AES 134th Convention Program

May 4 – 7, 2013

Fontana di Trevi Conference Centre, Rome, Italy

At recent AES conventions, authors have had the option of submitting complete 4- to 10-page manuscripts for peer-review by subject-matter experts. The following paper has been recognized as winner of the AES 134th Convention Peer-Reviewed Paper Award.

Loudness Measurement of Multitrack Audio Content Using Modifications of ITU-R BS.1770— *Pedro Duarte Pestana*, Catholic University of Oporto (CITAR-UCP), Oporto, Portugal and Lusíada University of Portugal (ILID), Oporto, Portugal; and Centro de Estatística e Aplicacoes (CEAUL-FCUL), Lisbon, Portugal; *Joshua D. Reiss*, Queen Mary University of London, London, UK; *Alvaro Barbos,* Catholic University of Oporto (CITAR-UCP), Oporto, Portugal
Convention Paper 8813

To be presented on Saturday, May 4, in Session P2 —Audio Signal Processing—Part 1

* * * *

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

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

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

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

(a) The paper was accepted for presentation at the AES 134th 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 134th AES Convention Student Paper Award is:

Parametric Spatial Audio Coding for Spaced Microphone Array Recordings—*Archontis Politis*, *Mikko-Ville Laitinen*, Aalto University, Espoo, Finland; *Jukka Ahonen*, Akukon Ltd., Helsinki, Finland; *Ville Pulkki,* Aalto University, Espoo, Finland
Convention Paper 8905

To be presented on Tuesday, May 7, in Session 16 —Spatial Audio—Part 2: 3-D Microphone and Loudspeaker Systems

* * * * *

Student/Career Event
STUDENT DELEGATE ASSEMBLY MEETING PLACE
Saturday 10:00 – 18:30
Sunday – Tuesday 9:00 – 18:30
Foyer

Come visit the SDA Booth to find out about AES student events at chapters around the world. This is also where you will see postings about finalists in the recording and design competitions.

Saturday, May 4 10:00 Sala Saba

Technical Committee Meeting on Perception and Subjective Evaluation of Audio Signals

Session P1 Saturday, May 4
10:30 – 12:30 Sala Carducci

EDUCATION AND SEMANTIC AUDIO

Chair: **Jörn Loviscach**, Bielefeld University of Applied Sciences, Bielefeld, Germany

10:30

P1-1 Alternate Software and Pedagogical Strategies for Teaching Audio Technology to Classically Trained Music Students— *Christopher J. Keyes*, Hong Kong Baptist University, Kowloon, Hong Kong

Despite the many benefits, teaching audio and audio technology to music students brings as many challenges as rewards. Their vastly different scientific and technical backgrounds often require fundamentally different pedagogical

strategies from other subjects in a music curriculum. The dissemination of information via lectures and/or lengthy technical readings is especially problematic. This paper addresses these challenges by presenting a new pedagogical approach centered on a free pedagogical software package: the *Interactive Workshop for Audio Intelligence and Literacy* (i-WAIL). When combined with inquiry-based learning approaches and projects based on the student's area of interest, it has so far proven highly effective with music students of a wide range of previous abilities.
Convention Paper 8809

11:00

P1-2 Auditory-Visual Attention Stimulator— *Adam Kupryjanow, Lukasz Kosikowski, Piotr Odya, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

A new approach to lateralization irregularities formation was proposed. The emphasis is put on the relationship between visual and auditory attention. In this approach hearing is stimulated using time scale modified speech, and sight is stimulated rendering the text of the currently heard speech. Moreover, displayed text is modified using several techniques, i.e., zooming, highlighting, etc. In the experimental part of the paper, results obtained for the reading comprehension training were presented. It was shown that usage of the proposed method could improve this skill in a group of children between the ages of 7 and 8 years.
Convention Paper 8810

11:30

P1-3 Evaluation of Acoustic Features for Music Emotion Recognition—*Chris Baume*, BBC Research and Development, London, UK

Classification of music by mood is a growing area of research with interesting applications, including navigation of large music collections. Mood classifiers are usually based on acoustic features extracted from the music, but often they are used without knowing which ones are most effective. This paper describes how 63 acoustic features were evaluated using 2,389 music tracks to determine their individual usefulness in mood classification, before using feature selection algorithms to find the optimum combination.
Convention Paper 8811

12:00

P1-4 Investigating Auditory Human-Machine Interaction: Analysis and Classification of Sounds Commonly Used by Consumer Devices—*Konstantinos Drossos,*[1] *Rigas Kotsakis,*[2] *Panos Pappas,*[3] *George M. Kalliris,*[2] *Andreas Floros*[1]
[1]Ionian University, Corfu, Greece
[2]Aristotle University of Thessaloniki, Thessaloniki, Greece
[3]Technological Educational Institute of Ionian Islands, Lixouri, Greece

Many common consumer devices use a short sound indication for declaring various modes of their functionality, such as the start and the end of their operation. This is likely to result in an intuitive auditory human-machine interaction, imputing a semantic content to the sounds used. In this paper we investigate sound patterns mapped to "Start" and "End" of operation manifestations and explore the possibility such semantics' perception to be based either on users' prior auditory training or on sound patterns that naturally convey appropriate information. To this aim, listening and machine learning tests were conducted. The obtained results indicate a strong relation between acoustic cues and semantics along with no need of prior knowledge for message conveyance.
Convention Paper 8812

Session P2 — Saturday, May 4
10:30 – 12:30 — Sala Foscolo

AUDIO SIGNAL PROCESSING—PART 1

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

10:30

P2-1 Loudness Measurement of Multitrack Audio Content Using Modifications of ITU-R BS.1770 —*Pedro Duarte Pestana,*[1,2] *Joshua D. Reiss,*[3] *Alvaro Barbosa*[1]
[1]Catholic University of Oporto (CITAR-UCP), Oporto, Portugal
[2]Lusíada University of Portugal (ILID), Oporto, Portugal; and Centro de Estatística e Aplicacoes (CEAUL-FCUL), Lisbon, Portugal
[3]Queen Mary University of London, London, UK

The recent loudness measurement recommendations by the ITU and the EBU have gained widespread recognition in the broadcast community. The material it deals with is usually full-range mastered audio content, and its applicability to multitrack material is not yet clear. In the present work we investigate how well the evaluated perception of single track loudness agrees with the measured value as defined by ITU-R BS.1770. We analyze the underlying features that may be the cause for this disparity and propose some parameter alterations that might yield better results for multitrack material with minimal modification to their rating of broadcast content. The best parameter sets are then evaluated by a panel of experts in terms of how well they produce an equal-loudness multitrack mix and are shown to be significantly more successful.
Convention Paper 8813

11:00

P2-2 Delayless Robust DPCM Audio Transmission for Digital Wireless Microphones—*Florian Pflug, Tim Fingscheidt*, Technische Universität Braunschweig, Braunschweig, Germany

The employment of digital wireless microphones in professional contexts requires ultra-low delay, strong robustness, and high audio quality. On the other hand, audio coding is required in order to comply with the restrictions on the bandwidth of the radio channel, making the resulting

source-coded audio signal more vulnerable to channel distortions. Therefore, in this paper we present a transmission system for differential pulse-code modulated (DPCM) audio with receiver-sided soft-decision error concealment exploiting channel reliability information, explicit redundancy by simple delayless parity-check codes, and residual redundancy within the source-coded audio signal. Simulations on frequency-shift keying (FSK)-modulated channels with additive white Gaussian noise show considerable gains in audio quality compared to hard-decision decoding and soft-decision decoding only exploiting reliability information and 0th-order a priori knowledge.
Convention Paper 8814

11:30

P2-3 Violin Sound Computer Classification Based on Expert Knowledge—*Adam Robak, Ewa Lukasik*, Poznan University of Technology, Poznan, Poland

The paper presents results of the analysis of violins recorded during the final stage of the international violin-makers competition held in Poznan in 2011. In the quest for attributes that are both efficient for machine learning and interpretable for human experts we referred to the research of violin acousticians: Duennwald, Buen, and Fritz and calculated violin sound power in frequency bands recommended by these researchers. The resulting features, obtained for the averaged spectra of the musical pieces played at the competition, were used for clustering and classification experiments. Results are discussed, and a notable experiment is presented where the classifier assigns each analyzed violin to an instrument from the precedent violin-makers' competition (2001) and compares their ranking.
Convention Paper 8815

12:00

P2-4 A Finite Difference Method for the Excitation of a Digital Waveguide String Model—*Leonardo Gabrielli*,[1] *Luca Remaggi*,[1] *Stefano Squartini*,[1] *Vesa Välimäki*[2]
[1]Universitá Politecnica delle Marche, Ancona, Italy
[2]Aalto University, Espoo, Finland

With Digital Waveguide modeling (DWG) a number of excitation methods have been proposed to feed the delay line properly. Generally speaking these may be based on signal models fitting recorded samples, excitation signals extracted from recorded samples or digital filter networks. While allowing for a stable, computationally efficient sound emulation, they may be unable to emulate secondary effects generated by the physical interaction of, e.g., distributed interaction of string and hammer. On the other hand, FDTD (Finite Difference Time Domain) models are more accurate in the emulation of the physical excitation mechanism at the expense of a higher computational cost and a complex coefficient design to ensure numerical stability. In this paper a mixed model is proposed composed of a two-step FDTD model, a commuted DWG, and an adaptor block to join the two sections. Properties of the model are provided and computer results are given for the case of the Clavinet tangent-string mechanism as an example application.
Convention Paper 8816

Workshop 1 — **Saturday, May 4**
10:30 – 12:30 — **Auditorium Loyola**

CAPTURING THE ACOUSTICS OF CONCERT HALLS WITH A LOUDSPEAKER ORCHESTRA

Chair: **Tapio Lokki**, Aalto University, Aalto, Finland

Panelists: *Jukka Pätynen*
Sakari Tervo

The sensory evaluation process to assess the quality of concert hall acoustics requires that the spatial sound in a concert hall can be reproduced accurately with a multichannel reproduction system. This workshop introduces methods to create a listening system, with which a subject can switch between the acoustics of different rooms, e.g., concert halls, in real time. The acoustics is captured by measuring spatial impulse responses from 34 loudspeakers on stage with a microphone array. The impulse responses are spatially encoded for a 24-loudspeaker 3-D setup. Finally, a convolution of anechoic symphony orchestra recordings with spatially decomposed convolution reverbs reproduces the spatial sound from the concert hall for listening and for comparison of acoustics. The workshop includes several sound samples from famous European concert halls.

Session EB1 — **Saturday, May 4**
11:00 – 12:30 — **Foyer**

ENGINEERING BRIEFS—POSTERS: PART 1

11:00

EB1-1 Timecode-Aware Loudness Monitoring: Accelerate Engineers' Everyday Workflow—*Arnaud Laborie, Samuel Tracol, Arnaud Destinay, Jacques Di Giovanni*, Trinnov Audio, Bry sur Marne, France

While EBU R-128 loudness normalization is in the process of being adopted by a majority of European countries, most real-time loudness meters aren't still completely adapted to mixing engineers' workflows, as continuous project measurements are always required to keep consistent loudness values. By slaving the loudness measurement to an incoming time code, every loudness and true peak values are constantly recorded and time stamped, allowing their calculation at any time. Engineers no longer need to manually pause, resume, or even start a measurement over to keep a relevant loudness monitoring.
Engineering Brief 78

11:00

EB1-2 Control of the Audio Signal Using Thermal Parameter for Protection of the Voice Coil—*Oanjin Kim, Keeyeong Cho, Jongwoo Kim*, Samsung Electronics Co., Ltd., Suwon, Korea

This engineering brief presents a procedure for a signal processing method to protect a voice coil from overheating. The basic concept is that of estimating temperature of the voice coil with heat transfer model and controlling the output level before the voice coil is burnt. This paper mainly focuses on the calculation method and various considerations in level control. A basic heat transfer model and a commonly used method in calculation of the thermal parameters are introduced.
Engineering Brief 79

11:00

EB1-3 Automatic Segmentation of Concert Recordings via a Heuristic Approach— *Andrew Ayers*, University of Miami, Coral Gables, FL, USA

In the age of digital recordings, many institutions maintain large databases of concert recordings. While segmentation of these concert recordings for mastering and production is a time-consuming task for humans, this paper presents a novel heuristic algorithm to automate that process. Building on other work in audio segmentation, technique from the music informatics community is used to detect events, classify them, and segment entire concert recordings unsupervised. A brief review of previous work and the methodology used in this approach are provided, as well as the results obtained on a corpus of sixteen concerts.
Engineering Brief 80

11:00

EB1-4 Simulation of a Near Field Loudspeaker System on Headphones—*Erich Meier*, amoenus audio by Erich Meier, Bern, Switzerland

The vast majority of recorded music is produced for reproduction via loudspeakers positioned at the standard 60° stereo triangle. To achieve the same sound impression on headphones with its near-field 180° transducer positions, the sound of the standard stereo 60° triangle has to be simulated. We describe a circuit using different serial and parallel delay-and-filter paths for direct and cross-feed channels, with the goal to achieve accurate near-field loudspeaker-sound and also a good externalized localization.
Engineering Brief 81

11:00

EB1-5 Influence of First Reflections in Listening Room on Subjective Listener Impression of Reproduced Sound—*Hidetaka Imamura*,[1] *Atsushi Marui*,[1] *Toru Kamekawa*,[1] *Masataka Nakahara*[2]
[1]Tokyo University of the Arts, Tokyo, Japan
[2]SONA Corp., Tokyo, Japan

This study investigated the perceptual factors regarding room acoustics such as spatial impression and timbre preferences, with focus on the arrival direction and pattern of the early reflections. Impulse responses were recorded with varying wall reflection and evaluated in a subjective test. Although no significant difference in timbre preferences and some evaluation terms were found in subjective listening test, the variation of early reflections did significantly influence listeners spatial impression. The method and results of the analysis is reported in the presentation.
Engineering Brief 82

11:00

EB1-6 Workload Estimation for Low-Delay Segmented Convolution—*Malte Spiegelberg*, HAW Hamburg, Hamburg, Germany

Zero-delay convolution usually follows a hybrid approach with convolution processing steps in both the time and the frequency domain [Gardner, *J. Audio Eng. Soc.*, vol. 43, 127-136 (1995 Mar.)]. Implementations are likely to ask for dynamic coding, and related workload estimations are focused on efficiency and are limited to the hybrid approach. This paper considers simpler implementations of segmented convolution that work in the frequency domain only and that achieve acceptable low delay for real-time applications when processing several seconds of impulse-response in FIR mode. Workload and memory demand are estimated for this approach in the context of likely application parameters.
Engineering Brief 88

Tutorial 1 — Saturday, May 4
11:00 – 12:30 — Sala Alighieri

MICROPHONE TECHNOLOGY

Presenter: **Ron Streicher**, Pacific Audio-Visual Enterprises, Pasadena, CA, USA

How do microphones work? What differentiates one operating type of transducer from another? How and why do they sound different? What are polar patterns, and how do they affect the way a microphone responds to sound? What is “proximity effect,” and why do some mics exhibit more of it than others? What’s the truth about capsule size—does it really matter? These are just a few of the topics covered in this in-depth tutorial on Microphone Technology.

Workshop 2 — Saturday, May 4
11:00 – 12:30 — Sala Manzoni

APPLICATIONS OF 3-D AUDIO IN AUTOMOTIVE

Chair: **Alan Trevena**, Jaguar Land Rover, Gaydon, UK

Panelists: *Oliver Hellmuth*, Fraunhofer IIS, Eerlangen, Germany
Michael Kelly, DTS, London, UK
Adam Sulowski, Audi AG, Ingolstadt, Germany
Bert Van Daele, Auro-3D, Mol, Belgium

While there are a number of technologies aimed at improving the spatial rendering of recorded sounds in automobiles, few are offer the advantages and challenges as 3-D surround. This workshop will explore theoretical applications and system configurations as well as limitations of 3-D surround applications in automotive. Questions such as what is the reference experience, and how is a system evaluated will be addressed.

Saturday, May 4 11:00 Sala Saba

Technical Committee Meeting on Human Factors in Audio Systems

Saturday, May 4 12:00 Sala Saba

Technical Committee Meeting on Loudspeakers and Headphones

Special Event
AWARDS PRESENTATION AND KEYNOTE ADDRESS
Saturday, May 4, 13:00 – 14:00
Auditorium Loyola

Opening Remarks:
- Executive Director Bob Moses
- President Frank Wells
- Convention Chair Umberto Zanghieri

Program:
- AES Awards Presentation
- Introduction of Keynote Speaker by Convention Chair Umberto Zanghieri
- Keynote Address by Stephen Webber

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:

BRONZE MEDAL AWARD
- Bob Schulein

FELLOWSHIP AWARD
- Zbigniew Kulka
- Lauri Savioja

CITATION AWARD
- Umberto Zanghieri

Keynote Speaker
This year's Keynote Speaker is **Stephen Webber**, internationally recognized author, professor of music production and engineering, and star hip-hop DJ. Author of such successful books as *Turntable Technique: The Art of the DJ*, Webber is an international clinician who presents workshops, master classes, and performance seminars on production, mixing techniques, and songwriting. An Emmy-winning composer who's works include the "Stylus Symphony," Webber was recently appointed Director of Music Technology Innovation at Berklee's new campus in Valencia, Spain. He addressed the Berklee Graduating Class of 2012 and has performed and lectured extensively in the US, China, Australia, Central America, and Europe. Webber is also an accomplished studio designer whose clients include actors Jack Black and Ben Stiller, and a producer/engineer who's worked with NAS, Mark O'Connor, and DJ Premier.

Entitled "Inventing the Album of 2025," Webber's AES Keynote will seek to consider what is technologically, economically, and artistically possible, to decipher what the record album should look, sound, and feel like come 2025.

Saturday, May 4 14:00 Sala Saba

Technical Committee Meeting on Automotive Audio

Saturday, May 4 14:00 Sala Montale

Standards Committee Meeting on SC-02-02, Digital Input/Output Interfacing

Session P3 Saturday, May 4
14:15 – 17:15 Sala Carducci

PERCEPTION

Chair: **Frank Melchior**, BBC R&D, London, UK

14:30

P3-1 The Relation between Preferred TV Program Loudness, Screen Size, and Display Format
—*Ian Dash*,[1] *Todd Power*,[2] *Densil A. Cabrera*[2]
[1]Consultant, Marrickville, NSW, Australia
[2]University of Sydney, Sydney, NSW, Australia

The effect of television screen size and display format on preferred TV program loudness was investigated by listening tests using typical program material. While no significant influence on preferred program loudness was observed from screen size or color level, other preference patterns related to soundtrack content type were observed that are of interest.
Convention Paper 8817

15:00

P3-2 Vibration in Music Perception—*Sebastian Merchel, M. Ercan Altinsoy*, Dresden University of Technology, Dresden, Germany

The coupled perception of sound and vibration is a well-known phenomenon during live pop or organ concerts. However, even during a symphonic concert in a classical hall, sound can excite perceivable vibrations on the surface of the body. This study analyzes the influence of audio-induced vibrations on the perceived quality of the concert experience. Therefore, sound and seat vibrations are controlled separately in an audio reproduction scenario. Because the correlation between sound and vibration is naturally strong, vibrations are generated from audio recordings using various approaches. Different parameters during this process (frequency and intensity modifications) are examined in relation to their perceptual consequences using psychophysical experiments. It can be concluded that vibrations play a significant role during the perception of music.
Convention Paper 8818

15:30

P3-3 An Assessment of Virtual Surround Sound Systems for Headphone Listening of 5.1 Multichannel Audio—*Chris Pike, Frank Melchior*, BBC Research and Development, Salford, UK

It is now common for broadcast signals to feature 5.1 surround sound. It is also increasingly common that audiences access broadcast content on portable devices using headphones. Binaural techniques can be applied to create a spatially enhanced headphone experience from surround sound content. This paper presents a subjective assessment of the sound quality of 12 state-of-the-art systems for converting 5.1 surround sound to a 2-channel signal for head- ➥

phone listening. A multiple stimulus test was used with hidden reference and anchors; the reference stimulus was an ITU stereo down-mix. Dynamic binaural synthesis, based on individualized binaural room impulse response measurements and head orientation tracking, was also incorporated into the test. The experimental design and detailed analysis of the results are presented in this paper.
Convention Paper 8819

16:00

P3-4 Effect of Target Signal Envelope on Direction Discrimination in Spatially Complex Sound Scenarios—*Olli Santala, Marko Takanen, Ville Pulkki*, Aalto University, Aalto, Finland

The temporal envelope of a sound signal has been found to have an effect on localization. Whether this is valid for spatially complex scenarios was addressed by conducting a listening experiment in which a spatially distributed sound source consisted of a target between two interfering noise-like sound sources, all emitting sound simultaneously. All the signals were harmonic complex tones with components within 2 kHz–8.2 kHz and were presented using loudspeaker reproduction in an anechoic chamber. The phases of the harmonic tones of the target signal were altered, causing the envelope to change. The results indicated that prominent peaks in the envelope of the target signal aided in the discrimination of its direction inside the widely distributed sound source.
Convention Paper 8820

16:30

P3-5 A Framework for Adaptive Real-Time Loudness Control—*Andrea Alemanno*,[1] *Alessandro Travaglini*,[2] *Simone Scardapane*,[1] *Danilo Comminiello*,[1] *Aurelio Uncini*[1]
[1]Sapienza University of Rome, Rome, Italy
[2]Fox International Channels Italy, Guidonia Montecelio (RM), Italy

Over the last few years, loudness control represents one of the most frequently investigated topics in audio signal processing. In this paper we describe a framework designed to provide adaptive real-time loudness measurement and processing of audio files and streamed content being reproduced by mobile players hosted in laptops, tablets, and mobile phones. The proposed method aims to improve the users' listening experience by normalizing the loudness level of the content in real-time, while preserving the original creative intent of the original soundtrack. The loudness measurement and adaptation is based on a customization of the High Efficiency Loudness Model algorithm described in the AES Convention Paper #8612 ("HELM: High Efficiency Loudness Model for Broadcast Content," presented at the 132nd Convention, April 2012). Technical and subjective tests were performed in order to evaluate the performance of the proposed method. In addition, the way the subjective test was arranged offered the opportunity to gather information on the preferred Target Level of streamed and media files reproduced on portable devices.
Convention Paper 8821

17:00

P3-6 The Perception of Masked Sounds and Reverberation in 3-D vs. 2-D Playback Systems—*Giulio Cengarle*,[1] Alexandre Pereda[2]
[1]Imm Sound S.A., a Dolby company, Barcelona, Spain
[2]Fundació Barcelona Media, Barcelona, Spain

This paper presents studies on perceptual aspects of spatial audio and their dependency on the playback format. The first study regards the perception of sound in the presence of a masker in stereo, 5.1, and 3-D. Psychoacoustic tests show that the detection threshold improves with the spread of the masker, which justifies the claim that individual elements of dense soundtracks are more audible when they are distributed in a wider panorama. The second study indicates that the perception of the reverberation level does not depend on the playback format. The joint interpretation of these results justifies the claim that in 3-D sound engineers can use higher levels of reverberation without compromising the intelligibility of the sound sources.
Convention Paper 8822

Workshop 3 **Saturday, May 4**
14:15 – 16:15 **Sala Alighieri**

SOUND MIXERS DISCUSS THEIR CRAFT

Chair: **Brian McCarty**, Coral Sea Studios Pty. Ltd., Clifton Beach, QLD, Australia

Panelists: *Marco Coppolecchia*
Simone Corelli
Carols Zarattini

Cinema sound has traditionally been limited in fidelity because optical soundtracks, used until recently, were incapable of delivering full-bandwidth audio to the theaters. As digital cinema replaces film in distribution, sound mixers are now delivering uncompressed lossless tracks to the audience. Leading sound mixers from both Hollywood and Europe will discuss their approach and methodology in producing award-winning soundtracks.

Workshop 4 **Saturday, May 4**
14:15 – 15:45 **Sala Manzoni**

SEMANTIC ANALYSIS FOR SPEECH SIGNALS

Co-chairs: **Jörn Loviscach**, University of Applied Sciences, Bielefeld, Germany
Christian Uhle, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Panelist: *Yves Raimond*, BBC R&D, London, UK

Semantic audio analysis is the automated extraction of meaning from audio signals, with applications in, e.g., music information retrieval, intelligent audio editing, and automated music transcription. The workshop focuses on the semantic analysis of signals containing speech. It discusses speech detection methods and their application to speech enhancement, broadcast monitoring, loudness control and audio coding, presents recent work on automatic tagging of speech in the BBC World Service radio archive, and gives an overview of possible future developments, e.g., speaker identification, emotion recognition, and automated editing of speech recordings. The panelists will discuss the state of the art in semantic audio analysis for speech and its challenges.

Session P4 **Saturday, May 4**
14:30 – 16:30 **Sala Foscolo**

AUDIO SIGNAL PROCESSING—PART 2

Chair: **Leonardo Gabrielli**, Università Politecnica delle Marche, Ancona, Italy

14:30

P4-1 User-Driven Quality Enhancement for Audio Signal—*Danilo Comminiello, Simone Scardapane, Michele Scarpiniti, Aurelio Uncini*, Sapienza University of Rome, Rome, Italy

Classical methods for audio and speech enhancement are often based on error-driven optimization strategies, such as the mean-square error minimization. However, these approaches do not always satisfy the quality requirements demanded by users of the system. In order to meet subjective specifications, we put forward the idea of a user-driven approach to audio enhancement through the inclusion in the optimization stage of an *interactive evolutionary algorithm* (IEA). In this way, performance of the system can be adapted to any user in a principled and systematic way, thus reflecting the desired subjective quality. Experiments in the context of echo cancellation support the proposed methodology, showing significant statistical advantage of the proposed framework with respect to classical approaches.
Convention Paper 8823

15:00

P4-2 Partial Spectral Flatness Measure for Tonality Estimation in a Filter Bank-Based Psychoacoustic Model for Perceptual Audio Coding—*Armin Taghipour,*[1] *Maneesh Chandra Jaikumar,*[1,2] *Bernd Edler,*[1] *Holger Stahl*[2]
[1]International Audio Laboratories Erlangen, Erlangen, Germany
[2]Hochschule Rosenheim, University of Applied Science, Rosenheim, Germany

Perceptual audio codecs use psychoacoustic models for irrelevancy reduction by exploiting masking effects in the human auditory system. In masking, the tonality of the masker plays an important role and therefore should be evaluated in the psychoacoustic model. In this study a partial Spectral Flatness Measure (SFM) is applied to a filter bank-based psychoacoustic model to estimate tonality. The Infinite Impulse Response (IIR) band-pass filters are designed to take into account the spreading in simultaneous masking. Tonality estimation is adapted to temporal and spectral resolution of the auditory system. Employing subjective audio coding preference tests, the Partial SFM is compared with prediction-based tonality estimation.
Convention Paper 8824

15:30

P4-3 A New Approach to Model-Based Development for Audio Signal Processing—*Carsten Tradowsky,*[1,2] *Peter Figuli,*[1] *Erik Seidenspinner,*[1] *Felix Held,*[1] *Jürgen Becker*[1]
[1]Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[2]CTmusic, Karlsruhe, Germany

Today, digital audio systems are restricted in their functionality. For example, a digital audio player still has a resolution of 16-bit and a sample rate of 44.1 kHz. This relatively low quality does not exhaust the possibilities given by modern hardware for music production. In most cases, the functionality is described in software. This abstraction is very common these days, as only few engineers understand the potential of their target hardware. The design-time increases significantly to develop efficiently for the target hardware. Because of the use of common compiler tool chains the software is statically mapped onto the hardware. This restricts the number of channels per processing core to a minimum when targeting high quality audio. One possibility to close the productivity gap, described above, is to use a high-level model-based development approach. The audio signal processing flow is described in a more abstract high level using the model-based development approach. This model is then platform-independently compiled including automatically generated simulation and verification input. Platform-dependent code can be automatically generated out of the verified model. This enables the evaluation of different target architectures and their trade-offs using the same model description. This paper presents a concept to use a model-based approach to describe audio signal processing algorithms. This concept is used to compile C- and HDL-code out of the same model description to evaluate different target platforms. The goal of this paper is to compare trade-offs for audio signal processing algorithms using a multicore Digital Signal Processor (DSP) target platform. Measurements using data parallelism inside the generated code show a significant speedup on the multicore DSP platform. A conclusion will be made regarding the usability of the proposed model-based tool flow as well as the applicability on the multicore DSP platform.
Convention Paper 8825

16:00

P4-4 Accordion Music and its Automatic Transcription to MIDI Format—*Tomasz Maciejewski, Ewa Lukasik*, Poznan University of Technology, Poznan, Poland

The paper is devoted to the problems related to the automatic transcription of the accordion sound. The accordion is a musical instrument from the free-reed aerophone family that is able to produce polyphonic, multi-chord music. First the playing modes are briefly characterized and problems related to the polyphonic nature of the sound is discussed. Then the results of the analysis and MIDI transcription of the right-side monophonic and polyphonic melodies are presented. Finally, an attempt to transcribe music ➥

generated by both sides of the instrument recorded in two channels is presented giving the foundation to further research.
Convention Paper 8827

Student/Career Event
OPENING AND STUDENT DELEGATE ASSEMBLY MEETING – PART 1
Saturday, May 4, 14:30 – 16:00
Auditorium Loyola

Chair: **Philip Waldenberger**

Vice Chair: **Marija Kovacina**

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. Students will also have the opportunity to hear from various AES officers about opportunities and scholarships available through the society.

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 the final afternoon of the convention.

Session P5 **Saturday, May 4**
15:00 – 16:30 **Foyer**

POSTERS: SPEECH PROCESSING

15:00

P5-1 A Speech-Based System for In-Home Emergency Detection and Remote Assistance—*Emanuele Principi,*[1] *Danilo Fuselli,*[2] *Stefano Squartini,*[1] *Maurizio Bonifazi,*[2] *Francesco Piazza*[1]
[1]Universitá Politecnica delle Marche, Ancona, Italy
[2]FBT Elettronica Spa, Recanti (MC), Italy

This paper describes a system for the detection of emergency states and for the remote assistance of people in their own homes. Emergencies are detected recognizing distress calls by means of a speech recognition engine. When an emergency is detected, a phone call is automatically established with a relative or friend by means of a VoIP stack and an Acoustic Echo Canceller. Several low-consuming embedded units are distributed throughout the house to monitor the acoustic environment, and one central unit coordinates the system operation. This unit also integrates multimedia content delivery services and home automation functionalities. Being an ongoing project, this paper describes the entire system and then focuses on the algorithms implemented for the acoustic monitoring and the hands-free communication services. Preliminary experiments have been conducted to assess the performance of the recognition module in noisy and reverberated environments and the out of grammar rejection capabilities. Results showed that the implemented Power Normalized Cepstral Coefficients extraction pipeline improves the word recognition accuracy in noisy and reverberated conditions, and that introducing a "garbage phone" in the acoustic model allows to effectively reject out of grammar words and sentences.
Convention Paper 8828

15:00

P5-2 Assessment of Speech Quality in the Digital Audio Broadcasting (DAB+) System—*Stefan Brachmanski, Maurycy Kin*, Wroclaw University of Technology, Wroclaw, Poland

The methods for assessment of speech quality fall into two classes: subjective and objective methods. This paper includes an overview of selected methods of subjective listening measurements (ACR, DCR) recommended by ITU-T. The influence of a bit-rate value on the sound quality was a subject of research presented in this paper. The influence of the Spectral Band Replication (SBR) process on the speech quality was also investigated. The tested samples were taken from the Digital Audio Broadcasting experimental emission in Poland as well as from an internet network. The subjective assessment for DAB speech signals has been performed with the use of both ACR and DCR methods. It turned out that SBR process influences significantly the speech quality at the lower bit-rates making it as good as for higher bit-rates. It was also found that for higher bit-rate values (96 kbit/s, or higher), the use of both methods causes the different results.
Convention Paper 8829

15:00

P5-3 Investigation on Objective Quality Evaluation for Heavily Distorted Speech—*Mitsunori Mizumachi*, Kyushu Institute of Technology, Kitakyushu, Fukuoka, Japan

Demand for evaluating speech quality is on the increase. It is advisable for evaluating the speech quality to employ the common objective measure for the wide variety of adverse speech signals. Unfortunately, current speech quality measures do not suit for heavily distorted speech signals. In this paper both the applicability and the limit of the perceptual evaluation of speech quality (PESQ) are investigated compared with the subjective mean opinion score (MOS) for noise-added and noise-reduced speech signals. It is found that the PEAQs are compatible with the MOSs for the noise-reduced speech signals in the non-stationary noise conditions.
Convention Paper 8830

15:00

P5-4 Novel 5.1 Downmix Algorithm with Improved Dialogue Intelligibility—*Kuba Lopatka, Bartosz Kunka, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

A new algorithm for 5.1 to stereo downmix is

introduced that addresses the problem of dialogue intelligibility. The algorithm utilizes proposed signal processing algorithms to enhance the intelligibility of movie dialogue, especially in difficult listening conditions or in compromised speaker setup. To account for the latter, a playback configuration utilizing a portable device, i.e., an ultrabook, is examined. The experiments are presented that confirm the efficiency of the introduced method. Both objective measurements and subjective listening tests were conducted. The new downmix algorithm is compared to the output of a standard downmix matrix method. The results of subjective tests prove that an improved dialogue intelligibility is achieved.
Convention Paper 8831

15:00

P5-5 Monaural Speech Source Separation by Estimating the Power Spectrum Using Multi-Frequency Harmonic Product Spectrum—*David Ayllon, Roberto Gil-Pita, Manuel Rosa-Zurera*, University of Alcala, Alcalá de Henares, Spain

This paper proposes an algorithm to perform monaural speech source separation by means of time-frequency masking. The algorithm is based on the estimation of the power spectrum of the original speech signals as a combination of a carrier signal multiplied by an envelope. A Multi-Frequency Harmonic Product Spectrum (MF-HPS) algorithm is used to estimate the fundamental frequency of the signals in the mixture. These frequencies are used to estimate both the carrier and the envelope from the mixture. Binary masks are generated comparing the estimated spectra of the signals. Results show an important improvement in the separation in comparison to the original algorithm that only uses the information from the HPS.
Convention Paper 8832

15:00

P5-6 The Effectiveness of Speech Transmission Index (STI) in Accounting for the Effects of Multiple Arrivals—*Timothy J. Ryan*,[1] *Richard King*,[2] *Jonas Braasch*,[3] *William L. Martens*[4]
[1]Webster University, St. Louis, MO, USA
[2]McGill University, Montreal, Quebec, Canada
[3]Rensselaer Polytechnic Institute, Troy, NY, USA
[4]University of Sydney, Sydney, NSW, Australia

The authors conducted concurrent experiments employing subjective evaluation methods to examine the effects of the manipulation of several sound system design and optimization parameters on the intelligibility of reinforced speech. During the course of these experiments, objective testing methods were also employed to measure the Speech Transmission Index (STI) associated with each of the variable treatments used. Included in this paper is a comparison of the results of these two testing methods. The results indicate that, while STI is capable of detecting many effects of multiple arrivals, it appears to overestimate the degradation to intelligibility caused by multiple arrivals with short delay times.
Convention Paper 8833

15:00

P5-7 Introducing Synchronization of Speech Mixtures in Blind Sparse Separation Problems—*Cosme Llerena, Lorena Álvarez, Roberto Gil-Pita, Manuel Rosa-Zurera*, University of Alcala, Alcalá de Henares, Spain

This paper explores the feasibility of using synchronization of speech mixtures prior to blind sparse source separation methods in order to improve their results. Broadly, methods that assume sparse sources use level and phase differences between mixtures as their features, and they separate signals from them. If each mixture is considerably delayed with respect to the rest of them, the information extracted from these differences can be wrong. With this idea in mind, this paper will focus on using Time Delay Estimation algorithms in order to synchronize the mixtures and observing the improvement that it provokes in a Blind Sparse Source Separation algorithm. The results obtained show the feasibility of using synchronization of the speech mixtures.
Convention Paper 8834

15:00

P5-8 An Embedded-Processor Driven Test Bench for Acoustic Feedback Cancellation in Real Environments—*Francesco Faccenda, Stefano Squartini, Emanuele Principi, Leonardo Gabrielli, Francesco Piazza*, Universitá Politecnica della Marche, Ancona (AN), Italy

In order to facilitate the communication among speakers, speech reinforcement systems equipped with microphones and loudspeakers are employed. Due to the acoustic couplings between them, the speech intelligibility may be ruined and, moreover, high channel gains could drive the system to instability. Acoustic Feedback Cancellation (AFC) methods need to be applied to keep the system stable. In this paper a new Test Bench for testing AFC algorithms in real environments is proposed. It is based on the TMS320C6748 processor, running the Suppressor-PEM algorithm, a recent technique based on the PEM-AFROW paradigm. The partitioned block frequency domain adaptive filter (PB-FDAF) paradigm has been adopted to keep the computational complexity low. A professional sound card and a PC, where an automatic gain controller has been implemented to prevent signal clipping, complete the framework. Several experimental tests confirmed the framework suitability to operate under diverse acoustic conditions.
Convention Paper 8835

Saturday, May 4 15:00 Sala Saba

Technical Committee Meeting on Hearing and Hearing Loss Prevention

Saturday, May 4 16:00 Sala Saba

Technical Committee Meeting on Network Audio Signals

Saturday, May 4 16:00 Sala Montale

Standards Committee Meeting on SC-04-01, Acoustics and Sound Source Modeling

➥

Tutorial 2 **Saturday, May 4**
16:30 – 18:30 **Auditorium Loyola**

ARE WE THERE YET? THE ULTIMATE ULTRA-PORTABLE PRODUCTION/RECORDING STUDIO. FROM IDEA TO FINAL MASTER: HOW TO WRITE, SEQUENCE, PRODUCE YOUR MUSIC, AND CONTROL YOUR STUDIO USING ONLY YOUR IPAD

Presenter: **Andrea Pejrolo**

In this highly interactive and hands-on presentation you will learn the tools, techniques, tips, and tricks required to write, produce, and mix a song using only your iPad.

Through practical examples and scenarios you will learn how to:

• Pick the best software for sequencing, producing, and mixing your music
• Pick the best iPad-compatible hardware tools (microphones, audio interface, MIDI interfaces, controller, etc.)
• Setup your mobile production studio
• Sketch your musical ideas
• Use your iPad as a creative inspirational tool for music composition and sound design
• Sequence and arrange your music ideas on your iPad
• Add real instruments and vocals
• Do a final professional mix
• Master your final mix

In addition, through practical examples and scenarios, you will learn: how to set up your studio in order to use your iOS device as a controller; configuration of iOS device for Logic Pro, Pro Tools, Digital Performer, Live, and Cubase/Nuendo; wireless (MIDI over WI-FI) and wired MIDI connection; proprietary versus open source (OSC, etc.) options; designing your own graphic interface and controllers; ergonomic aspects of using your iOS device in the studio.

Who should attend? Anyone who wants to create some music with their iPads, from beginners to advanced, Musicians, producers, recording engineers, home studio owners from intermediate to advanced.

Tutorial 3 **Saturday, May 4**
16:30 – 18:30 **Sala Alighieri**

MUSIC PRODUCTION FOR FILM

Presenters: **Simon Franglen**
Brian McCarty

Music production is the last piece of the production puzzle to be undertaken in a film production. It occurs after the film is edited and just prior to sound mixing. The composers have developed new tools to prepare the music including temp synth scores and other techniques. Simon Franglin is a Golden Globe nominee and has worked on *Avatar*, *Titanic*, and *Skyfall 007*, among many others.

Tutorial 4 **Saturday, May 4**
17:00 – 18:30 **Sala Foscolo**

UNDERSTANDING MICROPHONE SPECIFICATIONS

Presenter: **Eddy B. Brixen**, EBB-consult, Smorum, Denmark

There are lots and lots of microphones available to the audio engineer. The final choice is often made on the basis of experience or perhaps just habits. (Sometimes the mic is chosen because of the looks … .) Nevertheless, there is good information to be found in the microphone specifications. This tutorial will present the most important microphone specs and provide the attendee with up-to-date information on how these specs are obtained and understood. It also takes a critical look on how specs are presented to the user, what to look for and what to expect.

This tutorial is a follow up on the project X85, which has been taking place in the AES standards committee (SC 04-04).

Professional Training Session 1 **Saturday, May 4**
17:00 – 19:00 **Sala Manzoni**

THE USE OF ANALOG GEAR IN THE DIGITAL ERA

Presenter: **Alan Hyatt**

The history of PMI Audio Group and connected brands. Why we do what we do: a tour about the technical/philosophical differences that makes our brands special and what it is that we do differently from other companies.

Saturday, May 4 **17:00** **Sala Saba**

Technical Committee Meeting on Semantic Audio Analysis

Saturday, May 4 **17:00** **Sala Montale**

Standards Committee Meeting on SC-02-08, Audio File Transfer and Exchange

Professional Training Session 2 **Saturday, May 4**
17:30 – 18:30 **Sala Carducci**

WIDE BANDWIDTH EQUALIZATION IN THE ANALOG DOMAIN

Based on research related to the sensitivity of our ears throughout the full bandwidth of our hearing, the CharterOak PEQ-1 provides extremely musical equalization in a stereo tone control of unparalleled musicality. Utilizing a unique circuit, and power distribution, along with a unique configuration with respect to bandwidth and cut and boost around the center frequency choices, the PEQ-1 achieves extremely powerful and musical equalization with minimal phase shift at the high frequencies, and extremely high headroom and low distortion.

Special Event
OPEN HOUSE OF THE TECHNICAL COUNCIL AND THE RICHARD C. HEYSER MEMORIAL LECTURE
Saturday, May 4, 19:00 – 20:00
Auditorium Loyola

Lecturer: **Wolfgang Klippel**

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

134th AES Convention is Wolfgang Klippel. Klippel studied electrical engineering at the University of Technology of Dresden. After graduating in speech recognition, he joined a loudspeaker company in the eastern part of Germany. He was engaged in research on transducer modeling, acoustic measurement, and psychoacoustics. In 1987 he received a Ph.D. in technical acoustics. After spending a post-doctoral year at the Audio Research Group in Waterloo, Canada, and working at Harman International, Northridge, CA, he went back to Dresden in 1997 and founded Klippel GmbH which develops novel kinds of control and measurements systems dedicated to loudspeakers and other transducers. In 2007 he became professor of electro-acoustics at the University in Dresden. He is an AES Fellow, holds the AES publication award of 1993, and the ALMA Titanium Driver Award. The title of his lecture is, "Small, Loud-Speakers: Taking Physics to the Limit."

The loudspeaker is the weakest part in the audio chain. This statement might be true because the electro-acoustical transducer is one of the remains of the analog era still using a moving coil and a diaphragm, much as it was a century ago. A more important argument is the low efficiency of the loudspeaker generating more heat than sound power output while adding undesired distortion to the output signal. The transducer, enclosure, and other acoustical elements increasingly determine the size, weight, and cost of the audio system because electronic parts become smaller or are replaced by digital signal processing.

At the moment there is no alternative transduction principle that is mature enough to compete with the conventional techniques. Therefore the moving coil as the best practice is preserved and interlaced with new ideas provided by research, design, and manufacturing. This evolutionary process also results in significant progress in the absence of a revolutionary change. The development of smaller loudspeakers for personal audio equipment, automotive, and professional applications is a visible example of this process. The generation of sufficient acoustical output at acceptable quality requires higher amplitudes in the mechanical system and pushes the working range to the physical limits.

The cultivation of large signal transducer performance is the central topic in this lecture. In the eighties the nonlinear and thermal modeling of the transducer attracted the interest of more and more researchers eventually resulting in a reliable theory. The theory has been evaluated by fitting the models to real transducers and measuring the large signal parameters dynamically. This identification technique requires only the electrical signal at loudspeaker terminals and uses the back EMF for sensing the velocity of the voice coil. Adaptive measurement techniques using the loudspeaker itself as sensor have been developed to monitor the behavior of the loudspeaker in normal applications (e.g., car) while reproducing music or other audio signals. Recorded parameter variation and state variables (e.g. voice coil displacement) reveal the aging of the suspension and the impact of the climate.

The nonlinear and thermal parameters open a new way for loudspeaker diagnostics because these parameters describe the properties of the loudspeaker itself without the interaction with a stimulus (e.g., test signal, music). Numerical simulation tools (e.g., FEM) applied to magnetic and suspension systems show the relationship between material, geometry, and the nonlinear parameters. However, only a measurement can reveal an offset in the rest position of the voice coil caused by suspension parts made out of fabric, rubber, foam, and other visco-elastic material.

Operating loudspeakers at high amplitudes increases the risk of damaging the loudspeaker. The large signal model provides all state information (e.g., displacement and temperature of the voice coil) indicating a mechanical and thermal overload of the transducer. In woofers the motor and suspension nonlinearity is used to limit the maximum peak displacement and to prevent the voice coil from bottoming. New measurement techniques have been developed to detect a rubbing voice coil, loose particles, and other loudspeaker defects that produce impulsive distortion unacceptable to the human ear. The measurement of harmonic and intermodulation distortion cannot provide a comprehensive description of the large signal behavior as it reveals only symptoms of the nonlinearities depending on the choice of the stimulus. In contrast, the extended large signal model can be used to predict all state variables and the output signal for an arbitrary input signal. The effect of each nonlinearity and cooling mechanisms can be analyzed and design choices can be evaluated before the first prototype is finished. Novel auralization techniques make it possible to enhance or attenuate different kinds of signal distortion by a user defined scaling factor and to assess the audibility and impact on sound quality by listening tests or perceptual models. The closer link between physical and perceptual assessment opens new ways for defining the target performance of an audio product more accurately in marketing, developing products at optimal performance-cost ratio, and for ensuring sufficient quality in manufacturing. This optimization process leads to more or less nonlinear transducers having a clearly defined (regular) nonlinear characteristic. Distortion may become audible for a critical stimulus but is accepted in a trade-off between maximum output, efficiency, size, weight, and cost. The main objective in taming the transducer nonlinearities is the motor stability, to keep the voice coil in the gap, and to gain maximum peak to peak displacement. Asymmetrical curve shapes of the nonlinearities are reduced to avoid a DC-displacement of the coil generated by a nonlinear rectification process that reduces efficiency and may cause bottoming of the voice coil.

Portable applications where power consumption and battery capacity is an issue will require transducers with the highest efficiency and minimal use of resources (e.g., neodymium-magnet). Such a "green" speaker is a highly nonlinear speaker using a motor topology where the voice coil exploits the magnet field in the gap and gives the highest force factor value Bl(x) at the rest position x=0. Unfortunately, the varying force factor Bl(x) generates significant intermodulation distortion throughout the audio band if the voice coil is displaced and windings leave the gap. This undesired side-effect of an efficient motor structure can be compensated by an inverse nonlinear preprocessing of the electrical input signal. The control algorithms use the large signal transducer model and identify the parameters adaptively. By merging electro-acoustics and signal processing the loudspeaker becomes a self-learning system providing optimum performance over the lifetime of the audio product. Certainly, this is not the last step in the evolution of the loudspeaker ...

Student/Career Event
AES STUDENT MEET-UP!
Saturday, May 4, 22:00 – 24:00
Spanish Steps

Audio Students! We're all going to "meet-up" at the Spanish Steps for a late evening get together at one of Rome's most iconic locations. We'll hang out, talk, play some ukulele, and meet new audio student friends from around the globe. Wear your AES 134th Convention badge so you'll be easily identifiable.

Session P6 **Sunday, May 5**
09:00 – 13:00 **Sala Carducci**

RECORDING AND PRODUCTION

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

09:00

P6-1 Automated Tonal Balance Enhancement for Audio Mastering Applications—*Stylianos-Ioannis Mimilakis*,[1] *Konstantinos Drossos*,[2] *Andreas Floros*,[2] *Dionysios Katerelos*[1]
[1]Technological Educational Institute of Ionian Island, Lixouri, Greece
[2]Ionian University, Corfu, Greece

Modern audio mastering procedures are involved with the selective enhancement or attenuation of specific frequency bands. The main reason is the tonal enhancement of the original / unmastered audio material. The aforementioned process is mostly based on the musical information and the mode of the audio material. This information can be retrieved from a listening procedure of the original stimuli, or the correspondent musical key notes. The current work presents an adaptive and automated equalization system that performs the aforementioned mastering procedure, based on a novel method of fundamental frequency tracking. In addition to this, the overall system is being evaluated with objective PEAQ analysis and subjective listening tests in real mastering audio conditions.
Convention Paper 8836

09:30

P6-2 A Pairwise and Multiple Stimuli Approach to Perceptual Evaluation of Microphone Types—*Brecht De Man, Joshua D. Reiss*, Queen Mary University of London, London, UK

An essential but complicated task in the audio production process is the selection of microphones that are suitable for a particular source. A microphone is often chosen based on price or common practices, rather than whether the microphone actually sounds best in that particular situation. In this paper we perceptually assess six microphone types for recording a female singer. Listening tests using a pairwise and multiple stimuli approach are conducted to identify the order of preference of these microphone types. The results of this comparison are discussed, and the performance of each approach is assessed.
Convention Paper 8837

10:00

P6-3 Comparison of Power Supply Pumping of Switch-Mode Audio Power Amplifiers with Resistive Loads and Loudspeakers as Loads—*Arnold Knott, Lars Press Petersen*, Technical University of Denmark, Kgs. Lyngby, Denmark

Power supply pumping is generated by switch-mode audio power amplifiers in half-bridge configuration, when they are driving energy back into their source. This leads in most designs to a rising rail voltage and can be destructive for either the decoupling capacitors, the rectifier diodes in the power supply or the power stage of the amplifier. Therefore precautions are taken by the amplifier and power supply designer to avoid those effects. Existing power supply pumping models are based on an ohmic load attached to the amplifier. This paper shows the analytical derivation of the resulting waveforms and extends the model to loudspeaker loads. Measurements verify that the amount of supply pumping is reduced by a factor of four when comparing the nominal resistive load to a loudspeaker. A simplified and more accurate model is proposed and the influence of supply pumping on the audio performance is proven to be marginal.
Convention Paper 8838

10:30

P6-4 The Psychoacoustic Testing of the 3-D Multiformat Microphone Array Design and the Basic Isosceles Triangle Structure of the Array and the Loudspeaker Reproduction Configuration—*Michael Williams*, Sounds of Scotland, Le Perreux sur Marne, France

Optimizing the loudspeaker configuration for 3-D microphone array design can only be achieved with a clear knowledge of the psychoacoustic parameters of reproduction of height localization. Unfortunately HRTF characteristics do not seem to give us useful information when applied to loudspeaker reproduction. A set of psychoacoustic parameters have to be measured for different configurations in order to design an efficient microphone array recording system, even more so, if a minimalistic approach to array design is going to be a prime objective. In particular the position of a second layer of loudspeakers with respect to the primary horizontal layer is of fundamental importance and can only be based on the psychoacoustics of height perception. What are the localization characteristics between two loudspeakers situated in each of the two layers? Is time difference as against level difference a better approach to interlayer localization? This paper will try to answer these questions and also justify the basic isosceles triangle loudspeaker structure that will help to optimize the reproduction of height information.
Convention Paper 8839

11:00

P6-5 A Perceptual Audio Mixing Device—*Michael J. Terrell, Andrew J. R. Simpson, Mark B. Sandler*, Queen Mary University of London, London, UK

A method and device is presented that allows novice and expert audio engineers to perform mixing using perceptual controls. In this paper we use *Auditory Scene Analysis* [Bregman, 1990, MIT Press, Cambridge] to relate the multitrack component signals of a mix to the perception of that mix. We define the multitrack components of a mix as a group of audio streams, which are transformed into sound streams by the act of reproduction, and which

are ultimately perceived as auditory streams by the listener. The perceptual controls provide direct manipulation of loudness balance within a mixture of sound streams, as well as the overall mix loudness. The system employs a computational optimization strategy to perform automatic signal gain adjustments to component audio-streams, such that the intended loudness balance of the associated sound-streams is produced. Perceptual mixing is performed using a complete auditory model, based on a model of loudness for time-varying sound streams [Glasberg and Moore, *J. Audio Eng. Soc.*, vol. 50, 331-342 (2002 May)]. The use of the auditory model enables the loudness balance to be automatically maintained regardless of the listening level. Thus, a perceptual definition of the mix is presented that is listening-level independent, and a method of realizing the mix practically is given.
Convention Paper 8840

11:30

P6-6 On the Use of a Haptic Feedback Device for Sound Source Control in Spatial Audio Systems—*Frank Melchior, Chris Pike, Matthew Brooks, Stuart Grace*, BBC Research and Development, Salford, UK

Next generation spatial audio systems are likely to be capable of 3-D sound reproduction. Systems currently under discussion require the sound designer to position and manipulate sound sources in three dimensions. New intuitive tools, designed to meet the requirements of audio production environments, are needed to make efficient use of this new technology. This paper investigates a haptic feedback controller as a user interface for spatial audio systems. The paper will give an overview of conventional tools and controllers. A prototype has been developed based on the requirements of different tasks and reproduction methods. The implementation will be described in detail and the results of a user evaluation will be given.
Convention Paper 8841

12:00

P6-7 Audio Level Alignment—Evaluation Method and Performance of EBU R 128 by Analyzing Fader Movements—*Jon Allan, Jan Berg*, Luleå University of Technology, Piteå, Sweden

A method is proposed for evaluating audio meters in terms of how well the engineer conforms to a level alignment recommendation and succeeds to achieve evenly perceived audio levels. The proposed method is used to evaluate different meter implementations, three conforming to the recommendation EBU R 128 and one conforming to EBU Tech 3205-E. In an experiment, engineers participated in a simulated live broadcast show and the resulting fader movements were recorded. The movements were analyzed in terms of different characteristics: fader mean level, fader variability, and fader movement. Significant effects were found showing that engineers do act differently depending on the meter and recommendation at hand.
Convention Paper 8842

12:30

P6-8 Balance Preference Testing Utilizing a System of Variable Acoustic Condition—*Richard King*,[1,2] *Brett Leonard*,[1,2] *Scott Levine*,[1,2] *Grzegorz Sikora*[3]
[1]McGill University, Montreal, Quebec, Canada
[2]The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada
[3]Bang & Olufsen Deutschland GmbH, Pullach, Germany

In the modern world of audio production, there exists a significant disconnect between the music mixing control room of the audio professional and the listening space of the consumer or end user. The goal of this research is to evaluate a series of varying acoustic conditions commonly used in such listening environments. Expert listeners are asked to perform basic balancing tasks, under varying acoustic conditions. The listener can remain in position while motorized panels rotate behind a screen, exposing a different acoustic condition for each trial. Results show that listener fatigue as a variable is thereby eliminated, and the subject's aural memory is quickly cleared, so that the subject is unable to adapt to the newly presented condition for each trial.
Convention Paper 8843

Session P7 **Sunday, May 5**
09:00 – 11:30 **Sala Foscolo**

TRANSDUCERS—PART 1: LOUDSPEAKERS

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

09:00

P7-1 Distortion Improvement in the Current Coil of a Loudspeaker—*Gaël Pillonnet*,[1] *Eric Sturtzer*,[1] *Timothé Rossignol*,[2] *Pascal Tournier*,[2] *Guy Lemarquand*[3]
[1]University of Lyon, Lyon, France
[2]ONSemiconductor, Toulouse, France
[3]Universite du Mans, Le Mans Cedex 9, France

This paper deals with the comparison of voltage and current driving units in an active audio system. The effect of the audio amplifier control on the current coil of an electrodynamic loudspeaker is presented. In voltage control topology, the electromagnetic force linked to coil current is controlled through the load impedance. Thus, the electromechanical conversion linearity is decreased by the impedance variation, which implies a reduction of the overall audio quality. A current driving method could reduce the effect of the non-linear impedance by controlling the coil current directly, thereby the acceleration. Large signal impedance modeling is given in this paper to underline the non-linear effects of electrodynamic loudspeaker parameters on the coupling. As a result, the practical comparison of voltage and current driven methods proves that the current control reduces the voice coil current distortions in the three different loudspeakers under test.
Convention Paper 8844

➥

09:30

P7-2 Driving Electrostatic Transducers—*Dennis Nielsen, Arnold Knott, Michael A. E. Andersen*, Technical University of Denmark, Kgs. Lyngby, Denmark

Electrostatic transducers represent a very interesting alternative to the traditional inefficient electrodynamic transducers. In order to establish the full potential of these transducers, power amplifiers that fulfill the strict requirements imposed by such loads (high impedance, frequency depended load, and high bias voltage for linearization) must be developed. This paper analyzes power stages and bias configurations suitable for driving an electrostatic transducer. Measurement results of a ± 300 V prototype amplifier are shown. Measuring THD across a high impedance source is discussed and a high voltage attenuation interface for an audio analyzer is presented. THD below 0:1% is reported.
Convention Paper 8845

10:00

P7-3 Boundary Element Simulation of an Arrayable Loudspeaker Horn—*Tommaso Nizzoli*,[1] *Stefano Prati*[2]
[1]Acoustic Vibration Consultant, Reggio Emilia, Italy
[2]RCF S.p.A., Reggio Emilia, Italy

The Boundary Element method implemented in a commercial code is used to verify the acoustic directional characteristic in the far field of an arrayable loudspeaker's horn in comparison with the full space, far field, measured acoustic balloon. A simple model of the full arrayable loudspeaker horn's splay idealizes each source only by the calculated emission on a flat plane at each horn's mouth. This approach reduces significantly the BEM's calculation time with regard to having to model each source by its full geometry. Comparison with the full space, far field, predicted pressure from test data shows good agreement in all the frequency range of interest.
Convention Paper 8846

10:30

P7-4 Single Permanent Magnet Co-Axial Loudspeakers—*Dimitar Dimitrov*, BMS Production, Sofia, Bulgaria

Co-axial loudspeakers are designed with a single ring permanent magnetic structure assuring dual flux path for the two voice coil gaps. Internal magnet is used in two realizations with parallel flux division at both its diameters. These two varieties are convenient for Nd magnets and one of them has its local internal flux path crossing two gaps in series for dual membrane compression driver implementation. Another realization uses an external permanent magnetic structure with series flux through the gaps. Proposed co-axial loudspeaker types are very compact, simple, and lightweight. They all can use "Stepped Gap" designs for their low frequency voice coils. Comparative measurements with conventional co-axial loudspeakers reveal competitive performance with much reduced weight and production cost.
Convention Paper 8847
[Paper presented by Plamen Valtchev]

11:00

P7-5 Multiple Low-Frequency Corner Folded Horn Designs—*Rumen Artarski*,[1] *Plamen Valtchev*,[2] *Dimitar Dimitrov*,[3] *Yovko Handzhiev*[3]
[1]AVC, Sofia, Bulgaria
[2]Univox, Sofia, Bulgaria
[3]BMS Production, Sofia, Bulgaria

Low-Frequency horn loaded systems for pi/2 radiation are designed employing push-push operated dual membrane loudspeakers, closely mounted in parallel against each other. Tree Axis Symmetrical and efficient horn mouth loading is transformed to a symmetrical and uniform membrane cone loading. By doubling loudspeaker diaphragm, and respectively the horn throat, lower horn cut-off frequency was achieved with the same extension rate, besides acoustic power doubling. Two such corner horn systems could be stacked together for quarter space pi-loading, with important usage in front-of-stage subwoofer applications. Four horn systems are grouped together over the floor or on the ceiling for 2 pi-radiation. Finally, eight such systems could be united for full space low-frequency radiation.
Convention Paper 8848

Loudness 1 **Sunday, May 5**
09:00 – 10:00 **Sala Alighieri**

**LOUDNESS 101—
A HITCHHIKER'S GUIDE TO AUDIO NIRVANA**

Presenter: **Florian Camerer**, ORF - Austrian TV; Vienna, Austria; EBU - European Broadcasting Union

This session will bring participants up to speed regarding most aspects of loudness control and metering. It is targeted to sound engineers in general, giving a brief intro to the algorithm and the metering paradigms and then expanding to common misunderstandings, dangers as well as chances and challenges. Some new concepts like "gating" and "true peak level" will be explained in detail. As chairman of the European loudness group PLOUD and senior post-pro mixing engineer at ORF (Austrian TV), Florian Camerer has enough experience under his belt to provide a thorough workout!

Student/Career Event
STUDENT DESIGN COMPETITION
Sunday, May 5, 09:00 – 10:30
Auditorium Loyola

Moderator: **Colin Pfund**, Chair, AES SDA, North and Latin
American Regions

The three graduate level and three undergraduate level finalists of the AES Student Design Competition will present and defend their designs in front of a panel of expert judges. This is an opportunity for aspiring student hardware and software engineers to have their projects reviewed by the best minds in the business. It's also an invaluable career-building event and a great place for companies to identify their next employees. Students from both audio and non-audio backgrounds are encouraged to submit entries. Few restrictions are placed on the nature of the projects, but designs must be for use with audio. Examples include loudspeaker design, DSP plug-ins, analog hardware, signal analysis tools, mobile applications, and sound synthesis devices. Products should represent new, original ideas implemented in working-model prototypes.

The Student Design Competition Judges are: *Aurelio Uncini*, Sapienza University of Rome; *Stefano Daino*, DSP-Quattro; *Patrizio Pisani*, ZP Engineering; *Roberto Magalotti*, B&C Speakers.

Sunday, May 5 09:00 Sala Saba

Technical Committee Meeting on Acoustics and Sound Reinforcement

Sunday, May 5 09:00 Sala Montale

Standards Committee Meeting on SC-02-01, Digital Audio Measurement Techniques

Session P8 Sunday, May 5
09:30 – 11:00 Foyer

POSTERS: AUDIO PROCESSING AND SEMANTICS

09:30

P8-1 Combination of Growing and Pruning Algorithms for Multilayer Perceptrons for Speech/Music/Noise Classification in Digital Hearing Aids—*Lorena Álvarez, Enrique Alexandre, Cosme Llerena, Roberto Gil-Pita, Manuel Rosa-Zurera*, University of Alcala, Alcalá de Henares, Spain

This paper explores the feasibility of combining both growing and pruning algorithms in some way that the global approach results in finding a smaller multilayer perceptron (MLP) in terms of network size, which enhances the speech/music/noise classification performance in digital hearing aids, with the added bonus of demanding a lower number of hidden neurons, and consequently, lower computational cost. With this in mind, the paper will focus on the design of an approach that starts adding neurons to an initial small MLP until the stopping criteria for the growing stage is reached. Then, the MLP size is reduced by successively pruning the least significant hidden neurons while maintaining a continuous decreasing function. The results obtained with the proposed approach will be compared with those obtained when using both growing and pruning algorithms separately.
Convention Paper 8850

09:30

P8-2 Automatic Sample Recognition in Hip-Hop Music Based on Non-Negative Matrix Factorization—*Jordan L. Whitney, Colby N. Leider*, University of Miami, Coral Gables, FL, USA

We present a method for automatic detection of samples in hip-hop music. A sample is defined as a short extraction from a source audio corpus that may have been embedded into another audio mixture. A series of non-negative matrix factorizations are applied to spectrograms of hip-hop music and the source material from a master corpus. The factorizations result in matrices of base spectra and amplitude envelopes for the original and mixed audio. Each window of the mixed audio is compared to the original audio clip by examining the extracted amplitude envelopes. Several image-similarity metrics are employed to determine how closely the samples and mixed amplitude envelopes match. Preliminary testing indicates that, distinct from existing audio fingerprinting algorithms, the algorithm we describe is able to confirm instances of sampling in a hip-hop music mixture that the untrained listener is frequently unable to detect.
Convention Paper 8851

09:30

P8-3 Performance Optimization of GCC-PHAT for Delay and Polarity Correction under Real World Conditions—*Nicholas Jillings, Alice Clifford, Joshua D. Reiss*, Queen Mary University of London, London, UK

When coherent audio streams are summed, delays can cause comb filtering and polarity inversion can result in cancellation. The GCC-PHAT algorithm is a popular method for detecting (and hence correcting) the delay. This paper explores the performance of the Generalized Cross Correlation with Phase Transform (GCC-PHAT) for delay and polarity correction, under a variety of different conditions and parameter settings, and offers various optimizations for those conditions. In particular, we investigated the performance for moving sources, background noise, and reverberation. We consider the effect of varying the size of the Fourier Transform when performing GCC-PHAT. In addition to accuracy, computational efficiency and latency were also used as metrics of performance.
Convention Paper 8852

09:30

P8-4 Reducing Binary Masking Artifacts in Blind Audio Source Separation—*Toby Stokes, Christopher Hummersone, Tim Brookes*, University of Surrey, Guildford, Surrey, UK

Binary masking is a common technique for separating target audio from an interferer. Its use is often justified by the high signal-to-noise ratio achieved. The mask can introduce musical noise artifacts, limiting its perceptual performance and that of techniques that use it. Three mask-processing techniques, involving adding noise or cepstral smoothing, are tested and the processed masks are compared to the ideal binary mask using the perceptual evaluation for audio source separation (PEASS) toolkit. Each processing technique's parameters are optimized before the comparison is made. Each technique is found to improve the overall perceptual score of the separation. Results show a trade-off between interferer suppression and artifact reduction.
Convention Paper 8853

➥

09:30

P8-5 Detection of Sinusoids Using Statistical Goodness-of-Fit Test—*Pushkar P. Patwardhan, Ravi R. Shenoy*, Nokia India Pvt. Ltd., Bangalore, India

Detection of tonal components from magnitude spectrum is an important initial step in several speech and audio processing applications. In this paper we present an approach for detecting sinusoidal components from the magnitude spectrum using "goodness-of-fit" test. The key idea is to test the null-hypothesis that the region of spectrum under observation is drawn from the magnitude spectrum of an ideal windowed-sinusoid. This hypothesis is tested with a chi-square "goodness-of-fit" test. The outcome of this hypothesis test is a decision about the presence of sinusoid in the observed region of magnitude spectrum. We have evaluated the performance of the proposed approach using synthetically generated samples containing steady and modulated harmonics in clean and noisy conditions.
Convention Paper 8854

09:30

P8-6 Novel Designs for the Parametric Peaking EQ User Interface for Single Channel Corrective EQ Tasks—*Christopher Dewey, Jonathan Wakefield*, University of Huddersfield, Huddersfield, West Yorkshire, UK

This paper evaluates the suitability of existing parametric peaking EQ interfaces of analog and digital mixing desks and audio plugins for single channel corrective EQ tasks. It proposes novel alternatives based upon displaying FFT bin maximums for the full audio duration behind the EQ curve, automatically detecting and displaying the top five FFT bin maximum peaks to assist the engineer, an alternative numerical list display of top five FFT bin maximum peaks, and an interface that allows direct manipulation of the displayed FFT bin maximums. All interfaces were evaluated based on the time taken to perform a corrective EQ task, preference ranking, and qualitative comments. Results indicate that the novel EQ interfaces presented have potential over existing EQ interfaces.
Convention Paper 8855

09:30

P8-7 Drum Replacement Using Wavelet Filtering—*Robert Barañski,*[1] *Szymon Piotrowski,*[1] *Magdalena Plewa*[2]
[1]AGH University of Science and Technology, Kracow, Poland
[2]Gdansk University of Technology, Gdansk, Poland

The paper presents the solution that can be used to unify snare drum sound within a chosen fragment. The algorithm is based on the wavelet transformation and allows replacement of subbands of particular sounds, which are outside a certain range. Five experienced sound engineers put the algorithm under the test using samples of five different snare drums. Wavelet filtering seems to be useful in terms of drum replacement, while the sound engineers response was, in the most cases, positive.
Convention Paper 8856

09:30

P8-8 Collaborative Annotation Platform for Audio Semantics—*Nikolaos Tsipas, Charalampos A. Dimoulas, George M. Kalliris, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

In the majority of audio classification tasks that involve supervised machine learning, ground truth samples are regularly required as training inputs. Most researchers in this field usually annotate audio content by hand and for their individual requirements. This practice resulted in the absence of solid datasets and consequently research conducted by different researchers on the same topic cannot be effectively pulled together and elaborated on. A collaborative audio annotation platform is proposed for both scientific and application oriented audio-semantic tasks. Innovation points include easy operation and interoperability, on the fly annotation while playing audio content online, efficient collaboration with feature engines and machine learning algorithms, enhanced interaction, and personalization via state of the art Web 2.0 /3.0 services.
Convention Paper 8857

09:30

P8-9 Investigation of Wavelet Approaches for Joint Temporal, Spectral and Cepstral Features in Audio Semantics—*Charalampos A. Dimoulas, George M. Kalliris*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The current paper focuses on the investigation of wavelet approaches for joint time, frequency, and cepstral audio feature extraction. Wavelets have been thoroughly studied over the last decades as an alternative signal analysis approach. Wavelet-features have also been successfully implemented in a variety of pattern recognition applications, including audio semantics. Recently, wavelet-adapted mel-frequency cepstral coefficients have been proposed as applicable features in speech recognition and general audio classification, incorporating perceptual attributes. In this context, various wavelet configuration-schemes are examined for wavelet-cepstral audio features extraction. Additional wavelet parameters are utilized in the formation of wavelet-feature-vectors and evaluated in terms of salient feature ranking. Comparisons with classical time-frequency and cepstral audio features are conducted in typical audio-semantics scenarios.
Convention Paper 8858

Tutorial 5 **Sunday, May 5**
10:00 – 11:00 **Sala Manzoni**

3-D AUDIO—PRODUCE THE NEW DIMENSION

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

This workshop will have two parts. The first will be a presentation where new experiences and different ways to handle the different approaches like film, music, and radio plays will be shown. Complete sessions will be opened and production experiences, editing, as well as workflows and possibilities to delivery 3-D/spatial content into the common market will be shown and discussed.

The second part of this workshopwill be a Q & A session while a group of max. 10 participants can listen and work with 3-D content via a sophisticated 3D speaker virtualization on headphones.

Loudness 2 **Sunday, May 5**
10:00 – 10:30 **Sala Alighieri**

ALL LOUDNESS RECOMMENDATIONS ARE EQUAL —BUT SOME ARE MORE EQUAL THAN OTHERS

Presenter: **Andrew Mason**, BBC Research and Development, London, UK

This tutorial will give an explanation of the different loudness standards used in Europe (R 128), the US (A/85), and in other countries such as Australia and Japan.

Sunday, May 5 **10:00** **Sala Saba**

Technical Committee Meeting on Microphones and Applications

Loudness 3 **Sunday, May 5**
10:45 – 11:30 **Sala Alighieri**

LOUDNESS FOR COMMERCIALS— HOW ESTHETICS CHANGE(D)

Presenters: **Matteo Milani**, Freelance Sound Designer, Producer, Music Comoser
Alessandro Travaglini, Sound Designer/ Producer, Fox italy
Rubens Zambelli, Technical Service Manager, ADSTREAM Italia
Carlos Zarattini, Sound Designer/Music Composer, Discovery Channels Italy

Commercials have often been the number one complaint when it came to loudness problems and level jumps. The fear to be softer than the competition has led to overcompression of the spots and to an extremely narrow loudness range as well as the"perception of the audio signal being constantly smashed. 'Free transients" and the liberation from this loudness competition has now finally come with the transition to loudness normalization, and especially for commercials this is more than welcome. Four sound designers who have a lot of experience in producing detailed sound tracks for commercials will demonstrate how this new paradigm has changed and is still changing their approach and show how the new dynamic possibilities can be used to great effect. Among the points discussed will be:

- Use of headroom and dynamic processors
- Use of sound effects (in particular low-frequency sounds)
- Speech definition and sound artifacts, and
- Short-term loudness limitations

Examples of their work will be played and explained.

Tutorial 6 **Sunday, May 5**
11:00 – 13:00 **Auditorium Loyola**

DRUM PROGRAMMING

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

Drum programming has been evolving at the heart of many studio productions for some 30 years. Over this period, technological opportunities for enhanced creativity have multiplied in numerous directions. This tutorial will demonstrate a number of these directions as they are often implemented in contemporary professional practice, showing contrasting techniques used in the creation of both human emulation and the unashamedly synthetic. Alongside this, many of the studio production techniques often used to enhance such work will be discussed, ranging from dynamic processing to intricate automation.

The session will include numerous live demonstrations covering a range of approaches. Although introducing all key concepts from scratch, its range and hybridization should provide inspiration even for experienced practitioners.

Sunday, May 5 **11:00** **Sala Saba**

Technical Committee Meeting on Audio Forensics

Sunday, May 5 **11:00** **Sala Montale**

Standards Committee Meeting on SC-07-01, Metadata for Audio

Student/Career Event
STUDENT DESIGN EXHIBITION
Sunday, May 5, 11:15 – 12:45
Foyer

All accepted student entries to the AES Student Design Competition will have the opportunity to show off their designs at this poster/tabletop exhibition. This audio science fair is free and open to all convention attendees and is an opportunity for aspiring student hardware and software engineers to have their projects seen by the AES design community. It's also an invaluable career-building event and a great place for companies to identify their next employees. Students from both audio and non-audio backgrounds are encouraged to submit entries. Few restrictions are placed on the nature of the projects, but designs must be for use with audio. Examples include loudspeaker design, DSP plug-ins, analog hardware, signal analysis tools, mobile applications, and sound synthesis devices. Products should represent new, original ideas implemented in working-model prototypes.

Tutorial 7 **Sunday, May 5**
11:30 – 13:00 **Sala Manzoni**

ACOUSTIC ENHANCEMENT SYSTEMS—THE BASICS

Presenter: **Ben Kok**, SCENA acoustic consultants, Uden, The Netherlands

In the last decades multiple systems for electronic acoustic enhancement have been introduced. Some have disappeared over time, but others appear to be settled firmly. The claim for these systems is that the acoustics of avenue can be changed at the press of a ➥

button, at cost significantly lower than variable acoustics by structural means. Also, in situations where the natural acoustics did not work out properly, these systems are used to correct the flaws, again at lower costs than the structural alternatives. The question, of course, is do these systems really work and if so, what are the differences between these systems? And what would be the structural alternatives?

This tutorial will identify what acoustic properties can or should be influenced by an acoustic enhancement system. In relation to this, working principles and philosophies of some of the most popular systems are analyzed and similarities and differences are identified and related to specific acoustic situations.

Loudness 4 Sunday, May 5
11:45 – 13:15 Sala Alighieri

ARE MOVIES TOO LOUD? THE LOUDNESS RACE REACHES THE CINEMA

Presenters: **Florian Camerer**
Eelco Grimm

Cinema operators are on the receiving end of growing numbers of complaints from the audience of soundtracks that are "too loud." Just turning them down results in lowered dialog levels, which leaves the movie quieter but unintelligible. The EU has loudness standards that are now starting to be applied to cinemas. There is a possible need for standards to ensure the theaters are given soundtracks that meet EU laws. This workshop will investigate the issues involved and the work necessary to resolve the issues.

Tutorial 8 Sunday, May 5
12:00 – 13:00 Sala Foscolo

PERCEPTUAL TESTING OF SPEECH IN MODERN AUDIO PRODUCTS

Presenter: **Dan Foley**, Audio Precision

We are witnessing a convergence of audio technologies in modern compact consumer products, with small computing devices processing voice and general audio through a range of options that include analog, electroacoustic, Bluetooth, HDMI, and more. Whereas in the past sine and noise-based test signals have been used to verify performance, these new devices require both a mixture of connectivity options and the ability to assess the perceptual needs of voice communication.

Tests such as Perceptual Speech Quality (PESQ) and Perceptual Listening Quality (POLQA) are now being used to fully assess device performance where codecs and other non-linear processing is used. This presentation will focus on how these perceptual methods work and when to use them in conjunction with, or in lieu of, classic audio test methods and techniques.

Sunday, May 5 12:00 Sala Saba

Technical Committee Meeting on Studio Practices and Production

Student/Career Event
EDUCATION FAIR
Sunday, May 5, 13:00 – 14:30
Foyer

Institutions offering studies in audio (from short courses to graduate degrees) will be represented in a "table top" session. Information on each school's respective programs will be made available through displays and academic guidance. There is no charge for schools to participate.

Sunday, May 5 13:00 Sala Saba

Technical Committee Meeting on Audio for Telecommunications

Sunday, May 5 14:00 Sala Saba

Technical Committee Meeting on Spatial Audio

Sunday, May 5 14:00 Sala Montale

Standards Committee Meeting on SC-05-02, Audio Connectors

Loudness 5 Sunday, May 5
14:15 – 16:15 Sala Alighieri

MAKE LUFS NOT WAR

Presenter: **Thomas Lund**, TC Electronic A/S, Risskov, Denmark

Panelists: *Florian Camerer*, ORF - Austrian TV, Vienna, Austria
George Massenburg, McGill University, Montreal, Quebec, Canada

Newly produced pop/rock music rarely sounds good on fine loudspeakers, commercials on TV are annoyingly loud, and a visit to the cinema may be a deafening experience. This is audio's dark middle ages, from which there will be little content for future generations to enjoy.

However, 2013 could be the year where a renaissance again spreads from Italy. Transparent loudness normalization has arrived to radio, TV, and iPod; and the panel sets out to describe the far-reaching implications this will have on audio production at large. Hear about new quality-defining criteria, and save your next album for generations to come.

Session P9 Sunday, May 5
14:30 – 18:00 Sala Carducci

ROOM ACOUSTICS

Chair: **Chris Baume**, BBC Research & Development, UK

14:30

P9-1 Various Applications of Active Field Control —*Takayuki Watanabe, Masahiro Ikeda*, Yamaha Corporation, Hamamatsu, Shizuoka, Japan

The Active Field Control system is an acoustic enhancement system that was developed to improve the acoustic conditions of a space so as to match the acoustic conditions required for a variety of different types of performance programs. This system is unique in that it uses FIR filtering to ensure freedom of control and the concept of spatial averaging to achieve stability with a lower number of channels than compara-

tive systems. This system has been used in over 70 projects in both the U.S. and Japan. This paper will provide an overview of the characteristics of the system and examples of how the system has been applied.
Convention Paper 8859

15:00

P9-2 Comparative Acoustic Measurements: Spherical Sound Source vs. Dodecahedron—*Plamen Valtchev*,[1] *Denise Gerganova*[2]
[1]Univox, Sofia, Bulgaria
[2]Spherovox, Sofia, Bulgaria

Spherical sound source, consisting of a pair of coaxial loudspeakers and a pair of compression drivers and radiating into a common radially expanding horn, is used for acoustic measurements of rooms for speech and music. For exactly the same source-microphone pair positions, comparative measurements are made with a typical dodecahedron, keeping the same microphone technique, identical signals, and recording hardware under the same measuring conditions. Several software programs were used for evaluation of the acoustical parameters extracted from impulse responses. Parameters are presented in tables and graphics for better sound source comparisons. Spherical sound source reveals higher dynamic range and perfectly repeatable parameters with source rotation, which is in contrast to dodecahedron, where rotation steps resulted in some parameters' deviation.
Convention Paper 8860

15:30

P9-3 Archaeoacoustics: An Introduction—A New Take on an Old Science—*Lise-Lotte Tjellesen, Karen Colligan*, CLARP, London, UK

What is Archaeoacoustics and how is it defined? This paper will discuss the history and varying aspects of the discipline of archaeoacoustics, i.e., sound that has been measured, modeled, and analyzed with modern techniques in and around Ancient sites, temple complexes, and standing stones. Piecing together sound environments from a long lost past it is brought to life as a tool for archaeologists and historians. This paper will provide a general overview of some of the most prolific studies to date, discuss some measurement and modeling methods, and discuss where archaeoacoustics may be headed in the future and what purpose it serves in academia.
Convention Paper 8861

16:00

P9-4 Scattering Effects in Small-Rooms: From Time and Frequency Analysis to Psychoacoustic Investigation—*Lorenzo Rizzi, Gabriele Ghelfi*, Suono e Vita - Acoustic Engineering, Lecco, Italy

This work continues the authors' effort to optimize a DSP tool for extrapolating from R.I.R. information regarding mixing time and sound scattering effects with in-situ measurements. Confirming past thesis, a new specific experiment allowed to scrutinize the effects of QRD scattering panels over non-Sabinian environments, both in frequency and in time domain. Listening tests have been performed to investigate perception of scattering panels effecting small-room acoustic quality. The sound diffusion properties have been searched with specific headphone auralization interviews, convolving known R.I.R.s with anechoic musical samples and correlating calculated data to psychoacoustic responses. The results validate the known effect on close field recording in small rooms for music and recording giving new insights.
Convention Paper 8862

16:30

P9-5 The Effects of Temporal Alignment of Loudspeaker Array Elements on Speech Intelligibility—*Timothy J. Ryan*,[1] *Richard King*,[2] *Jonas Braasch*,[3] *William L. Martens*[4]
[1]Webster University, St. Louis, MO, USA
[2]McGill University, Montreal, Quebec, Canada
[3]Rensselaer Polytechnic Institute, Troy, NY, USA
[4]University of Sydney, Sydney, NSW, Australia

The effects of multiple arrivals on the intelligibility of speech produced by live-sound reinforcement systems are examined. Investigated variables include the delay time between arrivals from multiple loudspeakers within an array and the geometry and type of array. Subjective testing, using captured binaural recordings of the Modified Rhyme Test under various treatment conditions, was carried out to determine the first- and second-order effects of the two experimental variables. Results indicate that different interaction effects exist for different amounts of delay offset.
Convention Paper 8863

17:00

P9-6 Some Practical Aspects of STI Measurement and Prediction—*Peter Mapp*, Peter Mapp Associates, Colchester, Essex, UK

The Speech Transmission Index (STI) has become the internationally accepted method of testing and assessing the potential intelligibility of sound systems. The technique is standardized in IEC 60268-16, however, it is not a flawless technique. The paper discusses a number of common mechanisms that can affect the accuracy of STI measurements and predictions. In particular it is shown that RaSTI is a poor predictor of STI in many sound system applications and that the standard speech spectrum assumed by STI often does not replicate the speech spectrum of real announcements and is not in good agreement with other speech spectrum studies. The effects on STI measurements of common signal processing techniques such as equalization, compression, and AGC are also demonstrated and the implications discussed. The simplified STI derivative STIPA is shown to be a more robust method of assessing sound systems than RaSTI and when applied as a direct measurement method can have significant advantages over Impulse Response-based STI measurement techniques.
Convention Paper 8864

➥

17:30

P9-7 Combined Quasi-Anechoic and In-Room Equalization of Loudspeaker Responses— *Balazs Bank*, Budapest University of Technology and Economics, Budapest, Hungary

This paper presents a combined approach to loudspeaker/room response equalization based on simple in-room measurements. In the first step, the anechoic response of the loudspeaker, which mostly determines localization and timbre perception, is equalized with a low-order non-minimum phase equalizer. This is actually done using the gated in-room response, which of course means that the equalization is incorrect at low frequencies where the gate time is shorter than the anechoic impulse response. In the second step, a standard, fractional-octave resolution minimum-phase equalizer is designed based on the in-room response pre-equalized with the quasi-anechoic equalizer. This second step, in addition to correcting the room response, automatically compensates the low-frequency errors made in the quasi-anechoic equalizer design when we were using gated responses.
Convention Paper 8826

Session P10 **Sunday, May 5**
14:15 – 17:15 **Sala Foscolo**

TRANSDUCERS—PART 2: ARRAYS, MICROPHONES

Chair: **TBD**

14:30

P10-1 The Radiation Characteristics of an Array of Horizontally Asymmetrical Waveguides that Utilize Continuous Arc Diffraction Slots— *Soichiro Hayashi, Akira Mochimaru, Paul F. Fidlin*, Bose Corporation, Framinham, MA, USA

Previous work presented the radiation characteristics of a horizontally asymmetrical waveguide that utilizes a continuous arc diffraction slot. It showed good coverage control above 1 kHz, as long as the waveguide center line axis angle stays below a certain angle limit. This paper examines the radiation characteristics of an array of horizontally asymmetrical waveguides. Waveguides with different angular variations are developed, and several vertical arrays are constructed consisting of those waveguides. The radiation characteristics of the arrays are measured. Horizontally symmetrical and asymmetrical vertical arrays are compared, and the consistency inside of the coverage and limitations are discussed.
Convention Paper 8865

15:00

P10-2 Numerical Simulation of Microphone Wind Noise, Part 1: External Flow—*Juha Backman*, Nokia Corporation, Espoo, Finland

This paper discusses the use of the computational fluid dynamics (CFD) for computational analysis of microphone wind noise. The first part of the work, presented in this paper, discusses the behavior of the flow around the microphone. One of the practical questions answered in this work is the well-known difference between "pop noise," i.e., noise caused by transient flows, and wind noise generated by more stationary flows. It appears that boundary layer separation and related modification of flow field near the boundary is a significant factor in transient flow noise, while vortex shedding, emerging at higher flow velocities, is significant for steady state flow. The paper also discusses the effects of the geometrical shape and surface details on the wind noise.
Convention Paper 8866

15:30

P10-3 Listener Preferences for Different Headphone Target Response Curves—*Sean Olive, Todd Welti, Elisabeth McMullin*, Harman International, Northridge, CA, USA

There is little consensus among headphone manufacturers on the preferred headphone target frequency response required to produce optimal sound quality for reproduction of stereo recordings. To explore this topic further we conducted two double-blind listening tests in which trained listeners rated their preferences for eight different headphone target frequency responses reproduced using two different models of headphones. The target curves included the diffuse-field and free-field curves in ISO 11904-2, a modified diffuse-field target recommended by Lorho, the unequalized headphone, and a new target response based on acoustical measurements of a calibrated loudspeaker system in a listening room. For both headphones the new target based on an in-room loudspeaker response was the most preferred target response curve.
Convention Paper 8867

16:00

P10-4 Optimal Condition of Receiving Transducer in Wireless Power Transfer Based on Ultrasonic Resonance Technology—*WooSub Youm, Gunn Hwang, Sung Q Lee*, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea

Recently, wireless power transfer technology has drawn lots of attention because of wire reduction and charging convenience. Previous technologies such as magnetic resonance and induction coupling have drawbacks of short transfer distance and harmfulness to human health. As an alternative, ultrasonic resonance wireless power transfer technology is proposed. A pair of commercial ultrasonic transducer arrays for sensor application is used for wireless power transfer. There are two parameters that have to be decided to maximize the efficiency of power transfer. The first parameter is load resistance that is connected to receiving transducer. It should be matched to the impedance of receiving the transducer at resonance frequency. The second one is the operating frequency f transmitting transducer. It should be matched to the optimal frequency of the receiving transducer. In this paper this optimal load resistance and the frequency of the receiving transducer are analyzed based on circuit theory and verified through experiment.
Convention Paper 8868

16:30

P10-5 Design of a Headphone Equalizer Control Based on Principal Component Analysis—*Felix Fleischmann, Jan Plogsties, Bernhard Neugebauer*, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Unlike for loudspeakers, the optimal frequency response for headphones is not flat. A perceptually optimal target equalization curve for headphones was identified in a previous study. Moreover, strong variability in the frequency characteristics of 13 popular headphone models was observed. Model-specific equalization filters can be implemented but would require the headphone to be known. For most consumer applications this seems impractical. Principal component analysis was applied to the measured headphone equalization data, followed by a reduction of the degrees of freedom. The remaining error is identified objectively. It can be shown that using only one principal component, the sum of a fixed and a weighted filter curve can replace model-specific equalization.
Convention Paper 8869

17:00

P10-6 Effect of Elastic Gel Layer on Energy Transfer from the Source to Moving Part of Sound Transducer—*Minsung Cho, Elena Prokofieva, Mike Barker*, Edinburgh Napier University, Edinburgh, UK

The use of porous materials in sound is quite well-known. The materials can be used for absorption of unwanted radiation, for dissipation, and re-distribution of the waves while traveling through the thickness of material and for enhancement of the overall sound and therefore energy transfer within the material. A three layer system comprising of two rigid layers of material with a soft gel middle layer between them was investigated in this research work to establish the effect of the gel material on the system's energy performance. The role of the gel layer in transferring the energy to the panel was investigated. It was demonstrated by experiments that the gel layer between the top layer and the last layer enhances the performance of the overall construction and also minimizes the mechanical distortion by absorbing its bending waves. This effect enables to extend the radiating frequency of the construction up to high and down to lower frequencies.
Convention Paper 8849
[Paper Presented by Elena Prokofieva]

Workshop 5 **Sunday, May 5**
14:30 – 16:30 **Sala Manzoni**

CINEMA SOUND—ACOUSTICAL AND CALIBRATION ISSUES

Chair: **Brian McCarty**, Coral Sea Studios Pty. Ltd., Clifton Beach, QLD, Australia

Panelists: *David Murphy*
Philip Newell

First, we will take a theoretical look at the problems of cinema sound and shortcomings in the current SMPTE standards and testing methodology. Then we will explore a practical look at the perils of on-site measurements and the application of new standards to the calibration of sound systems in cinemas.

Tutorial 9 **Sunday, May 5**
14:45 – 16:15 **Auditorium Loyola**

ELECTRIC GUITAR—WHAT A RECORDIST OUGHT TO KNOW

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

Musicians obsess about every detail of their instrument. Engineers do the same with every detail of their studio. This tutorial merges those obsessions, so that a recording engineer can be more fully informed about the key drivers of tone for the electric guitar. Know the instrument first and let that drive your decisions for recording and mixing the instrument.

Sunday, May 5 **15:00** **Sala Saba**

Technical Committee Meeting on Signal Processing

Session P11 **Sunday, May 5**
15:30 – 17:00 **Foyer**

POSTERS: PERCEPTION AND EDUCATION

15:30

P11-1 Improve the Listening Ability Using E-Learning Methods—*Bartlomiej Kruk, Bartosz Zawieja*, Wroclaw University of Technology, Wroclaw, Poland

The main aim of this paper is to show the possibility of listening training using the methods of e-learning combined with the classical method of teaching. The described technique uses all available electronic media as well as traditional teaching methods. Theory and practice examples are discussed. However, e-learning allows the students to work independently and provide tests at a convenient place and time. The e-learning exercises have been designed to develop skills including memorization, timbre description, and improving hearing sensitivity for changes in sound.
Convention Paper 8870

15:30

P11-2 Optimizing Teaching Room Acoustics: Investigating the Exclusive Use of a Distributed Electroacoustic Installation to Improve the Speech Intelligibility—
Panagiotis Hatziantoniou,[1] *Nicolas Tatlas*,[2] *Stelios M. Potirakis*[2]
[1]University of Patras, Patras, Greece
[2]Technological Educational Institute of Piraeus, Athens, Greece

The possibility to improve speech intelligibility in classrooms of inadequate acoustic design, exclusively by using an electroacoustic installation is investigated in this paper. Measurement results derived from an overall six-loudspeaker arrangement response in different locations ➥

inside a test classroom are compared to those derived from a one-loudspeaker electroacoustic response in the speaker's location. Preliminary results indicate that well-established parameters such as C-50 and D-50 automatically calculated, show no significant improvement. Extended investigation of the responses as well as other criteria such as direct to reverb ratios (DRR) shown that there is noteworthy enhancement in addition to the subjective acoustic perception. Moreover, DRR is shown to further improve from the employment of a room correction technique based on smoothed responses.
Convention Paper 8871

15:30

P11-3 Software Techniques for Good Practice in Audio and Music Research—*Luis Figueira, Chris Cannam, Mark Plumbley*, Queen Mary University of London, London, UK

In this paper we discuss how software development can be improved in the audio and music research community by implementing tighter and more effective development feedback loops. We suggest first that researchers in an academic environment can benefit from the straightforward application of peer code review, even for ad-hoc research software; and second, that researchers should adopt automated software unit testing from the start of research projects. We discuss and illustrate how to adopt both code reviews and unit testing in a research environment. Finally, we observe that the use of a software version control system provides support for the foundations of both code reviews and automated unit tests. We therefore also propose that researchers should use version control with all their projects from the earliest stage.
Convention Paper 8872

15:30

P11-4 A Practical Step-by-Step Guide to the Time-Varying Loudness Model of Moore, Glasberg, and Baer (1997; 2002)—*Andrew J. R. Simpson, Michael J. Terrell, Joshua D. Reiss*, Queen Mary University of London, London, UK

In this tutorial article we provide a condensed, practical step-by-step guide to the excitation pattern loudness model of Moore, Glasberg, and Baer [*J. Audio Eng. Soc.*, vol. 45, 224–240 (1997 Apr.); *J. Audio Eng. Soc.*, vol. 50, 331–342 (2002 May)]. The various components of this model have been separately described in the well-known publications of Patterson et al. [*J. Acoust. Soc. Am.*, vol. 72, 1788–1803 (1982)], Moore [*Hearing*, 161-205 (Academic Press 1995)], Moore et al. (1997), and Glasberg and Moore (2002). This paper provides a consolidated and concise introduction to the complete model for those who find the disparate and complex references intimidating and who wish to understand the function of each of the component parts. Furthermore, we provide a consolidated notation and integral forms. This introduction may be useful to the loudness theory beginner and to those who wish to adapt and apply the model for novel, practical purposes.
Convention Paper 8873

15:30

P11-5 Subjective Evaluation of Sound Quality of Musical Recordings Transmitted via the DAB+ System —*Maurycy Kin*, Wroclaw University of Technology, Wroclaw, Poland

The results of research on the sound quality of various kinds of music transmitted via Digital Audio Broadcasting using Absolute Category Rating and Comparison Category Rating methods of scaling are presented. The results showed that bit-rate values influence significantly the results. A Spectral Band Replication processing of signals increases the sound quality higher for low bit-rates than for higher values, dependently on a kind of music. The spatial attributes of sound, as a perspective, spaciousness, localization stability and an accuracy of phantom source, also are dependent on the bit-rate, but these relations are different. It was also found that a method of evaluation gives different results, and a CCR method is more accurate for sound assessment for higher bit-rates.
Convention Paper 8874

15:30

P11-6 Music and Emotions: A Comparison of Measurement Methods—*Judith Liebetrau*,[1,2] *Sebastian Schneider*[2]
[1]Fraunhofer IDMT, Ilmenau, Germany
[2]Ilmenau University of Technology, Ilmenau, Germany

Music emotion recognition (MER) as a part of music information retrieval (MIR), examines the question which parts of music evoke what emotions and how can they be automatically classified. Classification systems need to be trained in terms of feature selection and prediction. Due to the subjectivity of emotions, the generation of appropriate ground truth data poses challenges for MER. This paper describes obstacles of defining and measuring emotions evoked by music. Two methods, in principle able to overcome problems in measuring affective states induced by music, are outlined and their results are compared. Although the results of both methods are in line with psychological theories of emotions, the question remains how good the perceived emotions are captured by either method and if these methods are sufficient for ground truth generation.
Convention Paper 8875

15:30

P11-7 Multidimensional Scaling Analysis Applied to Music Mood Recognition—*Magdalena Plewa, Bozena Kostek*, Gdansk University of Technology, Gdansk, Poland

The paper presents two experiments aimed at categorizing mood associated with music. Two parts of a listening test were designed and carried out with a group of students, most of whom were users of online social music services. The initial experiment was designed to evaluate the extent to which a given label describes the mood of the particular music excerpt. The second subjective test was conducted to collect the similarity data for the MDS (Multidimensional Scaling)

analysis. Results were subject of various MDS and correlation analysis. Obtained MDS representation is relevant and remains coherent with acclaimed 2-dimensional Thayer's model as well as with evaluation using six mood labels.
Convention Paper 8876

15:30

P11-8 Artificial Stereo Extension Based on Gaussian Mixture Model—*Nam In Park, Kwang Myung Jeon, Chan Jun Chun, Hong Kook Kim*, Gwangju Institute of Science and Technology (GIST), Gwangju, Korea

In this paper an artificial stereo extension method is proposed to provide the stereophonic sound. The proposed method employs a minimum mean squared error (MMSE) estimator based on a Gaussian mixture model (GMM) to produce stereo signals from a mono signal. The performance of the proposed stereo extension method is evaluated using a multiple stimuli with a hidden reference and anchor (MUSHRA) test and compared with that of the parametric stereo method. It is shown from the test that the mean opinion score of the signals extended by the proposed stereo extension method is around 5% higher than that of the conventional stereo extension method based on inter-channel coherence (ICC).
Convention Paper 8877

Sunday, May 5 15:30 Sala Montale

Standards Committee Meeting on SC-05-05, EMC

Workshop 6 Sunday, May 5
16:30 – 18:30 Sala Manzoni

DESIGN AND BUILD OF FIRST ATMOS-EQUIPPED DUBBING THEATER

Chair: **Brian McCarty**, Coral Sea Studios Pty. Ltd., Clifton Beach, QLD, Australia

Panelist: *Philip Newell*

With immersive sound being talked about as the "next generation sound" for cinema and TV, this workshop will discuss in detail the design, construction, and commissioning of the world's first Atmos-equipped dubbing stage in Moscow, Russia.

Loudness 6 Sunday, May 5
16:30 – 17:30 Sala Alighieri

GIVE PEAKS A CHANCE

Presenter: **Thomas Lund**, TC Electronic A/S, Risskov, Denmark

Hearing is our most acute temporal sense by far, but the terms we have for describing dynamic changes in audio are few and not well defined. This session is about micro-dynamics and macro-dynamics in music and in speech, what effect they have, and what it takes to actually register them as a listener. Engineers be warned. Though audio examples are given, the presentation will primarily be based on anatomy, physiology, and psychology.

[A follow-up session to the Loudness War panel.]

Student/Career Event
RECORDING COMPETITION—PART 1
Sunday, May 5, 16:30 – 18:30
Auditorium Loyola

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 Studio Recording
- Modern Recording

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). 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.

Sunday, May 5 17:00 Sala Saba

Technical Committee Meeting on Electro Magnetic Compatibility

Professional Training Session 3 Sunday, May 5
17:30 – 18:30 Sala Foscolo

WIDE BANDWIDTH EQUALIZATION IN THE ANALOG DOMAIN

Based on research related to the sensitivity of our ears throughout the full bandwidth of our hearing, the CharterOak PEQ-1 provides extremely musical equalization in a stereo tone control of unparalleled musicality. Utilizing a unique circuit, and power distribution, along with a unique configuration with respect to bandwidth and cut and boost around the center frequency choices, the PEQ-1 achieves extremely powerful and musical equalization with minimal phase shift at the high frequencies, and extremely high headroom and low distortion.

Loudness 7 Sunday, May 5
17:45 – 18:15 Sala Alighieri

LOUDNESS IN RADIO—THE NEXT STEP

Presenter: **Florian Camerer**, ORF - Austrian TV, Vienna, Austria

After the start of the switchover from peak normalization to loudness levelling in TV, the logical progression is the move into Radio. One could argue that due to the hypercompression used in music production and Pop/Rock-stations over the last years, loudness differences are not an issue in Radio.... but that somewhat cynical view is definitely not an excuse to leave out the vast world of Radio programming. On the contrary! Albeit from a different angle and with other strategies, loudness production has many benefits and advantages also for supercompressed Radio stations. In this session those differences and challenges will be examined, and an outlook on the ➥

forthcoming work of the EBU-loudness group PLOUD in that area will be given.

Special Event
BANQUET
Sunday, May 5, 20:30 – 23:00
Marriott Grand Hotel Flora
via Veneto, Rome

This year, the Banquet will be held on the Roof Garden of the charming and prestigious Marriott Grand Hotel Flora, located on via Veneto, at the very center of Rome and walking distance from the convention centee. The view is breathtaking, covering Villa Borghese and several noteworthy places of interest. A relaxed atmosphere and a carefully selected menu will provide the proper context to meet with AES friends after an intense day of scientific contributions (or tourist walking). And if you arrive early, you can have a seat at Harry's Bar, that has recently opened in Rome and is located on the opposite side of the street.

60 Euros per person AES members and nonmembers
Tickets will be available at the Registration desk.

Session P12 **Monday, May 6**
09:00 – 13:00 **Sala Carducci**

SPATIAL AUDIO—PART 1: BINAURAL, HEAD-RELATED-TRANSFER-FUNCTIONS

Chair: **Michele Geronazzo**, University of Padova, Padova, Italy

09:00

P12-1 Binaural Ambisonic Decoding with Enhanced Lateral Localization—*Tim Collins*, University of Birmingham, Birmingham, UK

When rendering an ambisonic recording, a uniform speaker array is often preferred with the number of speakers chosen to suit the ambisonic order. Using this arrangement, localization in the lateral regions can be poor but can be improved by increasing the number of speakers. However, in practice this can lead to undesirable spectral impairment. In this paper a time-domain analysis of the ambisonic decoding problem is presented that highlights how a non-uniform speaker distribution can be used to improve localization without incurring perceptual spectral impairment. This is especially relevant to binaural decoders, where the locations of the virtual speakers are fixed with respect to the head, meaning that the interaction between speakers can be reliably predicted.
Convention Paper 8878

09:30

P12-2 A Cluster and Subjective Selection-Based HRTF Customization Scheme for Improving Binaural Reproduction of 5.1 Channel Surround Sound —*Bosun Xie, Chengyun Zhang, Xiaoli Zhong*, South China University of Technology, Guangzhou, China

This work proposes a cluster and subjective selection-based HRTF customization scheme for improving binaural reproduction of 5.1 channel surround sound. Based on similarity of HRTFs from an HRTF database with 52 subjects, a cluster analysis on HRTF magnitudes is applied. Results indicate that HRTFs of most subjects can be classified into seven clusters and represented by the corresponding cluster centers. Subsequently, HRTFs used in binaural 5.1 channel reproduction are customized from the seven cluster centers by means of subjective selection, i.e., a subjective selection-based customization scheme. Psychoacoustic experiments indicate that the proposed scheme partly improves the localization performance in the binaural 5.1 channel surround sound.
Convention Paper 8879

10:00

P12-3 Spatially Oriented Format for Acoustics: A Data Exchange Format Representing Head-Related Transfer Functions—*Piotr Majdak,*[1] *Yukio Iwaya,*[2] *Thibaut Carpentier,*[3] *Rozenn Nicol,*[4] *Matthieu Parmentier,*[5] *Agnieszka Roginska,*[6] *Yoiti Suzuki,*[7] *Kankji Watanabe,*[8] *Hagen Wierstorf,*[9] *Harald Ziegelwanger,*[1] *Markus Noisternig*[3]
[1]Austrian Academy of Sciences, Vienna, Austria
[2]Tohoku Gakuin University, Tagajo, Japan
[3]UMR STMS IRCAM-CNRS-UPMC, Paris, France
[4]Orange Labs, France Telecom, Lannion, France
[5]France Television, Paris, France
[6]New York University, New York, NY, USA
[7]Tohoku University, Sendai, Japan
[8]Akita Prefectural University, Yuri-Honjo, Japan
[9]Technische Universität Berlin, Berlin, Germany

Head-related transfer functions (HRTFs) describe the spatial filtering of the incoming sound. So far available HRTFs are stored in various formats, making an exchange of HRTFs difficult because of incompatibilities between the formats. We propose a format for storing HRTFs with a focus on interchangeability and extendability. The spatially oriented format for acoustics (SOFA) aims at representing HRTFs in a general way, thus, allowing to store data such as directional room impulse responses (DRIRs) measured with a microphone-array excited by a loudspeaker array. SOFA specifications consider data compression, network transfer, a link to complex room geometries, and aim at simplifying the development of programming interfaces for Matlab, Octave, and C++. SOFA conventions for a consistent description of measurement setups are provided for future HRTF and DRIR databases.
Convention Paper 8880

10:30

P12-4 Head Movements in Three-Dimensional Localization —*Tommy Ashby, Russell Mason, Tim Brookes*, University of Surrey, Guildford, Surrey, UK

Previous studies give contradicting evidence as to the importance of head movements in localization. In this study head movements were shown to increase localization response accuracy in elevation and azimuth. For elevation, it was found that head movement improved localization

accuracy in some cases and that when pinna cues were impeded the significance of head movement cues was increased. For azimuth localization, head movement reduced front-back confusions. There was also evidence that head movement can be used to enhance static cues for azimuth localization. Finally, it appears that head movement can increase the accuracy of listeners' responses by enabling an interaction between auditory and visual cues.
Convention Paper 8881

11:00

P12-5 A Modular Framework for the Analysis and Synthesis of Head-Related Transfer Functions —*Michele Geronazzo, Simone Spagnol, Federico Avanzini*, University of Padova, Padova, Italy

The paper gives an overview of a number of tools for the analysis and synthesis of head-related transfer functions (HRTFs) that we have developed in the past four years at the Department of Information Engineering, University of Padova, Italy. The main objective of our study in this context is the progressive development of a collection of algorithms for the construction of a totally synthetic personal HRTF set replacing both cumbersome and tedious individual HRTF measurements and the exploitation of inaccurate non-individual HRTF sets. Our research methodology is highlighted, along with the multiple possibilities of present and future research offered by such tools.
Convention Paper 8882

11:30

P12-6 Measuring Directional Characteristics of In-Ear Recording Devices—*Flemming Christensen, Pablo F. Hoffmann, Dorte Hammershøi*, Aalborg University, Aalborg, Denmark

With the availability of small in-ear headphones and miniature microphones it is possible to construct combined in-ear devices for binaural recording and playback. When mounting a microphone on the outside of an insert earphone the microphone position deviates from ideal positions in the ear canal. The pinna and thereby also the natural sound transmission are altered by the inserted device. This paper presents a methodology for accurately measuring the directional dependent transfer functions of such in-ear devices. Pilot measurements on a commercially available device are presented and possibilities for electronic compensation of the non-ideal characteristics are considered.
Convention Paper 8883

12:00

P12-7 Modeling the Broadband Time-of-Arrival of the Head-Related Transfer Functions for Binaural Audio—*Harald Ziegelwanger, Piotr Majdak*, Austrian Academy of Sciences, Vienna, Austria

Binaural audio is based on the head-related transfer functions (HRTFs) that provide directional cues for the localization of virtual sound sources. HRTFs incorporate the time-of-arrival (TOA), the monaural timing information yielding the interaural time difference, essential for the rendering of multiple virtual sound sources. In this study we propose a method to robustly estimate spatially continuous TOA from an HRTF set. The method is based on a directional outlier remover and a geometrical model of the HRTF measurement setup. We show results for HRTFs of human listeners from three HRTF databases. The benefits of calculating the TOA in the light of the HRTF analysis, modifications, and synthesis are discussed.
Convention Paper 8884

12:30

P12-8 Multichannel Ring Upmix—*Christof Faller*,[1] *Lutz Altmann*,[2] *Jeff Levison*,[2] *Markus Schmidt*[1]
[1]Illusonic GmbH, Uster, Switzerland
[2]IOSONO GmbH, Erfurt, Germany

Multichannel spatial decompositions and upmixes have been proposed, but these are usually based on an unrealistically simple direct/ambient sound model, not capturing the full diversity offered by N discrete audio channels, where in an extreme case each channel can contain an independent sound source. While it has been argued that a simple direct/ambient model is sufficient, in practice such is limiting the achievable audio quality. To circumvent the problem of capturing multichannel signals with a single model, the proposed "ring upmix" applies a cascade of 2-channel upmixes to surround signals to generate channels for setups with more loudspeakers featuring full support for 360-degree panning with high channel separation.
Convention Paper 8908

Session EB2 — Monday, May 6
09:00 – 10:45 — Sala Foscolo

ENGINEERING BRIEFS—PAPERS: PART 1

Chair: **Etienne Corteel**, Sonic Emotion Labs, Paris, France

09:00

EB2-1 Measurement of Sound Quality Differences in Individual CD Media Using Residual Waveform Comparison—*Akira Nishimura*, Tokyo University of Information Sciences, Chiba-shi, Japan

To measure miniscule differences of sound quality that might exist between different CD media we compared residuals of two DA- and AD-converted waveforms from different discs on which the same data were recorded under clock synchronization conditions between DA and AD converters. This method clarifies the existence of differences in sound quality, except for sampling clock fluctuation. The results showed no media-dependent difference in sound quality. The main source of the residual waveform was change in the audio circuit transfer function associated with time since turning on the CD player.
Engineering Brief 83

09:15

EB2-2 Binaural Room Simulation for Acoustic Testing—*Scott Levine*,[1,2] *Brett Leonard*,[1,2] *Richard King*[1,2]
[1]McGill University, Montreal, Quebec, Canada
[2]The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada

Often, in testing with acoustic conditions as the independent variable, challenges arise with the ease and speed of altering acoustic conditions. This study compares two possibilities for testing different acoustic conditions. In this test, physically varied acoustic treatment is compared to binaural room simulation. Explorations of these two methods are conducted employing an in situ, task-based paradigm presented to highly trained listeners. Results indicate significant differences in acoustic conditions within binaural simulations; however do not provide corresponding data to actual acoustic alteration.
Engineering Brief 84

09:30

EB2-3 The Advantages of Using Active Crossovers in High-End Wireless Speakers—*David Jones*, CSR Limited, Manchester, UK

With the availability of standardized wireless interfaces and high performance codecs, wireless loudspeakers can be designed that suit the consumer demands of compactness and ease of use. This paper will examine the performance benefits of using active crossovers and digital equalization in an amplification subsystem based on a high performance digital input switching amplifier. Measurements of distortion and damping factors will be compared in an example signal chain and the influence these parameters have on the perceived audio quality of the speaker system will be discussed.
Engineering Brief 85

09:45

EB2-4 Comparative Analysis of Different Loudness Meters Based on Voice Detection and Gating—*Alessandro Travaglini*, Fox International Channels Italy, Guidonia Montecelio (RM), Italy

After decades of extensive investigation, the international broadcasting community, represented by technical associations and bodies, has set precise standards aimed to objectively assess loudness levels of programs. Although all standards rely on the same algorithm as described in ITU-R BS1770, there are still two possible ways to implement such metering, including voice detection and gating. These two different implementations might, in some cases, provide measurements that significantly differ from each other. Furthermore, while the gating feature is uniquely defined in the updated version of BS1770-3, voice detection is not currently specified in any standard and its implementation is the independent choice of manufacturers. This paper analyses this scenario by comparing the results and robustness provided by three different loudness meters based on voice detection. In addition, those values are compared with measurements obtained by using BS1770-3 compliant loudness meters, including tables, comments, and conclusions.
Engineering Brief 86

10:00

EB2-5 Assessing the Standardization of an Existing iOS Control Application to AES64-2012 Network Protocol—*Joan Amate*, Master Audio, Barcelona, Spain

The recent publication of AES64-2012 standard has motivated the comparison of Master Audio's own IP network control protocol against the new standard, in order to assess its interoperability or adaptability. This brief analyzes what AES64 means for manufacturers with existing control protocols who are willing to seek standardization. The protocol used for this assessment was developed for controlling self-amplified PA systems (built-in amplifier and processing), and is fully functional under Windows and iOS (iPad). Finally, a brief guide on how to face standardization is given from the manufacturer point of view.
Engineering Brief 87

10:15

EB2-6 Innovation in Audio: Update on Patent Activity in the Audio Field—*Elliot Cook, Joseph E. Palys*, Finnegan, Henderson, Farabow, Garrett & Dunner, LLP, Reston, VA, USA

This paper provides a sampling of recent patents relating to innovations in the audio field. The innovations come from a range of AES member companies and cover a diverse spectrum of technologies, such as music composition software, loudspeaker design, headphones, digital signal processing, microphones, and musical comprehension. In addition, unique statistical information regarding patent litigation in the audio field is provided. This information is based on original research regarding litigation involving audio patents. AES members may find this information helpful to better understand how audio innovations play a role in their industry.

Engineering Brief 89

10:30

EB2-7 An Examination of Early Analog and Digital Sampling—The Robb Wave Organ circa 1927—*Michael Murphy, Eric Kupp*, Ryerson University, Toronto, ON, Canada

This paper examines Frank Morse Robb's work in the late 1920s and early 1930s on his Wave Organ, the first successful electronic organ. The Robb Wave Organ originally functioned by creating a visual representation of an analog pipe organ waveform through means of an oscilloscope and engraving that representation onto metal tone wheels. Later versions of the organ featured a digital, almost PCM-style, waveform representation on the tone wheels. This predates the theoretical description of PCM by Alec Reeves, as well as the PCM patent filed by Oliver and Shannon in 1946. These sample-based methods of tone generation were unique to the Robb Wave Organ, and this paper serves to

place the organ within its contemporaries of that time period, most notably its primary competitor, the Hammond organ, launched in 1935.
Engineering Brief 90

Tutorial 10 **Monday, May 6**
09:00 – 10:30 **Auditorium Loyola**

CREATIVE DISTORTION—YOU ARE IN THE OVER-DRIVER'S SEAT

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

Distortion, strategically applied to elements of your mix, is a source of energy that lifts tracks up out of a crowded arrangement and adds excitement to the performance. Accidental distortion, on the other hand, is a certain sign that the production is unprofessional, dragging down its chance for success. Amps, stomp boxes, tubes, transformers, tape machines, the plug-ins that emulate them, and the plug-ins that create wholly new forms of distortion all offer a rich palette of distortion colors. Mix engineers must know how to choose among them, and how to tailor them to the music. This tutorial takes a close look at distortion, detailing the technical goings-on when things break-up, and defining the production potential of this, the caffeine of effects.

Workshop 7 **Monday, May 6**
09:00 – 12:00 **Sala Manzoni**

MULTICHANNEL IMMERSIVE AUDIO FORMATS FOR 3-D CINEMA AND HOME THEATER

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

Panelists *Brian Claypool*
Kimio Hamasaki
Charles Robinson
Wilfried Van Baelen

Several new immersive sound formats are under active consideration for cinema soundtrack production. Each was developed to create realistic sound "motion" and "immerse" the audience in a more realistic soundfield. This workshop is a repeat of the program presented at AES 133 in San Francisco, with the proponents of four of the leading immersive sound systems to discuss their specific technologies.

Monday, May 6 **09:00** **Sala Saba**

Technical Committee Meeting on High Resolution Audio

Monday, May 6 **09:00** **Sala Montale**

Standards Committee Meeting on SC-04-08, Sound Systems in Rooms

Session P13 **Monday, May 6**
10:00 – 11:30 **Foyer**

POSTERS: ROOM ACOUSTICS

10:00

P13-1 The Effect of Playback System on Reverberation Level Preference—*Brett Leonard*,[1,2] *Richard King*,[1,2] *Grzegorz Sikora*[3]
[1]McGill University, Montreal, Quebec, Canada
[2]The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada
[3]Bang & Olufsen Deutschland GmbH, Pullach, Germany

The critical role of reverberation in modern acoustic music production is undeniable. Unlike many other effects, reverberation's spatial nature makes it extremely dependent upon the playback system over which it is experienced. While this characteristic of reverberation has been widely acknowledged among recording engineers for years, the increase in headphone listening prompts further exploration of these effects. In this study listeners are asked to add reverberation to a dry signal as presented over two different playback systems: headphones and loudspeakers. The final reverberation levels set by each subject are compared for the two monitoring systems. The resulting data show significant level differences across the two monitoring systems.
Convention Paper 8886

10:00

P13-2 Adaptation of a Large Exhibition Hall as a Concert Hall Using Simulation and Measurement Tools— *Marco Facondini*,[1] *Daniele Ponteggia*[2]
[1]TanAcoustics Studio, Pesaro (PU), Italy
[2]Studio Ing. Ponteggia, Terni (TR), Italy

Due to the growing demand of multifunctional performing spaces, there is a strong interest in the adaptation of non-dedicated spaces to host musical performances. This leads to new challenges for the acousticians with new design constraints and very tight time frames. This paper shows a practical example of the adaptation of the "Sala della Piazza" of the *Palacongressi* of Rimini. Thanks to the combined use of prediction and measurement tools it has been possible to design the acoustical treatments with a high degree of accuracy, reaching all targets and at the same time respecting the tight deadlines.
Convention Paper 8887

10:00

P13-3 Digital Filter for Modeling Air Absorption in Real Time—*Carlo Petruzzellis, Umberto Zanghieri*, ZP Engineering S.r.L., Rome, Italy

Sound atmospheric attenuation is a relevant aspect of realistic space modeling in 3-D audio simulation systems. A digital filter has been developed on commercial DSP processors to match air absorption curves. This paper focuses on the algorithm implementation of a digital filter with continuous roll-off control, to simulate high frequency damping of audio signals in various atmospheric conditions, along with rules to allow a precise approximation of the behavior described by analytical formulas.
Convention Paper 8888

➥

10:00

P13-4 Development of Multipoint Mixed-Phase Equalization System for Multiple Environments—*Stefania Cecchi,*[1] *Marco Virgulti,*[1] *Stefano Doria,*[2] *Ferruccio Bettarelli,*[2] *Francesco Piazza*[1]
[1]Universitá Politecnica della Marche, Ancona, Italy
[2]Leaff Engineering, Ancona, Italy

The development of a mixed-phase equalizer is still an open problem in the field of room response equalization. In this context, a multipoint mixed-phased impulse response equalization system is presented taking into consideration a magnitude equalization procedure based on a time-frequency segmentation of the impulse responses and a phase equalization technique based on the group delay analysis. Furthermore, an automatic software tool for the measurement of the environment impulse responses and for the design of a suitable equalizer is presented. Taking advantage of this tool, several tests have been performed considering objective and subjective analysis applied in a real environment and comparing the obtained results with different approaches.
Convention Paper 8889
[Paper presented by Marco Virgulti]

10:00

P13-5 Acoustics Modernization of the Recording Studio in Wroclaw University of Technology —*Magdalena Kaminska, Patryk Kobylt, Bartlomiej Kruk, Jan Sokolnicki*, Wroclaw University of Technology, Wroclaw, Poland

The aim of this paper is to present results of the acoustic modernization at the Wroclaw University of Technology recording studio. During the project realization, the focus is on the problem arising in one part of the recording studio—the so-called flutter echoes phenomenon. To minimize this effect we present a several-stage process in which the studio is accommodated to expect this occurrence. The first step was to make some measurements of acoustic properties in the room with the concentration on the previously mentioned effect. Next, a one-dimension diffuser was designed and placed in the phenomenon incidence. The last stage of the research was an acoustic measurement after modification and comparison with the properties before the changes.
Convention Paper 8890

10:00

P13-6 Accurate Acoustic Echo Reduction with Residual Echo Power Estimation for Long Reverberation —*Masahiro Fukui,*[1] *Suehiro Shimauchi,*[1] *Yusuki Hioka,*[2] *Hitoshi Ohmuro,*[1] *Yoichi Haneda*[3]
[1]NTT Corporation, Musashino-shi, Tokyo, Japan
[2]University of Canterbury, Christchurch, New Zealand
[3]The University of Electro-Communications, Chofu-shi, Tokyo, Japan

This paper deals with the problem of estimating and reducing residual echo components that result from reverberant components beyond the length of FFT block. The residual echo reduction process suppresses the residual echo by applying a multiplicative gain calculated from the estimated echo power spectrum. However, the estimated power spectrum reproduces only a fraction of the echo-path impulse response and so all the reverberant component are not considered. To address this problem we introduce a finite nonnegative convolution method by which each segment of echo-impulse response is convoluted with a received signal in a power spectral domain. With the proposed method, the power spectra of each segment of echo-impulse response are collectively estimated by solving the least-mean-squares problem between the microphone and the estimated-residual-echo power spectra. The performance of this method was demonstrated by simulation results in which speech distortions were decreased compared with the conventional method.
Convention Paper 8891

Monday, May 6 11:00 Sala Montale

Standards Committee Meeting on SC-04-03, Loudspeaker Modeling and Measurement

Tutorial 11 Monday, May 6
11:30 – 13:00 Sala Foscolo

OPTIMIZING AND MEASURING THE SPEECH INTELLIGIBILITY OF SOUND SYSTEMS

Presenters: **Ben Kok**, BEN KOK acoustic consulting, Uden, The Netherlands
Peter Mapp, Peter Mapp Associates, Colchester, Essex, UK

Intelligibility is the most important parameter of any sound system—albeit a high-quality, theater sound reinforcement system, a church, or an emergency PA system operating deep underground in a road tunnel or an information system at an airport or rail station. The tutorial will discuss how to design intelligible sound systems for both sound reinforcement and emergency systems use. It will show how reverberation and noise and other acoustic characteristics of a building or space can affect speech and sound system intelligibility. We will also show how the choice and positioning of loudspeakers can have a significant impact on the potential intelligibility. The techniques will be illustrated with case histories and practical examples. Measurement of the potential intelligibility performance of a sound system will also be covered along with the equipment needed and the practical issues that need to be dealt with.

Monday, May 6 12:00 Sala Saba

Technical Committee Meeting on Coding of Audio Signals

Tutorial 12 Monday, May 6
12:15 – 14:15 Auditorium Loyola

TURNTABLE TECHNIQUE: THE ART OF THE DJ

Presenter: **Stephen Webber**, Berklee College of Music, Valencia, Spain

Mastering faders, pots, and switches as musical instruments? Turntable technique is where the audio engineer and the artist become one. Stephen Webber is the Director of Music Technology Innovation at Berklee College of Music in Valencia Spain, the composer or the "Stylus Symphony," and the author of *Turntable Technique: The Art of the DJ*. His tutorial, which has been featured around the world, challenges traditional notions of music and technology.

Tutorial 13 **Monday, May 6**
12:30 – 13:30 **Sala Alighieri**

DESIGNING FOR ULTRA-LOW THD+N IN ANALOG CIRCUITS

Presenter: **Bruce E. Hofer**, Audio Precision

A number of factors influence the distortion performance of analog circuits. Among the most important is the quality of components. Models of non-linearity and techniques for estimating distortion and noise will be discussed along with some tips on minimizing their contributions. This presentation is an updated subset of the design track seminar titled "Designing Audio in the 2010s" that was presented at the 2011 New York AES convention. It contains some significant new material that should be of interest to the serious analog design engineer.

Monday, May 6 **12:30** **Sala Montale**

Standards Committee Meeting on SC-04-04, Microphone Measurement and Characterization

Monday, May 6 **13:00** **Sala Saba**

Technical Committee Meeting on Fiber Optics for Audio

Monday, May 6 **14:00** **Sala Saba**

Technical Committee Meeting on Transmission and Broadcasting

Tutorial 14 **Monday, May 6**
14:15 – 16:15 **Sala Alighieri**

RUB & BUZZ AND OTHER IRREGULAR LOUDSPEAKER DISTORTION

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

Loudspeaker defects caused by manufacturing, aging, overload, or climate impact generate a special kind of irregular distortion commonly known as rub & buzz, which are highly audible and intolerable for the human ear. Contrary to regular loudspeaker distortions defined in the design process, the irregular distortions are hardly predictable and are generated by an independent process triggered by the input signal. Traditional distortion measurements such as THD fail in the reliable detection of those defects. This tutorial discusses the most important defect classes, new measurement techniques, audibility, and the impact on perceived sound quality.

Session P14 **Monday, May 6**
14:30 – 17:30 **Sala Carducci**

APPLICATIONS IN AUDIO

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

14:30

P14-1 Implementation of an Intelligent Equalization Tool Using Yule-Walker for Music Mixing and Mastering —*Zheng Ma, Joshua D. Reiss, Dawn A. A. Black*, Queen Mary University of London, London, UK

A new approach for automatically equalizing an audio signal toward a target frequency spectrum is presented. The algorithm is based on the Yule-Walker method and designs recursive IIR digital filters using a least-squares fitting to any desired frequency response. The target equalization curve is obtained from the spectral distribution analysis of a large dataset of popular commercial recordings. A real-time C++ VST plug-in and an off-line Matlab implementation have been created. Straightforward objective evaluation is also provided, where the output frequency spectra are compared against the target equalization curve and the ones produced by an alternative equalization method.
Convention Paper 8892

15:00

P14-2 On the Informed Source Separation Approach for Interactive Remixing in Stereo
—*Stanislaw Gorlow*,[1] *Sylvain Marchand*[2]
[1]University of Bordeaux, Talence, France
[2]Université de Brest, Brest, France

Informed source separation (ISS) has become a popular trend in the audio signal processing community over the past few years. Its purpose is to decompose a mixture signal into its constituent parts at the desired or the best possible quality level given some metadata. In this paper we present a comparison between two ISS systems and relate the ISS approach in various configurations with conventional coding of separate tracks for interactive remixing in stereo. The compared systems are Underdetermined Source Signal Recovery (USSR) and Enhanced Audio Object Separation (EAOS). The latter forms a part of MPEG's Spatial Audio Object Coding technology. The performance is evaluated using objective difference grades computed with PEMO-Q. The results suggest that USSR performs perceptually better than EOAS and has a lower computational complexity.
Convention Paper 8893

15:30

P14-3 Scene Inference from Audio—*Daniel Arteaga*,[1,2] *David García-Garzón*,[2] *Toni Mateos*,[3] *John Usher*[4]
[1]Fundacio Barcelona Media, Barcelona, Spain
[2]Universitat Pompeu Fabra, Barcelona, Spain
[3]imm sound, Barcelona, Spain
[4]Hearium Labs, San Francisco, CA, USA

We report on the development of a system to characterize the geometric and acoustic properties of a space from an acoustic impulse response measured within it. This can be thought of as the inverse problem to the common practice of obtaining impulse responses from either real-world or virtual spaces. Starting from an impulse response recorded in an original scene, the method described here uses a non-linear search strategy to select a scene that is perceptually as close as possible to the original one. Potential applications of this method include audio production, non-intrusive acquisition of room geometry, and audio forensics.
Convention Paper 8894

16:00

P14-4 Continuous Mobile Communication with Acoustic Co-Location Detection—*Robert Albrecht*,[1] *Sampo Vesa*,[2] *Jussi Virolainen*,[3] *Jussi Mutanen*,[4] *Tapio Lokki*[1]
[1]Aalto University, Espoo, Finland
[2]Nokia Research Center, Nokia Group, Finland
[3]Nokia Lumia Engineering, Nokia Group, Finland
[4]JMutanen Software, Jyväskylä, Finland

In a continuous mobile communication scenario, e.g., between co-workers, participants may occasionally be located in the same space and thus hear each other naturally. To avoid hearing echoes, the audio transmission between these participants should be cut off. In this paper an acoustic co-location detection algorithm is proposed for the task, grouping participants together based solely on their microphone signals and mel-frequency cepstral coefficients thereof. The algorithm is tested both on recordings of different communication situations and in real time integrated into a voice-over-IP communication system. Tests on the recordings show that the algorithm works as intended, and the evaluation using the voice-over-IP conferencing system concludes that the algorithm improves the overall clarity of communication compared with not using the algorithm. The acoustic co-location detection algorithm thus proves a useful aid in continuous mobile communication systems.
Convention Paper 8895

16:30

P14-5 Advancements and Performance Analysis on the Wireless Music Studio (WeMUST) Framework—*Leonardo Gabrielli, Stefano Squartini, Francesco Piazza*, Universitá Politecnica della Marche, Ancona (AN), Italy

Music production devices and musical instruments can take advantage of IEEE 802.11 wireless networks for interconnection and audio data sharing. In previous works such networks have been proved able to support high-quality audio streaming between devices at acceptable latencies in several application scenarios. In this work a prototype device discovery mechanism is described to improve ease of use and flexibility. A diagnostic tool is also described and provided to the community that allows to characterize average network latency and packet loss. Lower latencies are reported after software optimization and sustainability of multiple audio channels is also proved by means of experimental tests.
Convention Paper 8896

17:00

P14-6 Acoustical Characteristics of Vocal Modes in Singing—*Eddy B. Brixen*,[1] *Cathrine Sadolin*,[2] *Henrik Kjelin*[2]
[1]EBB-consult, Smorum, Denmark
[2]Complete Vocal Institute, Copenhagen, Denmark

According to the Complete Vocal Technique four vocal modes are defined: Neutral, Curbing, Overdrive, and Edge. These modes are valid for both the singing voice and the speaking voice. The modes are clearly identified both from listening and from visual laryngograph inspection of the vocal cords and the surrounding area of the vocal tract. In a recent work a model has been described to distinguish between the modes based on acoustical analysis. This paper looks further into the characteristics of the voice modes in singing in order to test the model already provided. The conclusion is that the model is too simple to cover the full range. The work has also provided information on singers' SPL and formants' repositioning in dependence of pitch. Further work is recommended.
Convention Paper 8897

Workshop 8 **Monday, May 6**
14:30 – 17:30 **Sala Foscolo**

CURRENT REFERENCE LISTENING ROOM STANDARDS: ARE THEY MEANINGFUL?

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

Panelists: *Sean Olive*
Francis Rumsey
Andreas Silzle
Thomas Sporer

The ITU BS.1116 standard "Methods for the Subjective Assessment of Small Impairments in Audio Systems … " contains guidelines for standard listening environments used for assessing subjective quality of audio systems. This includes factors like dynamics and frequency/directivity response for loudspeakers, physical, and acoustical properties of the room. This standardized playback environment should be consistent and representative of high quality systems, yet representative of systems that actually exist. This workshop takes a fresh look at the ITU BS.1116 standard and how it could be improved. In fact some changes that in theory might improve the specification might not actually be practical. Many aspects of the discussion would be relevant to reference listening room standards other then the BS.1116 standard, thus the more general workshop title.

Student/Career Event
EDUCATION FORUM
Monday, May 6, 14:30 – 16:30
Auditorium Loyola

Moderators: **Philip Waldenberger**, Chair, AES SDA,

Europe and International Regions
Colin Pfund, Chair, AES SDA, North and Latin American Regions
Marija Kovacina, Vice Chair, AES SDA, Europe and International Regions

Student Roundtable

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

Explore strategies for making connections with the professional world and discuss the curriculums and philosophies of your programs. Students, faculty, alumni, industry professionals, and anyone interested in commenting on the state of audio education are welcome to participate.

Session P15 **Monday, May 6**
15:00 – 16:30 **Foyer**

POSTERS: SPATIAL AUDIO

15:00

P15-1 Intelligent Acoustic Interfaces for Immersive Audio —*Danilo Comminiello, Michele Scarpiniti, Raffaele Parisi, Aurelio Uncini*, Sapienza University of Rome, Rome, Italy

Oncoming audio technologies privilege the perceptive quality of audio signals, thus offering users an *immersive audio experience*, which involves listening and acquisition of audio signals. In such a scenario a fundamental role is played by *intelligent acoustic interfaces* that aim at acquiring audio information, processing it, and returning the processed information under the fulfillment of quality requirements demanded by users. In this paper we introduce intelligent acoustic interfaces for immersive audio experience and we prove their effectiveness within the context of immersive speech communications. In particular, we introduce an intelligent acoustic interface composed of a combined adaptive beamforming scheme in conjunction with a microphone array, which is able to enhance the processed signals in immersive scenarios.
Convention Paper 8898

15:00

P15-2 The Effects of Spatial Depth in the Combinations of 3-D Imagery and 7-Channel Surround with Height Channels—*Toru Kamekawa*,[1] *Atsushi Marui*,[1] *Toshihiko 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

The effect of the speakers of the height direction in a 3-D imagery focused on the spatial depth were studied and conducted. The first experiment was carried out using a method of magnitude estimation asking how near or far the combination of perceived visual and auditory event is. In the second experiment, the subjects were asked to rate on suitability of the sound to the image using the same materials as the previous experiment. The results show that 7ch and 5ch surround were felt closely and 2ch stereo was felt far under the condition that there is no image. Regarding suitability of sound to an image, 3-D imagery with 7ch surround gives higher score in the near distance.
Convention Paper 8899

15:00

P15-3 Comparative Analysis on Compact Representation for Spatial Variation of Individual Head-Related Transfer Functions Based on Singular Value Decomposition— *Shouichi Takane*, Akita Prefectural University, Yurihonjo, Akita, Japan

In this paper the compact representation of the head-related transfer functions (HRTFs) or the Head-Related Impulse Responses (HRIRs) based on singular value decomposition (SVD) was investigated, focusing on the difference in the parameters to construct the average vectors and the covariance matrices in two points. One of them is on what the derived eigenvectors reflect the properties of HRTFs and/or HRIRs. As a result, high correlation between the parameters concerning SVD and the amplitude of HRTFs was found, and the high correlation was obtained in the wide region except the contralateral side. The second investigation is on the required number of the HRTFs or the HRIRs to construct the average vectors and the covariance matrices. It was found that the number of HRTFs are decreased to about 1/4 of the whole directions for the used HRTFs. Among three conditions of the HRIRs, the amplitude of HRTFs, and the log-amplitude of HRTFs, it was also shown that the amplitude of HRTFs is the most effective to construct the parameters of the SVD.
Convention Paper 8900

15:00

P15-4 Calculation of Individualized Near-Field Head-Related Transfer Function Database Using Boundary Element Method—*Yuanqing Rui, Guangzheng Yu, Bosun Xie, Yu Liu*, South China University of Technology, Guangzhou, China

Measurement is a common method to obtain the far-field HRTFs. Due to the difficulties in measurement, near-field HRTF databases for an artificial head are rare and an individualized database for human subjects is now unavailable. The present work adopts a laser 3-D scanner to acquire geometrical surfaces of human subjects and then uses boundary element methods to calculate the near-field HRTFs. At last, an individualized near-field HRTF database with 56 human subjects is established. To evaluate the accuracy of the database, the HRTFs for KEMAR are also calculated and compared to the measured ones.
Convention Paper 8901

➥

15:00

P15-5 A Standardized Repository of Head-Related and Headphone Impulse Response Data— *Michele Geronazzo, Fabrizio Granza, Simone Spagnol, Federico Avanzini*, University of Padova, Padova, Italy

This paper proposes a repository for storing full- and partial-body Head-Related Impulse Responses (HRIRs/pHRIRs) and Headphone Impulse Responses (HpIRs) from several databases in a standardized environment. The main differences among the available databases concern coordinate systems, sound source stimuli, sampling frequencies, and other important specifications. The repository is organized so as to consider all these differences. The structure of our repository is an improvement with respect to the MARL-NYU data format, born as an attempt to unify HRIR databases. The introduced information supports flexible analysis and synthesis processes and robust headphone equalization.
Convention Paper 8902

15:00

P15-6 Influence of Different Microphone Arrays on IACC as an Objective Measure of Spaciousness—*Marco Conceição,*[1,2] *Dermot Furlong*[1]
[1]Trinity College, Dublin, Ireland
[2]Instituto Politécnico do Porto, Porto, Portugal

Inter-Aural Cross Correlation measurements are used as physical measures that relate to listener spaciousness experience in a comparative study of the influence on spaciousness of different microphone arrays, thus allowing an objective approach to be adopted in the exploration of how microphone arrays affect the perceived spaciousness for stereo and surround sound reconstructions. The different microphone arrays recorded simulated direct and indirect sound components. The recorded signals were played back in three different rooms and IACC measurements were made for the reconstructed sound fields using a dummy head microphone system. The results achieved show how microphone array details influence the IACC peak and lead to a better understanding of how spaciousness can be controlled for 2 channel stereo and 5.1 presentations. Parametric variation of microphone arrays can therefore be employed to facilitate spaciousness control for reconstructed sound fields.
Convention Paper 8885

Monday, May 6 15:00 Sala Saba

Technical Committee Meeting on Sound for Digital Cinema and Television

Student/Career Event
RECORDING COMPETITION—PART 2
Monday, May 6, 16:30 – 18:30
Auditorium Loyola

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

• Traditional Acoustic Recording

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). 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.

Professional Training 4 Monday, May 6
16:30 – 18:30 Sala Alighieri

RECORDING AND MIXING WITH UNIVERSAL AUDIO

Presenter: **Wesley Bendall**

An overview is given of Universal Audio history from the 1950s to the present, detailing various well-known analog gear and hints on how and why you would, and should use them. Wesley will talk about the benefits that UA plugins can bring to a recording and a mix, and will explain and demonstrate the Apollo platform, especially new Apollo 16.

Monday, May 6 16:30 Sala Montale

Standards Committee Meeting on SC-02-12, Audio Applications of Networks

Special Event
AN IMMERSAV JAZZ CONCERT
FEATURING THE GREG BURK TRIO
Monday, May 7, 20:00 – 21:00
Saint Louis College of Music
Via Urbana 49/a
00184 Roma

A special music concert will be held Monday evening featuring a performance of the Greg Burk Jazz Trio. As a unique aspect of the concert, Robert Schulein will be making a combined binaural and HD video recording for downloading from the AES web site. following the convention. The audio/video recordings will be formatted for both headphone and loudspeaker listening (via cross talk cancellation) for viewing by AES members. This will allow attendees an opportunity to attend the event live and recreate the experience later by means of a surround audio with video recording process compatible with today's consumer viewer / listening technology.

Pianist/composer/educator **Greg Burk** has lived, studied, educated, and performed around the world. Over the last decade, living and working in Detroit, Boston, and now Rome, he has established himself not only as a vital sideman, but also as a leader with the rare ability to combine original compositions with memorable melodies and a desire to push the creative envelope as an improviser.

Session P16 **Tuesday, May 7**
09:00 – 12:30 **Sala Carducci**

SPATIAL AUDIO—PART 2: 3-D MICROPHONE AND LOUDSPEAKER SYSTEMS

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

09:00

P16-1 Recording and Playback Techniques Employed for the "Urban Sounds" Project—*Angelo Farina, Andrea Capra, Alberto Amendola, Simone Campanini*, University of Parma, Parma, Italy

The "Urban Sounds" project, born from a cooperation of the Industrial Engineering Department at the University of Parma with the municipal institution La Casa della Musica, aims to record characteristic soundscapes in the town of Parma with a dual purpose: delivering to posterity an archive of recorded sound fields to document Parma in 2012, employing advanced 3-D surround recording techniques and creation of a "musical" Ambisonics composition for leading the audience through a virtual tour of the town. The archive includes recordings of various "soundscapes," such as streets, squares, schools, churches, meeting places, public parks, train station, and airport, and everything was considered unique to the town. This paper details the advanced digital sound processing techniques employed for recording, processing, and playback.
Convention Paper 8903

09:30

P16-2 Robust 3-D Sound Source Localization Using Spherical Microphone Arrays—*Carl-Inge Colombo Nilsen,*[1,2] *Ines Hafizovic,*[2] *Sverre Holm*[1]
[1]University of Oslo, Oslo, Norway
[2]Squarehead Technology AS, Oslo, Norway

Spherical arrays are gaining increased interest in spatial audio reproduction, especially in Higher Order Ambisonics. In many audio applications the sound source detection and localization is of crucial importance, urging for robust localization methods suitable for spherical arrays. The well-known direction-of-arrival estimator, the ESPRIT algorithm, is not directly applicable to spherical arrays for 3-D applications. The eigenbeam ESPRIT (EB-ESPRIT) is based on the spherical harmonics framework and is especially derived for spherical arrays. However, the ESPRIT method is in general susceptible to errors in the presence of correlated sources and spatial decorrelation is not possible for spherical arrays. In our new implementation of spherical harmonics-based ESPRIT (SA2ULA-ESPRIT) the robustness against correlation is achieved by spatial decorrelation incorporated directly in the algorithm formulation. The simulated performance of the new algorithm is compared to EB-ESPRIT.
Convention Paper 8904

10:00

P16-3 Parametric Spatial Audio Coding for Spaced Microphone Array Recordings—*Archontis Politis,*[1] *Mikko-Ville Laitinen,*[1] *Jukka Ahonen,*[2] *Ville Pulkki*[1]
[1]Aalto University, Espoo, Finland
[2]Akukon Ltd., Helsinki, Finland

Spaced microphone arrays for multichannel recording of music performances, when reproduced in a multichannel system, exhibit reduced inter-channel coherence that translates perceptually to a pleasant "enveloping" quality, at the expense of accurate localization of sound sources. We present a method to process the spaced microphone recordings using the principles of Directional Audio Coding (DirAC), based on the knowledge of the array configuration and the frequency-dependent microphone patterns. The method achieves equal or better quality to the standard high-quality version of DirAC and it improves the common one-to-one channel playback of spaced multichannel recordings by offering improved and stable localization cues.
Convention Paper 8905

10:30

P16-4 Optimal Directional Pattern Design Utilizing Arbitrary Microphone Arrays: A Continuous-Wave Approach—*Symeon Delikaris-Manias, Constantinos A. Valagiannopoulos, Ville Pulkki*, Aalto University, Aalto, Finland

A frequency-domain method is proposed for designing directional patterns from arbitrary microphone arrays employing the complex Fourier series. A target directional pattern is defined and an optimal set of sensor weights is determined in a least-squares sense, adopting a continuous-wave approach. It is based on discrete measurements with high spatial sampling ratio, which mitigates the potential aliasing effect. Fourier analysis is a common method for audio signal decomposition; however in this approach a set of criteria is employed to define the optimal number of Fourier coefficients and microphones for the decomposition of the microphone array signals at each frequency band. Furthermore, the low-frequency robustness is increased by smoothing the target patterns at those bands. The performance of the algorithm is assessed by calculating the directivity index and the sensitivity. Applications, such as synthesizing virtual microphones, beamforming, binaural, and loudspeaker rendering are presented.
Convention Paper 8906

11:00

P16-5 Layout Remapping Tool for Multichannel Audio Productions—*Tim Schmele,*[1] *David García-Garzón,*[2] *Umut Sayin,*[1] *Davide Scaini,*[1,2] *Daniel Arteaga*[1,2]
[1]Fundació Barcelona Media, Barcelona, Spain
[2]Universitat Pompeu Fabra, Barcelona, Spain

Several multichannel audio formats are present in the recording industry with reduced interoperability among the formats. This diversity of formats leaves the end user with limited accessibility to content and/or audience. In addition, the preservation of recordings—that are made for a ➥

particular format—comes under threat, should the format become obsolete. To tackle such issues, we present a layout-to-layout conversion tool that allows converting recordings that are designated for one particular layout to any other layout. This is done by decoding the existent recording to a layout independent equivalent and then encoding it to desired layout through different rendering methods. The tool has proven useful according to expert opinions. Simulations depict that after several consecutive conversions the results exhibit a decrease in spatial accuracy and increase in overall gain. This suggests that consecutive conversions should be avoided and only a single conversion from the originally rendered material should be done.
Convention Paper 8907

11:30

P16-7 Discrimination of Changing Loudspeaker Positions of 22.2 Multichannel Sound System Based on Spatial Impressions—*Ikuko Sawaya, Kensuke Irie, Takehiro Sugimoto, Akio Ando, Kimio Hamasaki*, Science & Technology Research Laboratories, Japan Broadcasting Corp., Setagaya, Tokyo, Japan

On 22.2 multichannel reproduction, we sometimes listened to the sounds reproduced by a loudspeaker arrangement different from that on production, and we did not recognize the difference in spatial impression between them definitely. In this paper we discuss the effects of changing some of the loudspeaker positions from the reference on the spatial impressions in a 22.2 multichannel sound system. Subjective evaluation tests were carried out by altering the azimuthal and elevation angles from the reference in each condition. Experimental results showed that the listeners did not recognize the difference in spatial impression from the reference with some loudspeaker arrangements. On the basis of these results, the ranges keeping the equivalent quality of the spatial impressions to the reference are discussed when the reproduced material has the signals of all the channels of the 22.2 multichannel sound system.
Convention Paper 8909

12:00

P16-8 Modeling Sound Localization of Amplitude-Panned Phantom Sources in Sagittal Planes—*Robert Baumgartner, Piotr Majdak*, Austrian Academy of Sciences, Vienna, Austria

Localization of sound sources in sagittal planes (front/back and top/down) relies on listener-specific monaural spectral cues. A functional model approximating human processing of spectro-spatial information is applied to assess localization of phantom sources created by amplitude panning of coherent loudspeaker signals. Based on model predictions we investigated the localization of phantom sources created by loudspeakers positioned in the front and in the back, the effect of loudspeaker span and panning ratio on localization performance in the median plane, and the amount of spatial coverage provided by common surround sound systems. Our findings are discussed in the light of previous results from psychoacoustic experiments.
Convention Paper 8910

Tutorial 15 — Tuesday, May 7
09:00 – 11:00 — Sala Foscolo

PARALLELIZATION OF AUDIO SIGNAL PROCESSING ALGORITHMS ON MULTICORE DSPS

Presenters: **Carsten Tradowsky**, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany; CTmusic, Karlsruhe, Germany
Jochen Schyma, Texas Instruments

Today, digital audio systems for music production are restricted because they do not exhaust the possibilities given by modern hardware. One possibility to close the productivity gap is to use a high-level model-based development approach.

This tutorial presents a concept on using a model-based development approach to describe audio signal processing algorithms. This description is used to compile C-code out of the model to evaluate on the target platform. The generated C-code is then parallelized using the standard OpenMP compiler in TI's Code Composer Studio (CCS) development environment.

The goal of this tutorial is to present the development of audio signal processing algorithms using Matlab, CCS, and TMS320C6678—the eight-core Digital Signal Processor (DSP) from Texas Instruments (TI).

Workshop 9 — Tuesday, May 7
09:00 – 11:30 — Auditorium Loyola

ACOUSTIC ENHANCEMENTS SYSTEMS —IMPLEMENTATIONS

Chair: **Ben Kok**, SCENA acoustic consultants, Uden, The Netherlands

Panelists: *Steve Barbar*
Peter Mapp
Thomas Sporer
Takayuki Watanabe
Wieslaw Woszczyk

Acoustic enhancement systems offer the possibility to change the acoustics of a venue by electronic means. The way how this is achieved varies by the working principle and philosophy of the system implemented. In this workshop various researchers, consultants, and suppliers active in the field of enhancement systems will discuss working principles and implementations.

This workshop is in close relation with the tutorial on acoustic enhancement systems; people not yet too familiar with the applications and working principles of these systems are recommended to attend the tutorial before attending the workshop.

Tuesday, May 7 — 09:00 — Sala Saba

Technical Committee Meeting on Archiving, Restoration, and Digital Libraries

Session P17 Tuesday, May 7
10:00 – 11:30 Foyer

POSTERS: MEASUREMENTS AND MODELING

10:00

P17-1 Estimation of Overdrive in Music Signals— *Lasse Vetter, Michael J. Terrell, Andrew J. R. Simpson, Andrew McPherson*, Queen Mary University of London, London, UK

In this paper we report experimental and modeling results from an investigation of listeners' ability to estimate overdrive in a signal. The term overdrive is used to characterize the result of systematic, level-dependent nonlinearity typical of audio equipment and processors (e.g., guitar amplifiers). Listeners (N=7) were given the task of estimating the degree of overdrive in music signals that had been processed with a static, saturating nonlinearity to introduce varying degrees of nonlinear distortion. A statistical model is proposed to account for the data, which is based on a measure of time-variance in the summed frequency-response deviation introduced by the nonlinearity. This provides a useful "black-box" metric that describes the perceived amount of overdrive introduced by an audio processing device.
Convention Paper 8911

10:00

P17-2 Mobile Audio Measurements Platform: Toward Audio Semantic Intelligence into Ubiquitous Computing Environments— *Lazaros Vrysis, Charalampos A. Dimoulas, George M. Kalliris, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The current paper presents the implementation of a mobile software environment that provides a suite of professional-grade audio and acoustic analysis tools for smartphones and tablets. The suite includes sound level monitoring, real-time time-frequency analysis, reverberation time, and impulse response measurements, whereas feature-based intelligent content analysis is deployed in terms of long-term audio events detection and segmentation. The paper investigates the implementation of a flexible and user-friendly environment, which can be easily used by non-specialists, providing professional functionality and fidelity of specific-purpose devices and eliminating the mobile-interfacing and hardware limitations. Emphasis is given to the integration of additional capabilities that will offer valuable amenities to the user, having to do with the management of measurement sessions and intelligent cloud-based semantic analysis.
Convention Paper 8912

10:00

P17-3 System Identification Based on Hammerstein Models Using Cubic Splines—*Michele Gasparini, Andrea Primavera, Laura Romoli, Stefania Cecchi, Francesco Piazza*, Universitá Politecnica della Marche, Ancona (AN), Italy

Nonlinear system modeling plays an important role in the field of digital audio systems whereas most of the real-world devices show a nonlinear behavior. Among nonlinear models, Hammerstein systems are particular nonlinear systems composed of a static nonlinearity cascaded with a linear filter. In this paper a novel approach for the estimation of the static nonlinearity is proposed based on the introduction of an adaptive CatmullRom cubic spline in order to overcome problems related to the adaptation of high-order polynomials necessary for identifying highly nonlinear systems. Experimental results confirm the effectiveness of the approach, making also comparisons with existing techniques of the state of the art.
Convention Paper 8913

10:00

P17-4 Impulse Responses Measured with MLS or Swept-Sine Signals: A Comparison between the Two Methods Applied to Noise Barrier Measurements —*Paolo Guidorzi, Massimo Garai*, University of Bologna, Bologna, Italy

A sound source and a microphone grid are used for measuring a set of impulse responses with the purpose of estimating the in-situ acoustical characteristics of noise barriers (sound reflection and airborne sound insulation) following the CEN/TS 1793-5 European standard guidelines as improved by the European project QUIESST. The impulse responses are measured using MLS (Maximum Length Sequence) and Swept-sine signals. The acoustical characteristics of the noise barrier, obtained using the two signals, are generally equivalent, but in some critical measurement conditions a discrepancy can be found. Differences and advantages between measurements, obtained by means of MLS or Swept-sine methods are analyzed and discussed in this paper.
Convention Paper 8914

10:00

P17-5 Polar Measurements of Harmonic and Multitone Distortion of Direct Radiating and Horn Loaded Transducers—*Mattia Cobianchi,*[1] *Fabrizio Mizzoni,*[2] *Aurelio Uncini*[2]
[1]Lavoce Italiana, Colonna (RM), Italy
[2]Sapienza University of Rome, Rome, Italy

While extensive literature is available on the topic of polar pattern measurements and predictions of loudspeakers' fundamental SPL, only a single paper to our knowledge deals with the polar pattern of nonlinear distortions, in particular with harmonic distortion products of cone type loudspeakers. This paper contains the first results of a more thorough study intended as a complement to fill the gap both in measurement techniques and loudspeaker type. Relative and absolute harmonic distortion as well as relative and absolute multitone distortion, indeed, have been measured for cone, dome, and horn loaded transducers.
Convention Paper 8915

10:00

P17-6 An Efficient Nonlinear Acoustic Echo Canceller for Low-Cost Audio Devices— *Danilo Comminiello,*[1] *Antonio Grosso,*[2] *Fabio*

Cagnetti,[2] *Aurelio Uncini*[1]
[1]Sapienza University of Rome, Rome, Italy
[2]bdSound, Milan, Italy

One of the most challenging problems to address in the modeling of acoustic channels is the presence of nonlinearities generated by loudspeakers. This problem has become even more widespread due to the growing availability of low-cost devices that introduce a larger amount of distortions and decrease the quality of hands-free speech communication. In order to reduce the effect of loudspeaker nonlinearities on the speech quality, nonlinear acoustic echo cancellers are adopted. In this paper we present a new adaptive filtering structure for the reduction of nonlinearities in the acoustic path, based on a nonlinear transformation of the input audio signal by means of functional links. We use such a nonlinear model in conjunction with a linear filter providing a nonlinear adaptive architecture for acoustic applications. Experimental results prove the effectiveness of the proposed model in reducing loudspeaker nonlinearities affecting speech signals.
Convention Paper 8916

10:00

P17-7 Radiation Pattern Differences between Electric Guitar Amplifiers—*Justin Mathew, Stephen Blackmore*, New York University, New York, NY, USA

Various aspects of electric guitar amplifiers can differentiate them from one another. Two of the major differentiating characteristics are frequency response and unique radiation patterns. These characteristics are directly related to differences in shape, size, and circuit configuration between different guitar amplifier models. In this paper the differences in radiation patterns of multiple guitar amplifiers will be presented as well a method of classifying the differences.
Convention Paper 8917

Tuesday, May 7 10:00 Sala Saba

Technical Committee Meeting on Audio Recording and Mastering Systems

Tuesday, May 7 10:00 Sala Montale

Standards Committee Meeting, AESSC Plenary

Professional Training Session 5 Tuesday, May 7
10:30 – 12:30 Sala Manzoni

AUDIO DESIGN WORKSHOP LIVE—PART 1

Presenters: **Peter Larsen**, LOUDSOFT
John Richards, Oxford Digital

10:30 – 11:30 Designing Transducers for Compact Active Speakers, Peter Larsen, LOUDSOFT

The LOUDSOFT FINE software suite for Loudspeaker Design and Test & Measurement is a circle of design tools, where each component supports the other. With these powerful FINE programs you can design all parts of the speaker system. The FINE R+D analysis/measurement system and the FINE QC control system supplement the design software. The Rub & Buzz function is the best on the market finding 100% failures of both active USB speakers and micro speakers. These tools are rapid and effective in modern transducer and speaker system development and the simulations very closely matches measurements of the devices when built.

11:30 – 12:30 Optimizing Compact Loudspeaker Performance—The Role of DSP, John Richards, Oxford Digital

The form factor of all consumer devices from cell phones to flat panel TVs is shrinking and often the design is under the control of a stylist rather than anyone who understands acoustics. This, together with pressure on the Bill of Materials cost and increasing customer expectations, creates real challenges in obtaining good, differentiating audio quality in the product. DSP can mitigate many of the inherent deficiencies present. The tutorial shows how a combination of an integrated suite of audio effects together with automated frequency correction can be used to radically improve sound quality whilst providing rapid time to market.

Session EB3 Tuesday, May 7
11:15 – 13:00 Sala Foscolo

ENGINEERING BRIEFS—PAPERS: PART 2

Chair: **Brett Leonard**, McGill University, Montreal, Quebec, Canada

11:15

EB3-1 Spatial Sound Reinforcement Using Wave Field Synthesis—*Etienne Corteel*,[1] *Hubert Westkemper*,[2] *Cornelius Ihssen*,[3] *Khoa-Van Nguyen*[1]
[1]Sonic Emotion Labs, Paris, France
[2]Independent Tonmeister, Naples, Italy
[3]Sonic Emotion Labs, Oberglatt, Switzerland

Spatial audio in sound reinforcement remains an open topic, requiring good level coverage and at the same time good localization accuracy over a very large listening area, typically the entire audience. Wave Field Synthesis offers high localization accuracy over an extended listening area but the number of required loudspeakers, their placement on stage, and the level coverage that results from it can be problematic. The paper addresses these issues, presenting a case study of a sound reinforcement system based on Wave Field Synthesis for sound reinforcement for the play "The Panic" written by Rafael Spregelburd and directed by Luca Ronconi. The paper describes an improved Wave Field Synthesis rendering for sound reinforcement involving two arrays of loudspeakers at different heights. The paper addresses the practical implementation of the system in a theater and the overall installation: miking, real time tracking of actors, and loudspeakers used.
Engineering Brief 91

11:30

EB3-2 Sync-AV–Workflow Tool for File-Based Video Shootings—*Andreas Fitza*, University of Applied Science Mainz, Mainz, Germany

The Sync-AV workflow tool eases the sorting

and synchronization of video and audio footage without the need of expensive special hardware. It supports the preproduction, shooting, and postproduction. It consists of three elements: a script-information and metadata-gathering iOS app that is synchronized with a server-back-end and that can be used to exchange information on-set; a server database with a web-front-end that can sort files by their metadata and show dailies and that can be used to distribute and manage information during the pre-production; and a local import client that manages the footage ingest and sorts the files together. The client also takes care of the synchronization of the video that contains audio and separately recorded audio files and it renames the files and implements the metadata.
Engineering Brief 92

11:45

EB3-3 On the Optimum Microphone Array Configuration for Height Channels—*Hyunkook Lee, Christopher Gribben*, University of Huddersfield, Huddersfield, West Yorkshire, UK

To date no experimental data have been presented on the optimum microphone array configuration for new surround formats employing height channels. A series of subjective listening tests were conducted to investigate how the spacing between base and height microphones affects perceived spatial impression and overall preference. Four different spacings of 0, 0.5, 1, and 1.5 m were compared for various sound sources using a 9-channel loudspeaker setup. For sources with more continuous temporal characteristics, the spacing between the layers did not have any significant effect on spatial impression, whereas for more transient sources the 0 m layer appeared to produce a greater spatial impression than more spaced layers. Furthermore, the 0 m layer was more or similarly preferred to the spaced layers depending on source type.
Engineering Brief 93

12:00

EB3-4 Free Improv—The Hard Way—*Justin Paterson*, London College of Music, University of West London, London, UK

As a fringe genre, "Free Improvisation" does not normally attract large production budgets. Often time-constrained, the subsequent technological approach to the production tends to emphasize the naturalistic and neglects many of the tools and techniques that are commonplace in contemporary popular music. The author produced the album, *The Making of Quiet Things* by The Number (featuring Keith Tippett). This album consciously employed a range of contemporary approaches such as creative and corrective automation, reverberation-matching, audio editing, and extreme compression, while maintaining an overall impression of minimal mediation. This paper considers and contextualizes such an approach, reflecting on the practice and its implications for the genre.
Engineering Brief 94

12:15

EB3-5 International Experiences from a New Sound System Approach—*Thomas Lagö, Alan Boyer*, QirraSound Technologies LLC, Las Vegas, NV, USA

QirraSound's new sound system approach benefits from a high level of intelligibility and substantially lower feedback. These properties help in placing loudspeakers behind the performers and thus minimizing the need for monitor speakers. Substantial empiric testing has been done in applications in multiple countries and applications and results from these tests will be presented. Listeners and performers report increased feel and intelligibility and even people with hearing loss and/or sensitivity to high sound levels can enjoy the sound. It has also been noticed that the ability to talk while music is playing is much better than with classical systems. An overall outline of these test results and feedback will be reported.
Engineering Brief 95

12:30

EB3-6 Nonlinear Guitar Loudspeaker Simulation—*Thomas Schmitz, Jean J. Embrechts*, University of Liege, Liege, Belgium

In this study we simulated in real time the sound of a guitar amplifier loudspeaker, including its non-linear behavior. The simulation method is based on a non-linear convolution of the signal emitted by the instrument with the Volterra kernels, which were measured in anechoic conditions with a sine-sweep technique. The model has been implemented in a "VST" (Virtual Studio Technology) audio plugin. The loudspeaker simulation can be performed in real time with the Volterra kernels up to the third order and offers a good accuracy. Informal tests revealed that the simulated and the real sound were very close, although approximately 50 percent of the tested musicians were still able to hear a small difference.
Engineering Brief 96

12:45

EB3-7 Using Low-Latency Net-Based Solutions to Extend the Audio and Video Capabilities of a Studio Complex—*Paul Ferguson*, Edinburgh Napier University, Edinburgh, UK

Two low-latency IP-based systems, RedNet and LOLA, were selected by Edinburgh Napier University to link a new music building with existing broadcast and drama facilities and to allow low-latency audio and video collaboration with other institutions/organizations around the world. First to be examined will be their use of Focusrite RedNet and Audinate Dante to expand existing AES10 (MADI) point to point links. Second, an overview will be provided of the University's ongoing research with JANET and GARR (the UK and Italy National Research and Education Networks) into the use of the Italian LOLA system (LOw LAtency audio visual streaming system) to provide long-distance audio and video links for rehearsal and performance involving musicians in different countries.
Engineering Brief 97

➥

Tutorial 16 **Tuesday, May 7**
11:45 – 13:15 **Auditorium Loyola**

MICROPHONES: THE PHYSICS, METAPHYSICS, AND PHILOSOPHY

Presenter: **Ron Streicher**, Pacific Audio-Visual Enterprises, Pasadena, CA, USA

Before you can place the first microphone in the studio, you need to develop a clear understanding of the sound that you want to emanate from the loudspeakers when the project is finished. To do this, you need to determine what are the elements that are essential for creating the "sonic illusion," and then decide how to balance the often conflicting elements and competing demands of technology vs. art. Microphone techniques—although critical—are only a part of this process. Equally important are the criteria for monitoring and evaluating the results. Using recorded examples and practical demonstrations, the various aspects of this creative process are developed and brought into focus.

Professional Training Session 6 **Tuesday, May 7**
13:30 – 15:30 **Sala Manzoni**

AUDIO DESIGN WORKSHOP LIVE—PART 2

Presenters: **Anthony Waldron**, Audio EMC
Simon Wollard, Prism Sound

13:30 – 14:30 *Audio Power Amplifiers; EMC Best Practice Revealed*, Anthony Waldron, Audio EMC

The performance of analog and (now) digital audio systems is approaching an all time high—yet we still suffer bad bouts of noise, distortion, and interference on a regular basis. The physical reasons for poor audio performance is provided, along with practical solutions for designing audio amplifiers and systems to maximize signal integrity and be immune to noise and interference.

14:30 - 15:30 *Audio System Analysis—Tips and Tricks to Verify Your Designs*, Simon Wollard, Prism Sound

Active loudspeakers are becoming ever more complex in terms of their underlying architecture. The number of input interfaces has grown significantly in recent years; sophisticated signal processors have evolved to greatly enhance performance; amplifiers have grown more capable in terms of self-protection, output power and efficiency; transducers are ever-improving in terms of smoother response and lower distortion. This practical discussion will teach the audience how to take advantage of modern digital audio analysis techniques to provide far greater and more rapid insight into the performance of active loudspeakers than was previously possible with classical audio measurement techniques.

Session P18 **Tuesday, May 7**
14:30 – 16:30 **Sala Carducci**

SPATIAL AUDIO—PART 3: AMBISONICS, WFS

Chair: **Symeon Delikaris-Manias**, Aalto University, Helsinki, Finland

14:30

P18-1 An Ambisonics Decoder for Irregular 3-D Loudspeaker Arrays—*Daniel Arteaga*, Fundacio Barcelona Media, Barcelona, Spain; Universitat Pompeu Fabra, Barcelona, Spain

We report on the practical implementation of an Ambisonics decoder for irregular 3-D loudspeaker layouts. The developed decoder, which uses a non-linear search algorithm to look for the optimal Ambisonics coefficients for each loudspeaker, has a number of features specially tailored for reproduction in real-world 3-D audio venues (for example, special 3-D audio installations, concert halls, audiovisual installations in museums, etc.). In particular, it performs well even for highly irregular speaker arrays, giving an acceptable listening experience over a large audience area.
Convention Paper 8918

15:00

P18-2 The Effect of Low Frequency Reflections on Stereo Images—*Jamie A. S. Angus*, University of Salford, Salford, Greater Manchester, UK

This paper looks at the amount of absorption required to adequately control early reflections in reflection-controlled environments at low frequencies (< 700 Hz). This is where the Inter-aural Time Delay Cue (ITD) is important. Most work has focused on wideband energy time performance. In particular, it will derive the effect a given angle and strength of reflection has on the phantom image location using the Blumlein equations. These allow the effect of reflections as a function of frequency to be quantified. It will show that the effect of reflections are comparatively small for floor and ceiling reflections, but that lateral reflections depend on the phantom image location and get worse the more off-center the phantom image becomes.
Convention Paper 8919
[This paper was not presented but is available for purchase]

15:30

P18-3 Parametric Spatial Audio Reproduction with Higher-Order B-Format Microphone Input—
Ville Pulkki,[1] *Archontis Politis*,[1] *Giovanni Del Galdo*,[2] *Achim Kuntz*[3]
[1]Aalto University, Aalto, Finland
[2]Ilmenau University of Technology, Ilmenau, Germany
[3]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

A time-frequency-domain non-linear parametric method for spatial audio processing is presented here, which can utilize microphone input having directional patterns of any order. The method is based on dividing the sound field into overlapping or non-overlapping sectors. Local pressure and velocity signals are measured within each sector, and an individual Directional Audio Coding (DirAC) processing is performed for each sector. It is shown, that in certain acoustically complex conditions the sector-based processing enhances the quality compared to traditional first-order DirAC processing.
Convention Paper 8920

16:00

P18-4 Wave Field Synthesis of Virtual Sound Sources with Axisymmetric Radiation Pattern Using a Planar Loudspeaker Array—*Filippo*

Maria Fazi,[1] *Ferdinando Olivieri*,[1] *Thibaut Carpentier*,[2] *Markus Noisternig*[2]
[1]University of Southampton, Southampton, Hampshire, UK
[2]UMR STMS IRCAM-CNRS-UPMC, Paris, France

A number of methods have been proposed for the application of Wave Field Synthesis to the reproduction of sound fields generated by point sources that exhibit a directional radiation pattern. However, a straightforward implementation of these solutions involves a large number of real-time operations that may lead to very high computational load. This paper proposes a simplified method to synthesize virtual sources with axisymmetric radiation patterns using a planar loudspeaker array. The proposed simplification relies on the symmetry of the virtual source radiation pattern and on the far-field approximation, although a near-field formula is also derived. The mathematical derivation of the method is presented and numerical simulations validate the theoretical results.
Convention Paper 8921

Session EB4 — Tuesday, May 7
14:30 – 16:00 — Foyer

ENGINEERING BRIEFS—POSTERS: PART 2

14:30

EB4-1 Spatial Acoustic Synthesis—*Celambarasan Ramasamy*, Mind Theatre, Pasadena, CA, USA

This method leverages binaural sound synthesis to present a novel way for listeners to experience musical performances on a headphone. By separating out the notes from a musical instrument and synthesizing them individually in the space around the listener, the musical instrument is turned from a point source into a flowing musical volume that engulfs the listener. This can lead to interesting ways of thinking about a musical performance in terms of the distribution of the individual musical notes around the listener without being tied down to the notion of an individual musical instrument.
Engineering Brief 98

14:30

EB4-2 Enhancing the Learning of Stereo Microphone Techniques through the Use of a Simulated Learning Environment—*Colin Dodds*, Perth College, University of the Highlands and Islands, Perth, Scotland, UK

A key set of skills for aspiring recording engineers to acquire is that of good stereo microphone techniques. Within the world of education it is relatively straightforward to present the underpinning theory, but helping students gain the tacit knowledge necessary to achieve quality results can be difficult when access to suitable groups of musicians, spaces, and equipment is limited. A suitable learning environment was simulated within a computer program and tasks set to support the learning of stereo microphone techniques. A trial carried out on first year undergraduate sound production students revealed that using the simulated learning environment enhanced both the students' knowledge of and ability to apply stereo microphone techniques.
Engineering Brief 99

14:30

EB4-3 The Good Vibrations Problem—*Derry FitzGerald*, Dublin Institute of Technology, Dublin, Ireland

Many of the Beach Boys' records were mono only as this was Brian Wilson's preferred format. However, starting in the mid-1990s, stereo mixes of many of these classics were created by synchronizing the tracks from the instrumental multitrack with those of the vocal multitrack. Unfortunately, for a number of tracks, including "Good Vibrations," elements of the multitracks were missing, making a true stereo mix impossible. This paper deals with how stereo extraction mixes were created for a number of Beach Boys' songs using sound source separation techniques to separate sources from the original mono recordings, which were then panned to create stereo mixes. These mixes were used in reissues of Beach Boys albums in 2012.
Engineering Brief 100

14:30

EB4-4 Architectural Acoustics and Electroacoustics in the Asisium Theatre: An Integrated Construction Work—*Luca Quaranta*, CP Progetti S.r.l., Rome, Italy

Based on the multi-functional needs, from the conference room to the cinema with Dolby Surround sound effects, we performed an analysis of the room's acoustic response based on the model of a 3-D simulation, referenced on acoustic parameters according to ISO-3382, thus directing the choices of architectural design to obtain optimum acoustical conditions both on stage and in the audience. Design of the electroacoustic provides various control technologies for the equalization and a 6.1 Dolby Digital distribution of the amplified audio signal. In this contribution we present choices and design criteria, materials and audio devices used to obtain the final result, to underline the key role of architectural acoustics and electroacoustic integrated design.
Engineering Brief 101

14:30

EB4-5 Expressive Physical Modeling of Keyboard Instruments: From Theory to Implementation —*Stefano Zambon*,[1] *Leonardo Gabrielli*,[2] *Balázs Bank*[3]
[1]Viscount International S.p.A., Mondaino (RN), Italy
[2]Universitá Politecnica delle Marche, Ancona, Italy
[3]Budapest University of Technology and Economics, Budapest, Hungary

Physics-based algorithms for sound synthesis have been extensively studied in the past decades. Nevertheless, their use in commercial synthesizers is still limited due to the difficulty in achieving realistic and easily controllable sounds with current technology. In this Engineering Brief ➥

we present an overview of the models used in Physis Piano, a digital piano recently introduced in the market with dedicated physics-based algorithms for acoustic pianos, electric pianos (e.g., Rhodes, Wurlitzer, and Clavinet), and chromatic percussions (e.g. vibraphone, marimba, xylophone). The synthesis algorithms, which are based on standard techniques such as Modal Synthesis and Digital Waveguides, have been highly customized in order to faithfully reproduce the sound features of the original instruments and are easily controllable by a set of meaningful, user-friendly parameters.
Engineering Brief 102

Workshop 10 Tuesday, May 7
14:30 – 16:30 Sala Foscolo

MIKING FOR PA

Chair: **Eddy B. Brixen**, EBB-consult, Smørum, Denmark

Panelists: *Giacome De Caterini*, Casal Bauer, Italy
Igor Fiorini, VDM Group, Italy
Cathrine Sadolin, Complete Vocal Institute, Copenhagen, Denmark
Henrik Sadolin, Complete Vocal Institute, Copenhagen, Denmark

Miking for PA is a very important task. Providing amplification to the spoken or singing voice or the acoustical music instrument requires good knowledge about the sound source, about the PA-system, about the monitoring system—and about the microphones. This workshop takes you through some of the important issues and decisions when selecting the microphone with regards to peak level capacity, sensitivity, directivity, frequency response, sensitivity, etc. Getting balance, getting definition, getting the right timbre or "sound"—and still avoiding acoustical feedback, that's the thing. Recognized engineers and sound designers will generously share their experiences from their work on the stages. Warning: Some of the attendees may pick up ideas that will change their habits forever…

Student/Career Event
STUDENT DELEGATE ASSEMBLY MEETING —PART 2
Tuesday, May 7, 14:30 – 16:00
Auditorium Loyola

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.