microphone Multiple Source Audio Applications—*Elias Kokkinis,*[1] *Joshua Reiss,*[2] *John Mourjopoulos*[1]
[1]University of Patras, Patras, Greece
[2]Queen Mary University of London, London, UK

In this paper a simple and effective method is proposed to detect time intervals where only a single source is active (solo intervals) for multiple microphone, multiple source settings commonly encountered in audio applications, such as live sound reinforcement. The proposed method is based on the short term energy ratios between all available microphone signals, and a single threshold value is used to determine if and which source is solely active. The method is computationally efficient and results indicate that it is accurate and fairly robust with respect to reverberation time and amount of source interference.
Convention Paper 8328

11:00

P3-3 A Real-Time Sound Source Localization and Enhancement System Using Distributed Microphones—*Amparo Marti, Maximo Cobos, Jose J. Lopez*, Universidad Politecnica Valencia, Valencia, Spain

The Steered Response Power - Phase Transform (SRP-PHAT) algorithm has been shown to be one of the most robust sound source localization approaches operating in noisy and reverberant environments. A recently proposed modified SRP-PHAT algorithm has been shown to provide robust localization performance in indoor environments without the need for having a very fine spatial grid, thus reducing the computational cost required in a practical implementation. Sound source localization methods are commonly employed in many sound processing applications. In our case, we use the modified SRP-PHAT functional for improving noisy speech signals. The estimated position of the speaker is used to calculate the time-delay for each microphone and then the speech is enhanced by aligning correctly the microphone signals.
Convention Paper 8329

11:00

P3-4 Binaural Moving Sound Source Localization by Joint Estimation of ITD and ILD—*Cheng Zhou, Ruimin Hu, Weiping Tu, Xiaochen Wang, Li Gao*, Wuhan University, Wuhan, Hubei, China

Spatial cues ITD and ILD that provide sound localization information play a very important role in the binaural localization system. The efficient improvement of binaural moving sound source localization method by joint estimation of ITD and ILD based on Doppler effect is investigated. By removing Doppler effect influence, results show that the proposed binaural moving sound source localization method achieves 0.3%(velocity = 1m/s), 5.7%(velocity = 5m/s), and

10.5%(velocity = 10m/s) accuracy improvement in silent conditions. The performance of our method will be more effective as sound moves faster.
Convention Paper 8330
[Paper not presented but is available for purchase]

11:00

P3-5 Perceived Level of Late Reverberation in Speech and Music—*Jouni Paulus*,[1] *Christian Uhle*,[1] *Jürgen Herre*[1,2]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]International Audio Laboratories Erlangen, Erlangen, Germany

This paper presents experimental investigations on the perceived level of running reverberation in various types of monophonic audio signals. The design and results of three listening tests are discussed. The tests focus on the influence of the input material, the direct-to-reverberation ratio, and the reverberation time using artificially generated impulse responses for simulating the late reverberation. Furthermore, a comparison between mono and stereo reverberation is conducted. It can be observed that with equal mixing levels, the input material and the shape of the reverberation tail have a prominent effect on the perceived level. The results suggest that mono and stereo reverberation with identical reverberation times and mixing ratios are perceived as having equal level regardless of the material.
Convention Paper 8331

11:00

P3-6 Reverberation Enhancement in a Modal Sound Field—*Hugh Hopper, David Thompson, Keith Holland*, University of Southampton, Southampton, UK

The reverberation time of a room can be increased by using a reverberation enhancement system. These electronic systems have generally been installed in large rooms, where diffuse field assumptions are sufficiently accurate. Novel applications of the technology can be found by applying it to smaller spaces where isolated modal resonances will dominate at low frequency. An analysis of a multichannel feedback system within a rectangular enclosure is presented. To assess the performance of this system, metrics are defined based on the spatial and frequency variations of a diffuse field. These metrics are then used to optimize the parameters of the system using a genetic algorithm. It is shown that optimization significantly improves the performance of the system.
Convention Paper 8332

11:00

P3-7 An Advanced Implementation of a Digital Artificial Reverberator—*Andrea Primavera, Stefania Cecchi, Laura Romoli, Paolo Peretti, Francesco Piazza*, Universita Politecnica delle Marche, Ancona, Italy

Reverberation is a well known effect particularly important for listening to recorded and live music. In this paper we propose a real implementation of an enhanced approach for a digital artificial reverberator. Starting from a preliminary analysis of the mixing time, the selected impulse response is decomposed in the time domain considering the early and late reflections. Therefore, a short FIR is used to synthesize the first part of the impulse response, and a generalized recursive structure is used to synthesize the late reflections, exploiting a minimization criterion in the cepstral domain. Several results are reported taking into consideration different real impulse responses and comparing the results with those obtained with previous techniques in terms of computational complexity and reverberation quality.
Convention Paper 8333

11:00

P3-8 Evaluation of Spatial Impression Comparing 2-Channel Stereo, 5-Channel Surround, and 7-Channel Surround with Height Channels for 3-D Imagery—*Toru Kamekawa*,[1] *Atsushi Marui*,[1] *Toshikiko Date*,[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. A pairwise comparison test 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) asking better depth sense and better match between visual and audio images. The results show that 3-D imagery with 7 channel surround gives the highest sense of depth and match of visual and audio images.
Convention Paper 8334

Tutorial 1 **Friday, May 13**
11:00 – 12:30 **Room 2**

WHAT'S YOUR EQ IQ?

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

Equalization—we reach for it track by track and mix by mix, as often as any other effect. Easy enough, at first. EQ becomes more intuitive when you have a deep understanding of the parameters, types, and technologies used, plus deep knowledge of the spectral properties and signatures of the most common pop and rock instruments. Alex Case shares his approach for applying EQ and strategies for its use: fixing frequency problems, fitting the spectral pieces together, enhancing flattering features, and more.

Friday, May 13 **11:00** **Room Saint Julien**
Technical Committee Meeting on Automotive Audio

Special Event
AWARDS PRESENTATION AND KEYNOTE ADDRESS
Friday, May 13, 12:45 – 13:30
Room 1

Opening Remarks:
- Executive Director Roger Furness
- President Jim Kaiser
- Convention Chair Peter Mapp

Program:
- AES Awards Presentation
- Introduction of Keynote Speaker by Convention Chair Peter Mapp
- Keynote Address by Trevor Cox

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

Board of Governors Award
- **Jan Berg**
- **Kimio Hamasaki**
- **Toru Kamekawa**
- **Michael Kelly**

Fellowship Award
- **Andrzej Brzoska**
- **Christof Faller**
- **Masami "Sam" Toyoshima**

Bronze Medal Award
- **C. Robin Caine**
- **Yoshizo "Steve" Sohma**

Keynote Speaker

This year's Keynote Speaker is **Trevor Cox**. Trevor Cox is Professor of Acoustic Engineering at the University of Salford, a Senior Media Fellow funded by the Engineering and Physical Sciences Research Council, and President of the Institute of Acoustics (IOA).

One major strand of his research is room acoustics for intelligible speech and quality music production and reproduction. Cox's diffuser designs can be found in rooms around the world. He has co-authored a research book entitled *Acoustic Absorbers and Diffusers*. He was awarded the IOA's Tyndall Medal in 2004.

Cox has a long track record of communicating acoustic engineering to the public and has been involved in engagement projects worth over £1M. He was given the IOA award for promoting acoustics to the public in 2009. He has developed and presented science shows to 15,000 pupils including performing at the Royal Albert Hall, Purcell Rooms, and the Royal Institution. Cox has presented fifteen documentaries for BBC radio including: Life's soundtrack, Save our Sounds, and Science vs the Strad. His latest BBC radio 4 program due for broadcast in June is titled "Green Ears." The title of Cox's keynote speech is "Sounding Places, Past and Present: How Spaces Affect the Sounds We Make and Hear."

Architectural acoustics affect the perceived quality of live and recorded performances. Caves with ringing speleothem show that mankind has been exploiting venues with distinctive acoustics since the upper Palaeolithic era. This talk will explore unconventional places, both ancient and modern, with remarkable acoustics: a spherical room that allows you to whisper into one of your own ears; a water cistern with a reverberation time of over forty seconds, and the whispering arches of Grand Central in New York. What do these spaces say about our hearing abilities and the importance of sound to our lives?

Session P4 Friday, May 13 14:00 – 17:00
Room 1

MULTICHANNEL AND SPATIAL SIGNAL PROCESSING, PART 1

Chair: **Francis Rumsey**, Logophon, Ltd., Witney, UK

14:00

P4-1 3-D-Sound in Car Compartments Based on Loudspeaker Reproduction Using Crosstalk Cancellation—*Andre Lundkvist, Arne Nykänen, Roger Johnsson*, Luleå University of Technology, Luleå, Sweden

One way to enhance driving safety is to use signal sounds. Driver attention may further be improved by placing sounds in a 3-D space, using binaural synthesis. For correct loudspeaker reproduction of binaural signals, crosstalk between the channels needs to be cancelled out. In this study, a crosstalk cancellation algorithm was developed and tested. The algorithm was applied in a car compartment, and three loudspeaker positions were compared. A listening test was performed to determine the subjects' ability to correctly localize sounds. It was shown that loudspeakers placed behind the listener correctly reproduced sound sources in the back hemisphere. Loudspeakers located in front and above the listener gave a high number of front/back confusions for all angles.
Convention Paper 8335

14:30

P4-2 Design of a Compact Cylindrical Loudspeaker Array for Spatial Sound Reproduction—*Mihailo Kolundzija,*[1] *Christof Faller*,[1] *Martin Vetterli*[1,2]
[1]Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[2]University of California at Berkeley, Berkeley, CA, USA

Building acoustic beamformers is a problem whose solution is hindered by the wide-band nature of the audible sound. In order to achieve a consistent directional response over a wide range of frequencies, a conventional acoustic beamformer needs a high number of discrete loudspeakers and be large enough to achieve a desired low-frequency performance. The acoustic beamformer design described in this paper uses measurement-based optimal beamforming for loudspeakers located mounted on a rigid cylindrical baffle. Super-directional beamforming enables achieving desired directivity with multiple loudspeakers at low frequencies. High frequencies are reproduced with a single loudspeaker, whose highly directional reproduction—due to the cylindrical baffle—matches the design goals. In addition to the beamformer filter design procedure, it is shown how such a loudspeaker array can be used for spatial sound reproduction.
Convention Paper 8336

15:00

P4-3 A New Multichannel Microphone Technique for Effective Perspective Control—*Hyunkook*

Lee, University of Huddersfield, Huddersfield, UK

This paper introduces a new multichannel microphone technique that was designed to produce multiple listener perspectives. A listener perspective paradigm and the related spatial attributes are also proposed for the evaluation of spatial quality in acoustic recording and reproduction. The proposed microphone array employs five coincident microphone pairs, which can be transformed into virtual microphones with different polar patterns and directions depending on the mixing ratio. The results of interchannel correlation and informal subjective evaluations suggest that the proposed technique is able to offer an effective control over various spatial attributes.
Convention Paper 8337

15:30

P4-4 Experimental Analysis of Spatial Properties of the Sound Field Inside a Car Employing a Spherical Microphone Array—*Marco Binelli, Andrea Venturi, Alberto Amendola, Angelo Farina*, University of Parma, Parma, Italy

A 32-capsule spherical microphone array was employed for analyzing the spatial properties of the sound field inside a car. Both the background noise and the sound generated by the car's sound system were spatially analyzed, by superimposing false-color sound pressure level maps over a panoramic 360x180 degree image, obtained with a parabolic-mirror camera. The analysis of the noise field revealed the parts of the car body where more noise is leaking in, providing guidance for better soundproofing. The analysis of the impulse responses generated by the loudspeakers did show useful information on the reflection patterns, providing guidance for adding absorbent material in selected locations and for optimizing position and orientation of loudspeakers.
Convention Paper 8338

16:00

P4-5 Improved ITU and Matrix Surround Downmixing—*Christof Faller*,[1] *Pieter Schillebeeckx*[2]
[1]Illusonic LLC, St-Sulpice, Swtizerland
[2]Soundfield Ltd., Wakefield, UK

Improvements to ITU and matrix surround downmixing are proposed. The surround channels are often mixed into the downmix with reduced gain to prevent that the resulting stereo signal is overly ambient. A method is proposed that allows control of the amount of ambience in the downmix signal independently of direct sound. In a conventional matrix surround downmix, ambience from the surround channels appears impaired (negatively correlated). To address this issue, a technique is proposed that separates direct and ambient sound. Then, a matrix surround downmix is only applied to the direct sound, while ambient sound is treated like in an ITU downmix.
Convention Paper 8339

16:30

P4-6 Spaciousness Rating of 8-Channel Stereophony-Based Microphone Arrays—*Laurent Simon*,[1,2] *Russell Mason*[1]
[1]University of Surrey, Guildford, Surrey, UK
[2]now with INRIA – Rennes, Bretagne Atlantique, France

In previous studies, the localization accuracy and the spatial impression of 3-2 stereo microphone arrays were discussed. These showed that 3-2 stereo cannot produce stable images to the side and to the rear of the listener. An octagon loudspeaker array was therefore proposed. Microphone array design for this loudspeaker configuration was studied in terms of localization accuracy, locatedness, and image width. This paper describes an experiment conducted to evaluate the spaciousness of 10 different microphone arrays used in different acoustical environments. Spaciousness was analyzed as a function of sound signal, acoustical environment, and microphone array's characteristics. It showed that the height of the microphone array and the original acoustical environment are the two variables that have the most influence on the perceived spaciousness, but that microphone directivity and the position of sound sources is also important.
Convention Paper 8340

Session P5 Friday, May 13 14:00 – 17:30
Room 4

PRODUCTION AND BROADCAST

Chair: **Natanya Ford**

14:00

P5-1 Frequency Weighting and Ballistics for Program Loudness Modeling—*Ian M. Dash*,[1] *Benjamin Smith*,[2] *Densil Cabrera*[2]
[1]Australian Broadcasting Corporation, Sydney, NSW, Australia
[2]University of Sydney, Sydney, NSW, Australia

There has been both experimental and anecdotal evidence that the low-frequency performance in the ITU-R Recommendation BS.1770 program loudness measurement algorithm could be improved. A listening test with an emphasis on low frequency content was therefore conducted. An attempt was made to analyze the results in octave bands by multiple regression, but larger than expected variability precluded any useful outcome from this method. A simpler regression analysis was therefore performed using several fixed weighting curves and asymmetric integration with a range of time constants. Although the results largely support the present low frequency weighting curve, they also indicate that asymmetric integration provides better program loudness assessment than symmetric integration or high level gating.
Convention Paper 8341

14:30

P5-2 Evaluation of Live Meter Ballistics for Loudness Control—*Scott Norcross*,[1] *Félix Poulin*,[2] *Michel Lavoie*[1]

[1]Communications Research Centre, Ottawa, Ontario, Canada
[2]CBC/Radio-Canada, Montreal, Quebec, Canada

The broadcast community is transitioning from practices that revolved around audio peak normalization to one where the focus is on loudness consistency. Central to this effort is the loudness measurement algorithm described in ITU-R BS.1770. To assist in the mixing of audio that targets specific long-term loudness levels a need has been identified for some form of loudness-based live audio metering. The EBU has proposed meter specifications to indicate "momentary" and "short-term" loudness whose ballistics are defined, respectively, by a 400 ms and a 3000 ms integration window. In parallel with the EBU efforts, the CBC/Radio-Canada, in collaboration with the CRC, has been studying since 2008 various loudness meter ballistics that would be suitable in a production environment. This paper reports on a series of tests carried out by the CBC/Radio-Canada where various momentary meter ballistics were evaluated it was found that an IIR-based meter ballistic with a 400 ms time-constant was preferred.
Convention Paper 8342

15:00

P5-3 Adaptive Dynamics Enhancement—*Martin Walsh, Edward Stein, Jean-Marc Jot*, DTS, Inc., Scotts Valley, CA, USA

Modern recordings are being mastered with more and more aggressive dynamic range compression in an attempt to generate content that is louder than previous releases. This can often lead to large discrepancies in perceived loudness between tracks that were mastered at different periods of recent history. A commonly proposed solution to this problem involves the use of loudness normalization. While such normalization techniques help to reduce discrepancies in loudness, they cannot resurrect dynamics that were removed due to extreme levels of dynamic range compression. This paper outlines a technique for restoring the dynamics of modern music by continuously monitoring transient signal behavior together with the associated dynamic range levels. When dynamic range compression is likely, transients are restored to levels that are more expected for the type of material being played.
Convention Paper 8343

15:30

P5-4 Describing the Transparency of Mixdowns: The Masked-to-Unmasked-Ratio—*Philipp Aichinger*,[1,2] *Alois Sontacchi*,[2] *Berit Schneider-Stickler*[1]
[1]Medical University of Vienna, Vienna, Austria
[2]University of Music and Performing Arts Graz, Graz, Austria

In this paper a model that predicts the transparency of mixdowns is proposed. The masked-to-unmasked-ratio relates the original loudness of an instrument to its loudness in the mix. In order to assess this new measure a listening test is conducted. It is shown that instruments with a masked-to-unmasked-ratio of 10% or smaller are critical in mixdowns because most of them cannot be identified adequately. The newly suggested model is to be used in automatic mixdown algorithms and as an evaluating measure in future development whenever masking scenarios are to be described.
Convention Paper 8344

16:00

P5-5 The "Digital Solution": The Answer to a Lot of Challenges within New Production Routines at Today's Broadcasting Stations—*Stephan Peus*, Georg Neumann GmbH, Berlin, Germany

The introduction of networked production systems allows a very actual, topical, and efficient workflow especially within broadcast and TV applications. As a result production is moving closer to editorial work. Sound editing such as voice-over, etc., has to be done more and more by editors who naturally don't have that special knowledge as a sound engineer. Live recordings and interactive TV in future will call for further steps to simplify the production processes and to enhance reliability. We will explain practical examples of challenges from daily production routines, will give answers to solve the problems by using digital technology, and will show the effects on a more simple and reliable workflow.
Convention Paper 8345

16:30

P5-6 Automatic Mixing and Tracking of On-Pitch Football Action for Television Broadcasts—*Robert G. Oldfield, Benjamin G. Shirley*, University of Salford, Salford, UK

For the television broadcast of football in Europe, the sound engineer will typically have an arrangement of 12 shotgun microphones around the pitch to pick up on-pitch sounds such as whistle blows, players talking, and ball kicks, etc. Typically, during a match, the sound engineer will increase and decrease the levels of these microphones manually in accordance with where the action is on the pitch at a given time to prevent the final mix being awash with crowd noise. As part of the EU funded project, FascinatE, we have developed an automatic mixing algorithm that intelligently seeks key events on the pitch and turns on the corresponding microphones, the algorithm picks out the key events and automatically tracks the action eliminating the need for manual tracking.
Convention Paper 8346

17:00

P5-7 High Quality "Radio" Broadcasting over the Internet—*David Errock*, Wave Science Technology Ltd., London, UK

Broadcasting audio over the internet has dramatically improved the experience for listeners, including removing the geographic boundaries of terrestrial transmission, allowing playback on portable devices and time-shifting content. Internet audio has primarily increased choice with traditional broadcasters now competing alongside "internet only" music stations. Internet

distribution can also increase the audio quality, beyond that of terrestrial and satellite transmission channel constraints; unlike video streaming, which is of lower perceived quality than traditional transmission. With the sustainable data capacity of internet connections increasing, lossless audio can be streamed live to the consumer. This paper discusses the issues relating to these distribution techniques and the challenges that will be experienced by the broadcasters, the distribution network, receiver manufacturers, and listeners.
Convention Paper 8347

Session EB1 Friday, May 13 14:00 – 15:30
Foyer

DESIGN AND ASSESSMENT

14:00

EB1-1 Queen Mary's "Media and Arts Technology Studios" Audio System Design—*Martin James Morrell, Christopher A. Harte, Joshua D. Reiss,* Queen Mary University of London, London, UK

The Media and Arts Technology Studios is a new facility at Queen Mary University of London linking together a previous Listening Room, with our new rooms: Control Room, Performance Lab, and Plant Room. This engineering report discusses our design philosophy for our given brief to create a "world-class facility" for a space that is a "blank canvas" for researchers. We detail considerations for making an audio system that is simple for standard recording and playback while having a tremendous amount of routing options for users to create unique projects between all the connected spaces, featuring two separate spatial audio reproduction systems. The result is a 96 kHz/24 bit MADI based system using a multimode fiber optic network and dedicated wordclock throughout.
Engineering Brief 1

14:00

EB1-2 Revitalizing the Denis Arnold Hall for Multichannel Electroacoustic Sound Diffusion and Recording —*Duncan Williams*, University of Oxford, Oxford, UK

The Denis Arnold Hall, the flagship lecturing and performance space in the Faculty of Music, Oxford University, has recently been the beneficiary of a complete refurbishment, including dedicated design and specification for the performance and diffusion of electroacoustic composition and multichannel sonic art. The new configuration includes acoustic absorption and insulation, low frequency management, and eight flexible full range satellite loudspeakers. This diffusion system is complemented by a full range of playback formats (from 1/4" reel-to-reel, to Blu-ray, with line level patching for discrete or "stemmed" multichannel performance/playback), as well as 16 small diaphragm capacitor microphones, hung in 8 stereo pairs from the ceiling on an integrated winch system. The design and configuration of the space necessitated a lengthy consultation with composers, acousticians, electricians, and audio-visual specialists, and lessons learned along the way might be useful for those interested in adapting their own space for similar purposes.
Engineering Brief 2

14:00

EB1-3 Time Alignment of Subwoofers in Large PA Systems—*Natàlia Milán, Joan Amate*, Master Audio

In common PA systems the frequency range is divided into different ranges that are reproduced using different cabinets (subwoofers for the bass range and top cabinets for the mid-high range). This means different locations and positions of the sound sources and therefore notches and peaks in the crossover range. Time alignment is needed to adjust the arrival time of frequencies in the crossover area. But it's not only a matter of distance, since the crossover is modifying the phase, too. In this brief, a downscaled model is used to show how to measure the phase difference of a two-source PA system and correct it using FFT measuring software. Two situations are treated: two sources with overlapped frequency response and two sources sharing crossover frequency.
Engineering Brief 3

[Engineering Brief #4 was withdrawn]

14:00

EB1-5 Sound System Documentation for Construction Projects—*Thomas Knauss*, Marshall-KMK Acoustics, Ltd., Chappaqua, NY, USA

This engineering brief will provide sound system engineers, users, and operators with the opportunity to become familiar with the information, processes, and documentation required to successfully coordinate the electrical, rough-in, and general construction requirements of a professionally installed sound system with the needs of architects, engineers, and contractors responsible for construction projects. The brief will include a categorized list of the 24 points of information that every architect, engineer, and construction professional needs to know about each and every sound system device to be implemented on a construction or renovation project. The presenter maintains 30 years of audio experience including front of house engineer for national recording artists, large scale integration as a sound systems contractor to complete professional design services for recital halls, performing arts centers, corporate, entertainment, and house of worship facilities.
Engineering Brief 5

14:00

EB1-6 A Free Database of Head-Related Impulse Response Measurements in the Horizontal Plane with Multiple Distances—*Hagen Wierstorf, Matthias Geier, Alexander Raake, Sascha Spors*, T-Labs, Technische Universität Berlin, Berlin, Germany

A freely available database of Head-Related Impulse Response (HRIR) measurements is introduced. The impulse responses were measured in an anechoic chamber using a KEMAR manikin at four different loudspeaker distances—3 m, 2 m, 1 m, and 0.5 m—reaching from the far field to the near field. The loudspeaker was positioned at ear height and the manikin was rotated with a step motor in one degree increments. For the 3 m distance additional measurements have been carried out where the torso stayed fixed and only the artificial head was rotated. In addition to the raw impulse responses there are also data-sets available with a frequency response compensated for the use of several different headphones.
Engineering Brief 6

14:00

EB1-7 Loudness Measurement and Human Interpretation in Television Program Quality Checking—*Matthieu Parmentier*, France Television

Along the standardization of Loudness Measurement (EBU R128 and ITU-R BS.1770 update), France Television introduced loudness in 2010 to improve human interpretation during Quality Checking processes. After the defined measurement itself, done by a machine respecting the standard, studies have been conducted to fix objective limits according to the network's skills and viewer environments. By using the tools offered by the EBU R128, we have introduced specifications for audio acceptance beyond the loudness target of –23 LUFS, already taken into account. This work and results have been extended for common programs and short commercials in different ways. The measurement tools, developed according R128, have also been graphically developed to facilitate this additional reading.
Engineering Brief 7

14:00

EB1-8 Do Young People Actually Care about the Quality of Their MP3s?—*Ainslie Harris*, Robert Gordon University, Aberdeen, Scotland, UK

Ten focus groups were conducted in 2010 in Toronto, New York, and northern England, with participants ranging in age from 15 to 32 years old. Participants were asked about their file quality preferences for downloading music from online services, in the context of their existing positive/negative experiences, and being able to design a new service that would offer them the format(s) of their choice. This brief will outline some of the key qualitative findings, including: (1) Contrary to popular opinion, young people can tell the difference between a 128 k and 320 k MP3, but there appears to be an age threshold, below which listeners either do not notice or do not care. (2) Listening context and environment are important, and affect consumers' file quality preferences, even when only considering what type of MP3 to download. (3) Consumers still find that they must weigh their personal quality preferences against other practical requirements.
Engineering Brief 8

Workshop 2 **Friday, May 13**
14:00 – 16:00 **Room 2**

WHAT'S THE USE OF RECORDING?

Chair: **Sean Davies**, S.W. Davies Ltd., Aylesbury, UK

Panelists: *Ted Kendall*, Independent Engineer, Wales, UK
Robert Philip, Open University, Milton Keynes, UK
Gordon Reid, Cedar Audio Ltd., Cambridge, UK

As engineers we tend not to think too much about the eventual purpose of what we do, but this can influence not only our own techniques, but can have ramifications far afield. A panel of specialists will consider recording as evidence in legal cases; as propaganda; as a true record of musical styles for different periods; as a way of delivering drama; and as a means of transmitting secret intelligence.

Student/Career Development Event
THE CHANGING AUDIO PROFESSIONAL
Friday, May 13, 14:00 – 15:00
Room 5

Moderator: **Melvyn Toms**, JAMES (Joint Audio Media Education Services)

Panelists: *Samantha Bennett*, University of Westminster, UK
Dennis Weinreich, APRS/UK Screen/JAMES

Over the previous 30 years, the technical and operational demands placed upon working professionals in the audio industry have grown exponentially. Equally over this period the opportunities to learn and gain the necessary experience to become accomplished by sitting and working alongside established practitioners has fallen. To a large extent, public and private education has taken on this role but has it been successful?

Audio production and technology courses attract a large intake of industry hopefuls but will they be “sitting and working alongside established practitioners” and how can they be assured their course meets industry requirements?

Friday, May 13 **14:00** **Room Saint Julien**
Technical Committee Meeting on Archiving, Restoration, and Digital Libraries

Friday, May 13 **14:00** **Room Mouton Cadet**
Standards Committee Meeting SC-02-02 on Digital Input/Output Interfacing

Friday, May 13 **15:00** **Room Saint Julien**
Technical Committee Meeting on Human Factors in Audio Systems

Tutorial 2 **Friday, May 13**
16:00 – 18:00 **Room 5**

FORENSIC AUDIO ENHANCEMENT

Chair: **Eddy B. Brixen**, EBB Consult, Denmark

Panelists: *Andrey Barinov*, Speech Technology Center Ltd., Russia
Robin P. How, Metropolitan Police, London, UK
Gordon Reid, CEDAR Audio Ltd., Cambridge, UK

Attendees of this tutorial should expect to learn more about issues surrounding the enhancement of forensic audio recordings. Topics will include evidence handling and processing, problem-specific approaches, and the increasing enumeration of GSM modulated voice signals will be presented and discussed. The panel of presenters will include practitioners from law enforcement, software manufacturers, and the scientific community. Before and after audio examples will be demonstrated focusing on specific problems.

Friday, May 13 16:00 Room Saint Julien
Technical Committee Meeting on Audio for Telecommunications

Session P6 Friday, May 13 16:30 – 18:00
Foyer

POSTERS: AUDIO EQUIPMENT

16:30

P6-1 Study and Analysis of the Carrier Distortion Sources in PWM DCI-NPC Modulator— *Vicent M. Sala, Luis Romeral*, UPC-Universitat Politecnica de Catalunya, Terrassa, BCN, Spain

This paper studies and analyzes an analog PWM Modulator for a Half-Bridge DCI-NPC Amplifier (Diode Clamped Inverter – Neutral Point Clamped) as one of the sources of distortion in the switching amplification chain. The four types of error or distortion sources are studied and analyzed: Carriers Offset Error (COE), Carriers Phase Error (CPE), Carriers Symmetry Error (CSE), and Carriers Amplitude Error (CAE). This paper concludes that the two major sources of error or distortion in PWM modulation process for DCI-NPC topology are the Amplitude (CAE) and Offset (COE) Errors, the latter being largest contributor to the total distortion.
Convention Paper 8348
[Paper not presented but is available for purchase]

16:30

P6-2 Experimental Verification of an Electrostatic Transducer with a Partitioned Back Electrode— *Libor Husník*, Czech Technical University in Prague, Prague, Czech Republic

Electroacoustic transducers based on the condenser principle have usually a planar back electrode that is a one-piece entity. The previous work studied theoretically the possibility of making the back electrode partitioned. Such an arrangement can be used in a design of the so-called digital loudspeaker, in which the sizes of the partitioned electrode are in the ratio of the powers of 2, but this is only a special case. The aim of this work is to study performance of a system modeling the electrostatic transducer with the partitioned back electrode by way of measurement on an experimental sample. Various combinations of signals are applied on the partitioned electrode and the transducer response is measured.
Convention Paper 8349

16:30

P6-3 Capacitor "Sound" in Microphone Preamplifier DC Blocking and HPF Applications: Comparing Measurements to Listening Tests— *Robert-Eric Gaskell*, McGill University, Montreal, Quebec, Canada

The sonic effect of capacitors in various aspects of audio electronics design has long been discussed and speculated upon. Recent publications have tested many of these theories through rigorous distortion measurements of a variety of capacitor types under several test conditions. One particularly interesting result is a measurable increase in 2nd harmonic distortion for electrolytic and PET type capacitors when a DC bias is applied. This paper repeats these measurements for a set of capacitors commonly used in +48V blocking applications and high pass filters in microphone preamplifier designs. These physical measurements are then compared to the result of double blind listening tests in order to examine the audibility of these capacitor distortions as well as explore their sonic effect on various program materials.
Convention Paper 8350

16:30

P6-4 1W 104dBA SNR Filterless Fully-Digital Class D Audio Amplifier with EMI Reduction Technique— *Rossella Bassoli, Carlo Crippa, Federico Guanziroli, Germano Nicollini*, ST-Ericsson, Agrate Brianza, Monza Brianza, Italy

A 1W filter-less power DAC featuring 104 dBA SNR and EMI spreading is presented. Pre-distortion algorithms are used to reduce harmonic distortion inherent to the employed modulation process, and an oversampling noise shaper allows reducing modulator clock speed to facilitate hardware implementation while keeping high-fidelity quality. No analog circuits exist from I^2S interface to speaker, leading to zero output offset and good efficiency even at medium/low power levels. 1.2V digital and 2.7V-5V output supplies are used. Active area is $0.94mm^2$ in a 0.13micron CMOS process. Total harmonic distortion at maximum level is about 0.2%.
Convention Paper 8351

16:30

P6-5 Ultra-Low Power Audio Architecture for Portable Devices— *Kangeun Lee, Changyong Son, Dohyung Kim, Sihwa Lee*, Samsung Advanced Institute of Technology, Suwon, Korea

Current portable devices demand not only higher performance but also lower power consumption. For the reason, this paper describes a low power System-on-Chip (SoC) architecture that targets playback multimedia format. To significantly reduce power consumption of the SoC, the system exploits a DSP core to decode compressed audio. It is a key technique that a pre-buffer keeps compressed audio data. Therefore, the DSP can independently playback audio data, and other systems including CPU, and DRAM is able to be power off until all audio data remaining in the pre-buffer is exhausted by the DSP.

For seamless operation, we designed a new kernel driver that controls the DSP and CPU, and it is embedded into Linux kernel sets of the Google's Android 2.2. Energy efficiency is evaluated by using fourteen audio sequences encoded with mp3 of which format is 128 kbps stereo. The experimental results show that the proposed system required only 25% power of the conventional DVFS algorithm and guaranteed Quality of Service (QoS) for mp3 playing.
Convention Paper 8352

16:30

P6-6 Design of a Passive DGRC Column Loudspeaker with Wave Front Synthesis– *Xavier Meynial, Gilles Grégoire*, Active Audio, Saint-Herblain, France

The DGRC (Digital and Geometric Radiation Control) principle allows one to assign a large number of loudspeakers of a line array to a limited number of electronic channels. It was introduced in 2006 and is now used in digitally steerable column loudspeakers. In this paper we propose a very simple and straightforward implementation of this principle in a passive column, where delays are approximated with all-pass passive circuits. As a result, the column is placed vertically (not tilted) and radiates a wave front that ensures high speech intelligibility and homogeneous SPL coverage. We present the design of this column, as well as experimental results. Finally, we discuss its advantages in terms of visual integration, acoustic performances, and cost effectiveness.
Convention Paper 8353

16:30

P6-7 Design and Realization of a Reference Class Loudspeaker Panel for Wave Field Synthesis – *Stephan Mauer, Frank Melchior*, IOSONO GmbH, Erfurt, Germany

This paper describes the requirements, the design, nd the realization of a reference loudspeaker panel for Wave Field Synthesis (WFS). Beside the algorithm and the loudspeaker's acoustical performance the quality of Wave Field Synthesis is strongly dependent on the spatial aliasing frequency. Only below that frequency will the synthesized wave field be physically correct. The spatial aliasing frequency is related to the distance of adjacent speakers in the loudspeaker array. To raise the aliasing frequency to 2.8 kHz a loudspeaker panel was designed with a tweeter spacing of 6 cm. The transducers, electronics, and directivity were designed to obtain an excellent sound quality and SPL coverage. Simulations regarding the influence of speaker spacing and wave field artifacts have been made. Measurement results of the panel are given. The overall system design will be shown as an application example.
Convention Paper 8354

16:30

P6-8 Study and Analysis of Demodulation Filter Losses in DCI-NPC Multilevel Power Amplifiers– *Vicent M. Sala, Luis Romeral*, UPC-Universitat Politecnica de Catalunya, Terrassa, BCN, Spain

This paper presents a less studied source of efficiency losses in Multilevel Diode-Clamped-Inverter or Neutral-Point-Converter (DCI-NPC) Power Switching Amplifiers. Filter inductors generally add another significant contribution to the total power loss in Power Switching Amplifiers systems. This contribution is generally comparable to that of the switching power stage and it is important to obtain a reasonable accurate estimate of the losses. In this paper the four possible sources of losses in the demodulation filter are studied and analyzed and are defined by the expressions to calculate the value of the losses. Using these loss expressions, this paper analyzes the contribution of each of the sources. This work finishes up presenting the results of simulation and the conclusions.
Convention Paper 8355
[Paper not presented but is available for purchase]

Student/Career Development Event
OPENING AND STUDENT DELEGATE ASSEMBLY MEETING – PART 1
Friday, May 13, 16:30 – 18:00
Room 2

Chair: **Daniel Deboy**

Vice Chair: **Magdalena Plewa**

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 Monday, May 16.

Friday, May 13 17:00 Room Mouton Cadet
Standards Committee Meeting SC-02-08 on Audio-File Transfer and Exchange

Special Event
WELCOME RECEPTION AND PRESENTATION
Friday, May 13, 18:00 – 19:30
Room 1

Fraunhofer IIS invites all visitors to the AES 130th Convention to meet in a social atmosphere to catch up with friends and colleagues from the world of audio. There will be a short presentation of the "International Audio Laboratories Erlangen—The New Center of Audio Research" followed by a Welcome Drinks Reception.

Special Event
ORGAN RECITAL BY GRAHAM BLYTH
Friday, May 13, 20:00 – 21:30
Lincoln's Inn Chapel, London

Graham Blyth's traditional organ recital will be given on the new organ of the Lincoln's Inn Chapel. The Chapel has had a historic role in the life of the Inn. The present building was consecrated on Ascension Day, 1623, and there are services every Sunday in law terms. Lincoln's Inn Chapel had to wait almost 200 years until the first organ was installed in 1820. The modest organ of 1819-20 by Flight and Robson was replaced in 1856 by a three-manual and pedal organ by William Hill, one of the finest organ builders in the country. This organ was rebuilt on no less than nine occasions, most recently in 1969.

As the organ approached its 150th birthday it became clear that its increasing unreliability meant that a decision needed to be taken about its future. After exhaustive research it was concluded that so little of the original Hill organ had remained unaltered that it was impossible to restore it to any original condition.

After a lengthy tender process, a contract was signed with Kenneth Tickell in 2005 for a new three-manual organ. It was assembled in the Northampton workshops of Tickell in 2008-9, and arrived on site in July 2009. The case, of European oak, has been stained to match the woodwork of the Chapel, and the highly polished front pipes are complemented by shades of lime wood carved to a foliage design based on the surviving example of early sixteenth-century mural painting displayed in the room that now forms the passage between Gatehouse Court and Hardwicke Building. The new organ was installed by Kenneth Tickell in 2009-10.

The recital will include Bach's Great C major Prelude & Fugue, Franck's Fantasie in A, two recent pieces inspired by words of Sir John Betjeman by Dennis Wickens, and a complete performance of Widor's 5th Symphony. Programs and maps will be available from the Special Events Desk.

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.

Session P7 Saturday, May 14 09:00 – 11:00
Room 1

LIVE AND INTERACTIVE SOUND

Chair: **Michael Kelly**, Sony Computer Entertainment Europe, London, UK

09:00

P7-1 User Driven, Local Model, Reclassification of Drum Loop Audio Slices— *Henry Lindsay-Smith*,[1] *Skot McDonald*,[2] *Mark Sandler*[1]
[1]Queen Mary University of London, London, UK
[2]FXpansion Audio Ltd., London, UK

We present a method for significantly improving the results of drum loop slice classification. An onset detector is used to slice loops of percussion only audio. Low level features are extracted from the audio slices and the slices are classified into one of seven percussion classes by a previously trained PART decision table. This general classification algorithm shows only an adequate performance. The user is then allowed to correct incorrect classifications. Each corrected classification is combined with a subset of the original classifications and a nearest neighbor algorithm reclassifies the remaining slices according to the corrected local model. The resultant algorithm converges on a 100% correct solution, with nearly 40% fewer re-classifications than a non-assisted approach.
Convention Paper 8356

09:30

P7-2 Kick-Drum Signal Acquisition, Isolation and Reinforcement Optimization in Live Sound— *Adam J. Hill,*[1] *Malcolm O. J. Hawksford*,[1] *Adam P. Rosenthal,*[2] *Gary Gand*[2]
[1]University of Essex, Colchester, Essex, UK
[2]Gand Concert Sound, Glenview, IL, USA

A critical requirement for popular music in live-sound applications is the achievement of a robust kick-drum sound presented to the audience and the drummer while simultaneously achieving a workable degree of acoustic isolation for other on-stage musicians. Routinely a transparent wall is placed in parallel to the kick-drum heads to attenuate sound from the drummer's monitor loudspeakers, although this can cause sound quality impairment from comb-filter interference. Practical optimization techniques are explored, embracing microphone selection and placement (including multiple microphones in combination), isolation-wall location, drum-monitor electronic delay, and echo cancellation. A system analysis is presented augmented by real-world measurements and relevant simulations using a bespoke Finite-Difference Time-Domain (FDTD) algorithm.
Convention Paper 8357

10:00

P7-3 Development of a Virtual Performance Studio with Application of Virtual Acoustic Recording Methods— *Iain Laird*,[1] *Damian Murphy*,[2] *Paul Chapman*,[1] *Seb Jouan*[3]
[1]Glasgow School of Art, Glasgow, Scotland, UK
[2]University of York, Heslington, York, UK
[3]Arup, Glasgow, Scotland, UK

A Virtual Performance Studio (VPS) is a space that allows a musician to practice in a virtual version of a real performance space in order to acclimatize to the acoustic feedback received on stage before physically performing there. Traditional auralization techniques allow this by con-

volving the direct sound from the instrument with the appropriate impulse response on stage. In order to capture only the direct sound from the instrument, a directional microphone is often used at small distances from the instrument. This can give rise to noticeable tonal distortion due to proximity effect and spatial sampling of the instrument's directivity function. This work reports on the construction of a prototype VPS system and goes on to demonstrate how an auralization can be significantly affected by the placement of the microphone around the instrument, contributing to a reported "PA effect." Informal listening tests have suggested that there is a general preference for auralizations that process multiple microphones placed around the instrument.
Convention Paper 8358

10:30

P7-4 Interactive Audio Realities: An Augmented / Mixed Reality Audio Game Prototype— *Nikos Moustakas, Andreas Floros, Nicolas Grigoriou*, Ionian University, Corfu, Greece

Audio-games represent a game alternative based on audible feedback rather than on visual. They may benefit from parametric sound synthesis and advanced audio technologies (i.e., augmented reality audio), in order to effectively realize complex scenarios. In this paper a multiplayer game prototype is introduced that employs the concept of controlled mixed reality in order to augment the sound environment of each player. The prototype is realized as multiple user audiovisual installations, which are interconnected in order to communicate the status of the selected control parameters in real-time. The prototype reveals significant relevance to the-well known on-line multiplayer games, with its novelty originating from the fact that user interaction is realized in augmented reality audio environments.
Convention Paper 8359

Session P8 Saturday, May 14 09:00 – 12:30
Room 4

AUDIO EQUIPMENT

Chair: **John Dawson**

09:00

P8-1 Signal Level and Frequency Dependent Losses Inside Audio Signal Transformers and How to Prevent Those— *Menno van der Veen*, ir. bureau Vanderveen, Zwolle, The Netherlands

In an earlier work (Convention Paper 7125) a model was presented that explains that low voltage level audio signals are extra weakened when they are fed through a transformer. This extra weakening is caused by the signal level and frequency-dependent inductance of the transformer. Combining this extra weakening with the threshold of hearing curves, showed that noticeable loss of micro details occurs in the frequency band from 20 Hz to 1 kHz. This paper expands the previous work with measurements on several valve amplifiers, refines the model, and makes it applicable to macro signal levels close to saturation of the transformer. Methods are also given to minimize this extra weakening in transformers.
Convention Paper 8360

09:30

P8-2 Diaphonic Pump: A Sound-Activated Alternating to Static Pressure Converter— *Stephen D. Ambrose*,[1] *Robert Schulein*,[1] *Samuel Gido*[1,2]
[1]Asius Technologies LLC, Longmont, CO, USA
[2]University of Massachusetts, Amherst, MA, USA

This paper discusses the operating principles and basic construction of a diaphonic pump, which is a newly invented device for harvesting the energy inherent in sound waves and using it to pump air, thereby pressurizing a vessel. Although this device is of general utility, the embodiment discussed in this paper is used to harvest sound energy from the speaker (balanced armature transducer) of a personal listening device (headset or hearing aid), and use this as a power source to inflate a bubble in the listener's ear, thereby creating an acoustic seal. The diaphonic pump utilizes a natural asymmetry in the flow pattern when fluid is alternatingly pushed and pulled, back and forth through a small orifice known as a "synthetic jet." Sound waves provide the alternating flow pattern across the synthetic jet orifice. Prototype diaphonic pumps were built, which attach to a back volume of a balanced armature transducer and are small enough that the whole assembly, transducer, and pump, can fit in a human ear canal.
Convention Paper 8361
[Paper presented by Bob Schulein]

10:00

P8-3 Scanning the Magnetic Field of Electro-Dynamical Transducers— *Wolfgang Klippel*, University of Technology, Dresden Germany, Klippel GmbH, Dresden, Germany

The magnetic flux density in the magnetic gap and the geometry of the moving coil determine the force factor Bl, which is an important parameter of the electro-dynamical transducer. The paper presents a new measurement technique for scanning the flux density B(z, φ) on a cylindrical surface within and outside the magnetic gap using a Hall sensor and robotics changing the position of the sensor versus vertical position z and angle φ. The results derived from the scanning process reveal the real B field in the gap considering the fringe field and irregularities in the magnetization, which may initiate a rocking mode and rubbing of the voice coil at higher amplitudes. Using the geometry of the coil the static force factor Bl(x, i=0) can be calculated as a function of voice coil displacement x and compared with the dynamic force factor B(x,i) measured by a dynamic system identification techniques. Discrepancies between dynamic and static force factor characteristics are dis-

cussed and conclusions for loudspeaker design and manufacturing are derived.
Convention Paper 8362

10:30

P8-4 Comparison of Anemometric Probe and Tetrahedral Microphones for Sound Intensity Measurements—*Giulio Cengarle, Toni Mateos*, Fundació Barcelona Media, Barcelona, Spain

The measurement of sound intensity requires the acquisition of sound pressure and acoustic velocity in a coincident position. Various transducer topologies can be used to measure the acoustic velocity directly or indirectly. In this paper three transducers are compared: a pressure-velocity anemometric probe and two tetrahedral B-Format microphones from different manufacturers. The comparison has been carried out in different fields, ranging from anechoic to diffuse, reverberant field conditions. The analysis and comparison is based on intensimetric quantities such as the radiation index and the sound intensity vector. Strengths and limitations of the various approaches are reported, to suggest the preferred applications for each transducer.
Convention Paper 8363

11:00

P8-5 Prediction of Perceived Width of Stereo Microphone Setups—*Hans Riekehof-Böhmer*,[1] *Helmut Wittek*[2]
[1]HAW-Hamburg, Hamburg, Germany
[2]Schoeps Mikrofone, Karlsruhe, Germany

The diffuse-field correlation of the two signals generated by a stereophonic microphone setup has an effect on the perception of spatial width. A correlation meter is often used to measure the correlation coefficient. However, due to the frequency dependence of the correlation function, the correlation coefficient is not an appropriate value for predicting the perceived width when it comes to time-delay stereophony. By using the newly defined "Diffuse-Field-Image-Predictor" (DFI-Predictor) presented in this paper an attempt is made to reliably predict perceived width. Listening tests show that the DFI-Predictor is fairly suitable for this task. The aim of the study is to compare the spatial properties of different stereophonic microphone techniques by one calculated value.
Convention Paper 8364

11:30

P8-6 Synthesis of Polar Patterns as a Function of Frequency with a Twin Microphone: Audio Examples and Applications within the Creative Process of Music Mixing—*Matthias Kock*,[1] *Markus Kock*,[2] *Rainer Maillard,*[3] Malte Kob[1]
[1]Erich-Thienhaus-Institute, Detmold, Germany
[2]Leibniz Universität Hannover, Hannover, Germany
[3]Emil-Berliner Studios, Berlin, Germany

The directivity of a twin microphone can be chosen by variable weighting of the two output signals. In addition, the polar pattern can be adjusted as a function of frequency when controlled with a VST Plug-in in a modern DAW environment. A number of recordings were performed in rooms with variable size and quality. Presets with beneficial frequency-dependent directivities are compared to settings with constant directivity. It is discussed to what extend recordings can be further improved using the plug-in.
Convention Paper 8365

12:00

P8-7 Digital Microphones—What's it All About?—*John Willett*, Circle Sound Services, Oxfordshire, UK

It's now ten years since the first AES42 specification was published (AES42-2001) and the first AES42-compliant digital microphone came to the market. So this seems an opportune moment to look at AES42 digital microphones, their history, what they offer, the current market situation, and what the future may hold.
Convention Paper 8366

Session P9 Saturday, May 14 09:00 – 10:30
Foyer

POSTERS: PERCEPTION AND EVALUATION

09:00

P9-1 The Effect of Loudness Overflow on Equal-Loudness-Level Contours—*Andrew J. R. Simpson, Joshua D. Reiss*, Queen Mary University of London, London, UK

This paper presents a formal derivation of the Loudness Overflow Effect (LOE), which describes the impact of nonlinear distortion on loudness. Computational analysis is then performed, comprised of two experiments involving two compressive static nonlinearities and using two well-known time-varying loudness models. The results characterize the nonlinearities in terms of LOE as a function of frequency and of listening level in the case of 250-ms pure-tone stimuli, and in terms of the traditional equal-loudness-level contours. The analysis is then extended to synthesized wind instruments for one of the nonlinearities. The effect of the nonlinearity on loudness as a function of musical note fundamental frequency and listening level is described for various synthesized instruments.
Convention Paper 8367

09:00

P9-2 Evaluating the Use of Audio Smartphone Apps for Higher Education—*Anne Nortcliffe, Andrew Middleton, Ben Woodcock*, Sheffield Hallam University, Sheffield, UK

Digital audio technology has garnered interest in Education recently, being deployed by early adopter academics to provide audio feedback. Students have also used it, gathering audio notes on their personal devices to enhance their learning. However, the sharing and distributing of the recordings is time-consuming and requires separate technology. Smartphones with audio apps are able to support recording and distribu- ➥

tion/sharing of learning conversations more effectively because of their additional customizable and integrated functionality. This is attractive to Education now that it is clear that smartphones are becoming ubiquitous on campus. This paper describes an evaluation of audio apps for recording learning conversations by an academic and students and their experience in using smartphone audio apps to date.
Convention Paper 8368

09:00

P9-3 A Study of Human Perception of Temporal and Spectral Distortion Caused by Subwoofer Arrays— *Elena Shabalina*,[1] *Janko Ramuscak*,[2] *Michael Vorländer*[1]
[1]RWTH Aachen University, Aachen, Germany
[2]d&b audiotechnik GmbH, Backnang, Germany

The key task for a sound reinforcement system is to provide an even sound pressure level distribution over the whole listening area with possibly the same frequency response; and reduce radiation in wrong directions. For that a system should show a certain directivity, which for low frequencies can be achieved only by using multiple sound sources, for example, placed in a row in front of the stage. This technique can help to avoid strong interference and corresponding space sound pressure level variations of the conventional left/right setup of subwoofers. On the other hand, multiple sources cause changes of the impulse and frequency response of an array. Listening tests showed that these changes are audible for experienced listeners.
Convention Paper 8369

09:00

P9-4 Evaluation of the Psychoacoustic Perception of Geometric Acoustic Modeling-Based Auralization — *Aglaia Foteinou, Damian T. Murphy*, University of York, Heslington, York, UK

The subjective evaluation of the auralization of a simulated acoustic, with a view to establishing the success or otherwise of the results obtained, is usually best achieved using listening tests comparing the virtual environment with the actual measured space. As existing modeling methods still need to be improved, it is of critical importance to focus on the human perception of acoustic of the given space, rather than optimizing room acoustic parameters based only on objective measures. This paper uses a much simplified representation of a space with the resulting computer model giving the capability to control all of the examined acoustic or simulation parameters independently. A 3-D shoebox shape room is created and a variety of factors are changed every time in order to investigate their relevance and influence on human perception. These results are obtained from listening tests, and conclusions for the psychoacoustic perception of such a space are given.
Convention Paper 8370

09:00

P9-5 A Comparative Perceptual Evaluation of the Timbral Variations in Choral Location Recordings Created by Four Common Stereo Microphone Techniques— *Duncan Williams*, University of Oxford (Wolfson College), Oxford, UK

Choral recordings created on location were evaluated perceptually to determine the nature of the variations in timbre that might be elicited by the use of different stereo microphone techniques. Four stereo recordings were made simultaneously with coincident, near coincident, and spaced stereo microphone techniques. Listeners were invited to describe any perceived changes through a verbal elicitation experiment, informing an adjective “pool” of possible attributes. These attributes were reduced in number to six by verbal protocol analysis. The six remaining attributes were then scaled in a second listening experiment. Mean and standard deviation values in the results suggested that there was variation in three timbral attributes. This illustrated that the manipulation of timbral attributes by microphone technique, combined with perceptual analysis, is possible.
Convention Paper 8371

09:00

P9-6 Anchor Signals Validation for Two Dimensions of a Four-Dimensional Perceptive Space— *Yves Zango*,[1,2,3] *Régine Le Bouquin-Jeannès*,[2,3] *Nathalie Costet*,[2,3] *Catherine Quinquis*[1]
[1]Orange Labs Lannion Tech/Opera, Lannion Cedex, France,
[2]INSERM U642, Rennes France
[3]Université de Rennes, Rennes France

The subjective assessment of speech and sound codecs requires anchor signals to ensure its reliability. The reference system currently used is Modulated Noise Reference Unit (MNRU), which simulates only quantization noise. Now, the new generations of codecs present other impairments. In this study we consider speech quality as a multidimensional phenomenon and use dimensional reduction techniques to project codecs’ impairments in a four-dimensional space, each axis of the perceptive space corresponding to one of them. A verbalization test allowed characterizing two of these dimensions by the following attributes: “muffle” and “background noise.” Anchor signals were designed for these two dimensions, and a statistical analysis allowed validating the accuracy of at least one of these signals.
Convention Paper 8372

09:00

P9-7 Auditory Distance Perception: Criteria and Listening Room — *Jean-Christophe Messonnier*,[1] Alban Moraud[2]
[1]Conservatoire de Paris CNSMDP, Paris, France
[2]Altia, Paris, France

This paper is the result of a series of listening experiments carried out to investigate the correlation between auditory distance and two criteria: the ratio of direct to reverberant sound energy and the clarity C80. In the first section of this paper we will determine which of the two criteria is more efficient. The second section compares the values of these criteria when the same signal

is played on a well damped control room loudspeaker system and when it is played on a domestic stereophonic loudspeaker system. A second series of listening experiments shows how the auditory distance is perceived in both cases.
Convention Paper 8373

09:00

P9-8 Subjective Comparison between Stereo and Binaural Processing from B-Format Ambisonic Raw Audio Material—*Fábio W. Sousa*, University of York, York, North Yorkshire, UK

Using audio recorded in Ambisonic B-format from a sound field microphone and processed through both stereo and binaural tools, a subjective comparison is made. Hearing tests were performed taking into account personal experience and preference, as well as some spatial attributes concepts defined in previous works. Aiming to evaluate the real effectiveness of binaural processing, this paper considers the possibility of distributing contemporary music, originally developed for Ambisonic reproduction, through either conventional stereo or a method of high fidelity spatial processing directed specifically to reproduction through headphone systems, based on binaural technology. The sound images represented in both binaural and stereo processing are examined. Spatial attributes like wideness, depth, naturalness, and presence are evaluated.
Convention Paper 8374

Workshop 3 **Saturday, May 14**
09:00 – 11:00 **Room 3**

PANNING FOR MULTICHANNEL LOUDSPEAKER SYSTEMS

Chair: **Ville Pulkki**, Aalto University School of Science and Technology, Aalto, Finland

Panelists: *Filippo Fazi*, ISVR, UK
Matthias Frank, IEM, Austria
Florian Keiler, Technicolor DE, Germany

Amplitude panning is the most used method to position virtual sources over layouts where the number of loudspeakers is between two and about forty. The method is really simple, it provides a nice spatial effect, and does not color the sound prominently. This workshop reviews the working principle and psychoacoustic facts of amplitude panning for stereophony and for multichannel layouts. Panelists will describe some recent improvement suggestions to amplitude panning, which target some shortcomings of amplitude panning in spatial accuracy. A lively discussion is expected on the pros and cons of such processing.

Tutorial 3 **Saturday, May 14**
09:00 – 11:00 **Room 2**

MANAGING TINNITUS AS A WORKING AUDIO PROFESSIONAL

Presenters: **Neil Cherian**, Neurological Institute, Cleveland Clinic, Cleveland, OH, USA
Michael Santucci, Sensaphonics Hearing Conservation, Chicago, IL, USA

Tinnitus is a common yet poorly understood disorder where sound is perceived in the absence of an external source. Significant sound exposure, with or without hearing loss, is the most common risk factor. Tinnitus can be debilitating and can impair quality of life. Anxiety, depression, and sleep disorders are potential consequences. Most importantly for those in the audio industry, it can significantly impair auditory perception.

This tutorial will focus on methods in managing tinnitus in the life of an audio professional. Background information will be provided regarding the basic concept of tinnitus, pertinent anatomy and physiology, audiologic parameters of tinnitus, and an overview of current research. Suggestions for identifying and mitigating high risk behaviors will be covered. Elements of medical and audiologic evaluations of tinnitus will also reviewed.

Student/Career Development Event
EDUCATION FORUM PANEL
Saturday, May 14, 09:00 – 10:45
Room 5

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

Teaching the Teachers—A Round Table Discussion Among Audio Educators

While audio itself—in all her disciplines—advances at breakneck speed, the educators supporting it must make equivalent progress. AES conventions are reliable catalysts for earnest discussions among audio educators. However the convention, rich with so many activities, always seems to end too soon. Curriculum, personnel, and facilities must offer both time-proven fundamentals and cutting edge innovations. It happens in multiple modes: the classroom, the studio, the lab, online, campus committee meetings, and through industry relationships. We've all found solutions here, and nuggets of wisdom there; we wrestle with challenges and unknowns elsewhere. Join this discussion as we seek to define and prioritize the key issues facing educators and create a vision for the most effective way to address them in future AES activities—through conventions, conferences, online interactions, and more. Share your ideas for the most essential forms of research and the best ways to present the results: publications, tutorials, workshops, and other collaborations. What are the topics educators need to discuss, and what is the best format for sharing our advancements? AES provides the essential community for sharing and learning among audio educators. Help us design the next steps for increasing our productivity, accelerating our innovation, enriching our camaraderie, and enhancing our quality as educators.

Saturday, May 14 **09:00** **Room Saint Julien**
Technical Committee Meeting on Audio Forensics

Student/Career Development Event
STUDENT SCIENCE SPOT
Saturday, May 14 through Monday, May 16
Exhibits Area

Check out the Student Science Spot for student designs and projects! See the young creative geniuses demonstrate their knowledge in the flesh. Learn about other student events or just hang and meet some new audio friends. This is an amazing opportunity to share your ➥

hard work with convention participants! Hope to see you all there! http://www.aes.org/ students/blog/ 2011/3/130th-student-science-spot-call

Saturday, May 14 10:00 Room Saint Julien
Technical Committee Meeting on Microphones and Applications

Session P10 Saturday, May 14 11:00 – 13:00
Room 1

AUDIO CONTENT MANGEMENT

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

11:00

P10-1 A Comprehensive and Modular Framework for Audio Content Extraction, Aimed at Research, Pedagogy, and Digital Library Management—*Olivier Lartillot*, University of Jyväskylä, Jyväskylä, Finland

We present a framework for audio analysis and the extraction of low-level features, mid-level structures, and high-level concepts, altogether studied as a fully interwoven complex system. Composite operations are constructed via an intuitive programming language on top of Matlab. Datasets of any size can be processed thanks to implicit memory management mechanisms. The data structure enables a tight articulation between signal and symbolic layers in a unified framework. The resulting technology can be used as a pedagogical tool for the understanding of audio, speech, and musical processes and concepts, and for content-based discovery of digital libraries. Other applications includes intelligent browsing and structuring of digital library, information retrieval, and the design of content-based audio interfaces.
Convention Paper 8375

11:30

P10-2 Selected Playback Problems of Historical Grooved Media—*Nadja Wallaszkovits,*[1] *Franz Lechleitner*,[1] *Heinrich Pichler*[2]
[1]Phonogrammarchiv Austrian Academy of Sciences, Vienna, Austria
[2]Audio Consultant, Vienna, Austria

The paper discusses some selected playback problems of the replay and high quality archival transfer of historical grooved media, like cylinders, instantaneous discs and early coarse groove records. The topics outline the problems of noise reduction and the compensation of the horizontal tracking angle by means of stereo playback and modification of the sum and differential signals. A comparison between existing noise reduction methods and analog as well as digital phase and group delay compensation methods is given and discussed. Finally possible compensation methods for the change in noise spectrum caused by the groove velocity decrease at inner diameters with early discs are outlined. The authors propose a radius equalization in the digital domain by using a digital high pass filter without group delay distortion, using diameter dependent change of cut-off frequency.
Convention Paper 8376

12:00

P10-3 Automatic Recognition of Events in Audio Data Using Supercomputer Cluster—*Kuba Lopatka, Andrzej Czyzewski, Henryk Krawczyk*, Gdansk University of Technology, Gdansk, Poland

Dangerous events' automatic recognition by audio analysis employing parallel processing on a supercomputer cluster is described in the paper. Sound files recorded by microphones operating in a security surveillance system are processed by a sound event detection and classification algorithm. Because of the large amount of data, parallel computation is employed to speed up the analysis. The sound file recorded by the surveillance system is divided into chunks and processed by separate threads or processes. Several strategies for such parallel computation are introduced and discussed. Results obtained in tests using a supercomputer cluster are presented.
Convention Paper 8377

12:30

P10-4 Using Support Vector Machines for Automatic Mood Tracking In Audio Music—*Renato Panda, Rui Pedro Paiva*, University of Coimbra, Coimbra, Portugal

In this paper we propose a solution for automatic mood tracking in audio music, based on supervised learning and classification. To this end, various music clips with a duration of 25 seconds, previously annotated with arousal and valence (AV) values, were used to train several models. These models were used to predict quadrants of the Thayer's taxonomy and AV values, of small segments from full songs, revealing the mood changes over time. The system accuracy was measured by calculating the matching ratio between predicted results and full song annotations performed by volunteers. Different combinations of audio features, frameworks, and other parameters were tested, resulting in an accuracy of 56.3% and showing there is still much room for improvement.
Convention Paper 8378

Workshop 4 Saturday, May 14
11:00 – 13:00 Room 3

PRODUCTION FOR UPCOMING SPATIAL AUDIO SYSTEMS

Chair: **Frank Melchior**, IOSONO, Erfurt, Germany

Panelists: *Marc Emerit*, Orange Labs, France
Kimio Hamasaki, NHK, Tokyo, Japan
Jeff Levison, IOSONO GmbH, USA
Jörn Loviscach, University of Applied Sciences, Bielefeld, Bielefeld, Germany
Wieslaw Woszczyk, McGill University, Montreal, Quebec, Canada
Gregor Zielinsky, Sennheiser, Hannover, Germany

Driven by the force of 3-D pictures the demand for spatial audio experiences is increasing. Several systems have been proposed just adding more channels around the auditorium. While the integration of new channels seems to be simple from a workflow perspective, problems arise if the content needs to be adapted to different setups in different venues. On the other hand, and driven by new spatial audio reproduction methods, a paradigm shift toward object-based audio production is currently under discussion in the community. This workshop brings together the experience and tools in production for new multichannel formats and new spatial audio reproduction methods in general. Examples of related productions and installations are given.

Tutorial 4 Saturday, May 14
11:00 – 13:00 Room 5

IN-EAR MONITORING—PAST, PRESENT, AND EMERGING DEVELOPMENTS

Presenter: **Stephen Ambrose**, Asius Technologies

In the 1970s, sound engineer and performer Stephen Ambrose began pioneering work experimenting with the use of sound isolating earphones as a better means of hearing himself and others during live performances. This form of monitoring was in sharp contrast to the norm at the time, which consisted of an array of dedicated on-stage loudspeakers. Unlike with loudspeakers, performers with isolating earphones could hear their own unique mix, separate from others. As an additional benefit of in-ear monitoring, stage monitors are no longer present to leak output into the house sound system or recording mixes. Early work of Ambrose, with a wide range of well-known performers, established the basics of in-ear monitoring but was challenged by technical barriers and acceptance issues. Today, however, IEM (In-Ear-Monitoring) is in widespread use in many product forms. This tutorial, punctuated with a variety of demonstrations, will cover key developments responsible for the growing acceptance of IEM systems. Additionally, new improvements in the areas of sound isolation, comfort, occlusion effect suppression, and reduction of hearing fatigue, most of which are not yet commercially available, will be discussed.

Saturday, May 14 11:00 Room Saint Julien
Technical Committee Meeting on Coding of Audio Signals

Special Event
AES/MPG EVENT: WHERE DOES THE BUCK STOP? QC IN THE MANUFACTURING CHAIN
Saturday, May 14, 11:15 – 13:00
Room 2

Moderator: **Tony Platt**, MPG Director

Panelists: *Ray Staff*, Award-winning chief mastering engineer, Air Mastering
Pieter Stenekes, Founder of Sonoris Audio Engineering who are responsible for the DDP delivery software

We are all human and can make mistakes. It is also said that we do not recognize our own mistakes. In the video and broadcast world it is standard for the client to pay for an independent QC check at the studio and again at the broadcaster's facility. With vinyl it has always been normal to have a test pressing. All these procedures have been bypassed in CD mastering, and, in general, the passage of a master to manufacture is out of step with any QC that takes place. This can and has resulted in sometimes major and costly mistakes.

Our discussion will focus on the challenges posed by an emerging new shape of the industry where artists, small labels, and producers are dealing directly with manufacturers. Furthermore, these problems are also common to downloaded products. The incidences that have raised the issue of QC during the manufacturing stage of CDs can only be exacerbated if consumer delivery moves toward high quality downloads.

The growing tendency to third party agents arranging manufacture and distribution also raise issues in respect of delivery formats to manufacturers and aggregators and with the upsurge in delivery via FTP we need to establish some good practice for checking that what goes to manufacture is correct and most especially who is responsible.

This discussion launches the Mastering Section of Music Producers Guild which will be headed by Ray Staff. They propose to outline what quality control is and where the responsibilities lie and to publish a specification that will be posted on the MPG website so that anyone involved in the process can refer to for good practice and guidance.

Student/Career Development Event
EDUCATION AND CAREER/JOB FAIR
Saturday, May 14, 11:30 – 13:00
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.

Saturday, May 14 12:00 Room Saint Julien
Technical Committee Meeting on Spatial Audio

Saturday, May 14 12:00 Room Mouton Cadet
Standards Committee Meeting SC-02-12 on Audio Applications of Networks

Saturday, May 14 13:00 Room Saint Julien
Technical Committee Meeting on Studio Practices and Production

Session P11 Saturday, May 14 14:00 – 17:30
Room 1

ROOM ACOUSTICS

Chair: **Diemer de Vries**, Delft University of Technology, Delft, The Netherlands

14:00

P11-1 DTS Multichannel Audio Playback System: Characterization and Correction—*Zoran Fejzo*,[1] *James Johnston*[2]
[1]DTS, Inc., Calabasas, CA, USA
[2]DTS, Inc., Kirkland, WA, USA

Audio playback system correction methods are ➥

now commonplace in audio-video receivers. One goal of these systems is to correct deviation of loudspeaker/ room frequency response from some desired target curve. Unfortunately this correction may be inappropriate outside of a small area around the microphone location, and averaged measurements may provide unwanted timbre shifts. Some room correction algorithms capture the room responses at multiple locations and combine them to obtain a representative response that is used for frequency correction. We will present a loudspeaker/room correction system that attempts to achieve perceptually appropriate frequency correction in a wide listening area by using a closely spaced non-coincident multi-microphone array placed in a single location in the room. By use of special probe signals, this is achieved within a short measurement period.
Convention Paper 8379

14:30

P11-2 Evaluating the Auralization of Performance Spaces and Its Effect on Singing Performance—*Judith Brereton, Damian T. Murphy, David M. Howard*, University of York, Heslington, York, UK

Musicians alter their performance according to the acoustic environment in which they perform, but as yet a thorough parametric investigation of the effect of room acoustics on musical performance has not yet been achieved. A sufficiently "realistic" synthesized Room Impulse Response (RIR) will facilitate such a study, since this will allow the investigator greater control and knowledge of the room acoustic parameters involved. This paper reports the results of an experiment to evaluate a virtual acoustic space through the performance, interview, and audio analysis of the performance of a solo singer. Simulations of the same performance space using synthesized RIRs and measured RIRs were compared. In general, singers who took part in the trial could distinguish between the two simulations and rated the measured RIR simulation more highly in terms of warmth and reverberance.
Convention Paper 8380

15:00

P11-3 Enhancing the Configuration and Design of Sound Systems through Simulation—*Frederick Otten, Richard Foss*, Rhodes University, Grahamstown, South Africa

Audio Engineers are required to design and deploy large multichannel sound systems that meet a set of requirements and use networking technologies such as Firewire and Ethernet. Bandwidth utilization and latency need to be considered. Network Simulation can be used to accurately model a network and return such information. This paper discusses a software system that has been developed to create a simulation of a network using the AES-X170 protocol for command and control. This system shows information about bandwidth and latency and is able to detect problems with parameter relationships. It also provides the ability to perform offline editing. These features significantly enhance the audio engineer's ability to effectively design, configure, and evaluate their sound systems.
Convention Paper 8381

15:30

P11-4 What's Wrong with Scattering Theory?—*Ian M. Dash, Fergus R. Fricke*, University of Sydney, Sydney, NSW, Australia

The theory of long wave scattering from the side of a cylinder originated with Rayleigh and was extended to a general solution by Morse. Both solutions are based on an angular harmonic series expansion. This model is conceptually flawed. A number of physical paradoxes inherent in the model are outlined and discussed.
Convention Paper 8382

16:00

P11-5 New Proposals for the Calibration of Sound in Cinema Rooms—*Philip Newell*,[1] *Keith Holland*,[2] *Julius Newell*,[3] *Branko Neskov*[4]
[1]Acoustics Consultant, Moaña, Spain
[2]ISVR, University of Southampton, Southampton, UK
[3]Electroacoustics Engineer, Lisbon, Portugal
[4]Loudness Films, Lisbon, Portugal

The current practices for calibrating cinema rooms date back to the early 1970s. Much has been learned since then about the perception of sound, and measurement techniques have advanced greatly. Evidence has been growing that the present degree of room-to-room compatibility leaves much to be desired, and complaints about loudness and intelligibility problems persist. This paper looks at reassessing the whole process of the loudspeaker and room calibration from a modern perspective.
Convention Paper 8383
[Paper presented by Keith Holland]

16:30

P11-6 Some Preliminary Comparisons between the Diffusion Equation Model and Room-Acoustic Rendering Equation in Complex Scenarios—*Juan M. Navarro*,[1] *Jose Escolano*,[2] *Jose J. López*[3]
[1]San Antonio's Catholic University of Murcia, Guadalupe, Spain
[2]University of Jaén, Linares, Spain
[3]Universidad Politecnica de Valencia, Valencia, Spain

Recently, a model named acoustic radiative transfer equation has been proposed as a general theory to expand geometrical room acoustic modeling algorithms. This room acoustic modeling technique establishes the basis of two recently proposed algorithms, the acoustic diffusion equation model and the room acoustic rendering equation. This paper presents some comparisons of room-acoustic parameters in-situ measurements with prediction values from both methods in a real complex shape room in order to clarify advantages and limitations of both methods. Moreover, the memory requirements and computation time have been evaluated.
Convention Paper 8384

17:00

P11-7 Spatial Room Impulse Responses with a Hybrid Modeling Method—*Alex Southern, Samuel Siltanen, Lauri Savioja*, Aalto University, Aalto, Finland

The synthesis of an arbitrary enclosure room impulse response (RIR) may be performed using acoustic modeling. A number of acoustic modeling methods have been proposed previously each with their own advantages and limitations. This paper is concerned with mixing the RIRs from different modeling methods to synthesize a hybrid RIR. Low frequencies are modeled using the finite difference time domain method (FDTD), high frequencies are treated with geometric methods. A practical implementation for forming a hybrid RIR is discussed and further demonstrated in the context of a 2nd order B-Format spatial encode of the modeled sound field. The paper discusses the considerations and limitations of forming such hybrid RIRs using wave-based and geometric-based methods.
Convention Paper 8385

Session P12 Saturday, May 14 14:00 – 17:30
Room 4

BINAURAL SOUND

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

14:00

P12-1 Comparison of Speech Intelligibility in Artificial Head and Jecklin Disc Recordings—*Roger Johnsson, Arne Nykänen*, Luleå University of Technology, Luleå, Sweden

Binaural recordings are often done using artificial heads but can also be done with a Jecklin disc. In this study an experiment was designed that allowed evaluation of noise and reverberation suppression based on speech intelligibility measurements. Recordings of a voice and disturbing noise were done in a reverberant environment using one artificial head and four Jecklin discs of various sizes. A listening experiment using headphones was conducted to determine the speech intelligibility in the recordings and in a real life situation. It was found that there was no significant difference in the speech intelligibility between the artificial head and Jecklin disc with a diameter of 36 cm.
Convention Paper 8386

14:30

P12-2 A Comparison of Speech Intelligibility for In-Ear and Artificial Head Recordings—*Arne Nykänen, Roger Johnsson*, Luleå University of Technology, Luleå, Sweden

Good binaural reproductions should allow the listener to suppress noise and reverberation as when listening in real life. An experiment was designed where room properties and reproduction techniques were varied in a way that allowed evaluation of noise and reverberation suppression based on speech intelligibility measurements. Artificial head recordings were compared to in-ear recordings and real life listening. Artificial head recordings were found to be equivalent to real life listening. The speech intelligibility for in-ear recordings surpassed real life listening. A possible explanation may be inaccurate equalization. The equalization is critical for correct reproduction of binaural cues. The procedure used is convenient for validation of the performance of recording and reproduction equipment intended for sound quality studies.
Convention Paper 8387

15:00

P12-3 Perceptually Robust Headphone Equalization for Binaural Reproduction—*Bruno Masiero, Janina Fels*, RWTH Aachen University, Aachen, Germany

Headphones must always be adequately equalized when used for reproducing binaural signals if they are to deliver high perceptual plausibility. However, the transfer function between headphones and ear drums (HpTF) varies quite heavily with the headphone fitting for high frequencies, thus even small displacements of the headphone after equalization will lead to irregularities in the resulting frequency response. Keeping in mind that irregularities in the form of peaks are more disturbing than equivalent valleys, a new method for designing headphone equalization filters is proposed where not the average but an upper variance limit of many measured HpTFs is inverted. Such a filter yields perceptually robust equalization since the equalized frequency response will, with high chance, differ from the ideal response only by the presence of valleys in the high frequency range.
Convention Paper 8388

15:30

P12-4 Prediction of Perceived Elevation Using Multiple Pseudo-Binaural Microphones—*Tommy Ashby, Russell Mason, Tim Brookes*, University of Surrey, Guildford, Surrey, UK

Computational auditory models that predict the perceived location of sound sources in terms of azimuth are already available, yet little has been done to predict perceived elevation. Interaural time and level differences, the primary cues in horizontal localization, do not resolve source elevation, resulting in the "Cone of Confusion." In natural listening, listeners can make head movements to resolve such confusion. To mimic the dynamic cues provided by head movements, a multiple microphone sphere was created, and a hearing model was developed to predict source elevation from the signals captured by the sphere. The prototype sphere and hearing model proved effective in both horizontal and vertical localization. The next stage of this research will be to rigorously test a more physiologically accurate capture device.
Convention Paper 8389

16:00

P12-5 BRTF (Body Related Transfer Function) and Whole-Body Vibration Reproduction Systems—*M. Ercan Altinsoy, Sebastian Merchel*, Dresden University of Technology, Dresden Germany

If binaural recorded signals are played back via headphones, the transfer characteristic of the reproduction system has to be compensated for. Unfortunately, the transfer characteristic depends not only on the transducer itself, but also on mounting conditions and individual properties of the respective ear. This is similar with reproduction systems for whole-body vibrations. The transfer characteristic depends to a great extent on the individual body properties, e.g., weight or body mass index. In this study body related transfer functions of 60 subjects are measured using an electrodynamic excitation system. In addition anthropometric data of the subjects are collected. This paper reviews the existing whole-body vibration reproduction systems and discusses the importance of the individual transfer functions for whole-body vibration reproduction.
Convention Paper 8390

16:30

P12-6 HRTF-Enabled Microphone Array for Binaural Synthesis—*Malcolm O. J. Hawksford*, University of Essex, Colchester, Essex, UK

A synthesis technique incorporating a circular phased-array microphone is described where the horizontal polar response is matched to an arbitrary set of head-related transfer functions (HRTFs). The array can emulate the function of an artificial listener but without the need to embed physical anatomical features. Design techniques are described based upon polar response equalization computed over a discrete frequency space that together with dynamic coefficient processing enables spatial image manipulation and bespoke multi-listener environments with individual head tracking. A method of 2-D spatial filtering is described to scale the number of microphone signals. Applications include binaural recording optimally matched to an arbitrary number of listeners, distributed gaming, teleconferencing, multi-user interactive virtual reality, and remote surveillance.
Convention Paper 8391

17:00

P12-7 Interpolation and Range Extrapolation of Head-Related Transfer Functions Using Virtual Local Wave Field Synthesis—*Sascha Spors, Hagen Wierstorf, Jens Ahrens*, Deutsche Telekom Laboratories, Technische Universität Berlin, Berlin, Germany

Virtual environments that are based on binaural sound reproduction require datasets of head-related transfer functions (HRTFs). Ideally, these HRTFs are available for every possible position of a virtual sound source. However, in order to reduce measurement efforts, such datasets are typically only available for various source directions but only for one or very few distances. This paper presents a method for extrapolation of measured HRTF datasets from the source distance used in the measurements to other source distances. The method applies the concept of local Wave Field Synthesis to compute extrapolated HRTFs for almost arbitrary source positions with high accuracy. The method is computationally efficient and numerically stable.
Convention Paper 8392

Workshop 5 **Saturday, May 14**
14:00 – 15:00 **Room 3**

AURO 3D—HIGH QUALITY AND LOW LATENCY PCM ENCODING AND DATA REDUCTION

Chair: **Wilfried Van Baelen**, Galaxy Studios, Belgium

Auro-3D is first and foremost a family of surround formats including height information, starting with 4 overhead channels (and loudspeakers) and leading up to 13.1. The obvious use for such enlarged arrays is in the cinema, but the 9.1 Auro-3D solution is also applicable to the home theater, as the additional speakers are exactly above the existing ones (L+R front, L+R surround). Besides being a family of formats, also a codec has been developed that can fold down 9.1 to 5.1 or 5.1 to 2.0 purely in the PCM domain, with ultra-low latency decoding (1ms). The man behind Auro-3D, Wilfried Van Baelen from Galaxy Studios in Belgium, will give an introduction to the Auro-3D codec, talk about latest developments, and play some examples using the codec.

Special Event
AES/APRS/DV247 EVENT: TALKBACK-PRO 1 ACOUSTIC TREATMENT FOR SMALL SPACES
Saturday, May 14, 14:00 – 15:45
Room 2

Moderator: **Malcolm Atkin**

Panelists: *TBA*

As 3-D visuals hit the screens, there are more opportunities to provide surround-sound mixes. How can small rooms best be set-up to mix in surround and if mixing/control room space is terminally limited, can non-acoustic solutions offer successful fixes? This perennial problem has plagued aspiring producers and production rooms since audio began it's "democratizing" journey.

Hear studio acoustics experts and technologists debate and demonstrate the best way to exploit small production spaces for both conventional stereo and surround-sound mixes.

Saturday, May 14 14:00 Room Mouton Cadet
Standards Committee Meeting SC-05-02 on Audio Connectors

Student/Career Development Event
STUDENT RECORDING CRITIQUES
Saturday, May 14, 15:00 – 16:00
Sunday, May 15, 10:00 – 11:00
Monday, May 16, 11:00 – 12:00
Room Muscadet

Moderator: **Ian Corbett**, Kansas City Kansas Community College, Kansas City, KS, USA

Following the success of these events at the recent San Francisco convention—STUDENTS— bring a track to be critiqued in London!

These are non-competitive sessions, the idea being that you get "another" set of ears to listen to your work and make some constructive suggestions. Students are encouraged to bring in their stereo or surround projects for feedback and comments from a panel and audience. You will be able to sign-up for time slots at the first SDA meeting, on a first come, first served basis. Students who are finalists in the Recording Competition are excluded from participating in this event to allow the non-finalists an opportunity for feedback on their hard work. Bring your stereo or surround work on CD, DVD, or hard disc as clearly-labeled .wav files. The Student Recording Critiques are generously sponsored by PMC.

Saturday, May 14 15:00 Room Saint Julien
Technical Committee Meeting on Perception and Subjective Evaluation of Audio

Workshop 6 Saturday, May 14
15:15 – 16:15 Room 3

BOUNCE TO APP

Chair: **Michael Hlatky**, University of Applied Sciences Bremen, , Bremen, Germany

Panelists: *Jörn Loviscach*, University of Applied Sciences, Bielefeld, Germany
Martin MacMillian, Bounce Mobile
Martin Roth, RjDJ

Recorded music used to be a lean-back experience, but mobile devices have changed the game: Applications such as interactive 360-degree music videos begin to leverage the nonclassical man-machine interfaces of these devices—cameras, accelerometers, and GPS receivers—and employ their always-connectedness to access social music recommendation sites or music analysis and synthesis Web services. So far, the soundtracks of these apps have been produced from the final mixes. Can there be a better integration of audio production and the development of interactive music applications? What would that mean in terms of workflow and user interfaces? The issues that arise in this context are markedly different from those arising in the production of full-scale video games.

Workshop 7 Saturday, May 14
15:15 – 16:15 Room 5

ART OF RECORD PRODUCTION

Co-Chairs: **Katia Isakoff**
Simon Zagorski-Thomas

Panelists: *Haydn Bendall*, Engineer/Music Producer
Steve D'Agostino, Artist/Music Producer
Paschall de Paor, University of Glamorgan
Tony Platt, MPG and JAMES

This year will see the seventh international Art of Record Production (ARP) Conference in San Francisco and the re-launch of the *ARP Journal*. The Association for the Study of the Art of Record Production aims to provide an international and interdisciplinary organization for promoting the study of the production of recorded music that involves both academics and industry professionals. This workshop brings together the two directors of ARP with top record producers to discuss how the development of record production as an academic subject within the university system relates to the nuts and bolts of doing the job. What kind of approaches are researchers engaging in? What do producers think about these types of research and the emergence of practice led research in this field?

Saturday, May 14 15:30 Room Mouton Cadet
Standards Committee Meeting SC-05-05 on Grounding and EMC Practices

Workshop 8 Saturday, May 14
16:00 – 18:00 Room 2

HIGH RESOLUTION AUDIO PUBLISHING

Chair: **Stefan Bock**, MSM Studios, Germany

Panelists: *Morten Lindberg*, 2L, Norway
Darcy Proper, Wisseloord Studios, The Netherlands
Tokuyama Takeshi, Imagica, Japan
Neil Wilkes, Opus Productions
Yunichi Yoshio, Pioneer, Japan

There are a great deal of formats suitable for high resolution audio delivery. Which ones will succeed? What challenges do they need to overcome to become a standard? This panel will present a complete panorama of the available options, reviewing their advantages and disadvantages.

Saturday, May 14 16:00 Room Saint Julien
Technical Committee Meeting Advisory Group on Regulations

Session P13 Saturday, May 14 16:30 – 18:00
Foyer

POSTERS: PRODUCTION AND BROADCAST

16:30

P13-1 A Comparison of Kanun Recording Techniques as They Relate to Turkish Makam Music Perception – *Can Karadogan*, Istanbul Technical University, Istanbul, Turkey

This paper presents a quality comparison of microphone techniques applied on the kanun, a prominent traditional instrument of Turkish Makam music. Disregarding the effects of pre-amplifier color, A/D converter, compression, equalization, mixing, and mastering, only the studio recording step of music production is taken into focus. Microphone techniques were applied with varying placements and microphone types, and doing so, original Turkish Makam music etudes were recorded. Using short excerpts of these etudes, a survey comparing microphone techniques and placements as well as microphone types was prepared. Subjects were chosen from kanun players, sound engineers, and non-musicians who showed different perspectives, preferences, and descriptions to the sound samples of various microphone techniques.
Convention Paper 8393

16:30

P13-2 Objective Measurement of Produced Music Quality Using Inter-Band Dynamic Relationship Analysis— *Steven Fenton*,[1] *Bruno Fazenda*,[2] *Jonathan Wakefield*[1]
[1]University of Huddersfield, Huddersfield, UK
[2]University of Salford, Salford, UK

This paper describes and evaluates an objective measurement that grades the quality of a complex musical signal. 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 testing with real-world music excerpts and a greater number of listening subjects. A high degree of correlation is observed between the Mean Subject Scores (MSS) and the objective IBR descriptor suggesting it could be used as an additional model output variable (MOV) to describe produced music quality. The method lends itself to real-time implementation and therefore can be exploited within mixing, mastering, and monitoring tools.
Convention Paper 8394

16:30

P13-3 Evaluation of a New Algorithm for Automatic Hum Detection in Audio Recordings— *Matthias Brandt*,[1] *Thorsten Schmidt*,[2] *Joerg Bitzer*[1]
[1]Jade University of Applied Sciences, Oldenburg, Germany
[2]Cube-Tec International, Bremen, Germany

In this paper an evaluation of a recently published hum detection algorithm for audio signals is presented. To determine the performance of the method, large amounts of artificially generated and real-world audio data, containing a variety of music and speech recordings, are processed by the algorithm. By comparing the detection results with manually determined ground truth data, several error measures are computed: hit and false alarm rates, frequency deviation of the hum frequency estimation, offset of detected start and stop times, and the accuracy of the SNR estimation.
Convention Paper 8395

16:30

P13-4 Interactive Mixing Using Wii Controller— *Rod Selfridge, Joshua Reiss*, Queen Mary University of London, London, UK

This paper describes the design, construction, and analysis of an interactive gesture-controlled audio mixing system by the means of a wireless video game controller. The concept is based on the idea that the mixing engineer can step away from the mixing desk and become part of the performance of the piece of audio. The system allows full, live control of gains, stereo panning, equalization, dynamic range compression, and a variety of other effects for multichannel audio. The system and its implementation are described in detail. Subjective evaluation and listening tests were performed to assess usability and performance of the system, and the test procedure and results are reported.
Convention Paper 8396

16:30

P13-5 The Quintessence of a Waveform: Focus and Context for Audio Track Displays— *Jörn Loviscach*, Fachhochschule Bielefeld (University of Applied Sciences), Bielefeld, Germany

Oscilloscope-style waveform plots offer great insight into the properties of the audio signal. However, their use is impeded by the huge spread of timescales extending from fractions of a millisecond to several hours. Hence, waveform plots often require zooming in and out. This paper introduces a graphical representation through a synthesized quintessential waveform that shows the spectrum of the traditional waveform plot but does so at a much larger timescale. The quintessential waveform can reveal details about single periods at zoom levels where a regular waveform plot only indicates the signal's envelope. Compression renders the enormous ranges of frequencies and amplitudes more legible.
Convention Paper 8397

16:30

P13-6 Spatial Audio Processing for Interactive TV Services— *Johann-Markus Batke,*[1] *Jens Spille,*[1] *Holger Kropp,*[1] *Stefan Abeling*,[1] *Ben Shirley,*[2] *Rob G. Oldfield*[2]
[1]Technicolor, Research, and Innovation, Hannover, Germany
[2]University of Salford, Salford, UK

FascinatE is a European funded project that aims at developing a system to allow end users to interactively navigate around a video panorama showing a live event, with the accompanying audio automatically changing to match the selected view. The audiovisual content will be adapted to the users particular kind of device, covering anything from a mobile handset equipped with headphones to an immersive panoramic display connected with large loudspeaker setup. We describe how to handle audio content in the FascinatE context, covering simple stereo through to spatial sound fields. This will be performed by a mixture of Higher Order Ambisonics and Wave Field Synthesis. Our paper focuses on the greatest challenges for both techniques when capturing, transmitting, and rendering the audio scene.
Convention Paper 8398

16:30

P13-7 Wireless High Definition Multichannel Streaming Audio Network Technology Based on the IEEE 802.11 Standards— *Seppo Nikkila, Tom Lindeman, Valentin Manea*, ANT – Advanced Network Technologies Oy, Helsinki, Finland

A novel approach for the wireless distribution of uncompressed real-time multichannel streaming audio is presented. The technology is based on the IEEE 802.11 Point Coordination Function,

Contention Free Medium Access Control with size-optimized multicast frames. The implementation supports eight independent audio streams with simultaneous 24-bit audio samples at the rate of 192 kHz. Frame length alignment algorithm is developed for smooth, low jitter flow. An audio specific Forward Error Correction scheme and a low system latency inter-channel synchronization method are described. The clock drift is solved by a sample stuffing/stripping algorithm. Implemented hardware and software structures are presented and the technology is compared with other wireless audio networking concepts. Emerging multichannel content formats are briefly reviewed.
Convention Paper 8399

16:30

P13-8 Gestures to Operate DAW Software—*Wincent Balin*,[1] *Jörn Loviscach*[2]
[1]Universität Oldenburg, Oldenburg, Germany;
[2]Fachhochschule Bielefeld (University of Applied Sciences, Bielefeld, Germany

There is a noticeable absence of gestures—be they mouse-based or (multi-)touch-based—in mainstream digital audio workstation (DAW) software. As an example for such a gesture consider a clockwise O drawn with the finger to increase a value of a parameter. The increasing availability of devices such as smartphones, tablet computers, and touchscreen displays raises the question in how far audio software can benefit from gestures. We describe design strategies to create a consistent set of gesture commands. The main part of this paper reports on a user survey on mappings between 22 DAW functions and 30 single-point as well as multi-point gestures. We discuss the findings and point out consequences for user-interface design.
Convention Paper 8456

Student/Career Development Event
RECORDING COMPETITION—PART 1
Saturday, May 14, 16:30 – 18:30
Room 3

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:

- Sound for Visual Media 16:30 to 17:30
- Traditional Acoustic Recording 17:30 to 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 Monday 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.

Saturday, May 14 17:00 Room Saint Julien
Technical Committee Meeting on Electro Magnetic Compatibility

Saturday, May 14 17:00 Room Mouton Cadet
Standards Committee Meeting IEC 60268-3 ad-hoc meeting

Special Event
OPEN HOUSE OF THE TECHNICAL COUNCIL AND THE RICHARD C. HEYSER MEMORIAL LECTURE
Saturday, May 14, 18:30 – 20:00
Room 1

Lecturer: **Karlheinz Brandenburg**

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 130th AES Convention is Karlheinz Bandenburg. Brandenburg has been a driving force behind some of today's most innovative digital audio technology, notably the MP3 and MPEG audio standards. The research results of his dissertation are the basis of MPEG-1 Layer 3 (MP3), MPEG-2 Advanced Audio Coding (AAC), and most other modern audio compression schemes. He is acclaimed for pioneering work in digital audio coding, perceptual measurement techniques, Wave Field Synthesis (WFS), and psychoacoustics. His honors include the AES Silver Medal, the "IEEE Masaru Ibuka Consumer Electronic Award," the German Future Award, which he shared with his colleagues, and the Cross of the Order of Merit of the Federal Republic of Germany. Furthermore he is a member in the "Hall of Fame" of the Consumer Electronics Association and of the International Electrotechnical Commission. In 2009 he was appointed as Ambassador of the European Year of Creativity and Innovation. Brandenburg holds a Doctorate in Electrical Engineering from Friedrich-Alexander University Erlangen-Nuremberg and received honorary Doctorate degrees from the universities Koblenz-Landau and Lüneburg for his outstanding research work in the field of audio coding. He is professor at the Institute for Media Technology at Ilmenau University of Technology and director of the Fraunhofer Institute for Digital Media Technology IDMT in Ilmenau, Germany. He is married to Ines Rein-Brandenburg. They share their home in Ilmenau with two lovely cats. The title of his speech is "How to Provide High Quality Audio Everywhere: The MP3 Story and More …"

Once upon a time there was the challenge to transmit high quality audio over phone lines. While this seemed impossible, ideas from psychoacoustics and signal processing—work by many researchers—helped the seemingly impossible to become reality: MP3 and other audio codecs enabled seamless transport of audio over thin lines. However, the MP3 story did not end there. The Internet was changing shape, transforming from a text-based medium into a major carrier for sound of all kinds, including music. This meant changes not only for the payload (from text to audio), but also meant new dangers for the audio quality delivered to music lovers, and it changed business models for music sales dramatically, shaking up the music industry.

Today the challenges are different: we have a multitude of (legal) sources of music, and many of us have access to Terabytes of music. How do we find our way through this abundance of available content; how do we find the gems in the millions of medium quality music? Even then, audio reproduction still is far from perfect, so how can we really fulfill the dream of complete auditory illusion, and what are the problems even today? The talk will introduce some current work on both MIR (Music Information Retrieval), on audio reproduction and the psychoacoustic research needed to get further along toward the dream of perfectly reconstructed sound.

Brandenburg's presentation will be followed by a reception hosted by the AES Technical Council.

Student/Career Development Event
STUDENT PARTY
Saturday, May 14, 20:00 – 22:00

Join us for a fun and exciting evening at the Student Social. On Saturday night, May 14, at 8 pm the fun begins! Don't miss this chance to meet and engage with students from all over the world! Bring your own tracks and we will play them! Later, the local Student Section will guide us through the night. More information and details will be provided at SDA-1.

Session P14 Sunday, May 15 09:00 – 12:30
Room 1

MULTICHANNEL AND SPATIAL SIGNAL PROCESSING, PART 2

Chair: **Russell Mason**, University of Surrey, Guildford, Surrey, UK

09:00

P14-1 Spatial Analysis of Room Impulse Responses Captured with a 32-Capsule Microphone Array—*Angelo Farina, Alberto Amendola, Andrea Capra, Christian Varani*, University of Parma, Parma, Italy

The authors developed a new measurement system, which captures 32-channel impulse responses by means of a spherical microphone array and a matrix of FIR filters, capable of proving frequency-independent directivity patterns. This allows for spatial analysis with resolution much higher than what was possible with obsolete sum-and-delay beamforming. The software developed for this application creates a false-color video of the spatial distribution of energy, changing with running time along the impulse response duration. A virtual microphone probe allows extraction of the sound coming from any specific direction. The method was successfully employed in three concert halls, providing guidance for correcting some acoustical problems (echo, focusing) and for placing sound reinforcement loudspeakers in optimal positions.
Convention Paper 8400

09:30

P14-2 Control of the Beamwidth of a Beamformer with a Fixed Array Configuration—*Wan-Ho Cho,*[1] *Marinus M. Boone,*[2] *Jeong-Guon Ih,*[3] *Takeshi Toi*[1]
[1]Chuo University, Tokyo, Japan
[2]Delft University of Technology, Delft, The Netherlands
[3]Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea

The directional characteristic of the optimal beamformer of a transducer array depends not only on its hardware configuration but also on the stability factor. This parameter can be used to control the directivity of the array. In this paper a method, which is based on the proper selection strategy of the stability factor, is suggested to control the directional characteristics of the optimal beamformer without changing the array configuration. The selection method of the stability factor was investigated considering the trade-off relation between spatial resolution and noise amplification or array gain. The suggested method was applied to the problems of both microphone and loudspeaker arrays to obtain a specific directivity pattern of high resolution with a constant beamwidth.
Convention Paper 8401

10:00

P14-3 Design 3-D High Order Ambisonics Encoding Matrices Using Convex Optimization—*Haohai Sun, U. Peter Svensson*, Norwegian University of Science and Technology, Trondheim, Norway

In this paper we propose a convex optimization method for the design of 3-D High Order Ambisonics (3-D HOA) encoding matrices using spherical microphone arrays, which offers the possibility to impose spatial stop-bands in the directivity patterns of all the spherical harmonics while keeping the transformed audio channels still compatible with the 3-D HOA reproduction sound format. Using the proposed convex optimization method, the globally optimal encoding matrices can be obtained, and the suitable trade-off among several design factors, e.g., response distortions in the obtained spherical harmonic, the dynamic range of matrix coefficients, i.e., the system robustness, and frequencies, can be analyzed and illustrated. The proposed convex optimization is formulated as a form of second-order cone programming that can be efficiently solved. Numerical results validate the proposed method. This method can be easily generalized to the 2-D HOA cases.
Convention Paper 8402

10:30

P14-4 Principles in Surround Recordings with Height—*Günther Theile,*[1] *Helmut Wittek*[2]
[1]VDT, Geretsried, Germany
[2]Schoeps Mikrofone, Karlsruhe, Germany

New multichannel sound formats extending 5.1 with height channels are adding the third dimension to recordings. They provide a much wider range of spatial sound effects and allow more realism of spatial reproduction in terms of direct

sound, early and late reflections, reverberation, and ambience sound. Using the example of two upper layer front and two upper layer surround complementary loudspeakers (5.1+4, known as "Auro 3D 9.1") the psychoacoustic principles in the perception of elevated phantom sound sources, spatial depth, spatial impression, envelopment, ambient atmosphere, as well as directional stability within the sweet area are discussed. Concrete proposals for microphone configurations can evolve from these considerations.
Convention Paper 8403

11:00

P14-5 Efficient 3-D Sound Field Reproduction—*Mincheol Shin,*[1] *Filippo Fazi*,[1] *Jeongil Seo*,[2] *Philip A. Nelson*[1]
[1]ISVR, University of Southampton, Southhampton, UK
[2]Electronics Telecommunication Research Institute, Daejeon, Korea

A method is presented for efficient sound field reproduction with a loudspeaker array constituted by multiple sound sources that are three-dimensionally distributed. The physical reproduction of several target sound fields is investigated when the target sound fields are surrounded by multiple sound sources. A cost function to be minimized has been developed to obtain the optimal solution with reasonable energy distribution and sufficient sweet area when the distribution of loudspeakers is non-uniform. The performance of the proposed method is verified by the results of computer simulations and subjective tests in the cases of NHK 22.2 and ETRI 10.2 channel configurations.
Convention Paper 8404

11:30

P14-6 On the Scattering of Synthetic Sound Fields—*Jens Ahrens, Sascha Spors*, Deutsche Telekom Laboratories, Technische Universität Berlin, Berlin, Germany

In sound field synthesis a given arrangement of elementary sound sources is employed in order to synthesize a sound field with given desired physical properties over an extended region. The calculation of the driving signals of these secondary sources typically assumes free-field propagation conditions. The present paper investigates the scattering of such synthetic sound fields from unavoidable scattering objects like the head and body of a person apparent in the target region. It is shown that the basic mechanisms are similar to the scattering of natural sound fields. Though, synthetic sound fields can exhibit properties different to those of natural sound fields. Consequently, in such cases also the scattered synthetic sound fields exhibit properties different to those of scattered natural sound fields.
Convention Paper 8405

12:00

P14-7 Toward Mass-Customizing Up/Down Generic 3-D Sounds for Listeners: A Pilot Experiment Concerning Inter-Subject Variability—*John Au, Richard So, Andrew Horner*, The Hong Kong University of Science & Technology, Clearwater Bay, Kowloon, Hong Kong

The "delay-and-add" theory (Hebrank and Wright, 1974) was used to calculate 4068 matching scores between each of the 192 HRTFs and the dimensions of 12 pairs of ears for two incident sound directions (30 degrees up and 30 degrees down). Five HRTFs with 0, 25, 50, 75, and 100th percentile average matching scores were selected for two incident directions. These 10 HRTFs were used to produce 10 sound cues (5 from 30 degrees up and 5 from 30 degrees down). Ten listeners participated in a sound localization experiment to localize the 10 sound cues presented in random order and with 5 repetitions. Preliminary results indicated that matching scores can explain up to 22% of the inter-subject variations in localization errors. Potential applications are discussed.
Convention Paper 8406

Workshop 9 — **Sunday, May 15**
09:00 – 11:00 — **Room 2**

WHAT EVERY SOUND ENGINEER SHOULD KNOW ABOUT THE VOICE

Chair: **Eddy B. Brixen**, EBB Consult, Denmark, TC-MA and TC-AF

Panelists: *Evelyn Abberton*, Department of Phonetics and Linguistics University College, London, UK
Adrian Fourcin, Department of Phonetics and Linguistics University College, London, UK
Henrik Kjelin, Vocal Institute, Denmark
Julian McGlashan, Department of Otorhinolaryngology, Queen's Medical Centre Campus, Nottingham University Hospitals, UK
Mikkel Nymand, Timbre Music / DPA Microphones, Denmark
Cathrine Sadolin, Complete Vocal Institute, Denmark

The purpose of this workshop is to teach sound engineers how to listen to the voice before they even think of microphone picking and knob-turning.

The presentation and demonstrations are based on the "Complete Vocal Technique" (CVT) where the fundamental is the classification of all human voice sounds into one of four vocal modes named Neutral, Curbing, Overdrive, and Edge. The classification is used by professional singers within all musical styles, and has in a period of 20 years proved easy to grasp in both real life situations and also in auditive and visual tests (sound examples and laryngeal images/Laryngograph® waveforms). These vocal modes are found in the speaking voice as well.

Cathrine Sadolin, the developer of CVT, will involve the audience in this workshop, while explaining and demonstrating how to work with the modes in practice to achieve any sound and solve many different voice problems like unintentional vocal breaks, too much or too little volume, hoarseness, and much more. Julian McGlashan, Adrian Fourcin, and Evelyn Abberton will explain the physical aspects of the voice and demonstrate laryngograph waveforms and analyses. Mikkel Nymand will explain essential parameters in the recording chain—especially the microphone—to ensure reliable and natural recordings. ➥

Tutorial 5 **Sunday, May 15**
09:00 – 10:30 **Room 3**

AUDIO OVER IP— THE BASICS

Presenter: **Peter Stevens**, BBC R&D, London, UK

With the slow demise of ISDN connections and technology, audio over IP plays an increasingly important role for transporting audio signals from A to B. But this is only one area where audio over IP is used. Many more potential and also existing applications show where we are heading. This tutorial looks at the basics of this technology and pinpoints the advantages as well as the pitfalls and how they can be avoided. Practical examples will underline the concepts.

Tutorial 6 **Sunday, May 15**
09:00 – 10:30 **Room 4**

NO— IS AID "INTERACTIVE" MUSIC

Co-Chairs: **Dave Raybould**, Leeds Metropolitan University, Leeds, UK
Richard Stevens, Leeds Metropolitan University, Leeds, UK

This session will recap a number of common approaches to so called "interactive" music in games through a series of practical in-game demonstrations. We'll discuss what is understood by the terms reactive, adaptive, and interactive and put forward the argument that there's actually very little about the current use of music in games that is truly "interactive." The session will conclude with further examples of some potential future solutions of how we might better align the time-based medium of music with the nonlinear medium of games. This session is intended to be accessible to complete beginners but also thought provoking to old pros!

Sunday, May 15 **09:00** **Room Mouton Cadet**
Standards Committee Meeting SC-04-04 on Microphone Measurement and Characterization

Sunday, May 15 **10:00** **Room Saint Julien**
Technical Committee Meeting on Loudspeakers and Headphones

Workshop 10 **Sunday, May 15**
11:00 – 12:30 **Room 5**

ACOUSTICS FOR SOUND REINFORCEMENT

Chair: **Peter Mapp**, Peter Mapp Associates, Colchester, Essex, UK

Panelist: *Ben Kok*, Ben Kok – Acoustic Consulting

Traditionally, acoustic design for performance spaces focuses on optimum acoustics for non-reinforced performances. However, a fine concert hall for symphonic music can be very troublesome for reinforced sound. Even performance spaces designed with variable acoustics are mostly optimized for acoustic sources, little consideration is given to the needs for reinforced sound.

This session addresses the different acoustic requirements for reinforced sound as opposed to non-reinforced sound. Topics to be discussed will include:

– Reverberation, support or interference?

– Source directivity, positioning and aiming

– Discrete reflections and how to avoid them

– What to include in a new venue for optimum adaptation for reinforced sound

– Practical measures in existing situations

– Workarounds in acoustic harsh environments.

Tutorial 7 **Sunday, May 15**
11:00 – 12:30 **Room 3**

LOUDNESS AND EBU R 128— A BASIC TUTORIAL OF THE WHY AND WHAT

Presenter: **Florian Camerer**, ORF, Austrian TV, and Chairman of EBU PLOUD

The EBU recommendation R 128 "Loudness normalization and permitted maximum level of audio signals" is a milestone for establishing a new way to level audio signals. Together with four supporting documents it provides the framework for the transition into this new world of leveling audio based on loudness, not peaks. The chairman of the EBU group PLOUD will present R 128 in detail with the help of examples and will give an outlook of how and where the recommendation is already in practical use.

Sunday, May 15 **11:00** **Room Saint Julien**
Technical Committee Meeting on Hearing and Hearing Loss Prevention

Sunday, May 15 **11:00** **Room Mouton Cadet**
Standards Committee Meeting SC-02-01 on Digital Audio Measurement Techniques

Special Event
AES/APRS/DV247 EVENT: TALKBACK-PRO 2
LEAD ME DOWN THE SIGNAL PATH
Sunday, May 15, 11:15 – 13:00
Room 2

Moderator: **Gil Limor**

Panelists: *TBA*

"All you need is a great mic and a decent pre-amp and you can expect your home recordings to match any studio." A familiar claim but a pretty glib observation, especially when there are so many "affordable" options and such a wide diversity of "subjective" opinions.

"Listen, position, and condition" is the mantra of an engineer trying to capture a live sound accurately. How does each skill impact on the others? Can the flaws in a cheap microphone be rectified by cunning signal manipulation or does the notion of "less is more" apply to signal path?

A group of key experts from academics to manufacturers discuss the fundamentals of audio capture and refinement.

Session P15 **Sunday, May 15** **11:30 – 13:00**
Foyer

POSTERS: SPEECH AND CODING

11:30

P15-1 A New Robust Hybrid Acoustic Echo Cancellation Algorithm for Speech

Communication in Car Environments— *Yi Zhou, Chengshi Zheng, Xiaodong Li*, Chinese Academy of Sciences, Beijing, China

This paper studies a new robust hybrid adaptive filtering algorithm for acoustic echo cancellation (AEC) in a car speech communication system. The proposed algorithm integrates the affine projection algorithm (APA) and normalized least mean square (NLMS) algorithm. It can switch between them with a coefficient derivative-based technique to achieve overall optimum convergence performance. To help the algorithm combat deteriorating impulsive interferences widely encountered in car environments and enhance the system robustness in double talk period, robust statistics technique is also incorporated into the proposed algorithm. Experiments are conducted to verify the improved and robust overall AEC convergence performance achieved by the new algorithm.
Convention Paper 8407

11:30

P15-2 Speech Source Separation Using a Multi-Pitch Harmonic Product Spectrum-Based Algorithm — *Rehan Ahmed, Roberto Gil-Pita, David Ayllón, Lorena Álvarez*, University of Alcalá, Alcalá de Henares, Spain

This paper presents an efficient algorithm for separating speech signals by determining multiple pitches from mixtures of signals and assigning the sources to one of those estimated pitches. The pitch detection algorithm is based on Harmonic Product Spectrum. Since the pitch of speech signals fluctuates readily, a frame-based algorithm is used to extract the multiple pitches in each frame. Then, the fundamental frequency (pitch) for each source is estimated and tracked after comparing all the frames. The estimated fundamental frequency of the sources is then used to generate a set of binary masks that allow separating the signals in the Short Time Fourier Transform domain. Results show a considerable separation of the speech signals, justifying the feasibility of the proposed method.
Convention Paper 8408

11:30

P15-3 User-Orientated Subjective Quality Assessment in the Case of Simultaneous Interpreters— *Judith Liebetrau,*[1] *Thomas Sporer*,[1] *Paolo Tosoratti*,[2] *Daniel Fröhlich,*[1] *Sebastian Schneider,*[1] *Jens-Oliver Fischer*[1]
[1]Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany
[2]DG Interpretation, Brussels, Belgium

A study was conducted to explore various subjective quality aspects of audio-visual systems for interpreters. The study was designed to bring the existing requirements for on-site simultaneous interpretation up-to-date and to specify new requirements for remote interpretation (teleconferencing). The feasibility of using objective measurement methods in this context was examined. Several parameters influencing perceived quality, such as audio coding, video coding, room characteristics, and audio-visual latency were assessed. The results obtained are partially contradictory to previous studies. This leads to the conclusion that perceived quality is strongly linked to the focus, background, and abilities of the assessors. The test design, realization, and obtained results are shown, as well as a comparison to studies conducted with different types of users.
Convention Paper 8409

11:30

P15-4 Analysis of Parameters of the Speech Signal Loudness of Professional Television Announcers— *Mirjana Mihajlovic*,[1] *Dejan Todorovic*,[2] *Iva Salom*[3]
[1]Radio Television of Serbia, Belgrade, Serbia
[2]Radio Belgrade, Belgrade, Serbia
[3]Institute Mihajlo Pupin, Belgrade, Serbia

The paper analyzes speech loudness of professional announcers, who do large percentage of television audio signals. Over 100 files of speech signals recorded with same equipment were analyzed. For each file the authors calculated loudness and true peak, according to ITU-R BS 1770 and EBU R128 and RMS. There were criteria adopted for classification of records, and results were statistically analyzed. Analysis showed that the loudness of the speech signal depends on the gender of speakers, intonation, and type of program material, and also that files of greater loudness may not always have higher peak value to files of lower loudness. It was noted there was a significant overlap among the short-term loudness and RMS values calculated with the same time constant.
Convention Paper 8410

11:30

P15-5 Improved Prediction of Nonstationary Frames for Lossless Audio Compression— *Florin Ghido*, Tampere University of Technology, Tampere, Finland

We present a new algorithm for improved prediction of nonstationary frames for asymmetrical lossless audio compression. Linear prediction is very efficient for decorrelation of audio samples, however it requires segmentation of the audio into quasi-stationary frames. Adaptive segmentation tries to minimize the total compressed size, including the quantized prediction coefficients for each frame, thus longer frames that are not quite stationary may be selected. The new algorithm for computing the linear prediction coefficients improves compressibility of nonstationary frames when compared with the least squares method. With adaptive segmentation, the proposed algorithm leads to small but consistent compression improvements up to 0.56%, on average 0.11%. For faster encoding using fixed size frames, without including adaptive segmentation, it significantly reduces the penalty on compression with more than 0.21% on average.
Convention Paper 8411

11:30

P15-6 A Lossless/Near-Lossless Audio Codec for Low Latency Streaming Applications on Embedded Devices— *Neil Smyth*, Cambridge

➥

Silicon Radio, Belfast, Northern Ireland

Increasingly there is a need for high quality audio streaming in devices such as modular home audio networking systems, PMPs, and wireless loudspeakers. As wireless transmission capacities continue to increase it is desirable to utilize this increasing data bandwidth to perform real-time wireless streaming of audio content coded in a lossless or near-lossless format. In this paper we discuss the development of an adaptive audio coding algorithm that balances the design goals of low latency, low complexity, error robustness, and a dynamically-variable bit rate that scales to mathematically-lossless coding under suitable conditions. Particular emphasis is placed on the optimization of the algorithm structure for real-time audio processing applications and the mechanism by which "hybrid" lossless and near-lossless coding is achieved.
Convention Paper 8412

11:30

P15-7 Low-Delay Directional Audio Coding for Real-Time Human-Computer Interaction– *Tapani Pihlajamäki, Ville Pulkki*, Aalto University School of Electrical Engineering, Espoo, Finland

Games and virtual worlds require low-cost, good-quality, and low-delay audio engines. The Short-Time Fourier Transform-based (STFT) implementation of Directional Audio Coding (DirAC) for virtual worlds fulfills the two first demands. Unfortunately, the delay can be perceivably large. In this paper, a modification to DirAC is introduced that uses different time-frequency resolutions for STFT concurrently, which are not aligned in time in reproduction. This leads to the high-frequency non-diffuse content being reproduced as soon as possible and thus reduces the perceived delay. Informal tests show that the delay was short enough for a musician to play an instrument through the processing. Moreover, the results of formal listening tests show that the reduction in quality is perceptible but not annoying.
Convention Paper 8413

Workshop 11 **Sunday, May 15**
11:30 – 13:00 **Room 4**

EMERGING TRENDS IN AUDIO FOR GAMES

Presenters: **Michael Kelly**, Sony Computer Entertainment Europe, London, UK
Steve Martz, THX, CA, USA

This workshop looks at the current state of technology requirements for audio in game applications and discusses emerging trends and the technical requirements imposed by those trends. The event is presented by the Co-Chairs of the AES Technical Committee on Audio for Games. The workshop also summarizes some of the topics presented at the recent 41st International Conference on Audio for Games..

Sunday, May 15 **12:00** **Room Saint Julien**
Technical Committee Meeting on Signal Processing

Sunday, May 15 **13:00** **Room Saint Julien**
Technical Committee Meeting on High Resolution Audio

Sunday, May 15 **13:00** **Room Mouton Cadet**
Standards Committee Meeting SC-04-03 on Loudspeaker Modeling and Measurement

Session P16 **Sunday, May 15** **14:00 – 17:30**
Room 1

AUDIO SIGNAL PROCESSING AND ANALYSIS

Chair: **Jayant Datta**

14:00

P16-1 A New Approach to Designing Decimation Filters for Oversampled A/D Converters– *Jamie A. S. Angus*, University of Salford, Salford, Greater Manchester, UK

This paper presents a new approach to designing the necessary decimation filter in over-sampled noise shaping analog to digital converters. These filters are still finite impulse response designs but they are designed using novel window functions that produce a stop-band attenuation that increases with frequency thus matching the out of band noise characteristics of the noise-shaping modulator. As a result high quality decimation filters can be realized using shorter filter lengths.
Convention Paper 8414

14:30

P16-2 Warped IIR Filter Design with Custom Warping Profiles and its Application to Room Response Modeling and Equalization– *Balázs Bank*, Budapest University of Technology and Economics, Budapest, Hungary

In traditional warped FIR and IIR filters, the frequency-warping profile is adjusted by a single free parameter, leading to a less flexible allocation of frequency resolution. As an example, it is not possible to achieve a truly logarithmic frequency resolution, which would be often desired in audio applications. In this paper a new approach is presented for warped IIR filter design where the filter specification is transformed by any desired (e.g., logarithmic) frequency transformation, and a standard IIR filter is designed to this transformed specification. Then, the poles and zeros of this transformed filter are found and mapped back to the original frequency scale. Due to the approximations in mapping back the poles and zeros, the resulting transfer function may show some discrepancies from its optimal version. This is resolved by an additional optimization of the zeros of the final filter. Examples of loudspeaker-room response modeling and equalization are presented.
Convention Paper 8415

15:00

P16-3 Computationally Efficient Nonlinear Chebyshev Models Using Common-Pole Parallel Filters with the Application to

Loudspeaker Modeling—*Balázs Bank*, Budapest University of Technology and Economics, Budapest, Hungary

Many audio systems show some form of nonlinear behavior that has to be taken into account in modeling. For this, often a black-box model is identified, coming from the generality and simplicity of this approach. One such model is the simplified Volterra model, using parallel branches that have a polynomial-type nonlinearity and a linear filter in series. For example, Chebyshev models use Chebyshev polynomials as nonlinear functions, making the model identification a very straightforward procedure by using logarithmic sweep measurements. This paper proposes a highly efficient implementation of Chebyshev models by using fixed-pole parallel filters for the linear filtering part. The efficiency comes from the fact that parallel filters can have a logarithmic frequency resolution, which better fits the human hearing behavior than traditional FIR or IIR filters. Moreover, the branches can share the same denominators, leading to an additional performance benefit. The proposed model is particularly well suited for the real-time digital simulation of loudspeakers and other weakly nonlinear devices, such as tube guitar amplifiers.
Convention Paper 8416

15:30

P16-4 Event-Driven Real-Time Audio Processing with GPGPUs—*Tiziano Leidi,*[1] *Thierry Heeb,*[1] *Marco Colla*,[1] *Jean-Philippe Thiran*[2]
[1]ICIMSI-SUPSI, Manno, Switzerland
[2]EPFL, Lausanne, Switzerland

Development of real-time audio processing applications for GPGPUs is not without challenges. Parallel processing of audio signals is often constrained by serial dependencies within or between the algorithms. On GPGPUs, insufficient data pressure further limits the attainable performance improvements, as it causes inactivity of the GPU cores. In this paper we analyze the limits of audio processing on GPGPUs and present an approach based on event-driven scheduling, that maximizes data pressure to favor performance improvements. We also present recent enhancements of Audio n-Genie, an open-source development environment for audio-processing applications. By combining Audio n-Genie and the proposed approach, we show that it is possible to increase audio processing speed-up.
Convention Paper 8417

16:00

P16-5 A Comparison of Parametric Optimization Techniques for Tone Matching—*Matthew Yee-King*,[1] *Martin Roth*[2]
[1]Goldsmiths, University of London, London, UK
[2]Reality Jockey Ltd., London, UK

Parametric optimization techniques are compared in their abilities to elicit parameter settings for sound synthesis algorithms, which cause them to emit sounds as similar as possible to target sounds. A hill climber, a genetic algorithm, a neural net, and a data driven approach are compared. The error metric used is the Euclidean distance in MFCC feature space. This metric is justified on the basis of its success in previous work. The genetic algorithm offers the best results with the FM and subtractive test synthesizers but the hill climber and data driven approach also offer strong performance. The concept of sound synthesis error surfaces, allowing the detailed description of sound synthesis space, is introduced. The error surface for an FM synthesizer is described and suggestions are made as to the resolution required to effectively represent these surfaces. This information is used to inform future plans for algorithm improvements.
Convention Paper 8418

16:30

P16-6 On the Multichannel Sinusoidal Model for Coding Audio Object Signals—*Toni Hirvonen*,[1] *Athanasios Mouchtaris*[1,2]
[1]Institute of Computer Science, Foundation for Research and Technology–Hellas (FORTH-ICS) Heraklion, Crete, Greece (now with Dolby Laboratories, Stockholm, Sweden)
[2]University of Crete, Heraklion, Crete, Greece

This paper presents two improvements on a recently proposed multichannel sinusoidal modeling system for coding multiple audio object signals. The system includes extracting the sinusoidal components and an LPC envelope for each object signal, as well as transform coding of the residuals' downmix. The contributions of this paper are: (a) a psychoacoustic model for enabling the system to scale well with multiple object signals, and (b) an improved method to encode the common residual, tailored to the "white" nature of this signal. As a result, sound quality of around 90% on the MUSHRA scale is obtained for 10 simultaneous object signals coded with a total rate of 150 kbit/s, while retaining the individual object parametric representations.
Convention Paper 8419

17:00

P16-7 An Additive Synthesis Technique for Independent Modification of the Auditory Perceptions of Brightness and Warmth—*Asteris Zacharakis, Joshua Reiss*, Queen Mary University of London, London, UK

An algorithm that achieves independent modification of two low-level features that are correlated with the auditory perceptions of brightness and warmth was implemented. The perceptual validity of the algorithm was tested through a series of listening tests in order to examine whether the low-level modification was indeed perceived as independent and to investigate the influence of the fundamental frequency on the perceived modification. A Multidimensional Scaling analysis (MDS) on listener responses to pairwise dissimilarity comparisons accompanied by a verbal elicitation experiment examined the perceptual significance and independence of the two low-level features chosen. This is a first step for the future development of a perceptually based control of an additive synthesizer.
Convention Paper 8420

Session P17 Sunday, May 15 14:00 – 15:30
Foyer

POSTERS: BINAURAL AND SPATIAL AUDIO

14:00

P17-1 Repeatability of Localization Cues in HRTF Data Bases— *Elena Blanco-Martin, Silvia Merino Saez-Miera, Juan José Gomez-Alfageme, Luis Ignacio Ortiz-Berenguer*, Universidad Politecnica de Madrid, Madrid, Spain

Head-Related Transfer Function (HRTF) represents the time-spectral filtering that head and torso do to a sound that goes from a sound source to the ears. These transfer functions bring the localization cues that vary according to sound source position (azimuth and elevation). Human auditory system uses such localization cues for estimating the sound direction. HRTF are used in two ways. One way is for synthesizing binaural sound (virtual audio 3-D). The second way is for analyzing binaural sounds in order to estimate the localization of sound sources. Therefore HRTF are important and useful data for researchers. There are few public HRTF data bases. The most known is the Kemar Data Base from MIT. There is also the Cipic Data Base from UC Davis and the Itakura Data Base. In this paper we present a new public data base for researching use available on the web. The data base has been measured on the Head and Torso Simulator 4100 by Brüel & Kjær. In addition, a comparative study of localization cues from the data bases is carried out for showing their repeatability.
Convention Paper 8421

14:00

P17-2 Dynamic Head-Related Transfer Function Measurement Using a Dual-Loudspeaker Array— *Qinghua Ye, Qiujie Dong, Lingling Zhang, Xiaodong Li*, Institute of Acoustics, Chinese Academy of Sciences, Beijing, China

A dynamic Head-Related Transfer Function (HRTF) measurement method using a dual-loudspeaker array is presented, which reduces measuring requirements and increases the efficiency as well. First, the dual-loudspeaker array emits uncorrelated signals, while head size and head motion are obtained through short-time time-delay estimation. Second, in the approximation of linear time-invariant (LTI) within the testing time, multi-point continuous HRTF measurement is accomplished. Compared with FASTRAK head tracking system, the experimental results confirm the validity of the proposed method.
Convention Paper 8422

14:00

P17-3 Statistical Analysis of Binaural Room Impulse Responses— *Eleftheria Georganti, Alexandros Tsilfidis, John Mourjopoulos*, University of Patras, Patras, Greece

In previous work of the authors, the spectral magnitude of room transfer functions (RTFs) was analyzed using histograms and statistical quantities (moments), such as the kurtosis and skewness. In this paper the above analysis is extended to binaural room impulse responses (BRIRs) and the dependence of the statistical measures on the room acoustical properties, such as the reverberation time, the room size, and the source-receiver distance is discussed. Emphasis is given on the binaural measure of the magnitude squared coherence (MSC), which is considered to be an important cue for binaural hearing related to perceptual aspects such as the source width, the envelopment and the spaciousness. After a brief overview of the existing MSC models, a perceptually-motivated MSC implementation is examined, based on a gammatone filterbank. MSC results for various rooms and source-receiver positions are presented and related to the existing MSC models.
Convention Paper 8423

14:00

P17-4 Spatial Sound and Stereoscopic Vision— *Paul Mannerheim*, University of California, Santa Barbara, Santa Barbara, CA, USA

This paper presents a technique for reproducing coherent audio visual images for multiple users, only wearing 3-D glasses and without utilizing head racking. The recent emergence of 3-D content has increased the demand for technology that can display visual images that are coherent with sound images for multiple users. Audio visual object difference is here investigated for analyzing the size of the sweet spot of a system that combines a visual display technique named stereoscopy with a sound reproduction technique called wave field synthesis. The sweet spot of such a configuration is limited due to differences in characteristics between the sound reproduction system and the visual display; however as a consequence, it is found that the number sources in the wave field synthesis array can be reduced.
Convention Paper 8424

14:00

P17-5 Designing Ambisonic Decoders for Improved Surround Sound Playback in Constrained Listening Spaces— *David Moore*,[1] *Jonathan Wakefield*[2]
[1]Glasgow Caledonian University, Glasgow, Scotland, UK
[2]University of Huddersfield, Huddersfield, UK

Much research has been undertaken to optimize irregular 5-speaker Ambisonic decoders for idealized listening environments. In such environments loudspeaker placement is not restricted and can conform to the ITU 5.1 standard. In domestic settings, the room shape, furniture, and television positioning often restrict speaker placement. It is often the case that a compromised speaker layout is enforced by other domestic requirements. This paper seeks to derive Ambisonic decoders to optimize perceived localization performance for these constrained asymmetrical speaker layouts. This work uses a heuristic search algorithm to derive decoder coefficients and simultaneously optimize loud-

speaker angle within specified bounds. Theoretical results are shown for different orders of newly derived Ambisonic decoders for typical domestic scenarios.
Convention Paper 8425

14:00

P17-6 Decoding for 3-D—*Johannes Boehm*, Technicolor, Research & Innovation, Hannover Germany

Three dimensional spatial sound reproduction using irregular loudspeaker layouts requires a special decoder design. We present the fundamentals of Higher Order Ambisonics (HOA) decoding. Then we focus on beam forming techniques to derive solutions for irregular spaced setups. Panning functions are created by vector base amplitude panning or least-squares methods as patterns and are modeled by spherical harmonics. The required HOA order for effective beam forming proves to be in inverse proportion to the minimal angular spacing of speakers. We demonstrate decoder design using an example setup of 16 speakers, and evaluate objective performance criteria. We conclude that decoders for irregular setups require higher HOA orders compared to decoders for regular setups using the same number of speakers and discuss the consequences.
Convention Paper 8426

14:00

P17-7 Real-Time Reproduction of Moving Sound Sources by Wave Field Synthesis: Objective and Subjective Quality Evaluation—*Michele Gasparini, Paolo Peretti, Stefania Cecchi, Laura Romoli, Francesco Piazza*, Universita Politecnica delle Marche, Ancona, Italy

Wave Field Synthesis (WFS) is an audio rendering technique that allows the reproduction of an acoustic image over an extended listening area. In order to achieve a realistic sensation, true representation of moving sound sources is essential. In this paper a real time algorithm that significantly reduces computational efforts in the synthesis of WFS driving functions in presence of moving sound sources is presented. High efficiency can be obtained taking into account a model based on phase approximation and Short Time Fourier Transform (STFT). The influence of the streaming frame size and of the source velocity on well known sound field artifacts has been studied, considering PC simulations and listening tests.
Convention Paper 8427

14:00

P17-8 Assessing Diffuse Sound Field Reproduction Capabilities of Multichannel Playback Systems—*Andreas Walther, Christof Faller*, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

The generation of subjectively diffuse sound fields is an essential part of creating pleasing synthetic sound fields using loudspeaker playback. A number of studies have been published presenting subjective evaluations of the diffuse sound field reproduction capabilities of different loudspeaker setups. We present a model, based on interaural coherence and interaural level difference, for estimating perceived diffuseness of synthetic sound fields evoked by an arbitrary number of transducers at different positions. The results of different loudspeaker setups and listener orientations are compared.
Convention Paper 8428

Workshop 12 — **Sunday, May 15**
14:00 – 16:00 — **Room 2**

DEAF AID LOOPS AND ASSISTIVE LISTENING SYSTEMS—UNDERSTANDING THE NEEDS AND THE TECHNOLOGY

Chair: **Peter Mapp**, Peter Mapp and Associates, Colchester, UK

Panelists: *Conny Anderson*, Univox
Doug Edworthy, DGE Associates
Ken Hollands, Ampetronic
John Woodgate, J M Woodgate & Associates

Although deaf aid loop systems have been around for a considerable period of time, they are still not well understood, and many systems do not achieve an acceptable standard. Deaf aid loops—or AFILS as they are formally known—are installed in all forms of venue or public building raging from churches, theaters, and cinemas to railway and underground stations. Loop systems can vary from large systems covering many hundreds of seats to counter-top systems intended for individual users. The workshop will review the theory and practice of Audio Frequency Induction Loop Systems (AFILS) as well as discussing other technologies such as infra red. Ways of optimizing system performance and day-to-day problems will be discussed.

Tutorial 8 — **Sunday, May 15**
14:00 – 15:30 — **Room 4**

VIDEO FOR AUDIO ENGINEERS

Presenter: **David Tyas**, Ikon, Worcester, UK

Incorporating video into a system can be a good additional source of revenue for audio professionals. It may be simply adding a projector to a theatre or digital signage to augment a VA system. But, with multiple analogue video formats and a transition to an equally confusing array of incompatible digital formats, what connects together and works, how to you convert between the formats and how do you cable and connect.

This short tutorial will cover the basics of video with the emphasis on the newer formats, how to interface these and how to incorporate legacy products into systems.

Sunday, May 15 **14:00** **Room Saint Julien**
Technical Committee Meeting on Audio for Games

Special Event
AES/APRS EVENT:
THE RECORDING LEGACY OF PHIL SPECTOR
Sunday, May 15, 14:30 – 16:15
Room 3

Moderator: **Barry Marshall**

The First Producer? Phil Spector and the Legacy of the *Svengali* Producer

This lecture will explore the career of Phil Spector, the legendary figure who arguably created the role of the record producer as we know it today. It will trace Spector's historical development from his early career as an artist, his apprenticeship with Atlantic Records and Leiber & Stoller (the first credited producers), his flowering as a producer-mogul with his Philles record label, his initial "retirement" in 1967, his re-invention as the producer of John Lennon and George Harrison to his later sporadic (yet sometimes still seminal) work with artists like the Ramones. The emphasis of the lecture will be on the development of the role of the producer—through Spector's initial creation of his "Wall of Sound" and his *Svengali* approach to artist development—and then through his sublimation of the *Svengali* role into a more measured approach with his work with the gigantic talents of artists like Lennon and Harrison. The lecture will include lots of examples of Spector's work.

Sunday, May 15 15:00 Room Saint Julien
Technical Committee Meeting on Acoustics and Sound Reinforcement

Sunday, May 15 15:00 Room Mouton Cadet
Standards Committee Meeting SC-03-06 and SC-03-07 Digital Library and Archive Systems and Audio Metadata (combined meeting)

Workshop 13 Sunday, May 15
15:45 – 16:45 Room 4

IMPLEMENTATION OF INTERACTIVE MUSIC IN GAMES

Chair: **Erasmus Talbot**, Black Rock Studio, Brighton, UK

This workshop looks at the implementation of interactive music in current video games. It focuses on the approach taken in "Split/Second," an arcade racing game packed with action of an epic scale. To reflect its allusions to Hollywood blockbusters in audio, the team turned their backs to licensed music tracks and opted for a bespoke score. Initially all music was delivered externally, but the talk explains how the team became increasingly involved in the music production process. This approach allowed the team to realize creative possibilities that could only emerge from strong familiarity with the game.

The workshop gives a detailed explanation of the novel production method and playback system used in Split/Second and highlights how bold choices, informed by a particular outlook on interactive music, can lead to a simple, yet effective smoke-and-mirror type solution. It will also highlight some pitfalls that were encountered, which elicited future improvements to the system.

Session EB2 Sunday, May 15 16:00 – 17:30
Foyer

ENGINEERING TOOLS AND METHODS

16:00

EB2-1 Creative Abuse in Time Stretching—*Justin Paterson*, University of West London, London, UK

An area of digital audio manipulation currently in flux is that of time stretching. Following the emergence of real-time granular synthesis as a compositional tool, early sampler-based implementations were pushed beyond "authenticity" to create new timbres in the commercial music of the 1990s. As the algorithms improve, allowing more flexible and transparent implementation today, even more opportunities for a new "creative abuse" exist. This brief will first contextualize through consideration of the metaphor of authenticity in the tape recording of the 1940s and its soon-parallel abuse, which offered new pathways into multi-tracking and Musique Concrète. The brief will chronologize, and continue by examining potential for exploitation of stretching artifacts in some contemporary algorithms, and propose a quantification of this effect.
Engineering Brief 9

16:00

EB2-2 Activity Flow in Music Equalization: The Cognitive and Creative Implications of Interface Design—*Joshua Mycroft,*[1] *Justin Paterson*[2]
[1]Queen Mary University of London, London, UK
[2]University of West London, London, UK

The mixing desk metaphor found in many Digital Audio Workstations (DAW) creates a quantitative visual display that is highly structured and segmented. While this is useful for transmitting large amounts of quantitative data it can inhibit the more intuitive and performative aspects inherent in music mixing. This paper's focus is the cognitive and creative issues encountered using current music production software equalizers and the influence they exert on the initial approach, task workflow, and final output of the user. Equalizers have been chosen to exemplify this, due to their pivotal balance between aural and visual modalities. The paper draws conclusions as to the effectiveness of current software equalizer designs and proposes modifications to design.
Engineering Brief 10

16:00

EB2-3 The Preset Is Dead; Long Live the Preset—*Justin Paterson*, University of West London, London, UK

The use of preset sounds in audio production has long been scorned by professional producers, some of whom have cited a lack of originality or integrity, or perhaps a proliferation of homogenization in productions. Despite this, manufacturers have continued to develop ever-larger ranges of presets, now extending beyond instruments and effects, to EQ and even whole channel strips. Developmental work continues to further automate aspects of the mixing process itself. This brief will examine the implications of presets from yesterday to today, and using Logic Pro as a case study, offer some insight into the relevance of this evolving arena to the professional, and the implications to the enthusiast. It will conclude with some conjecture for the future.
Engineering Brief 11

16:00

EB2-4 Automated Pure-Tone Audiometry Software Tool with Extended Frequency Range—*Thomas Bisitz*,[1] *Andreas Silzle*[2]
[1]HörTech GmbH, Oldenburg, Germany
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

In research institutes and other areas it is necessary to check the hearing of test subjects for listening tests (e.g., for assessing the quality of audio processing algorithms, etc.) by measuring pure tone audiograms. Often, neither a necessary skilled operator nor a special audiometer for high frequencies (>8kHz, important for the evaluation of a lot audio processing algorithms) are available. As an alternative, a software for automated audiogram measurements is presented here. This new kind of software runs on a standard PC with a high-quality sound card and audiological headphones and is operated by the test subject him/herself (self-screening). The implemented adaptive procedure allows fast standard audiogram measurements also for high frequencies up to 16 kHz. The challenges with respect to dynamic range and calibration are discussed.
Engineering Brief 12

16:00

EB2-5 A Simple Reliable Power Amplifier with Minimal Component Count—*John Vanderkooy*, *Kevin B. Krauel*, University of Waterloo, Waterloo, Ontario, Canada

We study an audio power amplifier that has three essential active components: two power MOSFETS and one operational amplifier. Such an amplifier will be reliable because MOSFETS have good safe-operating area properties, and there are no small semiconductors that require high voltage ratings. The topology is that of an op-amp directly driving a grounded-source complementary Class-B MOSFET output stage. The center-tapped power supply for such an amplifier is floating, so each channel must have a separate supply, and there must be a small ±15-volt supply for the op-amp as well. We discuss the design and study the amplifier with simulations and an experimental prototype. It achieves good performance.
Engineering Brief 13

16:00

EB2-6 Parametric Study of Magnet System Topologies for Microspeakers—*Holger Hiebel*, NXP Semiconductors Austria GmbH, "Sound Solutions," Vienna, Austria

A parametric simulation study was done in order to compare 3 different electrodynamic magnet system topologies: center-magnet, ring-magnet, and double-magnet (combined inner and outer magnet ring). The study is based on a row of simulations of the BL-factor (FEM and coil winding calculation), moving mass and effective radiating area of a microspeaker design where the inner coil diameter was changed. The dependency of the sound pressure level, the electrical quality factor, and the resonance frequency in a closed box on the inner coil diameter were derived to yield comparison charts for the 3 topologies.
Engineering Brief 14

16:00

EB2-7 Signal Processing Applications for Automotive Audio—*Robert Cadena*, Visteon Corporation, Van Buren Township, MI, USA

The automotive environment presents unique challenges to audio playback. Modern technology such as smart phones and high-efficiency power trains have created a new generation of audio sources and the need for audio systems capable of dealing with them. This Engineering Brief will introduce emerging automotive audio algorithms (listed below) and their design considerations.
Engineering Brief 15

Sunday, May 15 16:00 Room Saint Julien
Technical Committee Meeting on Network Audio Systems

Workshop 14 **Sunday, May 15**
16:30 – 18:00 **Room 2**

CHAMPAGNE LIFESTYLE ON A BEER BUDGET

Chairs: **Simon Bishop**, UK
Richard Merrick, UK

Simon Bishop and Richard Merrick contrast location audio acquisition from both ends of the budget spectrum. Discussing and comparing techniques and tricks collectively acquired from 60 man-years of experience, from being awash with money and equipment to begging, borrowing, and hunting on eBay. Light-hearted, but informative, both will prove that it's not the size of your nail, but the skill of the guy with the hammer!

Regardless of your budget, we'll prove you still have to have STANDARDS!!

Student/Career Development Event
RECORDING COMPETITION—PART 2
Sunday, May 15, 16:30 – 18:30
Room 3

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 Multi-Track Studio Recording 16:30 to 17:30

• Modern Multi-Track Studio Recording 17:30 to 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 Monday 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.

Workshop 15 **Sunday, May 15**
17:00 – 18:00 **Room 4**

GREEN ISSUES AND THE FUTURE OF TOURING SOUND

Chair: **Phil Anthony**, Martin Audio

Panelists: *Catherine Langabeer*, Julie's Bicycle
Claudio Lastrucci, Powersoft
Andy Mead, Firefly Solar

A 2007 study by the University of Oxford's Environmental Change Institute, commissioned by music industry "greening" body Julie's Bicycle, found that the annual overall Greenhouse gas emissions for the UK music industry alone were approximately half a million tons of carbon. Of this total, three quarters of the impact were derived from activities associated with live music performances.

With such a large share of the UK music industry's emissions, finding environmentally friendly solutions to the live sector's carbon and environmental impacts is an ongoing concern—and an opportunity to get ahead of artist and audience demand, more volatile energy costs with all the knock-on impacts, and emerging regulation.

Phil Anthony from Martin Audio, manufacturers of professional audio sound products, will chair this session and will also be explaining how they reduced the carbon footprint of operating their new touring sound system. Alongside Phil, panelists from Julie's Bicycle, a non-profit company working with the creative industries to coordinate best practice in sustainability and greenhouse gas emission reduction; Powersoft, the developers of efficient "Green Audio Power" amplifiers; and Firefly Solar, whose provision of AV equipment and solar generators aims to move the industry towards sustainable power solutions, will present information on how members of the live community are implementing green initiatives in their events, tours, and technology, many at low or no additional cost.

Tutorial 9 **Sunday, May 15**
17:00 – 18:00 **Room 5**

POWER LIMITATIONS IN MICRO TRANSDUCERS

Presenter: **Bo Rohde Pedersen**, Rohde Consulting
Andreas Eberhardt Sørensen, Pulse HVT

Voice coil temperature is typically a HOT topic when discussing power handling of transducers and many manufactures use IEC tests to specify and rank the input power a specific transducer can endure. For micro speakers this situation is also the case but as we are dealing with small coils and thin wires, the need to know and map the thermal factors are even more important.

The factors are the input power, losses, and cooling effects that happen around the voice coil. In this session these factors will be mapped and investigated as well as demonstrated. The first part of the session holds a mini tutorial in using an infrared thermal camera to measure the voice coil temperature as well as a discussion on what must be in focus when setting up such a system to measure micro transducer thermal aspects.

As the temperature increases our tests shows temperatures of up to 200°C in the voice coil and as a result the Rdc (Dc resistance) also increases and the resulting power dissipated decreases. The theoretical thermal model of the transducer will be examined and revised to fit the findings of our research. The used thermal model is of second order and includes nonlinearities to describe the change of the voice coil resistance and the velocity depending cooling of the voice coil.

Under the test are two micro loudspeakers one with a round and one with a rectangular voice coil because the request for higher output and lower resonance frequencies have moved the development from small round coils to larger (relative to the transducer) rectangular coils. The effect of this transition will be documented and the differences as well as the change in the thermal factors will be explained. We want to prove with this session that the trend has an overall positive effect on the overall thermal behavior of the voice coil and the micro transducer's capability to endure higher input power.

Finally this tutorial should end with a discussion of the factual situation in the micro transducer market that the customers are asking for more power from the driving amplifier and first question is; will the micro transducer be able to handle that input power? Today the amplifiers deliver 1-1.5 W in 8 Ohms with a build in voltage step up but customers are asking for 2-3 W and micro transducers will have to be designed to handle input abuse of this scale. Once this power handling is in place the next question will be; what is gained on the output side of the micro transducer?

Sunday, May 15 **17:00** **Room Saint Julien**
Technical Committee Meeting on Semantic Audio Analysis

Sunday, May 15 **17:00** **Room Mouton Cadet**
Standards Committee Meeting SC-04-01 on Acoustics and Sound Source Modeling

Special Event
BANQUET
Sunday, May 15, 18:45 – 22:30
The Musical Museum
399 High Street
Brentford, Middlesex

This year the Banquet will take place at The Musical Museum containing one of the world's foremost collections of automatic instruments. From the tiny clockwork Musical Box to the self playing "Mighty Wurlitzer," the collection embraces an impressive and comprehensive array of sophisticated reproducing pianos, orchestrions, orchestrelles, residence organs, and violin players. On the ground floor there are 4 galleries to display working instruments; upstairs is a concert hall seating 230, complete with stage and, of course, an orchestra pit from which the Wurlitzer console will rise to entertain you, just as it did in the cinema in the 1930s.

While having a pre-dinner drink, you will be able to wander round the exhibits and even listen to some of them. There will also be time at the end of the evening to look round if you missed it before dinner.

The ticket price includes all food and drinks and the bus to the museum and back. There are limited spaces at this event, so book now to guarantee a place! The bus will leave the Novotel at 18:45 and return approximately half an hour later for the second group. The bus will be available at 22:30 for those needing an early night and will return to collect the final group approximately half-an-hour later.

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

Session P18 Monday, May 16 09:00 – 13:00
Room 1

SOURCE ENHANCEMENT

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

09:00

P18-1 Dereverberation in the Spatial Audio Coding Domain—*Markus Kallinger,*[1] *Giovanni Del Galdo,*[1] *Fabian Kuech*,[1] *Oliver Thiergart*[2]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]International Audio Laboratories, Erlangen, Germany

Spatial audio coding techniques are fundamental for recording, coding, and rendering spatial sound. Especially in teleconferencing, spatial sound reproduction helps in making a conversation feel more natural reducing the listening effort. However, if the acoustic sources are far from the microphone arrangement, the rendered sound may easily be corrupted by reverberation. This paper proposes a dereverberation technique, which is integrated efficiently into the parameter domain of Directional Audio Coding (DirAC). Utilizing DirAC's signal model we derive a parametric method to reduce the diffuse portion of the recorded signal. Instrumental quality measures and informal listening tests confirm the efficiency of the proposed method to render a spatial sound scene less reverberant without introducing noticeable artifacts.
Convention Paper 8429

09:30

P18-2 Blind Single-Channel Dereverberation for Music Post-Processing—*Alexandros Tsilfidis, John Mourjopoulos*, University of Patras, Patras, Greece

Although dereverberation can be useful in many audio applications, such techniques often introduce artifacts that are unacceptable in audio engineering scenarios. Recently, the authors have proposed a novel dereverberation approach, suitable for both speech and music signals, based on perceptual reverberation modeling. Here, the method is fine-tuned for sound engineering applications and tested for both natural and artificial reverberation. The results show that the proposed technique efficiently suppresses reverberation without introducing significant processing artifacts and the method is appropriate for the post-processing of music recordings.
Convention Paper 8430

10:00

P18-3 Joint Noise and Reverberation Suppression for Speech Applications—*Elias K. Kokkinis, Alexandros Tsilfidis, Eleftheria Georganti, John Mourjopoulos*, University of Patras, Patras, Greece

An algorithm for joint suppression of noise and reverberation from speech signals is presented. The method requires a handclap recording that precedes speech activity. A running kurtosis technique is applied in order to extract an estimation of the late reflections of the room impulse response from the clap while a moving average filter is employed for the noise estimation. Moreover, the excitation signal derived from the Linear Prediction (LP) analysis of the noisy speech along with the estimated power spectrum of the late reflections are used to suppress late reverberation through spectral subtraction while a Wiener filter compensates for the ambient noise. A gain magnitude regularization step is also implemented to reduce overestimation errors. Objective and subjective results show that the proposed method achieves significant speech enhancement in all tested cases.
Convention Paper 8431

10:30

P18-4 System Identification for Acoustic Echo Cancellation Using Stepped Sine Method Related to FFT Size—*TaeJin Park, Seung Kim, Koeng-mo Sung*, Seoul National University, Seoul, Korea

A stepped sine method was applied for system identification to cancel acoustic echoes in a speaker phone system that has been widely used in recent mobile devices. We applied the stepped sine method by regarding Discrete Fourier Transform (DFT) as a uniform-DFT filter bank. By using this stepped sine method, we were able to obtain more accurate and detailed characteristics of non-linearity, dependent on the amplitude and frequency of speech. We stored the non-linearity information into linear transform matrices and estimated the responses of the mobile device speaker. The proposed method exhibits higher echo return loss enhancement (ERLE) and increased correlation when compared to the conventional method.
Convention Paper 8432

11:00

P18-5 Using Spaced Microphones with Directional Audio Coding—*Mikko-Ville Laitinen*,[1] *Fabian Kuech*,[2] *Ville Pulkki*[1]
[1]Aalto University, Espoo, Finland
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Directional audio coding (DirAC) is a perceptually motivated method to reproduce spatial sound, which typically uses input from first-order coincident microphone arrays. This paper presents a method to additionally use spaced microphone setups with DirAC. It is shown that since diffuse sound is incoherent between spatially separated microphones at certain frequencies, no decorrelation in DirAC processing is needed, which improves the perceived quality. Furthermore, the directions of sound sources are perceived to be more accurate and stable.
Convention Paper 8433

11:30

P18-6 Parameter Estimation in Directional Audio Coding Using Linear Microphone Arrays—*Oliver Thiergart*,[1] *Michael Kratschmer,*[2] *Markus* ➥

Kallinger,[2] *Giovanni Del Galdo*[2]
[1]International Audio Laboratories, Erlangen, Germany
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Directional Audio Coding (DirAC) provides an efficient description of spatial sound in terms of few audio downmix signals and parametric side information, namely the Direction Of Arrival (DOA) and the diffuseness of the sound. Traditionally, the parameters are derived based on the active sound intensity vector that is often determined via 2-D or 3-D microphone grids. Adapting this estimation strategy to linear arrays, which are preferred in various applications due to form factor constraints, yields comparatively poor results. This paper proposes to replace the intensity based DOA estimation in DirAC by a specific estimation of signal parameters via rotational invariant techniques, namely ESPRIT. Moreover, a diffuseness estimator exploiting the correlation between the array sensors is presented. Experimental results show that the DirAC concept can be applied in practice also in conjunction with linear arrays.
Convention Paper 8434

12:00

P18-7 Extraction of Voice from the Center of the Stereo Image—*Aki Härmä, Munhum Park*, Philips Research Laboratories Eindhoven, Eindhoven, The Netherlands

Detection and extraction of the center vocal source is important for many audio format conversion and manipulation applications. First, we study some generic properties of stereo signals containing sources panned exactly to the center of the stereo image and propose an algorithm for the separation of a stereo audio signal into a center and side channels. In the 128th AES Convention a paper was presented (Convention Paper 8071) on listening tests comparing the perception of the widths of the stereo images of synthetic signal combinations. In this paper the same experiment is repeated with real stereo audio content using the proposed center separation algorithm. The main observation is that there are clear differences in the results. The reasons for the differences are discussed in light of the literature and analysis of the test signals and their binaural characteristics in the listening test setup.
Convention Paper 8435

12:30

P18-8 Directional Segmentation of Stereo Audio via Best Basis Search of Complex Wavelet Packets—*Jeremy Wells*, University of York, York, North Yorkshire, UK

A system for dividing time-coincident stereo audio signals into directional segments is presented. The purpose is to give greater flexibility in the presentation of spatial information when two-channel audio is reproduced. For example, different inter-channel time shifts could be introduced for segments depending on their direction. A novel aspect of this work is the use of complex wavelet packet analysis, along with “best basis” selection, in an attempt to identify time-frequency atoms that belong to only one segment. The system is described, with reference to the relevant underlying theory, and the quality of its output for the best bases from complex wavelet packets is compared with methods based on more established analysis and processing methods.
Convention Paper 8436

Session P19 Monday, May 16 09:00 – 10:30
Foyer

POSTERS: ROOM ACOUSTICS

09:00

P19-1 System Identification of Equalized Room Impulse Responses by an Acoustic Echo Canceller Using Proportionate LMS Algorithms—*Stefan Goetze,*[1] *Feifei Xiong,*[1] *Jan Ole Jungmann,*[2] *Markus Kallinger,*[3] *Karl-Dirk Kammeyer,*[4] *Alfred Mertins*[2]
[1]Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany
[2]University of Lübeck, Lübeck, Germany
[3]University of Oldenburg, Oldenburg, Germany (now with Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany)
[4]University of Bremen, Bremen, Germany

Hands-free telecommunication systems usually employ subsystems for acoustic echo cancellation (AEC), listening-room compensation (LRC), and noise reduction in combination. This contribution discusses a combined system of a two-stage AEC filter and an LRC filter to remove reverberation introduced by the listening room. An inner AEC is used to achieve initial echo reduction and to perform system identification needed for the LRC filter. An additional outer AEC is used to further reduce the acoustic echoes. The performance of proportionate filter update schemes such as the so-called proportionate normalized least mean squares algorithm (PNLMS) or the improved PNLMS (IPNLMS) for system identification of equalized impulse response (IR) are shown and the mutual influences of the subsystems are analyzed. If the LRC filter succeeds in shaping a sparse overall IR for the concatenated system of LRC filter and room impulse response (RIR) the PNLMS performs best since it is optimized for the identification of sparse IRs. However, the equalization may be imperfect due to channel estimation errors in periods of convergence and due to the so-called tail-effect of AEC, i.e., the fact that only the first part of a RIR is identified in practical systems. The IPNLMS is more appropriate in this case to identify the equalized IR.
Convention Paper 8437

09:00

P19-2 Virtual Room Acoustics: A Comparison of Techniques for Computing 3-D–FDTD Schemes Using CUDA—*Craig Webb, Stefan Bilbao*, University of Edinburgh, Edinburgh, Scotland, UK

High fidelity virtual room acoustics can be

approached through direct numerical simulation of wave propagation in a defined space. Three-dimensional Finite Difference Time Domain schemes can be employed and adapt well to a parallel programming model. This paper examines the various approaches for calculating these schemes using the Nvidia CUDA architecture. We test the different possibilities for structuring computation, based on the available memory objects and thread-blocking model. A standard test simulation is computed at double precision under different arrangements. We find that a 2-D extended tile blocking system, combined with shared memory usage, produces the fastest computation for our scheme. However, shared memory usage is only marginally faster than direct global memory access, using the latest FERMI GPUs.
Convention Paper 8438

09:00

P19-3 Acoustic Parameters of Chosen Orthodox Churches Overview and Preliminary Psychoacoustic Tests Using Choral Music— *Pawel Maleck, Jerzy Wiciak*, AGH University of Science and Technology, Kraków, Poland

A lot of acoustic research was done for Roman Catholic churches and because of some differences in traditions and culture comparing to Orthodoxy, it is not the best idea to use its acoustic estimators for Orthodox churches. The paper shows results of measurements in some Orthodox churches in Poland and the proposal of psychoacoustic tests using the convolution technique, which would allow formulating the new acoustic outlines for Orthodox churches. The research has been done especially considering choral music, which is an inseparable part of Eastern Christian culture; so as a test, there were Orthodox choir sound samples recorded in an anechoic chamber convoluted with impulse response of measured churches.
Convention Paper 8439

09:00

P19-4 A Comparative Study of Various "Optimum" Room Dimension Ratios— *John Sarris*, Aretaieio University Hospital, Athens, Greece

Various "optimum" room dimension ratios that have been proposed in the literature are studied and compared. Since each proposal is based on a different criterion, independent objective measures of the acoustic quality of the various room ratios are applied in this paper. The most straightforward metric is the flatness of a room's corner to corner frequency response, but since this is not representative of the variations of the sound pressure within the closed space, different metrics are employed to quantify these variations. Simulation results are presented to evaluate the effectiveness of the various ratios for the case of a small and a larger room.
Convention Paper 8440

09:00

P19-5 Perception of Spatial Distribution of Room Response Reproduced Using Different Loudspeaker Setups— *Javier Gomez*,[1] *Rafael C. D. Paiva*,[1,2] *Kai Saksela*,[1] *Thomas Svedström*,[1] *Ville Pulkki*[1]
[1]Aalto University, Espoo, Finland
[2]Nokia Technology Institute INdT, Brazil

A listening test was conducted to assess the effect that different direct-to-reverberant ratios, loudspeaker setups, and reverberation times have on the directional perception of a synthetic room response. The results show that the perceived spatial distribution of reverberation is dependent on the speaker setup, on the reverberation time, and on the listener. Additionally, they show that a small amount of direct sound, even when barely masked by reverberation, shifts the overall perception to the front.
Convention Paper 8441

09:00

P19-6 Modal Analysis and Sound Field Simulation of Vibration Panels in a Free Sound Field and in a Rectangular Enclosure— *Kazuhiko Kawahara*,[1] *Akihiro Sonogi*,[1] *Shin-ya Sato*[1,2]
[1]Kyushu University, Fukuoka, Japan
[2]Nittobo Acoustic Engineering Co., Ltd., Tokyo, Japan

Several types of implementation of panel loudspeakers, for example, distributed mode loudspeaker (DML), are proposed. The diaphragm of a DML is thought to behave like a bending plate. The diaphragm of an electrostatic loudspeaker is thought to move as a rigid plate. What is the difference of an acoustical sound field generation feature with each implementation? In this paper we used models of same dimensional rectangular panel. One is a rigid diaphragm. Another is a bending panel that has an actuator in the center of it. We made simulations of directivity and sound pressure distribution in a rectangular room. We made simulations of sound pressure distribution in a rectangular room. We could show less coherent SPL distribution in the case of a bending panel.
Convention Paper 8442

09:00

P19-7 Simple Room Acoustic Analysis Using a 2.5 Dimensional Approach— *Patrick Macey*, PACSYS Limited, Nottingham, UK

Cavity modes of a finite bounded region with rigid boundaries can be used to compute the steady state harmonic response for excitation of an acoustic cavity by a point source. In Cuboid domains, with analytical modes, this is straightforward. For more general regions, determining a set of orthonormal modes is more difficult. The finite element method could be used for arbitrary 3-D regions. The current work investigates a hybrid numerical/analytical method, applicable to rooms of constant height, but arbitrary cross section. A 2-D FE analysis is used to compute the cross-section modes/frequencies. Three-dimensional modes are constructed as an outer product of the cross section modes and 1-D modes for the height. Comparison is made with BEM computations.
Convention Paper 8443

09:00

P19-8 Early Reflections Design for Natural Stereo Sound Listening—*Chul-Jae (Jay) Yoo*, RADSONE, Inc., Gyeonggi-do, Korea

For natural stereo sound in rooms, some considerations must be followed. For example, to maximize time density, Feedback Delay Network (FDN) is used in which the feedback matrix has every element consisting of absolute value 1. To model faster high frequency decay according to the increasing number of reflections, absorption filters after delay lines in the FDN are generally used. Mixing matrix for uncorrelated L/R output is also used. In this paper an elaborate design scheme of the first early reflection portion in early reflections was proposed to obtain higher IACC values than those of conventional systems, resulting in a more precise sound source image with ambience.
Convention Paper 8444

Workshop 16 **Monday, May 16**
09:00 – 11:00 **Room 3**

UP AND DOWNMIXING—APPROACHES, SOLUTIONS

Chair: **Florian Camerer**, ORF – Austrian TV

Panelists: *Tom Allom*, Penteo
Antoine Hurtado, Isostem
Clemens Par, Swissaudec
Pieter Schillebeeckx, Soundfield
Christian Struck, Lawo

Upmix solutions are coming to the market with increasing speed. Broadcasters need those solutions especially for HD services to fulfill the expectations of the audience for consistent surround sound presentation. Various approaches can be found that try to provide a convincing illusion of a discrete surround sound signal. In this workshop, 6 company representatives will present their products in a short introduction. Then their upmix results of 2 stereo tracks previously given to them will be played and discussed. Also, downmix aspects will be touched as this is also a vital ingredient of a suitable surround sound signal.

Tutorial 10 **Monday, May 16**
09:00 – 10:30 **Room 4**

ANALOG STANDARDS—WHY ARE THERE SO MANY?

Presenter: **Sean Davies**, S.W. Davies Ltd., Aylesbury, UK

"We like standards, that's why we have so many."
(variously attributed)

Although most current audio work is in the digital domain there remains a vast legacy of recordings made in analog, access to which will be required for some considerable time to come. Knowledge is therefore required of the standards applicable to each recording, also the relevant technical literature.

This tutorial will seek to explain how and why the seemingly perverse plurality of standards arose. Fields covered will be: (1) Impedances; why 30 ohms ("R"), 200R, 300R, 600R; (2) Levels; the TU, the Neper, the Decibel; (3) Metering, the VU versus the PPM; (4) Disc recordings; speeds, equalization curves; (5) Tape recordings, speeds, equalization curves; (6) A bag of miscellaneous horrors such as color codes, connectors, and so on.

Tutorial 11 **Monday, May 16**
09:00 – 10:15 **Room 5**

ROADIE—HOW TO GAIN AND KEEP A CAREER IN THE LIVE MUSIC INDUSTRY

Presenter: **Andy Reynolds**, Owner, LiveMusicBusiness.com

Live music is a huge industry—65 million people around the world attended a concert in 2010, with gross ticket sales of over $3 billion. Every single one of those concerts needs a sound reinforcement system and audio engineers to rig the system and help enhance the sound coming from stage.

But how do you gain, and more importantly keep, an audio job in the live music business? How do you differentiate between the different audio roles on a show or event? What specialized skills are relevant to concert audio engineering?

This tutorial will examine the role of the audio engineer in modern concert production, look at routes into the industry, offer advice on finding and dealing with clients as well as the technical skills and knowledge that every live audio engineer should be aware of.

Special Event
AES/APRS/DV247 EVENT: TALKBACK-PRO 3
THE MYTHS AND MYSTERIES OF DSP
Monday, May 16, 09:00 – 10:45
Room 2

Moderator: **Wes Maebe**

Where would we be without DSP? Given the colossal and ubiquitous influence DSP processing has on all our functionality, the answer probably would be "way back in the analog world" but is there good DSP and bad DSP? What should our priorities be when it comes to spending on A/D – D/A, soft auxiliaries, tracking and mixing programs, and master/finalizing tools.

Issues with interfaces, latency, compression, and plain taste will always plague the subjective judgments of engineers and producers. Is there a "normal" digital base upon which creativity can grow unhindered by the pitfalls of mismatched technologies, uneven standards, and deliberate hurdles to interoperability.

Clearing the cloudy mists that envelop proprietary toys will be a panel of technologists, enthusiasts, and practitioners.

Monday, May 16 **09:00** **Room Saint Julien**
Standards Committee Meeting AESSC Plenary

Tutorial 12 **Monday, May 16**
10:30 – 12:00 **Room 5**

YOU, A ROOM, AND A PAIR OF HEADPHONES: A LESSON IN BINAURAL AUDIO

Presenter: **Ben Supper**, Focusrite

Stereo recordings that are intended for loudspeaker reproduction do not work properly over headphones. Methods for converting stereo for headphone monitoring have become increasingly elaborate as the availability of digital signal processing has increased. The challenge is to arrive at a stereophonic effect that is as the composer, engineer, and producer intended: stable, with correct perspective, out-of-head localization, and no excessive polarization of the stereo image.

To do these things successfully is not as trivial as may first be assumed. No two listeners are identical; choice of headphones are a matter of taste. Moreover, the illusion needs reverberation, but this must not get in the way of the music.

This tutorial picks a path through the peculiarities of the human hearing system, discussing aspects of our perception of loudspeakers and rooms, and how close we can come to understanding and modeling them completely. Insights will be given into how the human auditory system can be made to believe a virtually-rendered sound field, how to do this cheaply, and why we don't need to cram the entire gamut of reality onto a tiny piece of silicon.

Monday, May 16 10:30 Room Mouton Cadet
Technical Committee Meeting on Fiber Optics for Audio

Workshop 17 Monday, May 16
11:00 – 13:00 Room 2

(WHY) DOES LIVE SOUND HAVE A BAD NAME?

Chair: **Mark Bailey**, UK

Panelists: *Nichola Bailey*, Nichola Bailey School of Voice
Bob Lee, QSC Audio Products
David Millward, Freelance FOH Engineer
Ian Nelson, FOH Engineer, Adlib Audio
Bruce Olson, Acoustic Consultant, ADA Acoustics
Tuomo Tolonen, Shure Distribution UK

In this session we will explore the aspects that contribute to good sound and those that conspire against it. Our panel of practitioners, users and manufacturers will discuss and debate matters determined to be the scourge (and saviour) of live audio.

In addition to discussion, we will attempt to practically demonstrate some of the points. We look forward to seeing you there!

Tutorial 13 Monday, May 16
11:00 – 13:30 Room 4

HOT AND NONLINEAR—LOUDSPEAKERS AT HIGH AMPLITUDES

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

Nonlinearities inherent in electro-dynamical transducer and the heating of the voice coil and magnetic system limit the acoustical output, generate distortion and other symptoms at high amplitudes. The large signal performance is the result of a deterministic process and predictable by lumped parameter models comprising nonlinear and thermal elements. The tutorial gives an introduction into the fundamentals, shows alternative measurement techniques and discusses the relationship between the physical causes and symptoms depending on the properties of the particular stimulus (test signal, music). Selection of meaningful measurements, the interpretation of the results and practical loudspeaker diagnostic is the main objective of the tutorial, which is important for designing small and light transducers producing the desired output at high efficiency and reasonable cost.

Workshop 18 Monday, May 16
11:30 – 13:00 Room 3

THE EUROVISION SONGCONTEST 2010 IN OSLO —A CASE STUDY

Co-chairs: **Gaute Nistov**, NRK Oslo
Ina von Lucas, NRK Oslo

The Eurovision Songcontest is the biggest live transmission in Europe. The demands on all aspects of production are daunting and the bar has been raised in past years to challenging heights. Sound Supervisor Gaute Nistov tells how he stepped up to the plate, and Ina van Lucas will elaborate on the complex communication and routing system installed.

Monday, May 16 12:30 Room Mouton Cadet
Technical Council Meeting

Session P20 Monday, May 16 14:00 – 16:00
Room 1

SUBJECTIVE EVALUATION

Chair: **Bruno Fazenda**, University of Salford, Salford, UK

14:00

P20-1 Selection of Audio Stimuli for Listening Tests —*Jonas Ekeroot,*[1] *Jan Berg,*[1] *Arne Nykänen*[2]
[1]Luleå University of Technology, Piteå, Sweden
[2]Luleå University of Technology, Luleå, Sweden

Two listening test methods in common use for the subjective assessment of audio quality are the ITU-R recommendations BS.1116-1 for small impairments and BS.1534-1 (MUSHRA) for intermediate quality. They stipulate the usage of only critical audio stimuli (BS.1116-1) to reveal differences among systems under test, or critical audio stimuli that represents typical audio material in a specific application context (MUSHRA). A poor selection of stimuli can cause experimental insensitivity and introduce bias, leading to inconclusive results. At the same time this selection process is time-consuming and labor-intensive, and is difficult to conduct in a systematic way. This paper reviews and discusses the selection of audio stimuli in listening test-related studies.
Convention Paper 8445

14:30

P20-2 A Listening Test System for Automotive Audio: PART 5 – The Influence of Listening Environment on the Realism of Binaural Reproduction—*Francois Postel*,[1,2] *Patrick Hegarty*,[3] *Søren Bech*[3]
[1]Bang & Olufsen A/S, Struer, Denmark (now at Arkamys, Paris, France)
[2]Department of Mechanical Vibrations and Acoustics, UTC, Compiegne, France
[3]Bang & Olufsen A/S, Struer, Denmark

Binaural technology is used to capture elements of an in-car sound field and reproduce them over ➥

headphones at another place and time. An experiment to test the influence of the listening environment on the realism of such a binaural reproduction is described. A panel of 12 trained listeners rated a range of stimuli for 6 elicited attributes of sound quality. Ratings are made for the actual sound field in the test vehicle, for a binaural reproduction in the same vehicle, and for a binaural reproduction in a listening room. The results show that the tested binaural reproduction system is able to preserve either the rank order or the perceived magnitudes of the impressions of the sound field for the attributes Precision, Treble, Stereo impression, Bass, and Reverberation, independent of the listening environment.
Convention Paper 8446

15:00

P20-3 Differences in Preference for Noise Reduction Strength between Individual Listeners— *Rolph Houben*,[1] *Tjeerd M. H. Dijkstra*,[2,3] *Wouter A. Dreschler*[1]
[1]Academic Medical Center, Amsterdam, The Netherlands
[2]Radboud University, Nijmegen, The Netherlands
[3]Technical University Eindhoven, The Netherlands

There is little research on user preference for different settings of noise reduction, especially for individual users. We therefore measured individual preference for pairs of audio streams differing in noise reduction strength. Data was analyzed with a logistic probability model that is based on a quadratic preference utility function. This allowed for an estimate of the optimal setting for each individual subject. For five out of ten subjects the optimized setting differed significantly from the optimum obtained for the grouped data. However, the predicted preference for the individual optimum (60%) was only slightly higher than chance level (50%), which can be considered as too weak to advocate individualization of noise reduction for these normally hearing subjects. However, in hearing-impaired subjects this may be different.
Convention Paper 8447

15:30

P20-4 Assessment of Stereo to Surround Upmixers for Broadcasting— *David Marston*, BBC R&D, London, UK

Broadcasters are now transmitting 5.1-channel surround sound as part of their HD TV services. However, since much of the audio content in original program material is 2-channel stereo, broadcasters are required to switch between the two formats. Switching between 2- and 5.1-channel formats can cause problems in decoding the content, including switching artifacts and loudness changes. In general it is preferable to transmit all program audio in 5.1 surround, and this can be achieved by automatically upmixing any 2-channel stereo content to 5.1 format prior to broadcasting. This paper reports on tests designed to assess the subjective performance of a selection of upmixers for use in the broadcast chain.
Convention Paper 8448

Student/Career Development Event
STUDENT DELEGATE ASSEMBLY MEETING— PART 2
Monday, May 16, 12:15 – 13:45
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 P21 Monday, May 16 14:00 – 15:30
Foyer

POSTERS: PROCESSING AND ANALYSIS

14:00

P21-1 Automatic Classification of Musical Audio Signals Employing Machine Learning Approach— *Pawel Zwan, Bozena Kostek, Adam Kupryjanow*, Gdansk University of Technology, Gdansk, Poland

This paper presents a thorough analysis of automatic classification applied to musical audio signals. The classification is based on a chosen set of machine learning algorithms. A database of 60 music composers/performers was prepared for the purpose of the described research. For each of the musicians, 15 to 20 music pieces were collected. All the pieces were partitioned into 20 segments and then parameterized. The feature vector consisted of 171 parameters, including MPEG-7 low-level descriptors and mel-frequency cepstral coefficients (MFCC) complemented with time-related dedicated parameters. The task of the classifier was to recognize the composer/performer and to properly categorize a selected piece of music. The paper also presents and discusses the results of classification.
Convention Paper 8449

14:00

P21-2 Evaluation of Onset Detection Algorithms in Popular Polyphonic Music on a Large Scale Database— *Stephan Hübler, Rüdiger Hoffmann*, Technische Universität Dresden, Dresden Germany

This paper introduces a large database of popular polyphonic music containing drums (10,238 onsets) for the evaluation of onset detection algorithms. The database has been manually annotated by expert listeners. The inter-rater variability leads to an understanding of inter-human variations. Four common detection functions are investigated: spectral difference, high

frequency content, phase deviation, and the psychoacoustic one of Klapuri. We present an additional detection function based on the MPEG7 feature audio spectrum envelope. An adaptive peak picker determines the onsets that are compared with the manual labels. Results show that detection functions based on spectral difference obtain observable better results. The study provides a thorough investigation of onset detection algorithms in popular polyphonic music.
Convention Paper 8450
[Paper presented by Till Moritz Essinger]

14:00

P21-3 Using the Viterbi Algorithm for Error Correction in an Autocorrelation-Based Pitch Detector– *Bob Coover*, Gracenote, Inc., Emeryville, CA, USA

An autocorrelation-based method for detecting the fundamental pitch of an audio signal is presented in which the Viterbi algorithm is used in place of the error correction portion of the detector. The Viterbi algorithm is used to locate the most likely pitch path through the audio file. This method is compared to a typical heuristic and median filtering-based error correction approach that has been historically used in this type of algorithm. The Viterbi algorithm results are significantly better than the typical error correction methods for choosing the best and most plausible path through the pitch estimates.
Convention Paper 8451

14:00

P21-4 Spectral Equalization for GHA-Applied Restoration to Damaged Historical 78 rpm Records– *Teruo Muraoka, Takahiro Miura, Tohru Ifukube*, The University of Tokyo, Tokyo, Japan

The authors have been engaged in the research of In-harmonic Frequency Analysis "GHA," which enables the separation of desired signal-components and noise. Its primary purpose has been noise-reduction. Recently, the authors succeeded in conducting GHA in practical time length and carried out many sound restorations of historical 78 rpm records. Thanks to GHA's sufficient separation of target signal-component from noisy objects, the restored signal is noise-less, however its tone quality is unnatural when it is reproduced using current audio equipment. This is due to fact that the recorded sounds were tuned to match to audio equipment in that age, therefore spectral equalization is necessary. In practice, extreme frequency emphases are required, but it had been impossible because of the existences of scratch noise. GHA-applied restoration removed theses difficulties, and equalization curve was obtained by comparing long-term spectrum of restored music with that of the same recorded music by current musicians. Generally equalizations are very complicated and were done utilizing a parametric equalizer.
Convention Paper 8452

14:00

P21-5 Selection of Approximated Activation Functions in Neural Network-Based Sound Classifiers for Digital Hearing Aids– *Lorena Álvarez, Cosme Llerena, Enrique Alexandre, Roberto Gil-Pita, Manuel Rosa-Zurera*, University of Alcalá, Alcalá de Henares, Spain

The feasible implementation of signal processing techniques on hearing aids is constrained by the limited number of instructions per second to implement the algorithms on the digital signal processor the hearing aid is based on. This adversely limits the design of a neural network-based classifier embedded in the hearing aid. Aiming at helping the processor achieve accurate enough results, and in the effort of reducing the number of instructions per second, this paper focuses on exploring the most adequate approximations for the activation function. The experimental work proves that the approximated neural network-based classifier achieves the same efficiency as that reached by the "exact" networks (without these approximations), but, this is the crucial point, with the added advantage of extremely reducing the computational cost on the digital signal processor.
Convention Paper 8453

14:00

P21-6 Development of Multiband Dynamic Range Compressor Regarding Noise Characteristics – *Hoon Heo, Mingu Lee, Seokjin Lee, Koeng-Mo Sung*, Seoul National University, Seoul, Korea

It is hard to hear sounds from digital TVs or mobile phones in noisy environments because of the masking effect. It could be solved by a simple amplification; however, special process for masking bands will be a further solution in some restricted situation. We proposed an algorithm named "perceptual irrelevant component elimination" using a modified multiband dynamic range compressor, which does not increase the signal level and enhances its perceptual signal-to-noise ratio by about 1 dB for speech signals and about 3 dB for music signals.
Convention Paper 8454

14:00

P21-7 Designing Sets of N Doubly Complementary IIR Filters– *Alexis Favrot, Christof Faller*, Illusonic LLC, Lausanne, Switzerland

A filter design procedure is described for obtaining sets of N doubly complementary IIR filters for any N and any bandpass frequencies. The N doubly complementary IIR filters are built following a tree-like structure based on pairs of doubly complementary IIR filters and additional all-pass filters. The sum of all band signals is an all-pass filtered version of the original audio signal. The complementary IIR filters can be used instead of an analysis filterbank (full rate). The corresponding synthesis filterbank is simply the sum of all band signals. The proposed filters enable high quality delay critical audio processing.
Convention Paper 8455

Workshop 19 **Monday, May 16**
14:00 – 15:30 **Room 4**

AUTOMIXING AND ARTIFICIAL INTELLIGENCE

Chair: **Michael "Bink" Knowles**, Independent Live Sound Mixer

Panelists: *Alice Clifford*, Queen Mary University of London, London, UK
Ron Lorman, Busy Puppy Productions
Josh Reiss, Queen Mary University of London, London, UK

This workshop will cover two aspects of automixing: the current use of automixers for live sound mixing, and a projection of the future of artificial intelligence and automatic functions applied to live sound.

The first part of the workshop will focus on the problems facing the operator who is mixing panel discussions and improvisational theater, and the benefits of automixers under the control of human hands. The second part will consider the aspects of live sound mixing that may be totally automated in the future, including the setting of input gain, high-pass filtering, channel equalization, and the more distant goal of musical balance.

Tutorial 14 **Monday, May 16**
14:00 – 15:30 **Room 3**

SURROUND SOUND FORMATS

Presenter: **Hugh Robjohns**, Technical Editor and Trainer, Project Manager

After many years of Surround Sound becoming mainstream in the broadcast business, this tutorial looks back at the basics of formats and technical details of surround sound, covering diverse matrix techniques like Dolby Surround, compression systems like DTS or Dolby AC-3, as well as broadcast-specific coding schemes like Dolby E. Metadata, lineup tones, and mixing issues with respect to the coding systems will be touched upon, too.

Special Event
AES/APRS EVENT: I DIDN'T GET WHERE I AM TODAY . . .
Monday, May 16, 14:00 – 15:45
Room 2

Moderator: **Peter Filleul**, APRS

Panelists: *David Fisher*
Elliott Randall
Dennis Weinreich

Careers are magically unpredictable things. So many successful icons in the audio world will blame serendipity, a chance meeting, or the whims of luck for how they ended up. Others had a clear direction, a specific focus, and have stuck to their original ambition and some swear by training and qualifications.

Prominent audio characters describe what influenced their "career moves" and whether flexibility, a broad outlook and the "maverick" spirit are essential elements to achieving longevity in the audio business and muse on the expectations generated by audio education and training.

A 90 minute seminar moderated by Peter Filleul of the APRS exploring the differences between the theory and the practice of the career ambitions of a set of accomplished audio addicts.

Workshop 20 **Monday, May 16**
15:45 – 17:45 **Room 4**

LOUDNESS IN BROADCASTING—EBU RECOMMENDATION R 128 STEPS ON THE GAS

Chair: **Florian Camerer**, ORF – Austrian TV, and Chairman of EBU PLOUD

Panelists: *Yannick Dumartineix*, TSR Geneva
Matthieu Parmentier, France Television
Alessandro Travaglini, Fox Italia

The EBU group PLOUD has now published all their intended documents, centering around the loudness recommendation R 128. Implementation of this concept is underway, and the loudness normalization paradigm is gaining ground. The end of the "loudness war" is nigh! At least in broadcasting that seems to be the prospect. . . .

In this workshop, members of PLOUD who already implemented the concepts of R 128 will talk about their experiences and illustrate it with examples.

Tutorial 15 **Monday, May 16**
16:00 – 17:30 **Room 2**

DELAY FX—LET'S NOT WASTE ANY MORE TIME

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

The humble delay is an essential effect for creating the loudspeaker illusion of music that is better than real, bigger than life. A broad range of effects—comb filtering, flanging, chorus, echo, reverb, pitch shift, and more—are born from delay. One device, one plug-in, and you've got a vast pallet of production possibilities. Join us for this thorough discussion of the technical fundamentals and production strategies for one of the most powerful signal processes in the audio toolbox: delay.