AES 138th Convention Program

May 7 – 10, 2015

Sofitel Victoria Hotel, Warsaw, Poland

At 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 136th Convention Peer-Reviewed Paper Award.

* * * * *

The Winner of the 138th AES Convention Best Peer-Reviewed Paper Award is: Discussion of the Wavefront Sculpture Technology Criteria for Straight Line Arrays—
Frank Schultz, University of Rostock, Rostock, Germany; *Florian Straube*, Technical University Berlin, Berlin, Germany; *Sascha Spors*, University of Rostock, Rostock, Germany
Convention Paper 9323

To be presented on Sunday, May 10, in Session 15—Spatial Audio—Part 2

* * * * *

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 138th 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 138th AES Convention Student Paper Award is: Auditory Adaptationin Spatial Listening Tasks
—Florian Klein, Stephan Werner, Technical University Ilmenau, Ilmenau, Germany
Convention Paper 9281

To be presented on Friday, May 8, in Session 9 —Perception—Part 1

* * * * *

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

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

Thursday, May 7 09:00 Room 609B

Technical Committee Meeting on Acoustics and Sound Reinforcement

Session P1 Thursday, May 7
10:00 – 12:30 Room Belweder

SPATIAL AUDIO—PART 1

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

10:00

P1-1 Subjective Loudness of 22.2 Multichannel Programs—
Tomoyasu Komori,[1,2] *Satoshi Oode,*[1] *Kazuho Ono,*[1] *Kensuke Irie,*[1] *Yo Sasaki,*[1] *Tomomi Hasegawa,*[1] *Ikuko Sawaya*[1]
[1]NHK Science and Technology Research Laboratories, Setagaya-ku, Tokyo, Japan
[2]Waseda University, Shinjuku-ku, Tokyo, Japan

NHK is planning 8K Super Hi-Vision (SHV) broadcasting with 22.2 multichannel sound as a new broadcasting service. The current loudness measurement algorithm, however, are only standardized up to 5.1 channels in Recommendation ITU-R BS.1770. To extend the algorithm beyond 5.1 ch, we conducted a subjective loudness evaluation of various program materials and formats. The results showed that different formats differed only slightly. Furthermore, we measured objective loudness values on the basis of an algorithm compatible with the current algorithm and found that the objective loudness values had a good correlation with the subjective loudness values.
Convention Paper 9219

10:30

P1-2 MPEG-D Spatial Audio Object Coding for Dialogue Enhancement (SAOC-DE)—*Jouni Paulus,*[1] *Jürgen Herre,*[1,2] *Adrian Murtaza,*[2] *Leon Terentiv,*[2] *Harald Fuchs,*[2] *Sascha Disch,*[1] *Falko Ridderbusch*[2]
[1]International Audio Laboratories Erlangen, Erlangen, Germany
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

The topic of Dialogue Enhancement and personalization of

audio has recently received increased attention. Both hearing-impaired and normal-hearing audience benefit, for example, from the possibility of boosting the commentator speech to minimize listening effort, or to attenuate the speech in favor of sports stadium atmosphere in order to enhance the feeling of being there. In late 2014, the ISO/MPEG standardization group made available a new specification, Spatial Audio Object Coding for Dialogue Enhancement (SAOC-DE), which was closely derived from the well-known MPEG-D Spatial Audio Object Coding (SAOC). This paper describes the architecture and features of the new system. The envisioned applications will be discussed and the performance of the new technology is demonstrated in subjective listening tests.
Convention Paper 9220

11:00

P1-3 Multichannel Systems: Listeners Choose Separate Reproduction of Direct and Reflected Sounds—
Piotr Kleczkowski, Aleksandra Król, Pawel Malecki, AGH University of Science and Technology, Krakow, Poland

Arguments can be put forward for the separation of direct and reflected components of the sound field and reproducing them through appropriate transducers, but there is no definite opinion about that. In this work the perceptual effect of separation in commonly used 5.0 and 7.0 multichannel systems was investigated. Four listening experiments were performed involving several schemes of separation and a variety of experimental conditions. The listeners consistently preferred some schemes involving separation to schemes without separation.
Convention Paper 9221

11:30

P1-4 On the Influence of Headphone Quality in the Spatial Immersion Produced by Binaural Recordings—*Pablo Gutierrez-Parera, Jose J. Lopez, Emanuel Aguilera*,
Universidad Politecnica de Valencia, Valencia, Spain

The binaural recordings obtained using an acoustic manikin produce a realistic sound immersion played through high quality headphones. However, most people commonly use headphones of inferior quality as the ones provided with smartphones or music players. Factors such as frequency response, distortion, and the disparity between the left-right transducers could be some of the degrading factors. This work lays the foundation for a strategy for studying what are the factors that affect the end result and what level do. A first experiment focuses on the analysis of how the disparity in levels between the two transducers affects the final result. A second test studies the influence of the frequency response. A third test analyzes the effects of distortion using a Volterra kernels scheme for the simulation of the distortion using convolutions. The results of this work reveal how disparity between both transducers can affect the perception of direction. In the case of frequency response the results are more difficult to quantify and further work will be necessary. Finally the study reveals that the distortion produced by the range of headphones tested does not affect to the perception of binaural sound.
Convention Paper 9222

12:00

P1-5 Binaural Audio with Relative and Pseudo Head Tracking
—Christof Faller,[1] *Fritz Menzer,*[2] *Christophe Tournery*[1]
[1]Illusonic GmbH, Zurich, Switzerland
[2]Technische Universität München, Munich, Germany

While it has been known for years that head-tracking can significantly improve binaural rendering, it has not been widely used in consumer applications. The goal of the proposed techniques is to leverage head tracking, while making it more usable for mobile applications, where the sound image shall not have an absolute position in space. Relative head tracking keeps the sound image in front, while reducing the effect of head movements to only small fluctuations. Relative head tracking can be implemented with only a gyrometer; there is no need for absolute direction. An even more economical technique with the goal to improve binaural rendering is pseudo head tracking. It generates small head movements using a random process without resorting to a gyroscope. The results of a subjective test indicate that both relative and pseudo head tracking can contribute to spaciousness and front/back differentiation.
Convention Paper 9223

Thursday, May 7 10:00 Room 609B

Technical Committee Meeting on High Resolution Audio

Session P2 Thursday, May 7
10:30 – 12:30 Foyer

POSTERS: EDUCATION AND PERCEPTION

10:30

P2-1 Effects of Ear Training on Education on Sound Quality of Digital Audio for Non-Technical Undergraduates—
Akira Nishimura, Tokyo Univeristy Information Sciences, Chiba-shi, Japan

This paper demonstrates the effectiveness of ear training in lectures on audio processing conducted over the 2013 and 2014 academic terms. Student understanding of the lecture content was assessed by comparing scores of written tests that covered the sound quality of perceptual audio codecs and other topics, which were administered after lectures with and without ear training on identifying bit rates of sound files. The same approach was applied to a lecture on audio digitizing and ear training on identifying sampling frequencies. The test scores of assessments that focused on the sound quality of perceptual audio codecs were significantly higher among students who had participated in ear training compared to those who had not participated in such training. In contrast, no significant difference was found in the group scores of participants tested after ear training on identifying sampling frequency. The effectiveness of ear training being limited to perceptual codings was investigated in terms of prior knowledge of the technical terms.
Convention Paper 9224

10:30

P2-2 Evaluation of the Low-Delay Coding of Applause and Hand-Clapping Sounds Caused by Music Appreciation—
Kazuhiko Kawahara,[1] *Yutaka Kamamoto,*[2] *Akira Omoto,*[1] *Takehiro Moriya*[2]
[1]Kyushu University, Fukuoka, Japan
[2]NTT Communication Science Laboratories, Kanagawa, Japan

Recently, the improvement of network resources enables us to distribute the contents in real-time. This paper presents the low-delay coding of applause sound and hand-clapping sound with less parameters by means of synthesizing these sounds at the receiver site. We found that number of people clapping their hands were corresponding to a sound volume of applause. In other words, no one considers who is clapping. Additionally, on the hand-clapping

sound, the time interval of clapping also should be important. Based on such information, preliminary experiments confirm that our approach, which synthesize applause and hand-clapping sound from a few parameters, successfully generates natural applause and hand-clapping sounds.
Convention Paper 9225

10:30

P2-3 Subjective Evaluation of High Resolution Audio under In-Car Listening Environments—*Mitsunori Mizumachi,*[1] *Ryuta Yamamoto,*[2] *Katsuyuki Niyada*[3]

[1]Kyushu Institute of Technology, Kitakyushu, Fukuoka, Japan
[2]Digifusion Japan Co., Ltd., Hiroshima, Japan
[3]Hiroshima Cosmopolitan University, Hiroshima, Japan

High resolution audio (HRA) becomes increasingly popular both for music production and the consumers. It enables to record a music performance in a wide-band and precise digital audio format. It is, however, unclear in its perceptual advantage under some listening environments. In this study listening tests were carried out inside cars where 34 participants listened to the same music in four different audio formats. The participants chose an audio format with better quality in paired comparison among 192 kHz/24 bits PCM, 48 kHz/16 bits PCM, and two kinds of lossy-compressed MPEG audio formats. The participants, who are familiar with HRA and live music performance, could significantly discriminate among the audio formats.
Convention Paper 9226

10:30

P2-4 Investigating Factors that Guitar Players to Perceive Depending on Amount of Distortion in Timbre—*Koji Tsumoto, Atsushi Marui, Toru Kamekawa*, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

Typical electric guitar timbre could be classified into three classes according to amount of distortion. Timbre with less distortion is called "Clean" and heavily distorted timbre is called "Distorted." Timbre between "Clean" and "Distorted" is called "Crunch." To investigate the factors that guitar players perceive depending on amount of distortion, semantic differential analysis using eight bipolar adjective scales was employed. Twenty guitar players including six professionals played their instruments through a guitar amp with nine different distortion level settings. Two factors were found in factor analysis, and "Clean" and "Distorted" were located opposite to each other. "Crunch" was located in the middle of latent factors and each anchoring adjectives used in the evaluation. Also the result of regression analysis indicated "Activeness Factor" was the reliable factor corresponding to the amount of distortion.
Convention Paper 9227

10:30

P2-5 Perception of Timbre Changes vs. Temporary Threshold Shift—*Bartlomiej Kruk, Maurycy Kin*, Wroclaw University of Technology, Wroclaw, Poland

This paper presents results of research on an influence of Temporary Threshold Shift (TTS) on the detection of changes in timbre of musical samples. The experiment was carried out with conditions that normally exist in a studio when sound material is recorded and mixed. The level of sound exposure that represents the noise signal is 90 dB, and this is an average value of sound level existing in control room. This musical material may be treated as a noise so TTS phenomenon may occur after several time durations: 60, 90, and 120 minutes. Ten subjects participated in the main part of the experiment and all of them have the normal hearing thresholds. The stimuli contained the musical material with introduced changes in timbre up to +/–6 dB in low (100 Hz), middle (1 kHz), and high frequency (10 kHz) regions. It turned out that listening to the music with an exposure of 90 dB for 1 hour influences the hearing thresholds for middle frequency region (about 1–2 kHz); and this has been reflected in a perception of timbre changes: after 1 hour listening the changes of spectrum in middle-frequencies region are perceived with a threshold of 3 dB while the changes of low and high ranges of spectrum were perceived with the thresholds of 1.8 and 1.5 dB, respectively. After the longer exposure, the thresholds shifted up to 3.5 dB for the all investigated stimuli.
Convention Paper 9228

10:30

P2-6 Hybrid Multiresolution Analysis of "Punch" in Musical Signals—*Steven Fenton, Hyunkook Lee, Jonathan Wakefield*, University of Huddersfield, Huddersfield, UK

This paper presents a hybrid multi-resolution technique for the extraction and measurement of attributes contained within a musical signal. Decomposing music into simpler percussive, harmonic, and noise components is useful when detailed extraction of signal attributes is required. The key parameter of interest in this paper is that of punch. A methodology is explored that decomposes the musical signal using a critically sampled constant-Q filterbank of quadrature mirror filters (QMF) before adaptive windowed short term Fourier transforms (STFT). The proposed hybrid method offers accuracy in both the time and frequency domains. Following the decomposition transform process, attributes are analyzed. It is shown that analysis of these components may yield parameters that would be of use in both mixing/mastering and also audio transcription and retrieval.
Convention Paper 9229

10:30

P2-7 Five Aspects of Maximizing Objectivity from Perceptual Evaluations of Loudspeakers: A Literature Study—*Christer Volk,*[1,2] *Søren Bech,*[2,3] *Torben H. Pedersen,*[1] *Flemming Christensen*[2]

[1]DELTA SenseLab, Hørsholm, Denmark
[2]Aalborg University, Aalborg East, Denmark
[3]Bang & Olufsen a/s, Struer, Denmark

A literature study was conducted focusing on maximizing objectivity of results from listening evaluations aimed at establishing the relationship between physical and perceptual measurements of loudspeakers. The purpose of this study was to identify and examine factors influencing the objectivity of data from the listening evaluations. This paper addresses the following subset of aspects for increasing the objectivity of data from listening tests: The choice of perceptual attributes, relevance of perceptual attributes, choice of loudness equalization strategy, optimum listening room specifications, as well as loudspeaker listening in-situ vs. listening to recordings of loudspeakers over headphones.
Convention Paper 9230

10:30

P2-8 Modding Game Audio for Education—*Ricardo Bragança*, United Arab Emirates University, Al Ain, Abu Dhabi, UAE

Worldwide there is no formal curriculum for game audio. This paper will dwell on what can be done to change the current status quo. We intend to shed some light on possible solutions and guidelines that can be used by ➥

schools in order to achieve a higher awareness on how to implement game audio successfully in a university's curriculum. We believe that due to its interdisciplinary nature, cross faculty cooperation and corporate partnerships are advised and will promote a better understanding on how to tackle the topic. Constructivist teaching methods and a student centric inquiry based learning approach is suggested to enhance the learning experience and insure adequate content absorption.
Convention Paper 9231

10:30

P2-9 The Acoustic Properties of Different Types of Earplug Used by Sound Engineers—*Bartlomiej Kruk, Michal Luczynski*, Wroclaw University of Technology, Wroclaw, Poland

The main aim of this paper is to test various types of earplugs used by sound engineers. At live events, when sound engineers need to use earplugs for health reasons, it is very important that they maintain correct hearing perception abilities. The linear frequency response allows engineers to avoid mistakes when working with sound. Earplugs were tested for attenuation depending on frequency. The authors tested earplugs in the different methods: subjectively using pure tone audiometry and objectively using the designed and created ear canal model. Research allowed to choose the appropriate earplugs for sound engineering purposes.
Convention Paper 9233

10:30

P2-10 Psychoacoustic Annoyance Monitoring with WASN for Assessment in Urban Areas—*Jaume Segura,*[1] *Santiago Felici,*[1] *Maximo Cobos,*[1] *Ana Torres,*[2] *Juan M. Navarro*[3]
[1]Universitat de Valencia, Burjassot, Spain
[2]Polytechic University School of Cuenca, Cuenca, Spain
[3]Universidad Católica San Antonio, Murcia, Guadalupe (Murcia), Spain

The assessment of the subjective annoyance caused by noise pollution in cities is a matter of major importance as its influence is growing-up in urban areas. Different methods and techniques have been used to model this annoyance in terms of several psychoacoustic parameters, which define different aspects of the acoustic affection from noise pollution in the human behavior. In this paper we describe a monitoring system based on a wireless acoustic sensor network that measures and computes the psychoacoustic metrics following the Zwicker's annoyance model, in a distributed way and at different points simultaneously in urban areas. The nodes of this network run complex algorithms to find out these metrics. These nodes are Single-Board Computer platforms, in particular Raspberry Pi.
Convention Paper 9234

10:30

P2-11 The Advanced Sound System Listening Room at Dolby—*Sunil G. Bharitkar*, Dolby Laboratories, San Francisco, CA, USA

A listening room at Dolby has been designed to test the spatial and timbre performance of next generation audio formats recommended in the new ITU-R BS.2051-0 (Advanced sound system for program production). The room has been best designed to conform to the new ITU-R BS.1116-2 (Methods for the subjective assessment of small impairments in audio systems) specification for testing the performance of next-generation audio codecs. Detailed physical and acoustical measurements have been conducted using international standards that demonstrate satisfying elements in both these international recommendations and that are presented in the paper. Subjective testing is ongoing and some preliminary feedback is included as well.
Convention Paper 9337

Workshop 1 **Thursday, May 7**
11:00 – 12:45 **Room Opera**

**AUTOMOTIVE SOUND—
THE MAKING OF THE SOUND SYSTEM IN A CAR**

Chair: **Grzegorz Sikora**, Bang & Olufsen Deutschland GmbH, Pullach, Germany

Panelists: *Thomas Bachmann*, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
Edwin Cappleman, Bang & Olufsen
Peter John Chapman, Bang & Olufsen a/s, Struer, Denmark

In this workshop we look behind the scenes of developing an OEM sound system, "the making of." It's a unique opportunity to understand the key fundamentals of the Automotive Audio perspective, from the project team behind many successful OEM sound systems. This workshop will cover a range of topics, from the initial idea to the finished product. The experts from the respective fields will discuss car cabin acoustics, loudspeaker selection and placement, mechanics, signal flow and amplifier characteristics, sound tuning, audio design philosophy and creation of the 3D sound algorithm. Apart from the engineering part, project planning and communication between OEM and supplier will also be explained. Each topic will be discussed in general and in the context of actual projects.

Key Technology Briefings 1 **Gerriets**
Thursday, May 7 **11:00 – 11:45**
PSE Stage

**FLEXIBLE ACOUSTICS IN MULTI-PURPOSE VENUES
(BY THE USE OF FABRICS AND MEMBRANE ABSORBERS)**

Presenter: **Jonas Schira**, Gerriets

The acoustics requirements for theaters, opera houses, multipurpose halls, or community centers have changed dramatically over the recent years. A few years ago it was still acceptable to only reduce the overall reverberation time to meet the performance requirements, but today the acoustics of a venue need a more specific scope of requirements unique to the space. Using fabrics and inflatable membrane absorbers can be an efficient solution for flexible and adjustable room acoustics in all frequency bands.

Thursday, May 7 **11:00** **Room 609B**

Technical Committee Meeting on Archiving, Restoration and Digital Libraries

Tutorial 1 **Thursday, May 7**
11:15 – 12:45 **Room Królewski**

MUSIC INFORMATION RESEARCH FROM A CULTURAL PERSPECTIVE

Presenter: **Xavier Serra**, Universitat Pompeu Fabra, Barcelona, Spain

A major challenge in Music Information Research (MIR) is the automatic generation of musically meaningful information with which to better describe and exploit audio music recordings. But a piece of music makes sense especially within a particular social and cultural context and its analysis and description has to take that into account.

In this tutorial I will introduce the research being done in the project CompMusic (http://compmusic.upf.edu) related to the extraction of musically relevant features from audio music recordings related to melody and rhythm, and on the semantic analysis of the contextual information of those recordings, by studying five world music traditions that have a quite distinct personality: Hindustani (North India), Carnatic (South India), Turkish-makam (Turkey), Arab-Andalusian (Maghreb), and Beijing Opera (China).

Key Technology Briefings 2 — **Prism Sound**
Thursday, May 7 — **12:00 – 12:45**
PSE Stage

SOUND QUALITY IN AUDIO INTERFACES – POPULAR MYTHS AND SOME PRACTICAL ADVICE

Presenter: **Graham Boswell**, Prism Sound

[abstract not available]

Thursday, May 7 — **12:00** — **Room 609B**

Technical Committee Meeting on Fiber Optics for Audio

Special Event
AWARDS PRESENTATION AND KEYNOTE ADDRESS
Thursday, May 7, 13:00 – 14:00
Room Balowa AB

Opening Remarks:
- Executive Director Bob Moses
- President Andres Mayo
- Convention Chairs Bozena Kostek & Umberto Zanghieri

Program:
- AES Awards Presentation by Frank Wells, Awards Chair
- Introduction of Keynote Speaker by Convention Chair
- Keynote Address by Florian Camerer

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

BOARD OF GOVERNORS AWARD
- Lauri Savioja
- Sascha Spors
- Umberto Zanghieri

Keynote Speaker
This year's Keynote Speaker is **Florian Camerer.** Camerer joined the Austrian Broadcasting Corporation (ORF) in 1990. In 1995 he became a staff-sound-engineer ("Tonmeister") mainly in the field of production sound and post-production. Already in 1993 Camerer started to get interested in multichannel audio. He mixed the first program of the ORF in Dolby Surround ("Arctic Northeast") and is since then involved in all aspects of multichannel audio at ORF. Since autumn 2008 Camerer chairs the EBU group PLOUD, successfully introducing loudness normalization instead of peak levelling in Europe. He is lecturing on an international basis especially in dramaturgical aspects of surround sound productions, microphones for surround sound, multichannel audio for HD, and loudness. His keynote address is entitles "Zen and the Art of Listening."

The world in 2015 is a noisy place. Car traffic, airplanes, construction sites, cellphone ringtones, music, and muzak everywhere, … There is little chance to escape the constant beating of our eardrums. This has many consequences, not only for professional audio engineers, but for daily life, for relationships, for communication, for our sense of detail, for attentiveness, for subtlety. In this talk, focus will be put on raising the awareness for listening instead of hearing, for timing out instead of burning out, for opening the senses to receive something that is at the heart of communication since millennia: a story.

Key Technology Briefings 3 — **Roland**
Thursday, May 7 — **14:00 – 14:30**
PSE Stage

UNAVAILABLE

Presenter: **Unavailable**

[abstract not available]

Thursday, May 7 — **14:00** — **Room 609B**

Technical Committee Meeting on Automotive Audio

Thursday, May 7 — **14:00** — **Room 609A**

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

Tutorial 2 — **Thursday, May 7**
14:30 – 16:00 — **Room Królewski**

AUDIO FORENSICS— WHAT'S IT ALL ABOUT

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

Panelists: *Gordon Reid*, CEDAR Audio Limited, Cambridge, UK
TBA

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

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

Tutorial 3 — **Thursday, May 7**
14:30 – 16:00 — **Room Opera**

FOUNDATIONS AND PRACTICAL ASPECTS OF SOUND FIELD SYNTHESIS

Presenters: **Sascha Spors**, University of Rostock, Rostock, Germany
Franz Zotter, IEM, University of Music and Performing Arts, Graz, Austria

Sound Field Synthesis (SFS) aims at the physical synthesis of sound fields within a predefined listening area by an ensemble of loudspeakers. Wave Field Synthesis (WFS) and higher-order Ambisonics (HOA) are two prominent representatives of Sound Field Synthesis. This tutorial outlines the history of SFS, its acoustic foundations, and discusses how practical solutions can be achieved in terms of loudspeaker arrays and digital signal processing. The practical realization by such systems involves various approximations and limitations. The technical and psychoacoustic impact of these is reviewed. This tutorial concludes with an outlook on recent research results and developments.

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

Tutorial 4 **Thursday, May 7**
14:30 – 15:45 **Room Balowa AB**

MASTERING FOR MUSICIANS AND HOME STUDIO OWNERS

Presenter: **Andres Mayo**, Andres Mayo Mastering & Audio Post, Buenos Aires Argentina

With technology becoming cheaper and easier to use than ever, most artists are willing to create at least a part of their records at home. This event shows the difficulties that will appear and how to handle them in the most efficient way, taken from real-life examples at the mastering studio. With over 25 years in the mastering field, Mayo has first-hand experience in the dramatic changes of the recording process. This workshop aims to show what you should know before giving the final touch to a mix.

Spatial Audio Demo 1 **Thursday, May 7**
14:30 – 15:15 **Room Saski**

HEADPHONE ENTERTAINMENT: HOW THE FUTURE CAN SOUND USING VIRTUAL SPATIAL AUDIO PRODUCTION APPROACHES

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

Headphones are commonly used when listening to music. Games and mobile film entertainment are a growing headphone application too. Furthermore virtual reality is evolving where headphones certainly will have a basic role. Simple stereo on headphones shouldn't be the end but the beginning of involving audio applications with the flexibility of headphone applications. The tutorial will show different applications and tools and will provide a lot of listening examples.

Session P3 **Thursday, May 7**
15:00 – 17:30 **Room Belweder**

RECORDING AND PRODUCTION

Chair: **Richard King**, McGill University, Montreal, Quebec, Canada; Centre for Interdisciplinary Research in Music Mediaand Technology (CIRMMT), Montreal, Quebac, Canada

15:00

P3-1 **Perceptual Evaluation of Music Mixing Practices—**
Brecht De Man,[1] *Matthew Boerum,*[2,3] *Brett Leonard,*[4] *Richard King,*[2,3] *George Massenburg,*[2,3] *Joshua D. Reiss*[1]
[1]Queen Mary University of London, London, UK
[2]McGill University, Montreal, Quebec, Canada
[3]Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Montreal, Quebac, Canada
[4]University of Nebraska at Omaha, Omaha, NE, USA

The relation of music production practices to preference is still poorly understood. Due to the highly complex process of mixing music, few studies have been able to reliably investigate mixing engineering, as investigating one process parameter or feature without considering the correlation with other parameters inevitably oversimplifies the problem. In this paper we present an experiment where different mixes of different songs, obtained with a representative set of audio engineering tools, are rated by experienced subjects. The relation between the perceived mix quality and sonic features extracted from the mixes is investigated, and we find that a number of features correlate with quality.
Convention Paper 9235

15:30

P3-2 **Automated Equalization of Mobile Device's Microphones**
—*Przemek Maziewski*, Intel Technology Poland, Gdansk, Poland

To achieve high and uniform audio quality in mobile devices their microphones must be equalized. The equalization is typically done manually, requiring lab time, costly equipment, and experienced engineers. This paper presents an automated equalization procedure. It is done using a reference microphone and an external loudspeaker. Each internal microphone is tuned to match the reference microphone's response to the excitations generated via the external loudspeaker. Additionally, each internal microphone's equalization is amended with the inverse equalization characteristic of the reference microphone calculated against the chosen reference, e.g., specific certification requirements. This way the final equalization includes both the internal vs reference microphone delta and the correction required for the reference microphone to pass the chosen certification.
Convention Paper 9236

16:00

P3-3 **Use of Audio Editors in Radio Production—**
Chris Baume,[1,2] *Mark D. Plumbley,*[2] *Janko Calic*[2]
[1]BBC Research and Development, London, UK
[2]University of Surrey, Guildford, Surrey, UK

Audio editing is performed at scale in the production of radio, but often the tools used are poorly targeted toward the task at hand. There are a number of audio analysis techniques that have the potential to aid radio producers, but without a detailed understanding of their process and requirements, it can be difficult to apply these methods. To aid this understanding, a study of radio production practice was conducted on three varied case studies—a news bulletin, drama, and documentary. It examined the audio/metadata workflow, the roles and motivations of the producers, and environmental factors. The study found that producers prefer to interact with higher-level representations of audio content like transcripts and enjoy working on paper. The study also identified opportunities to improve the work flow with tools that link audio to text, highlight repetitions, compare takes, and segment speakers.
Convention Paper 9237

16:30

P3-4 **Cross-Adaptive Polarity Switching Strategies for Optimization of Audio Mixes—***Pedro Duarte*
Pestana,[1,2] *Joshua D. Reiss,*[3] *Alvaro Barbosa*[1]
[1]Catholic University of Oporto, CITAR, Oporto, Portugal
[2]Universidade de Lisboa, Lisbon, Portugal
[3]Queen Mary University of London, London, UK

Crest factor is an often overlooked part of audio production, yet it acts as an important limit to overall loudness. We propose a technique to optimize relative polarities in order to yield the lowest possible peak value. We suggest this is a way of addressing loudness maximization that is more sonically transparent than peak limiting or compression. We also explore additional uses that polarity analysis may have in the context of mixing audio. Results show this is a fairly effective strategy, with average crest factor reductions of 3 dB, resulting in equivalent values for loudness enhancement. While still not comparable to the amount of reduction peak limiters are typically used for, the approach is seen as more transparent via subjective evaluation, through a multi-stimulus test.
Convention Paper 9238

17:00

P3-5 Adaptation and Varying Acoustical Condition and the Resulting Effect on Consistency of High Frequency Preference—*Richard King,[1,2] Brett Leonard,[3] Stuart Bremner,[1,2] Grzegorz Sikora[4]*
[1]McGill University, Montreal, Quebec, Canada
[2]The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada
[3]University of Nebraska at Omaha, Omaha, NE, USA
[4]Bang & Olufsen Deutschland GmbH, Pullach, Germany

The ability to consistently evaluate the frequency response of a music program in a listening room is one of the most fundamental tasks required of an audio engineer. This study requires expert listeners to adjust the high frequency content of audio program under the influence of three different acoustic conditions. The length of exposure is varied to test the role of adaptation on such a task. Results show that there is not a significant difference in the variance of participants' results when exposed to one condition for a longer period of time. However, some individual subjects exhibit adaptive tendencies within the temporal range tested.
Convention Paper 9239

Key Technology Briefings 4 **Studer**
Thursday, May 7 **15:00 – 15:45**
PSE Stage

AUDIO NETWORKING – CHALLENGES & SOLUTIONS

Presenter: **Roger Heiniger**, Studer Product Manager

Using Audio over IP technology to distribute networked audio sources to targets was recently a very hot topic. Now this technology has to be integrated into the systems and although there are many advantages and new possibilities using this technology, there are also some challenges to face and solve. We have several AoIP protocols in the market: Does this mean, an integrator has to choose in the beginning on what protocol he builds up his system and stick with it forever? There is the AES67 standard that promises the ability to interconnect but currently this standard handles only the transport of the audio streams? What about Mic-Pre settings, stream announcement, etc.? How about bandwidth handling in more complex network topologies, incorporating several switches in series?

This session considers these challenges and looks at possible solutions offering ease of use and reliability to the end user.

Thursday, May 7 **15:00** **Room 609B**

Technical Committee Meeting on Semantic Audio Analysis

Spatial Audio Demo 2 **Thursday, May 7**
15:15 – 16:00 **Room Saski**

HEADPHONE ENTERTAINMENT: HOW THE FUTURE CAN SOUND USING VIRTUAL SPATIAL AUDIO PRODUCTION APPROACHES

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

Headphones are commonly used when listening to music. Games and mobile film entertainment are a growing headphone application too. Furthermore virtual reality is evolving where headphones certainly will have a basic role. Simple stereo on headphones shouldn't be the end but the beginning of involving audio applications with the flexibility of headphone applications. The tutorial will show different applications and tools and will provide a lot of listening examples.

Session P4 **Thursday, May 7**
16:00 – 18:00 **Foyer**

POSTERS: SPATIAL AUDIO

16:00

P4-1 Variation of Interaural Time Difference Caused by Head Offset Relative to Coordinate Origin —*Guangzheng Yu, Yuye Wu*, South China University of Technology, Guangzhou, China

Interaural time difference (ITD) is related with spatial position (distance and direction) of the sound source and head size. Assuming the sound source and the coordinate system are fixed, the position relationship between the sound source and head center will be influenced by the head offset relative to the coordinate origin, which may lead to the spatial distribution distortion of ITD in measuring head-related transfer functions (HRTFs). In this paper the variation of ITD caused by head offset is analyzed using the conventional Woodworth ITD model consisting of a spherical head and a point sound source. Results show that the forward (or backward) offsets of head result in small variation of ITDs, however, the spatial distribution distortion of ITDs introduced by the rightward (or leftward) offset of head is unacceptable.
Convention Paper 9240

16:00

P4-2 Functional Representation for Efficient Interpolations of Head Related Transfer Functions in Mobile Headphone Listening—*Joseph Sinker, Jamie Angus*, University of Salford, Salford, Greater Manchester, UK

In this paper two common methods of HRTF/HRIR dataset interpolation, that is simple linear interpolation in the time and frequency domain, are assessed using a Normalized Mean Square Error metric. Frequency domain linear interpolation is shown to be the superior of the two methods, but both suffer from poor behavior and inconsistency over interpolated regions. An alternative interpolation approach based upon the Principal Component Analysis of the dataset is offered; the method uses a novel application of the Discrete Cosine Transform to obtain a functional representation of the PCA weight vectors that may be queried for any angle on a continuous scale. The PCA/DCT method is shown to perform favorably to the simple time domain method, even when applied to a dataset that has been heavily compressed during both the PCA and DCT analysis.
Convention Paper 9241

16:00

P4-3 Binaural Hearing Aids with Wireless Microphone Systems including Speaker Localization and Spatialization—*Gilles Courtois,[1] Patrick Marmaroli,[1] Hervé Lissek,[1] Yves Oesch,[2] William Balande[2]*
[1]Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland;
[2]Phonak Communciations AG, Murten, Switzerland

The digital wireless microphones systems for hearing aids have been developed to provide a clean and intelligible speech signal to hearing-impaired listeners for, e.g., school or teleconference applications. In this technology the voice of the speaker is picked up by a body-worn microphone, wirelessly transmitted to the hearing aids and rendered in a diotic way (same signal at both ears), preventing any speaker localization clues from being provided. The reported algorithm performs a real-time binaural localization and tracking of the talker so that the clean speech signal

can then be spatialized, according to its estimated position relative to the aided listener. This feature is supposed to increase comfort, sense of immersion, and intelligibility for the users of such wireless microphone systems.
Convention Paper 9242

16:00

P4-4 On the Development of a Matlab-Based Tool for Real-Time Spatial Audio Rendering—*Gabriel Moreno,*[1] *Maximo Cobos,*[1] *Jesus Lopez-Ballester,*[1] *Pablo Gutierrez-Parera,*[1] *Jaume Segura,*[1] *Ana Torres*[2]
[1]Universitat de Valencia, Burjassot, Spain
[2]Polytechic University School of Cuenca, Cuenca, Spain

Spatial audio has been a topic of intensive research in the last decades. Although there are many tools available for developing real-time spatial sound systems, most of them work under audio-oriented frameworks. However, despite a significant number of signal processing researchers and engineers who develop their algorithms in MATLAB, there is not currently any MATLAB-based tool for rapid spatial audio system prototyping and algorithm testing. This paper presents a tool for spatial audio research and education under this framework. The presented tool provides the user with a friendly graphical user interface (GUI) that allows to move freely a number of sound sources in 3D and to develop specific functions to be used during their reproduction.
Convention Paper 9243

16:00

P4-5 Psychoacoustic Investigation on the Auralization of Spherical Microphone Array Processing with Wave Field Synthesis—*Gyan Vardhan Singh*, Technische Universität Ilmenau, Ilmenau, Germany

In the present work we have investigated the perceptual effects induced by various errors and artifacts that arise when spherical microphone arrays are used on the recording side. For spatial audio it is very important to characterize the acoustic scene in three-dimensional space. In order to achieve this three dimensional characterization of a sonic scene, spherical microphone arrays are employed. The use of these spherical arrays has some inherent issues because of some errors and by virtue of mathematics involved the in processing. In this paper we analyzed these issues on recording side (spherical microphone array) that plague the audio quality on the rendering side and did a psychoacoustic investigation to access the extent to which the errors and artifacts produce a perceivable affect during auralization when the acoustic scene is reproduced using wave field synthesis.
Convention Paper 9244

16:00

P4-6 Evaluation of a Frequency-Domain Source Position Estimator for VBAP-Panned Recordings—*Alexander Adami,*[1] *Jürgen Herre*[1,2]
[1]International Audio Laboratories Erlangen, Erlangen, Germany
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

A frequency-domain source position estimator is presented that extracts the position of a VBAP-panned directional source by means of a direct-ambience signal decomposition. The directional signal components are used to derive an estimate of the panning gains that can be used to derive the estimated source position. We evaluated the mean estimated source positions as a function of the ideal source position as well as of different ambience energy levels using simulations. Additionally, we analyzed the influence of a second directional source to the estimated source positions.
Convention Paper 9245

16:00

P4-7 A Listener Position Adaptive Stereo System for Object-Based Reproduction—*Marcos F. Simón-Gálvez, Dylan Menzies, Filippo Maria Fazi, Teofilo de Campos, Adrian Hilton*, University of Surrey, Guildford, Surrey, UK

Stereo reproduction of spatial audio allows the creation of stable acoustic images when the listener is placed in the sweet spot, a small region in the vicinity of the axis of symmetry between both loudspeakers. If the listener moves slightly towards one of the sources, however, the images collapse to the loudspeaker the listener is leaning to. In order to overcome such limitation, a stereo reproduction technique that adapts the sweet spot to the listener position is presented here. This strategy introduces a new approach that maximizes listener immersion by rendering object-based audio, in which several audio objects or sources are placed at virtual locations between the stereo span. By using a video tracking device, the listener is allowed to move freely between the loudspeaker span, while loudspeaker outputs are compensated using conventional panning algorithms so that the position of the different audio objects is kept independent from that of the listener.
Convention Paper 9246

16:00

P4-8 Optimization of Reproduced Wave Surface for Three-Dimensional Panning—*Akio Ando, Hiro Furuya, Masafumi Fujii, Minoru Tahara*, University of Toyama, Toyama, Japan

Three-dimensional panning is an essential tool for production of 3D sound material. The typical method is an amplitude panning. The amplitude panning generates the weighting coefficients on the basis of the direction of virtual sound source (desired direction) and the directions of loudspeakers, or the distances between the virtual source and each loudspeaker. It then distributes the weighted signal of the corresponding sound into loudspeakers. The amplitude panning sometimes brings blurred image and deteriorates the timbre of sound. In this paper we propose the new method that optimizes the shape of the wave surface synthesized by multiple loudspeakers. The computer simulation with the frontal six-loudspeaker system showed that the new method achieved the improvement of the reproduced wave surface of sound and its frequency response.
Convention Paper 9247

16:00

P4-9 Estimation of the Radiation Pattern of a Violin During the Performance Using Plenacoustic Methods—*Antonio Canclini, Luca Mucci, Fabio Antonacci, Augusto Sarti, Stefano Tubaro*, Politecnico di Milano, Milan, Italy

We propose a method for estimating the 3D radiation pattern of violins during the performance of a musician. A rectangular array of 32 microphones is adopted for measuring the energy radiated by the violin in the observed directions. In order to gather measurements from all the 3D angular directions, the musician is free to move and rotate in front of the array. The position and orientation of the violin is estimated through a tracking system. As the adopted hardware is very compact and non-invasive, the musician plays in a natural fashion, thus replicating the radiation conditions of a real scenario. The

experimental results prove the accuracy and the effectiveness of the method.
Convention Paper 9248

16:00

P4-10 An Evaluation of the IDHOA Ambisonics Decoder in Irregular Planar Layouts—*Davide Scaini*,[1,2] *Daniel Arteaga*[1,2]
[1] Universitat Pompeu Fabra, Barcelona, Spain
[2]Dolby Iberia S.L., Barcelona, Spain

In previous papers we presented an algorithm for decoding higher order Ambisonics for irregular real-world 3D loudspeaker arrays, implemented in the form of IDHOA, an open source project. IDHOA has many features tailored for the reproduction of Ambisonics in real audio venues. In order to benchmark the performance of the decoder against other decoding solutions, we restrict the decoder to 2D layouts, and in particular to the well studied 5.1 and 7.1 surround layouts and in particular to the well studied stereo, 5.1,and 7.1 surrounds. We report on the results of the objective evaluation of the IDHOA decoder in these layouts and of the subjective evaluation in 5.1 by benchmarking IDHOA against different decoding solutions.
Convention Paper 9249

16:00

P4-11 A General Purpose Modular Microphone Array for Spatial Audio Acquisition—*Jesus Lopez-Ballester, Maximo Cobos, Juan J. Perez-Solano*, *Gabriel Moreno, Jaume Segura*, Universitat de Valencia, Burjassot, Spain

Sound acquisition for spatial audio applications usually requires the use of microphone arrays. Surround recording and advanced reproduction techniques such as Ambisonics or Wave-Field Synthesis usually require the use of multi-capsule microphones. In this context, a proper sound acquisition system is necessary for achieving the desired effect. Besides spatial audio reproduction, other applications such as source localization, speech enhancement or acoustic monitoring using distributed microphone arrays are becoming increasingly important. In this paper we present the design of a general-purpose modular microphone array to be used in the above application contexts. The presented system allows performing multichannel recordings using multiple capsules arranged in different 2D and 3D geometries.
Convention Paper 9250

16:00

P4-12 Immersive Content in Three Dimensional Recording Techniques for Single Instruments—*Brett Leonard*,[3] *David Benson*,[1,2] *Will Howie*[1,2]
[1]McGill University, Montreal, Quebec, Canada
[2]Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Montreal, Quebec, Canada
[3]University of Nebraska at Omaha, Omaha, NE, USA

"3D Audio" has become a popular topic in recent years. A great deal of research is underway in spatial sound reproduction through computer modeling and signal processing, while less focus is being placed on actual recording practice. This study is a preliminary test in establishing effective levels of height-channel information based on the results of a listening test. In this case, an acoustic guitar was used as the source. Eight discrete channels of height information were combined with an eight-channel surround sound mix reproduced at the listener's ear height. Data from the resulting listening test suggests that while substantial levels of height channel information increase the effect of immersion, more subtle levels fail to provide increased immersion over the conventional multichannel mix.
Convention Paper 9251

Workshop 2 — Thursday, May 7
16:00 – 17:30 — Room Balowa AB

ISO/MPEG-H AUDIO—THE NEW STANDARD FOR UNIVERSAL SPATIAL / 3D AUDIO CODING

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

Panelists: *Alexander Krüger*, Technicolor
Nils Peters, Qualcomm, San Diego, CA, USA; University of California Berkeley, Berkeley, CA, USA
Jan Plogsties, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
Ulli Scuda, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Recently, the ISO/MPEG standardization group created the MPEG-H 3D Audio specification to go along with Ultra High Definition Television (UHDT) video. The specification features several unique elements, such as handling of channel-based content, object-based content, and higher order ambisonics (HOA) content or the capability of rendering encoded high-quality content on a wide range of loudspeaker setups (22.2 ... 5.1 ... stereo / headphones). This workshop provides an overview of the MPEG-H 3D Audio standard regarding its underlying architecture, technology, performance and how to produce immersive content for it.

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

Key Technology Briefings 5 — Wroclaw Univ. of Tech.
Thursday, May 7 — 16:00 – 16:45
PSE Stage

NEW HOME FOR POLISH NATIONAL RADIO SYMPHONY ORCHESTRA IN KATOWICE

Presenter: **Piotr Z. Kozłowski**, Wrocław University of Technology, Chair of Acoustics and Multimedia

The Polish National Radio Symphony Orchestra in Katowice (NOSPR = Narodowa Orkiestra Symfoniczna Polskiego Radia w Katowicach, in polish) is one of the greatest and most famous orchestras throughout the world. This group had no concert hall of their own that matched its level of quality in the past. They moved from one venue to another with no success toward improving the quality of the acoustics. Last period they spent at Congress Hall, where the acoustics were completely different from their needs and expectations.

As a result of the decision that a new concert hall should be built to become the new home for the orchestra, an international architectural competition for the complex design was announced in 2008.

The design stage started in March 2009 and was finished in January 2011. The general contractor started the building process in March 2012 and finished September 2014. The acoustical tuning process took five months—from May to September 2014. The first concert was played on October 1, 2014. This evening belongs of course to Polish National Radio Symphony Orchestra and for one and only Krystian Zimerman.

In the presentation we will discuss various topics from the many technological aspects: • noise control and the building acoustics; • room acoustics of the acoustic qualified interiors; • electro-acoustics system of the entire building with the multitrack recording system; • stage management system; • stage lighting; • upper and lower stage mechanics with the control system; • digital signage information system for the viewers (agenda).

Technological topics are described separately for the most ➥

important areas of building such as: two concert halls, four team rehearsal rooms, 37 solo rehearsal rooms, recording studio, backstage, OB vans docks, ventilation, and electrical powering rooms. We will also present design stage dilemmas. Final results of acoustical measurements done during tuning process are presented to show why the two concert halls and the entire complex has received the highest praise from artists and auditors.

Thursday, May 7 16:00 Room 609B

Technical Committee Meeting on Audio Forensics

Tutorial 5 Thursday, May 7
16:30 – 18:00 Room Królewski

DIGITAL TRANSFER OF THE ARMANDO LEÇA FOLK MUSIC COLLECTION

Presenter: **Nadja Wallaszkovits**, Phonogrammarchiv, Austrian Academy of Science, Vienna, Austria

Nadja Wallaszkovits will discuss the restoration, transfer, and digitization of a unique collection of folk music recorded during 1939–1940 in the rural areas and mountain villages of Portugal by the folklorist Armando Leça, in collaboration with National Radio (Emissora Nacional). Her presentation will open with a short introduction of this important field research project that resulted in the first known collection of recordings documenting rural musical practices from nearly all regions of Portugal. Thereafter, she will present a historical overview of early magnetic tape developments and the birth of audio tape recorder technology, focusing on the characteristics of the individual tape machine used by Armando Leça. The problems of carrier handling, restoration, and transfer of these valuable original tapes will be discussed along with the judicial use of signal enhancement during the playback process.

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

Student Event and Career Development
OPENING AND STUDENT DELEGATE ASSEMBLY MEETING – PART 1
Thursday, May 7, 16:30 – 18:00
Room Opera

Presenters: **Simon-Claudius Wystrach**
Steven Van Dyne
Brecht De Man
Zach Bloomstein

The first Student Delegate Assembly (SDA) meeting is the official opening of the Convention's student program and a great opportunity to meet fellow students from all corners of the world. The SDA Officers will provide a top-level view of the AES' student activities, and announce the candidates for the election of the SDA's new Vice Chair, the Student Competition finalists, and all upcoming student events at the Convention.

Students and Student Sections will be given the opportunity to introduce themselves and talk about their respective activities to stimulate international communication, enabled and guided by the Student Delegate Assembly.

The event is open to all, and we invite students and educators in particular to come along to this meeting.

Spatial Audio Demo 3 Thursday, May 7
17:00 – 17:30 Room Saski

3D AUDIO FOR GAMES

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

Today 3D audio is a well known task for film as well for music production. Approaches and technologies have been discussed and presented at AES conventions for some years. But involving 3D audio for games is just beginning. A first tool, approaches, and a game demo will be shown and visions and possibilities for the future will be discussed.

Thursday, May 7 17:00 Room 609B

Technical Committee Meeting on Microphones and Applications

Thursday, May 7 17:00 Room 609A

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

Spatial Audio Demo 4 Thursday, May 7
17:30 – 18:00 Room Saski

3D AUDIO FOR GAMES

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

Today 3D audio is a well known task for film as well for music production. Approaches and technologies have been discussed and presented at AES conventions for some years. But involving 3D audio for games is just beginning. A first tool, approaches, and a game demo will be shown and visions and possibilities for the future will be discussed.

Special Event
PARIS LIVE RECORDING VERSUS WARSAW LIVE PERFORMANCE
Thursday, May 7, 18:15 – 20:00
Balowa A/B

Presenter: **Andrew Lipinski**, Lipinski Sound Corporation

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

Artists: **Ewa Izykowska**, Soprano
Wojciech Switala, Piano

Should live performance be a reference for creating a sound illusion in recording? To address this question Andrew Lipinski—a renowned Polish engineer/producer—will present his 8.0 channel high-resolution surround sound recording with height information, made during a live public concert in Paris.

Both the Paris recording and the Warsaw live concert will feature the same artists as well as the same keyboard instrument. Chopin Songs will be performed by world-class Polish artists: Ewa Izykowska (soprano) and Wojciech Switala (piano).

In 2010, the 200 anniversary of Chopin's birth, Lipinski took upon himself the task of recording Chopin Songs in one of only two remaining authentic concert venues where Chopin had originally performed: the Conservatory of Music in Paris. This Paris recording features the single existing original instrument on which Chopin performed, the Pleyel 1843 grand piano, which is still in absolutely perfect condition. For the Warsaw concert, an identical Pleyel 1843 will be used, although the instrument in Warsaw has been renovated. The same piano tuner prepares both instruments.

The Paris recording was filmed in HD and published on Blu-ray disc with 96 kHz/24 bit, 8.0 surround sound containing three channels of height information. Andrew employed his own analog electronics with only seven discrete transistors working in pure class A between the microphone diaphragm and the analog-to-digital conversion.

During this A/B presentation Andrew will share with the audience his best practices and will speculate on where a purist recording technology will likely evolve.

PROGRAM OF THE CONCERT
FRYDERYK CHOPIN [1810–1849]: SONGS

- "A Young Girl's Wish, Op. 74 No 1," Lyrics by S. Witwicki
- "Sad River, Op. 74 No 3," Lyrics by S. Witwicki
- "A Drinking Song, Op. 74 No 4," Lyrics by S. Witwicki
- "A Girl's Desire, Op. 74 No 5," Lyrics by S. Witwicki
- "Out of My Sight, Op. 74 No 6," Lyrics by A. Mickiewicz
- "The Messenger, Op. 74 No 7," Lyrics by S. Witwicki
- "The Handsome Lad, Op. 74 No 8," Lyrics by B. Zaleski
- "Elegy, Op. 74 No 9," Lyrics by Z. Krasinski
- "The Warrior, Op. 74 No 10," Lyrics by S. Witwicki
- "Death's Divisions, Op. 74 No 11," Lyrics by B. Zaleski
- "My Sweetheart, Op. 74 No 12," Lyrics by A. Mickiewicz
- "The Ring, Op. 74 No 14," Lyrics by S. Witwicki

Student Event and Career Development
AES STUDENT PARTY
Thursday, May 7, 20:30 – 24:00
Klub Harenda
Krakowskie Przedmieście 4-6

The AES Student Party is open to any 138th Convention participant with a STUDENT BADGDE. It will be held at Klub Harenda located at Krakowskie Przedmieście 4-6 (5 min. walk from Sofitel). Come meet fellow students from around the world. Live performances by Grammy Nominated David Miles Huber, Chris Calcutt, MLDV, Marek Walaszek, and students from Berklee Valencia will be accompanied by technical descriptions of the tools and processes used. Doors will open at 8:30pm. Performances will run from 9:30/10pm–Midnight.

Session P5 **Friday, May 8**
09:00 – 11:00 **Room Belweder**

AUDIO SIGNAL PROCESSING

Chair: **Christoph Musialik**, Sennheiser Audio Labs, Waldshut-Tiengen Germany

09:00

P5-1 Multi-Rate System for Arbitrary Audio Processing— *Daekyoung Noh*, DTS, Inc., Santa Ana, CA, USA

An efficient multi-rate system for arbitrary audio processing is proposed. In order to minimize computational complexity, high sampling rate signals are decimated and split into two subbands. The process can be repeated in the low band to obtain a maximally decimated system. Then, only the lowest band is being processed with arbitrary audio processing. Amplitude and group delay compensation are performed to the rest of the bands to minimize aliasing noise and amplitude distortion that can be caused when the bands are recombined due to arbitrary audio processing performed in the low bands. The Goertzel algorithm transition band addition/subtraction method is introduced for group delay correction in real-time processing. Once arbitrary processing is done in the lowest band the subbands are then up-sampled and recombined. Finally, test results and computational advantages are discussed.
Convention Paper 9252

09:30

P5-2 A Short-Term Analysis of a Digital Sigma-Delta Modulator with Nonstationary Audio Signals—*Marcin Lewandowski*, Warsaw University of Technology, Warsaw, Poland

Signal conversion quality of sigma-delta (ΣΔ) digital-to-analog audio converters (DACs) mainly depends on the ΣΔ modulator's parameters. Conventional quality examination of ΣΔ audio DACs has been performed in the frequency domain and can be considered indicative of quality only in the case of linear and stationary systems. However, highly nonlinear and nonstationary ΣΔ modulators create errors that depend on the input signal. In this study, a method for evaluating ΣΔ modulators in the time-domain is presented. Simulations and analysis were performed with the use of music signals. Results showed that the short-term performance of digital ΣΔ modulators is highly correlated with the variation of the input signal. This is particularly important as ΣΔ modulators are commonly used in DACs of both consumer and professional audio equipment.
Convention Paper 9253

10:00

P5-3 Application of Sinusoidal Coding for Enhanced Bandwidth Extension in MPEG-H USAC—*Tomasz Zernicki,*[1] *Maciej Bartkowiak,*[2] *Lukasz Januszkiewicz,*[1] *Marcin Chryszczanowicz*[1]
[1]Zylia sp. z o.o., Poznan, Poland
[2]Poznan University of Technology, Poznan, Poland

A new audio coding technique applicable to very low bit rates is proposed. The existing MPEG-D standard of Unified Speech and Audio Coding (USAC) is enhanced by a new High Frequency Sinusoidal Coder (HFSC), employed for improving the subjective quality of high frequency spectral content. The paper gives an overview of the new technique as well as it offers some insight into the operation modes and delay issues. A statistically evidenced significant improvement of the audio quality resulting from applying this technique is demonstrated.
Convention Paper 9254

10:30

P5-4 Practical Considerations of Time-Varying Feedback Delay Networks—*Sebastian J. Schlecht, Emanuël A. P. Habets*, International Audio Laboratories Erlangen, Erlangen, Germany

Feedback delay networks (FDNs) can be efficiently used to generate parametric artificial reverberation. Recently, the authors proposed a novel approach to time-varying FDNs by introducing a time-varying feedback matrix. The formulation of the time-varying feedback matrix was given in the complex eigenvalue domain, whereas this contribution specifies the requirements for real valued time-domain processing. In addition, the computational costs of different time-varying feedback matrices, which depend on the matrix type and modulation function, are discussed. In a performance evaluation, the proposed orthogonal matrix modulation is compared to a direct interpolation of the matrix entries.
Convention Paper 9255

Session P6 **Friday, May 8**
09:00 – 11:30 **Room Królewski**

ROOM ACOUSTICS

Chair: **Lauri Savioja**, Aalto University, Espoo, Finland

09:00

P6-1 Radio Studio Acoustics Part 1: Subjective Evaluation— *Ian Dash, Mark Bowry*, Australian Broadcasting Corporation, Sydney, NSW, Australia

A subjective evaluation exercise to determine the perceived acoustical quality of 12 small acoustic spaces used for radio production and presentation was conducted using two recording methods. One set of recordings used multiple

microphones and the other set used a single microphone. A common listening test was run using both sets of recordings. The novel approach using multiple microphones was found to be a more sensitive and reliable method, and leads to different conclusions on acoustic performance criteria compared with a more conventional assessment. Group differences due to the reader, listener gender, listener location, and listener skill level were performed using ANOVA.
Convention Paper 9256

09:30

P6-2 Radio Studio Acoustics Part 2: Correlation of Objective Measurements with Subjective Assessment—*Ian Dash, Mark Bowry*, Australian Broadcasting Corporation, Sydney, NSW, Australia

A subjective evaluation was made of the acoustic quality of 12 small- to medium-sized acoustic spaces used for radio production and presentation. Correlation analysis was used to relate the results from that evaluation to measured objective acoustical parameters of those rooms. The results suggest that low early reverberation time at lower frequencies is of high importance to listeners, and that listeners prefer consistent reverberation time in all bands from 125 Hz upwards.
Convention Paper 9257

10:00

P6-3 Estimation of Room Reflection Parameters for a Reverberant Spatial Audio Object—*Luca Remaggi, Philip Jackson, Philip Coleman*, University of Surrey, Guildford, Surrey, UK

Estimating and parameterizing the early and late reflections of an enclosed space is an interesting topic in acoustics. With a suitable set of parameters, the current concept of a spatial audio object (SAO), which is typically limited to either direct (dry) sound or diffuse field components, could be extended to afford an editable spatial description of the room acoustics. In this paper we present an analysis/synthesis method for parameterizing a set of measured room impulse responses (RIRs). RIRs were recorded in a medium-sized auditorium, using a uniform circular array of microphones representing the perspective of a listener in the front row. During the analysis process, these RIRs were decomposed, in time, into three parts: the direct sound, the early reflections, and the late reflections. From the direct sound and early reflections, parameters were extracted for the length, amplitude, and direction of arrival (DOA) of the propagation paths by exploiting the dynamic programming projected phase-slope algorithm (DYPSA) and classical delay-and-sum beamformer (DSB). Their spectral envelope was calculated using linear predictive coding (LPC). Late reflections were modeled by frequency-dependent decays excited by band-limited Gaussian noise. The combination of these parameters for a given source position and the direct source signal represents the reverberant or "wet" spatial audio object. RIRs synthesized for a specified rendering and reproduction arrangement were convolved with dry sources to form reverberant components of the sound scene. The resulting signals demonstrated potential for these techniques, e.g., in SAO reproduction over a 22.2 surround sound system.
Convention Paper 9258

10:30

P6-4 Sound Radiation Control for Reducing the Effect of Strong Reflections—*Jiho Chang, Jaeyoun Cho, Yoonjae Lee*, Samsung Electronics Co. Ltd., Suwon, Gyeonggi-do, Korea

This paper is concerned with sound radiation control of a loudspeaker that has omni-directional radiation in free field condition. When the loudspeaker is placed very close to flat surfaces, the reflections from the surfaces are as strong as the direct sounds and deteriorate the omni-directionality. This study assumes a proximate wall and attempts to analyze the effect of a reflection from the wall in terms of the directivity and to improve the omni-directionality by using a circular array of loudspeakers. For a given distance, weights for loudspeakers are calculated that make the sound radiation omni-directional in spite of the wall.
Convention Paper 9259

11:00

P6-5 Acoustical Measurements of Warsaw's Chamber Opera House Using Two Types of Sound Sources for Subsequent Auralization—*Wieslaw Woszczyk,*[1] *Tadeusz Fidecki,*[2] *Jung Wook (Jonathan) Hong,*[1,3] *Tomasz Rudzki,*[2] *David Benson*[1,4]
[1]McGill University, Montreal, Quebec, Canada
[2]F. Chopin University of Music, Warsaw, Poland
[3]GKL Audio Inc., Montreal, QC, Canada
[4]The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada

Impulse response measurements using log sine sweeps were made in the Warsaw's Chamber Opera House in eight microphone locations on the floor area and at two locations on the balcony, with two microphone elevations, using two types of sound sources having different directional radiation characteristics. The Opera House having only 159 seats originates from 1775 and is renowned for its excellent acoustics fitting for Mozart operas. The measurements show how within this relatively small venue, an opera director can create a wide range of acoustic perspectives for voices and instruments, and achieve a desired dramatic effect. In a subsequent multichannel auralization, anechoic instrumental and vocal sounds were placed virtually in the opera house, and a listening panel compared the renderings. The experiment underlines the importance of choosing directional characteristics of sound sources used in the measurements of room impulse responses intended for subsequent applications
Convention Paper 9260

Workshop 3 **Friday, May 8**
09:00 – 11:00 **Room Balowa AB**

MIXING MEETS MASTERING:
WHERE THE LINE BECOMES BLURRED

Chair: **Rob Toulson**, Anglia Ruskin University, Cambridge, UK

Panelists: *George Massenburg*, Schulich School of Music, McGill University, Montreal, Quebec, Canada; Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Montreal, Quebec, Canada
Mandy Parnell, Black Saloon Studios, London, UK
Darcy Proper, Wisseloord Studios, Hilversum, The Netherlands
Michael Romanowski, Michael Romanowski Mastering, San Francisco, CA, USA; Owner Coast Recorders
Ronald Prent, Wisseloord Studios, Hilversum, The Netherlands

This workshop will discuss the evolution of mixing and mastering, particularly with reference to scenarios when the two practices become merged into one. There are a number of arguments for and against the use of mastering techniques at the mixing stage. For example, it can be argued that mix engineers need to take a greater responsibility towards dynamics and noise cancellation. Whereas, in contrast, the use of mix bus limiting when generating draft listening copies can confuse and falsify the sign-off process. Furthermore, mastering engineers might be tempted to work from mix stems, but does that mean they are effectively mixing as well as mastering the songs? We will also explore professional and home studio practices, innovative tools, and educational approaches.

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

Workshop 4 — **Friday, May 8**
09:00 – 10:30 — **Room Opera**

RECENT DEVELOPMENTS IN SOUND FIELD SYNTHESIS

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

Panelists: *Filippo Maria Fazi*, University of Southampton, Southampton, Hampshire, UK
Ville Pulkki, Aalto University, Espoo, Finland
Thomas Sporer, Fraunhofer IDMT, Ilmenau, Germany

Wave Field Synthesis (WFS) and higher-order Ambisonics (HOA) are two prominent representatives of Sound Field Synthesis techniques. Even after several decades research and development for both techniques is still very actively conducted. This workshop discusses recent activities and trends in WFS and HOA.

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

Friday, May 8 — **09:00** — **Room 609B**

Technical Committee Meeting on Hearing and Hearing Loss Prevention

Friday, May 8 — **10:00** — **Room 609B**

Technical Committee Meeting on Coding of Audio Signals

Spatial Audio Demo 5 — **Friday, May 8**
10:15 – 10:45 — **Room Saski**

ISO/MPEG-H AUDIO—THE NEW STANDARD FOR UNIVERSAL SPATIAL / 3D AUDIO CODING

Presenter: **Ulli Scuda**, Fraunhofer IIS, Erlangen, Germany

Recently, the ISO/MPEG standardization group created the MPEG-H 3D Audio specification to go along with Ultra High Definition Television (UHDT) video. The specification features several unique elements, such as handling of channel-based content, object-based content, and higher order ambisonics (HOA) content or the capability of rendering encoded high-quality content on a wide range of loudspeaker setups (22.2 ... 5.1 ... stereo / headphones). This demonstration showcases some of the features and sonic performance of the MPEG-H 3D Audio standard.

This demonstration supports the workshop of the same name given on Thursday.

Workshop 5 — **Friday, May 8**
10:45 – 12:15 — **Room Opera**

CINEMA LOUDNESS WORKING GROUP—DRAFT RECOMMENDED PRACTICE

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

Panelists: *Florian Camerer*, ORF - Austrian TV, Vienna, Austria; EBU - European Broadcasting Union
Eelco Grimm, Grimm Audio, Utrecht, The Netherlands

The AES 57th International Conference in Hollywood discussed the plethora of audio mixes that are being created for use across multiple distribution platforms. It is vital for the industry to simplify the mixing process and develop techniques and technology to allow one master mix to service all these platforms. We will review the Draft Recommended Practice as proposed by the Standards Working Group.

This session is presented in association with the AES Technical Committee on Sound for Digital Cinema and Television

Spatial Audio Demo 6 — **Friday, May 8**
10:45 – 11:15 — **Room Saski**

ISO/MPEG-H AUDIO—THE NEW STANDARD FOR UNIVERSAL SPATIAL / 3D AUDIO CODING

Presenter: **Ulli Scuda**, Fraunhofer IIS, Erlangen, Germany

Recently, the ISO/MPEG standardization group created the MPEG-H 3D Audio specification to go along with Ultra High Definition Television (UHDT) video. The specification features several unique elements, such as handling of channel-based content, object-based content, and higher order ambisonics (HOA) content or the capability of rendering encoded high-quality content on a wide range of loudspeaker setups (22.2 ... 5.1 ... stereo / headphones). This demonstration showcases some of the features and sonic performance of the MPEG-H 3D Audio standard.

This demonstration supports the workshop of the same name given on Thursday.

Session P7 — **Friday, May 8**
11:00 – 13:00 — **Foyer**

POSTERS: AUDIO SIGNAL PROCESSING

11:00

P7-1 Feature Learning for Classifying Drum Components from Nonnegative Matrix Factorization—*Matthias Leimeister*, Native Instruments GmbH, Berlin, Germany

This paper explores automatic feature learning methods to classify percussive components in nonnegative matrix factorization (NMF). To circumvent the necessity of designing appropriate spectral and temporal features for component clustering, as usually used in NMF-based transcription systems, multilayer perceptrons and deep belief networks are trained directly on the factorization of a large number of isolated samples of kick and snare drums. The learned features are then used to assign components resulting from the analysis of polyphonic music to the different drum classes and retrieve the temporal activation curves. The evaluation on a set of 145 excerpts of polyphonic music shows that the algorithms can efficiently classify drum components and compare favorably to a classic "bag-of-features" approach using support vector machines and spectral mid-level features.
Convention Paper 9261

11:00

P7-2 Blind Bandwidth Extension System Utilizing Advanced Spectral Envelope Predictor—*Kihyun Choo*,[1] *Anton Porov*,[2,3] *Eunmi Oh*[1]
[1]Samsung Electronics Co., Ltd., Suwon, Korea
[2]Samsung R&D Institute Russia, Moscow, Russia
[3]ITMO University, Saint-Petersburg, Russia

We propose a blind bandwidth extension (BWE) technique ➥

that improves the quality of a narrow-band speech signal using time domain extension and spectral envelope prediction in the frequency domain. In the time domain, we use a spectral double shifting method. Further, a new spectral envelope predictor is introduced in the frequency domain. We observe less distortion when the attribute is transferred from low to high frequency, instead of reflecting the original high band. The proposed blind BWE system is applied to the decoded output of an adaptive multi-rate (AMR) codec at 12.2 kbps to generate a high-frequency spectrum from 4 to 8 kHz. The blind BWE was objectively evaluated with the AMR and AMR wideband codecs and subjectively evaluated by comparing it with the AMR.
Convention Paper 9262

11:00

P7-3 Time Domain Extrapolative Packet Loss Concealment for MDCT Based Voice Codec—*Shen Huang, Xuejing Sun*, Dolby Laboratories, Beijing, China

A novel low latency packet loss concealment technique for transform-based codecs is proposed. The algorithm combines signals from Inverse Modulated Discrete Cosine Transform (IMDCT) domain and the previous reconstructed signal from time domain with aligned phase, with which a pitch-synchronized concealment is performed. This minimizes aliasing artifacts that occur in MDCT domain concealment for voiced speech signals. For unvoiced speech, speech-shaped comfort noise is inserted. When there is a burst loss, a position-dependent concealment process is performed for different stages of packet losses. Subjective listening tests using both naïve and expert listeners suggest that the proposed algorithm generates fewer artifacts and offers significantly better performance against legacy packet repetition based approaches.
Convention Paper 9263

11:00

P7-4 Scalable Parametric Audio Coder Using Sparse Approximation with Frame-to-Frame Perceptually Optimized Wavelet Packet Based Dictionary—*Alexey Petrovsky, Vadzim Herasimovich, Alexander Petrovsky*, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

This paper is devoted to the development of a scalable parametric audio coder based on a matching pursuit algorithm with a frame-based psychoacoustic optimized wavelet packet dictionary. The main idea is to parameterize audio signal with a minimum number of non-negative elements. This can be done by applying sparse approximation such as matching pursuit algorithm. In contrast with current approaches in audio coding based on sparse approximation we introduce a model of dynamic dictionary forming for each frame of input audio signal individually based on wavelet packet decomposition and dynamic wavelet packet tree transformation with psychoacoustic model. Experimental results of developed encoder and comparison with modern popular audio encoders are provided.
Convention Paper 9264

11:00

P7-5 General-Purpose Listening Enhancement Based on Subband Non-Linear Amplification with Psychoacoustic Criterion—*Elias Azarov, Maxim Vashkevich, Vadzim Herasimovich, Alexander Petrovsky*, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Near end listening enhancement is an effective approach for speech intelligibility improvement in noisy conditions that is applied mainly for telecommunications. However potential application field of the concept of near end listening enhancement is much wider and can be extended for listening of any audio content (including music and other sounds) in quiet and noisy conditions. This paper proposes an algorithm for near end listening enhancement designed for processing both speech and music that, considering subjective listening tests, significantly improves the listening experience. The algorithm is based on subband nonlinear amplification of the audio signal in accordance with noise spectral characteristics and personal hearing thresholds of the listener. The algorithm is experimentally implemented as an application for smartphones.
Convention Paper 9265

11:00

P7-6 Arbitrary Trajectory Estimation of a Moving Acoustic Source—*Sai Gunaranjan Pelluri, Thippur V. Sreenivas*, Indian Institute of Science, Bangalore, India

The state-of-the-art passive methods for estimating the trajectory of a moving acoustic source involve computing the cross-correlation function either directly or indirectly (as in the case of the Beam Forming Approach) between pairs of microphones. Also there have been several Active Acoustic techniques such as SONAR, RADAR, etc., which have been used to estimate the source parameters such as velocity, trajectory, etc. They involve pinging the source with a known signal. In this paper, given the fact that the moving source itself generates a signal, we propose a technique by which we estimate the source trajectory using only the signal captured at the receiver thereby avoiding the need to ping the source and without computing the cross-correlation function.
Convention Paper 9266
This Poster was presented on Sunday as part of Poster Session 16.

11:00

P7-7 Speech Analysis Based on Sinusoidal Model with Time-Varying Parameters—*Elias Azarov, Maxim Vashkevich, Alexander Petrovsky*, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Extracting speech-specific characteristics from a signal such as spectral envelope and pitch is essential for parametrical speech processing. These characteristics are used in many speech applications including coding, parametrical text-to-speech synthesis, voice morphing, and others. This paper presents some original estimation techniques that extract these characteristics using a sinusoidal model of speech with instantaneous parameters. The analysis scheme consists of two steps: first the parameters of sinusoidal model are extracted from the signal, and then these parameters are transformed to required characteristics. Some evaluations of the presented techniques are carried out on synthetic and natural speech signals to show potential of the presented approach.
Convention Paper 9267

11:00

P7-8 A Low-Delay Algorithm for Instantaneous Pitch Estimation—*Elias Azarov, Maxim Vashkevich, D. Likhachov, Alexander Petrovsky*, Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

Estimation of instantaneous pitch provides high accuracy for

frequency-modulated pitches and can be beneficial compared to conventional pitch extraction techniques for unsteady voiced sounds. However, applying an estimator of instantaneous pitch to a practical real-time speech processing application is a hard problem because of high computational cost and high inherent delay. The paper presents an algorithm for instantaneous pitch estimation specifically designed for real-time applications. The analysis scheme is based on the robust algorithm for instantaneous pitch tracking (IRAPT) featuring an efficient processing scheme and low inherent delay. The paper presents some evaluation results using synthesized and natural speech signals that illustrate actual performance of the algorithm.
Convention Paper 9268

11:00

P7-9 Content-Based Music Structure Analysis Using Vector Quantization—*Nikolaos Tsipas, Lazaros Vrysis, Aristotle Charalampos A. Dimoulas, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

Music structure analysis has been one of the challenging problems in the field of music information retrieval during the last decade. Past years advances in the field have contributed toward the establishment and standardization of a framework covering repetition, homogeneity, and novelty based approaches. With this paper an optimized fusion algorithm for transition points detection in musical pieces is proposed, as an extension to existing state-of-the-art techniques. Vector-Quantization is introduced as an adaptive filtering mechanism for time-lag matrices while a structure-features based self-similarity matrix is proposed for novelty detection. The method is evaluated against 124 pop songs from the INRIA Eurovision dataset and performance results are presented in comparison with existing state-of-the-art implementations for music structure analysis.
Convention Paper 9269

11:00

P7-10 Clock Skew Compensation by Adaptive Resampling for Audio Networking—*Leonardo Gabrielli,*[1] *Michele Bussolotto,*[1] *Stefano Squartini,*[1] *Fons Adriaensen*[2]
[1]Universitá Politecnica delle Marche, Ancona, Italy
[2]Huawei European Research Center, Munich, Germany

Wired Audio Networking is an established practice since years based on both proprietary solutions or open hardware and protocols. One of the most cost-effective solutions is the use of a general purpose IEEE 802.3 infrastructure and personal computers running IP based protocols. One obvious shortcoming of such setups is the lack of synchronization at the audio level and the presence of a network delay affected by jitter. Two approaches to sustain a continuous audio flow are described, implemented by the authors in open source projects based on a relative and absolute time adaptive resampling. A description of the mechanisms is provided along with simulated and measured results, which show the validity of both approaches.
Convention Paper 9270

11:00

P7-11 Analysis of Onset Detection with a Maximum Filter in Recordings of Bowed Instruments—*Bartlomiej Stasiak, Jedrzej Monko*, Lódz University of Technology, Lódz, Poland

This work presents a new approach to assessment of the quality of onset detection functions on the example of bowed instruments recordings. Using this method, we test a vibrato suppression technique based on a maximum filter. The results, obtained with the aid of a specially constructed database of audio recordings, reveal problems connected with certain qualities of the sound signal generated by a bowed instrument and with the effectiveness of the onset detection process.
Convention Paper 9271

11:00

P7-12 An FPGA-Based Virtual Reality Audio System—*Wolfgang Fohl, David Hemmer*, Hamburg University of Applied Sciences, Hamburg, Germany

A distributed system for mobile virtual reality audio is presented. The system consists of an *audio server* running on a PC or Mac, a *remote control app* for an iOS6 device, and the *mobile renderer* running on a system-on-chip (SoC) with a CPU core and signal processing hardware. The server communicates with the renderer via WLAN. It sends audio streams over a self-defined lightweight protocol and exchanges status and control data as OSC (Open Sound Control) messages. On the mobile renderer, HRTF filters are applied to each audio signal according to the relative positions of the source and the listener's head. The complete audio signal processing chain has been designed in Simulink. The VHDL code for the SoC's FPGA hardware has been automatically generated by Xilinx's System Generator. The system is capable of rendering up to eight independent virtual sources.
Convention Paper 9328

Project Studio Expo 1
Friday, May 8 **11:00 – 11:45**
PSE Stage

PLUG-INS: CONSIDERATIONS

Presenter: **Maciej Polański**, Musoneo.pl, Warsaw, Poland

A practical presentation that explains the sorts of things users should be considering when employing plug-ins—such as CPU power, different plug-in types, legacy and latency issues. Maciej will outline a sensible and practical approach to working in the box and will address the problems that will be encountered and the solutions and compromises that should be aimed for.

Friday, May 8 **11:00** **Room 609B**

Technical Committee Meeting on Signal Processing

Friday, May 8 **11:00** **Room 609A**

Standards Committee Meeting SC-05-02 on Audio Connectors

Tutorial 6 **Friday, May 8**
11:15 – 13:15 **Room Belweder**

MICROPHONES—CAN YOU HEAR THE SPECS?

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

Panelists: *Jürgen Breitlow*, Georg Neumann GmbH, Berlin, Germany
David Josephson, Josephson Engineering, Inc., Santa Cruz, CA, USA
Helmut Wittek, SCHOEPS GmbH, Karlsruhe, Germany

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 and how it relates to perceived sound. It takes a critical look on how specs are presented to the user, what to look for, and what to expect.

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

Tutorial 7 **Friday, May 8**
11:45 – 13:15 **Room Królewski**

LIVE SOUND SUBWOOFER SYSTEM CALIBRATION

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

Achieving even low-frequency coverage across a large audience area while simultaneously minimizing sound energy on stage is easier said than done in large-scale live sound reinforcement. Non-ideal subwoofer system calibration causes severe seat-to-seat variability in frequency response, which often goes unnoticed by engineers firmly planted to the mix position. This tutorial presents practical approaches to system calibration which provide improved low-frequency performance. Source placement, polar pattern control, clustering and DSP techniques are explained using simulations to highlight their respective advantages and disadvantages. All solutions offered consider practical restrictions including: system efficiency, stage layout, truck space, work timeframe, budget, etc. Perceptual considerations are also touched upon, inspecting whether the increasingly-common method of driving subwoofers with a mono auxiliary send is an imperfect approach.

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

Workshop 6 **Friday, May 8**
11:15 — 12:15 **Room Balowa AB**

3D 9.1 FOR CREATING AN IMPRESSIVE MOVIE SOUND

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

A 3D 9.1 reproducing system provides much more possibilities to create an impressive sound of a movie in comparison to stereo and Surround. The most important improvements are more loudspeakers for creating real sound sources, an incredible envelopment, and realistic sounding soundscapes. The sound of movies contains several sound layers. Each sound of those layers can be extended from mono or stereo to 3D by using psychoacoustical rules. Those psychoacoustical rules are divided into different groups according to the above-described improvements. During the presentation, a couple of rules will be described and demonstrated with example recordings (synchronized audio and video).

This session is presented in association with the AES Technical Committee on Sound for Digital Cinema and Television

Spatial Audio Demo 7 **Friday, May 8**
12:00 – 13:00 **Room Saski**

THE PRODUCTION OF INVOLVING 3D AUDIO EXPERIENCES FOR MUSIC APPLICATIONS

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

It's one thing to move audio objects around the listener like a spaceship in a movie application. It's a totally different approach to create musical landscapes that involve the listener emotionally in 3D audio. Strategies, tools, and visions will be shown and a lot of listening examples will be provided.

Project Studio Expo 2
Friday, May 8 **12:00 – 12:45**
PSE Stage

TRACKING IN THE BOX

Presenter: **Sebastian Witkowski**, Sound Engineer, Warsaw, Poland

While mixing in the box is something that many do, having a disciplined strategic approach to tracking in the box pays enormous creative and efficiency dividends if you end up mixing a project yourself or if it will be mixed by someone else. Sebastian gives a practical presentation on a smarter way to work towards the mix.

Friday, May 8 **12:00** **Room 609B**

Technical Committee Meeting on Human Factors in Audio Systems

Tutorial 8 **Friday, May 8**
12:30 – 14:00 **Room Opera**

LISTENING TESTS—
UNDERSTANDING THE BASIC CONCEPTS

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

Listening tests and other forms of data collection methods that rely on human responses are important tools for audio professionals as these methods assist our understanding of audio quality. There are numerous examples of tests, either formally recommended and widely used, or specially devised for a single occasion. In order to understand listening tests and related methods, and also to potentially design and fully benefit from their results, some basic knowledge is required. This tutorial aims to address audio professionals without prior knowledge of listening test design and evaluation. The fundamentals of what to ask for, how to do it, whom to engage as listeners, what sort of results that may be expected, and similar issues will be covered— preferably in co-operation with the audience. The goal is to create an understanding of the basic concepts used in experimental design in order to enable audio professionals to appreciate the feasibility of listening tests.

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

Workshop 7 **Friday, May 8**
12:30 – 14:00 **Room Balowa AB**

LOUDNESS WARS: GIVE PEAKS A CHANCE

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

Panelists: *Florian Camerer*, ORF - Austrian TV, Vienna, Austria; EBU - European Broadcasting Union
George Massenburg, McGill University, Montreal, Quebec, Canada

Music production, distribution, and consumption has been caught in a vicious spiral rendering two decades of our music heritage irreversibly damaged. Today, new tracks and remastered ones typically sound worse than what could even be expected from compact cassette. As a professional society, do we just sit by and let that happen on our watch?

Florian and Thomas are at the helm of two European initiatives to reverse this spiral: EBU R 128 and EU legislation to prevent early hearing loss from listening to personal music players. That combination will soon make even the most ignorant A&R manager realize that it's futile to master music louder than –16 LUFS, as more and more platforms as well as Radio are implementing loudness normalization per default.

Spatial Audio Demo 8 — **Friday, May 8**
13:00 – 14:00 — **Room Saski**

THE PRODUCTION OF INVOLVING 3D AUDIO EXPERIENCES FOR MUSIC APPLICATIONS

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

It's one thing to move audio objects around the listener like a spaceship in a movie application. It's a totally different approach to create musical landscapes that involve the listener emotionally in 3D audio. Strategies, tools, and visions will be shown and a lot of listening examples will be provided.

Project Studio Expo 3
Friday, May 8 — **13:00 – 13:45**
PSE Stage

MIXING IN THE BOX

Presenter: **Jarek Yaro Płocica**, Jazzda Music, Warsaw, Poland

Leading mixer Jarek Yaro Plocica explains the types of project he has handled, how he approaches them, what guides his decisions on what you use and how you use it, the problems he encounter and the solutions and compromises he arrives at to solve them. A masterclass on refining processes and workflows.

Friday, May 8 — **13:00** — **Room 609B**

Technical Committee Meeting on Loudspeakers and Headphones

Student Event and Career Development
RECORDING CRITIQUES
Friday, May 8, 13:30 – 14:15, Room Belweder
Saturday, May 9, 13:00 – 14:00, Room Opera
Sunday, May 10, 12:30 – 13:30, Room Opera

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

The Student Recording Critiques are non-competitive listening sessions, designed for students to listen to their recordings and productions on a world-class playback system, and receive feedback on their work. Students are invited to bring along their mixes and have them critiqued by a panel renowned industry professionals in order to get pointers as to how they can push their skills to the next level.

It is vital that interested students sign up at Convention's the student booth immediately after the first SDA meeting and bring their work on CD, DVD, memory stick, or hard disc as clearly labelled 44.1 kHz WAVE or AIFF files.

Please note that finalists of the Student Recording Competition are excluded to submit their work to the Recording Critiques to give non-finalists a way of having their work assessed.

Session P8 — **Friday, May 8**
14:00 – 16:00 — **Room Królewski**

TRANSDUCERS—PART 1

Chair: **Finn Agerkvist**, Technical University of Denmark, Kgs. Lyngby, Denmark

14:00

P8-1 **Assessing Influence of a Headphone Type on Individualized Ear Training**—*Sungyoung Kim*,[1] *Sean Olive*[2]
[1]Rochester Institute of Technology, Rochester, NY, USA
[2]Harman International, Northridge, CA, USA

Technical ear training has been provided for two groups of engineering students. The treatment group received and conducted the training using a professional-level headphone and the control group did same training with their own consumer-level earphone or headphone. To investigate a possible influence of a headphone type, both groups took two standardized matching tests before and after 15-week technical ear training. The comparison of two test results shows that the headphone type significantly differentiated trainees' matching performance of the treatment group from the control group.
Convention Paper 9272

14:30

P8-2 **GaN Power Stage for Switch Mode Audio Amplification**—*Rasmus Overgaard Ploug*,[1] *Arnold Knott*,[1] *Søren Bang Poulsen*[2]
[1]Technical University of Denmark, Kongens Lyngby, Denmark
[2]Texas Instruments Denmark A/S, Kongens Lyngby, Denmark

Gallium Nitride (GaN) based power transistors are gaining more and more attention since the introduction of the enhancement mode eGaN Field Effect Transistor (FET), which makes an adaptation from Metal-Oxide Semiconductor (MOSFET) to eGaN based technology less complex than by using depletion mode GaN FETs. This project seeks to investigate the possibilities of using eGaN FETs as the power switching device in a full bridge power stage intended for switch mode audio amplification. A 50 W 1 MHz power stage was built and provided promising audio performance. Future work includes optimization of dead time and investigation of switching frequency versus audio performance.
Convention Paper 9273

15:00

P8-3 **Characterizing the Frequency Response of Headphones—A New Paradigm**—*Ulrich Horbach*, Harman Advanced Technology Group, Northridge, CA, USA

Traditional headphone measurements suffer from large variations if carried out on human subjects with probe microphones, and standardized couplers introduce additional biases, as concluded in a recent paper. Beyond that, there is no clear indication in literature about what the actual perceived frequency response of a headphone might be. This paper explores new measurement methods that avoid the human body as much as possible by measuring the headphone directly, in an attempt to overcome these restrictions and gain more accuracy. Design principles are described in the second part. A novel, DSP controlled, high-quality headphone is introduced that offers the ability to auto-calibrate its frequency response to the individual who is wearing it.
Convention Paper 9274

15:30

P8-4 **Improved Measurement of Leakage Effects for Circum-Aural and Supra-Aural Headphones**—*Todd Welti*, Harman International Inc., Northridge, CA, USA

Headphone leakage effects can have a profound effect on low frequency performance of headphones. A large survey, including over 2000 individual headphone measurements, was undertaken in order to compare leakage effects on test subjects and leakage effects of the same headphones measured on a test fixture. Ten different commercially available headphones were used, each measured on eight different test subjects and a test fixture with several sets of ➥

pinnae. Modifications to the pinnae were investigated to see if the leakage effects measured on the test fixture could be made to better match the real word leakage effects measured on human test subjects.
Convention Paper 9275

Project Studio Expo 4
Friday, May 8 14:00 – 14:45
PSE Stage

UNDERSTANDING MICROPHONES

Presenter: **Udo Wagner**, Microtech-Gefell, Germany

All you ever wanted to know about microphones: how they work at an electronic and mechanical level; different types of microphone and their pros and cons; how microphone types have evolved; the relative merits of old and new designs; how to look after mics and when to have them serviced.

Friday, May 8 14:00 Room 609B

Technical Committee Meeting on Sound for Digital Cinema and Television

Tutorial 9 Friday, May 8
14:15 – 15:45 Room Opera

INTERACTIVE MUSIC: PAST, PRESENT, AND FUTURE…

Presenters: **Justin Paterson**, London College of Music, University of West London, London, UK
Rob Toulson, Anglia Ruskin University, Cambridge, UK

Listeners have long been inspired to interact with commercial music and create new representations of popular releases. Vinyl offered many opportunities to reappropriate chart music, from scratching and tempo manipulation to mixing multiple songs. Nowadays, artists can engage their audience to interact with the music by offering mix stems for experimentation, a trend started by Nine Inch Nails in 2005 continuing to artists such as U2 in 2014. With the extended processing power of mobile devices, the opportunities for interactive music are limitless; both Bjork and Peter Gabriel have explored these new platforms in an interactive manner. In this session the presenters will offer a history and context of interactive music, demonstrate their own funded research activities, and discuss future possibilities.
This session is presented in association with the AES Technical Committee on Recording Technology and Practices

Workshop 8 Friday, May 8
14:15 – 15:45 Room Balowa AB

HOW ARE WE LEARNING MASTERING, TEACHING MASTERING—THE NEXT WAVE

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

Panelist: *Darcy Proper*, Wisseloord Studios, Hilversum, The Netherlands

Traditionally mastering has been learned by apprenticing. Now with the proliferation of educational resources and the evolution of affordable high quality in-the-box processing, more people are practicing mastering in more places than ever before. Teaching a young engineer to become a top flight mastering engineer can be challenging. Have you wondered: What does "Experienced Mastering Engineer" mean? What's the secret of mastering? In this workshop, seasoned mastering engineers and educators discuss how the craft is being taught and learned and how the next generation of mastering engineers will learn from their contemporaries. Topics will include what time tested practices remain essential and what is new in the discipline of mastering. Attendees of this workshop will walk away with a clearer understanding of what it takes to thrive in today's mastering market, how to assess internship/mentorship over "going solo" early in a mastering career, and how to grow/build your mastering skills in today's market.

Session P9 Friday, May 8
14:30 – 18:00 Room Belweder

PERCEPTION—PART 1

Chair: **Jürgen Herre**, International Audio Laboratories, Erlangen, Germany; Fraunhoffer Institute for Integated Circuits IIS, Erlangen, Germany

14:30

P9-1 Exposure of Music Students to Sound in Large Music Ensembles—*Maciej Jasinski*,[1] *Agnieszka Pietrzak*,[1] *Jun Ho Shin*,[1,2] *Jan Zera*[1],
[1]Warsaw University of Technology, Warsaw, Poland
[2]Kyungpook National University, Daegu, Korea

Exposure of musicians to sounds on stage has been a topic of numerous studies over the past 50 years. Nevertheless, the problem is still being researched as inconsistent conclusions have been obtained as to the risk of hearing loss among musicians. In this study exposure of music students to sound was measured during their activity as members of large ensembles: a student symphony orchestra, a wind orchestra, and a big-band. The measurements showed that critical conditions exceeding the permissible daily sound exposure level of 85 dB (A) occurred in the case of musicians playing brass, woodwinds, and percussion instruments with a high exposure of the neighboring groups of musicians directly exposed to the sound thereof.
Convention Paper 9276

15:00

P9-2 Effects of Psychoacoustical Factors on the Perception of Musical Signals in the Context of Environmental Soundscape—*Zhiyong Deng*,[1] *Jian Kang*,[2] *Aili Liu*[1]
[1]Capital Normal University, Beijing, China
[2]University of Sheffield, Sheffield, South Yorkshire, UK

Increasing attention is paid to the benefits of music or music-like signals in soundscape and the soundscape design projects. The perception and awareness of musical signals in the context of environmental soundscape has been suggested with a number of psychoacoustical factors involved. In this paper sound pressure level, noisiness, listeners' hearing training background, and interaction of the sound sources have been found to influence the perception of consonance, extraction, and awareness of the musical signals in the context of environmental soundscape. This paper also gives a brief discussion on the theoretic definition of music perception and consonance or pleasance.
Convention Paper 9277

15:30

P9-3 Directional Bands Revisited—*Rory Wallis, Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

Listening tests were undertaken as part of a comprehensive analysis of directional bands. The effects of frequency, loudspeaker position, signal duration, and bandwidth were all considered. The results confirmed the existence of directional bands for 1, 4, and 8 kHz 1/3-octave band bursts. A relationship between pitch and height was also observed, leading to the suggestion that the pitch-height effect and directional bands are part of the same localiza-

tional mechanism. Bandwidth was found to have a variable effect on localization, depending on frequency, indicating that the spectral cues used in vertical localization are not of equal bandwidth. Loudspeaker position and signal duration also had some influence on localization judgments although this was found to be somewhat erratic.
Convention Paper 9278

16:00

P9-4 The Effect of Dynamic Range Compression on Perceived Loudness for Octave Bands of Pink Noise in Relation to Crest Factor—*Mark Wendl, Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

A listening test was performed to find the changes in perceived loudness for differing crest factors of octave bands of pink noise as a result of limiting. Each octave band had a continuous and a transient sample of which both had five samples ranging from an uncompressed to a compressed with a difference of 4 dB FS crest factor calculation with increments of 1 dB FS. Two playback levels of 50 dB SPL and 70 dB SPL were used. The perceived loudness followed the RMS change within the pink noise; however certain octave bands appeared to have a non-linear relationship between loudness perception and crest factor changes.
Convention Paper 9279

16:30

P9-5 Interaction of Perceived Distance and Depth Comparing Audio Playback System and Musical Context—*Toru Kamekawa, Atsushi Marui*, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

The effects of the audio playback system and musical context were studied focusing on the perceived distance and the spatial depth. The experiments were carried out using a method of magnitude estimation comparing how near or far the combination of perceived visual and auditory event is, between two performers where one is fixed and another is moved back and forward. In the first experiment, the participants compared the difference among seven distances. In the second experiment, the participants compared direct-to-reverberant ratio (DR ratio) and differences of sound pressure level (SPL) among several musical context such as melody and accompaniment, precedence and chase of almost identical phrases, and non musical stimulus (pulse sound). The results showed that the perceived distance and depth were affected by the existence of image, DR ratio, and SPL. Furthermore these effects are different from musical context and playback system such as the existence of center, rear, and height loudspeakers.
Convention Paper 9280

17:00

P9-6 Auditory Adaptation in Spatial Listening Tasks—*Florian Klein, Stephan Werner*, Technische Universität Ilmenau, Ilmenau, Germany

This paper investigates auditory adaptation processes in spatial listening tasks for normal hearing people. The auditory adaptation process to altered auditory cues of thirteen participants is monitored and compared to their normal hearing listening performance. Binaural room impulse responses are measured for each participant and for an artificial head. Listeners are trained to artificial binaural room impulse responses in an audio-visual training task. Nine out of thirteen listeners could increase their elevation perception significantly and two of these listeners performed better with trained artificial binaural room impulse responses than with their individual measured room impulse responses regarding elevation error in the median plane. The listening test is supported by an interview that asks for externality.
Convention Paper 9281

17:30

P9-7 Discrimination of Formant Amplitude in Noise—*Tomira Rogala,[1] Piotr Sliwka[2]*
[1]Fryderyk Chopin University of Music, Warsaw, Poland
[2]Stefan Cardinal Wyszynski University, Warsaw, Poland

The paper reports the results of an experiment carried out to determine the just noticeable difference in timbre of noise. The variations of timbre were obtained through modification of the spectrum envelope of a pink noise—the formant amplitude was increased. The listeners were asked to indicate which one of three noise bursts in a trial sounded different than the remaining two. The results of listeners without musical experience were only a little worse than those obtained from tonmeister students. The skill of detecting slight changes in noise is easy to train and the jnd for formant amplitude change is very low.
Convention Paper 9282

Friday, May 8 14:30 Room 609A

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

Spatial Audio Demo 9 Friday, May 8
14:45 – 16:15 Room Saski

RECORDING MUSIC IN IMMERSIVE AUDIO

Presenters: **Morten Lindberg**, 2L (Lindberg Lyd AS), Oslo, Norway; Westerdals - Oslo School of Arts, Communication and Technology
Daniel Shores, Sono Luminus, Boyce, VA, USA; Shenandoah Conservatory Music Production and Recording Technology

In this Demo we will be playing music recorded in 9.1 immersive audio and discussing microphone techniques and production workflow.

Project Studio Expo 5
Friday, May 8 15:00 – 15:45
PSE Stage

A MIC FOR EVERY APPLICATION"

Presenter: **Helmut Wittek**, Schoeps, Germany

An essential guide to microphone use: dispelling/supporting ideas on specific mics for specific applications; why we have different mic types and what they are for; misconceptions about mics; historical attitudes towards using mics; practical technical issues about mic use; the mic 'types' that every mic cupboard should have; and examples and approaches to miking a variety of different instrument types.

Friday, May 8 15:00 Room 609B

Technical Committee Meeting on Electro Magnetic Compatibility

Session P10 Friday, May 8
16:00 – 18:00 Foyer

POSTERS: TRANSDUCERS

16:00

P10-1 Loudspeaker Systems by Linear Motion Type Piezoelectric Ultrasonic Actuators—*Daichi Nagaoka,[1]* ➡

Juro Ohga,[2,3] *Hirokazu Negishi,*[3] *Ikuo Oohira,*[4] *Kazuaki Maeda,*[5] *Kunio Oishi*[1]
[1]Tokyo University of Technology, Hachioji-shi, Tokyo, Japan
[2]Shibauro Institute of Technology, Tokyo, Japan
[3]Mix Corporation, Kanagawa, Japan
[4]Consultant, Kanagawa-ken, Japan
[5]TOA Corporation, Hyogo, Japan

The research group of authors have been developing completely new loudspeaker constructions that are driven by piezoelectric ultrasonic motors. This paper proposes two sorts of applications of piezoelectric linear actuators to both direct radiator and horn loudspeakers. A direct-radiator loudspeaker with a cone radiator driver by piezoelectric actuators shows smooth frequency characteristics in low frequency region because its radiating motion includes no significant resonance in the working frequency region. Therefore, it is useful for radiation of the lowest frequency part of audio signal. A horn loudspeaker by the same actuators works in rather moderate frequency region that is higher than the cut-off frequency of hones of ordinary size.
Convention Paper 9283

16:00

P10-2 Low Frequency Nonlinear Model for Loudspeaker Transducers—*Shaolin Wei, Tony Xie, Hunter Huang*, Guo Guang Electric Corporation (GGEC). Guangzhou, China

In this paper a nonlinear loudspeaker transducer model and its solution are presented. A simple and effective iteration procedure to obtain the solution of the nonlinear equation is proposed. This procedure is a powerful tool for determination of a periodic solution of a non-linear equation of motion. Further, the sound pressure of fundamental, second order, and third order harmonic distortion are also calculated. The solutions obtained using the present iteration method can give the directions to how to lower the second and third harmonics.
Convention Paper 9284
[This poster was presented.]

16:00

P10-3 Active Control of a String Instrument Bridge Using the Posicast Technique—*Liam B. Donovan, Andrew McPherson*, Queen Mary University of London, London, UK

This paper presents an active bridge allowing for precise audio-rate manipulation of a string's termination for the purposes of modifying string instrument timbre. The design of the bridge actuator and height sensor is discussed, and the benefits of using feedforward posicast control over a feedback compensator for controlling the dynamics of the severely underdamped bridge actuator system are established.
Convention Paper 9285

16:00

P10-4 Efficiency Optimization in Class-D Audio Amplifiers—*Akira Yamauchi, Arnold Knott, Ivan H. H. Jørgensen, Michael A. E. Andersen*, Technical University of Denmark, Kgs. Lyngby, Denmark

This paper presents a new power efficiency optimization routine for designing Class-D audio amplifiers. The proposed optimization procedure finds design parameters for the power stage and the output filter, and the optimum switching frequency such that the weighted power losses are minimized under the given constraints. The optimization routine is applied to minimize the power losses in a 130 W class-D audio amplifier based on consumer behavior investigations, where the amplifier operates at idle and low power levels most of the time. Experimental results demonstrate that the optimization method can lead to around 30% of efficiency improvement at 1.3 W output power without significant effects both on the audio performances and on the efficiency at high power levels.
Convention Paper 9286

16:00

P10-5 Investigation of Energy Consumption and Sound Quality for Class-D Audio Amplifiers Using Tracking Power Supplies—*Akira Yamauchi, Henrik Schneider, Arnold Knott, Ivan H. H. Jørgensen, Michael A. E. Andersen*, Technical University of Denmark, Kgs. Lyngby, Denmark

The main advantage of Class-D audio amplifiers is high efficiency that is often stated to be more than 90% but, at idle or low power levels the efficiency is much lower. The waste energy is an environmental concern, a concern in mobile applications where long battery operation is required and a concern in other applications where multiple amplifier channels are generating heat problems. It is found that power losses at low power levels account for close to 78% of energy consumption based on typical consumer behavior investigations. This paper investigates the theoretical limits of ideal stepless power supply tracking and its influence on power losses, audio performance, and environmental impact for a 130 W class-D amplifier. Both modeled and experimental results verify that a large improvement of efficiency can be achieved with a new challenge for a self-oscillating controller to keep the audio quality in such a system. The energy consumption may be reduced by up to 72%. The investigation is extended to a commercialized class-D amplifier as well.
Convention Paper 9287

16:00

P10-6 Soundbar System with Embedded Multichannel Digital Amplifier SoC—*Jeongil Seo,*[1] *Jae-Hyoun Yoo,*[1] *Taejin Park,*[1] *Taejin Lee,*[1] *Myunggeun Yoo,*[2] *Geunho Jang,*[2] *Jae-Hee Won,*[2] *Yeongha Choi*[2]
[1]Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea
[2]NeoFidelity, Inc., Seongnamsi, Kyunggido, Korea

This paper presents collaboration approach results between an audio signal processing algorithm and a digital amplifier structure for efficient implementation of soundbar applications. If we want to provide a virtual surround sound image with a linear loudspeaker array, high performance DSP and multichannel digital amplifiers are required. The required performance for DSP is depended on the algorithm complexity for creating a virtual surround sound image. However the final computation before generating the output loudspeaker signals (e.g., 16 channels) is generally composed of simple delay and sum computations from an input audio signal (e.g., 5.1 channels) to output multichannel loudspeaker signals. Therefore we redesigned the audio signal processing software by two-block processing, and the second block for simple delay and sum computation is implemented at the multichannel digital amplifier, which has an independent DSP core. Through computational simulation and hardware implementation, the proposed system showed equivalent performance with the conventional one block processing system.
Convention Paper 9288

16:00

P10-7 Loudspeaker Impedance Emulator for Multi Resonant Systems—*Niels Elkjær Iversen, Arnold Knott*, Technical University of Denmark, Kgs. Lyngby, Denmark

Specifying the performance of audio amplifiers is typically

done by playing sine waves into a pure ohmic load. However real loudspeaker impedances are not purely ohmic but characterized by its electrical, mechanical, and acoustical properties. Therefore a loudspeaker emulator capable of adjusting its impedance to that of a given loudspeaker is desired for measurement purposes. An adjustable RLC-based emulator is implemented with switch controlled capacitors, air gap controlled inductors, and potentiometers. Calculations and experimental results are compared and show that it is possible to emulate the loudspeaker impedance infinite baffle-, closed box-, and the multi resonant vented box-loudspeaker by tuning the component values in the proposed circuit. Future work is outlined and encouraged that the proposed impedance emulator is used as part of a control circuit in a switch-mode based impedance emulator.
Convention Paper 9289

16:00

P10-8 How "Green" Is My Amp?—*Jamie Angus*, University of Salford, Salford, Greater Manchester, UK

This paper examines the potential threat of power restriction legislation on audio power amplifier design. By considering the interaction between the amplitude distribution of audio signals with the efficiency characteristics of the different amplifier classes it shows that some of the linear classes can perform well as regards energy consumption and thus can compete with switching class D systems. Furthermore, it discusses the possibility of optimizing some of the linear amplifier classes in conjunction with the amplitude probabilities of real audio signals to effect a further reduction in average power consumption. Thus resulting in the "greenest possible" amplifier for a given class of power amplification.
Convention Paper 9290

16:00

P10-9 An Investigation into Utilizing Opto-Sensors to Function as Parts of MIDI Controllers—*Richard Corke, Andrew J. Horsburgh*, Southampton Solent University, Southampton, UK

The presented research focuses on the application of MIDI controllers utilizing opto-sensors to read and translate physical contact into controllable MIDI information. The described technology was found to provide improved interaction and degree of movement translation with other MIDI-capable devices. To demonstrate this, a projection controller using existing infrared technology will be used in conjunction with a microcontroller allowing for communication between analog and digital control signals.
Convention Paper 9291

16:00

P10-10 Investigation of Current Driven Loudspeakers —*Henrik Schneider, Finn T. Agerkvist, Arnold Knott, Michael A. E. Andersen*, Technical University of Denmark, Kgs. Lyngby, Denmark

Current driven loudspeakers have previously been investigated but the literature is limited and the advantages and disadvantages are yet to be fully identified. This paper makes use of a non-linear loudspeaker model to analyze loudspeakers with distinct non-linear characteristics under voltage and current drive. A multi tone test signal is used in the evaluation of the driving schemes since it resembles audio signals to a higher degree than the signals used in total harmonic distortion and intermodulation distortion test methods. It is found that current drive is superior over voltage drive in a 5" woofer where a copper ring in the pole piece has not been implemented to compensate for eddy currents. However the drive method seems to be irrelevant for a 5" woofer where the compliance, force factor as well as the voice coil inductance has been optimized for linearity.
Convention Paper 9292

16:00

P10-11 Design and Evaluation of Accelerometer-Based Motional Feedback—*Henrik Schneider, Emilio Pranjic, Finn T. Agerkvist, Arnold Knott, Michael A. E. Andersen*, Technical University of Denmark, Kgs. Lyngby, Denmark

The electro dynamic loudspeaker is often referred to as the weakest link in the audio chain due to low efficiency and high distortion levels at low frequencies and high diaphragm excursion. Compensating for loudspeaker nonlinearities using feedback or feedforward methods can improve the distortion and enable radical design changes in the loudspeaker that can lead to efficiency improvements. In combination this has motivated a revisit of the accelerometer-based motional feedback technique. Experimental results on an 8-inch subwoofer show that the total harmonic distortion can be significantly reduced at low frequencies and large displacements.
Convention Paper 9293

Workshop 9 **Friday, May 8**
16:00 – 18:00 **Room Opera**

OBJECTIVE EVALUATION IN SEMANTIC AUDIO ANALYSIS AND PROCESSING

Chairs: **György Fazekas**, Queen Mary University of London, London, UK
Christian Uhle, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Panelists: *Colin Raffael*, Columbia University, New York, USA
Xavier Serra, Universitat Pompeu Fabra, Barcelona, Spain
Bob Sturm, Queen Mary University of London, London, UK

Semantic audio analysis and processing facilitate methods to interact with digital audio either at the producer or consumer end of the production value chain. Techniques include note onset detection, beat tracking, chord estimation, i.e., the extraction of meaningful information from music signals, as well as interactive techniques such as the use of audio source separation to enable remixing or improved audio editing. There have been significant advancements in this field over the last few years, but in many areas of semantic audio objective evaluation remains elusive due partly to the difficulties of mapping between signal features and high-level concepts meaningful to humans, and partly to the difficulties in identifying relevant factors, such as appropriate perceptual attributes in specific user-facing applications. Other emerging and industry relevant needs include scaling up evaluation to large audio collections, support research reproducibility possibly via the use of ontologies, and the inclusion of problems related to non-Western music. This workshop will discuss cross cutting issues in the evaluation of semantic audio technologies. The discussion will be centered around case studies related to objective evaluation, discuss emerging needs in industry and academia, and aim to help practitioners in the adaptation of semantic audio techniques throughout the field of audio.

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

Project Studio Expo 6
Friday, May 8 **16:00 – 16:45**
PSE Stage

GETTING THE MOST FROM YOUR MIC AND THE PERFORMANCE

Presenter: **Mietek Felecki**, Sound Engineer, Warsaw

Mietek has experience in lots of different types of recording and in his presentation looks particularly at vocals and band recording scenarios. He explains what he found hardest when he started out and how he has made things 'easier' for himself.

Student Event and Career Development
RECORDING COMPETITION—PART 1
Friday, May 8, 16:30 – 18:30
Room Balowa AB

Presenters: **Simon-Claudius Wystrach, Steven Van Dyne**
Brecht De Man, Zach Bloomstein

The Student Recording Competition is a highlight at each convention. A panel of renowned industry professionals will critique and judge the best productions submitted by students from all around the world and make the final decision on the awards they will receive.

Finalists will present their work and detail production intentions and key decisions in front of the Convention audience, after which the judges will ask questions and give their feedback.

The Student Recording Competition Finales are an excellent opportunity for anyone interested to hear the work of the next generation of audio engineers, listen to the comments of distinguished professionals, and pick up new tricks for their own productions. Come along to this great event, learn about what your fellow students are up to and get connected with your international colleagues!

16:30 Category 4, **Sound for Visual Media**—*Judges:* Florian Camerer; Elizabeth Fauske; Kris Górski; Liz Teusch

17:30 Category 1, **Traditional Acoustic Recording** —*Judges:* Malgorzata Albinska; Ian Corbett; Ian Kagey; Morten Lindberg

Friday, May 8 16:30 Room 609A

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

Spatial Audio Demo 10 Friday, May 8
17:00 – 17:30 Room Saski

3D 9.1 FOR CREATING AN IMPRESSIVE MOVIE SOUND

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

A 3D 9.1 reproducing system provides much more possibilities to create an impressive sound of a movie in comparison to stereo and Surround. The most important improvements are more loudspeakers for creating real sound sources, an incredible envelopment, and realistic sounding soundscapes. The sound of movies contains several sound layers. Each sound of those layers can be extended from mono or stereo to 3D by using psychoacoustical rules. Those psychoacoustical rules are divided into different groups according to the above-described improvements. During the presentation, a couple of rules will be described and demonstrated with example recordings (synchronized audio and video). This 9.1 demonstration accompanies the workshop "3D 9.1 for Creating an Impressive Movie Sound."

Friday, May 8 17:00 Room 609B

Technical Committee Meeting on Network Audio Signals

Spatial Audio Demo 11 Friday, May 8
17:30 – 18:00 Room Saski

3D 9.1 FOR CREATING AN IMPRESSIVE MOVIE SOUND

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

A 3D 9.1 reproducing system provides much more possibilities to create an impressive sound of a movie in comparison to stereo and Surround. The most important improvements are more loudspeakers for creating real sound sources, an incredible envelopment, and realistic sounding soundscapes. The sound of movies contains several sound layers. Each sound of those layers can be extended from mono or stereo to 3D by using psychoacoustical rules. Those psychoacoustical rules are divided into different groups according to the above-described improvements. During the presentation, a couple of rules will be described and demonstrated with example recordings (synchronized audio and video). This 9.1 demonstration accompanies the workshop "3D 9.1 for Creating an Impressive Movie Sound."

Special Event
ORGAN CONCERT
Friday, May 8, 20:00 – 21:00
Field Cathedral of the Polish Army
Dluga Street 13/15, Warsaw

Francis Rumsey will be giving an organ recital during the 138th Convention in Warsaw at the Field Cathedral of the Polish Army.

The Field Cathedral of the Polish Army (Polish: Katedra Polowa Wojska Polskiego, also known as the Church of Our Lady Queen of the Polish Crown) is the main garrison church of Warsaw and the representative cathedral of the entire Polish Army. In the past the church served a variety of communities and roles: it used to be the church of the Collegium Nobilium and in the 19th century was also turned into a Russian Orthodox church. Currently all major military religious feasts in Warsaw are held there.

It is located on Długa Street 13/15, opposite the Heroes of the Warsaw Uprising Monument and Krasiński Square, close to Warsaw's Old Town.

Musical selections will include: Bach (from Clavierübung III, Prelude in E flat (St Anne), Christe alle welt trost, Fugue (St Anne)); Mendelssohn (Sonata VI); Karg-Elert (Nun Danket Alle Gott).

Session P11 Saturday, May 9
09:00 – 12:00 Room Belweder

SOUND LOCALIZATION AND SEPARATION

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

09:00

P11-1 Classification of Spatial Audio Location and Content Using Convolutional Neural Networks—*Toni Hirvonen*, Dolby Labortories, Stockholm, Sweden

This paper investigates the use of Convolutional Neural Networks for spatial audio classification. In contrast to traditional methods that use hand-engineered features and algorithms, we show that a Convolutional Network in combination with generic preprocessing can give good results and allows for specialization to challenging conditions. The method can adapt to, e.g., different source distances and microphone arrays, as well as estimate both spatial location and audio content type jointly. For example, with typical single-source material in a simulated reverberant room, we can achieve cross-validation accuracy of 94.3% for 40-ms frames across 16 classes (eight spatial directions, content type speech vs. music).
Convention Paper 9294

09:30

P11-2 A Theoretical Analysis of Sound Localization, with Application to Amplitude Panning—*Dylan Menzies, Filippo Maria Fazi*, University of Southampton, Southampton, Hampshire, UK

Below 700 Hz sound fields can be approximated well over a region of space that encloses the human head, using the

acoustic pressure and gradient. With this representation convenient expressions are found for the resulting Interaural Time Difference (ITD) and Interaural Level Difference (ILD). This formulation facilitates the investigation of various head-related phenomena of natural and synthesized fields. As an example, perceived image direction is related to head direction and the sound field description. This result is then applied to a general amplitude panning system and can be used to create images that are stable with respect to head direction.
Convention Paper 9295

10:00

P11-3 Audio Object Separation Using Microphone Array Beamforming—*Philip Coleman, Philip Jackson, Jon Francombe*, University of Surrey, Guildford, Surrey, UK

Audio production is moving toward an object-based approach, where content is represented as audio together with metadata that describe the sound scene. From current object definitions, it would usually be expected that the audio portion of the object is free from interfering sources. This poses a potential problem for object-based capture, if microphones cannot be placed close to a source. This paper investigates the application of microphone array beamforming to separate a mixture into distinct audio objects. Real mixtures recorded by a 48-channel microphone array in reflective rooms were separated, and the results were evaluated using perceptual models in addition to physical measures based on the beam pattern. The effect of interfering objects was reduced by applying the beamforming techniques.
Convention Paper 9296

10:30

P11-4 Limits of Speech Source Localization in Acoustic Wireless Sensor Networks—*David Ayllón, Roberto Gil-Pita, Manuel Rosa-Zurera, Guillermo Ramos-Auñón*, University of Alcalá, Alcalá de Henares, Madrid, Spain

Acoustic Wireless Sensor Networks (AWSN) have become very popular in the last years due to the drastic increment in the number of wireless nodes with microphones and computational capability. In such networks accurate knowledge of sensor node locations is often not available, but this information is crucial to process the collected data by means of array processing techniques. In this paper we consider the error in the estimation of the position of the nodes as a traditional microphone mismatch with large values, and we perform a detailed study of the effect that a large microphone mismatch has on the accuracy of TDOA- based source localization techniques.
Convention Paper 9297

11:00

P11-5 Improving Speech Mixture Synchronization in Blind Source Separation Problems—*Cosme Llerena-Aguilar, Guillermo Ramos-Auñón, Francisco J. Llerena-Aguilar, Héctor A. Sánchez-Hevia, Manuel Rosa-Zurera*, University of Alcalá, Alcalá de Henares, Madrid, Spain

The use of wireless acoustic sensor networks carry many advantages in the speech separation framework. Since nodes are separated by greater distances than a few centimeters, they can cover rooms completely, although these new distances involve certain problems to be solved. For instance, important time differences of arrival between the speech mixtures captured at the different microphones can appear, affecting the performance of classical sound separation algorithms. One solution consists in synchronizing the speech mixtures captured at the microphones. Following with this idea, we put forward in this paper a new time delay estimation method that outperforms classical methods in order to synchronize speech mixtures. The results obtained show the feasibility of using our proposal aiming at synchronizing speech mixtures.
Convention Paper 9298

11:30

P11-6 Direction of Arrival Estimation of Multiple Sound Sources Based on Frequency-Domain Minimum Variance Distortionless Response Beamforming—
Seung Woo Yu,[1] *Kwang Myung Jeon,*[1] *Dong Yun Lee,*[1] *Hong Kook Kim*[1,2]
[1]Gwangju Institute of Science and Technology, Gwangju, Korea
[2]City University of New York, New York, NY, USA

In this paper a method for estimating the direction-of-arrivals (DOAs) of multiple non-stationary sound sources is proposed on the basis of a frequency-domain minimum variance distortionless response (FD-MVDR) beamformer. First, an FD-MVDR beamformer is applied to multiple sound sources, where the beamformer weights are updated according to the surrounding environments for the reduction of the sidelobe effect of the beamformer. Then, multistage DOA estimation is performed to reduce computational complexity regarding the beam search. Finally, a median filter is applied to improve the DOA estimation accuracy. It is demonstrated that the average DOA estimation error of the proposed method is smaller than those of the methods based on conventional GCC-PHAT, MVDR-PHAT, and FD-MVDR, with lower computational complexity than that of the conventional FD-MVDR-based DOA estimation method.
Convention Paper 9299

Session P12 **Saturday, May 9**
09:00 – 12:30 **Room Królewski**

APPLICATIONS IN AUDIO

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

09:00

P12-1 Reconstruction of Mechanically Recorded Audio Signals Using White-Light Interferometry—*Khac Phuc Hung Thai,*[1,2] *Philippe Gournay,*[2] *Roch Lefebvre,*[2] *Serge Charlebois*[2]
[1]INSA Centre Val de Loire, Blois, France
[2]Université de Sherbrooke, Sherbrooke, QC, Canada

This paper presents a method to reconstruct a digital audio signal from a physical and analog sound-recording medium such as the Edison cylinder. A non-contact 3D optical profilometer based on white-light interferometry provides topographic information about small overlapping sections of the recording medium. For each of these sections, the natural curvature of the medium is compensated, grooves on the surface are detected, and short depth trajectories representing the audio signal are extracted. These trajectories are then concatenated to produce a digital audio signal that is post-processed to simulate the distinctive frequency response of an actual mechanical player. The effectiveness of this method is demonstrated on both tonal and vocal recordings using time, frequency, and time-frequency features as well as informal listening.
Convention Paper 9300

09:30

P12-2 Recognition of Hazardous Acoustic Events Employing

Parallel Processing on a Supercomputing Cluster— *Kuba Lopatka, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

A method for automatic recognition of hazardous acoustic events operating on a super computing cluster is introduced. The methods employed for detecting and classifying the acoustic events are outlined. The evaluation of the recognition engine is provided: both on the training set and using real-life signals. The algorithms yield sufficient performance in practical conditions to be employed in security surveillance systems. The specialized framework for parallel processing of multimedia data streams KASKADA, in which the methods are implemented, is briefly introduced. An experiment intended to assess outcomes of parallel processing of audio data on a supercomputing cluster is featured. It is shown that by employing supercomputing services the time needed to analyze the data is greatly reduced.
Convention Paper 9301

10:00

P12-3 Perceptual Evaluation of an Audio Film for Visually Impaired Audiences—*Mariana Lopez*, Anglia Ruskin University, Cambridge, UK

This paper explores a format of sonic art referred to as *audio film* that was developed to study different ways in which film sound production and postproduction techniques could be applied to the enhancement of Audio Description (AD) for visually impaired film and television audiences. A prototype of this format was tested with a group of nine volunteers with sight loss in order to test its effectiveness. The perceptual evaluation demonstrated the potential of this format for conveying a clear narrative as well as providing an entertaining experience. Future work will include the investigation of conventions to indicate scene changes within audio-only formats as well as studying the impact of object-based mixing on *audio films*.
Convention Paper 9302

10:30

P12-4 Reproduction of Realistic Background Noise for Testing Telecommunications Devices—*Juan David Gil Corrales,*[1] *Wookeun Song,*[2] *Ewen MacDonald*[1]
[1]Technical University of Denmark, Lyngby, Denmark
[2]Brüel & Kjær Sound and Vibration Measurement A/S, Nærum, Denmark

A method for reproduction of sound, based on crosstalk cancellation using inverse filters, was implemented in the context of testing telecommunications devices. The effect of the regularization parameter, number of loudspeakers, type of background noise, and a technique to attenuate audible artifacts, were investigated. The quality of the reproduced sound was evaluated both objectively and subjectively with respect to the reference sounds, at points where telecommunications devices would be potentially placed around the head. The highest regularization value gave the best results, the performance was equally good when using eight or four loudspeakers, and the reproduction method was shown to be robust for different program materials. The proposed technique to reduce audible artifacts increased the perceived similarity.
Convention Paper 9303

11:00

P12-5 Simulation of Parameters of Tube Audio Circuits Using Web Browsers—*Grzegorz Makarewicz*, Warsaw University of Technology, Warsaw, Poland

The paper describes the program/simulator for computer-aided design of audio amplifiers using electron tubes. It was developed in JavaScript scripting language and thanks to its embedding in a web browser it does not require installation on the user's computer. The simulator can be used for the design and education purposes, without any limitations, by multiple users simultaneously. It is based on mathematical models of the triode and pentode and allows parameters of electron tubes as well as the most important parameters of tube amplifiers to be simulated for both unbalanced (single-ended) and balanced (push-pull) configuration.
Convention Paper 9304

11:30

P12-6 Improvements and User Preferences in Auralization for Multi-Party Teleconferencing Systems Using Binaural Audio—*Emanuel Aguilera, Jose J. Lopez, Pablo Gutierrez-Parera*, Universitat de Valencia, Valencia, Spain

The introduction of spatial audio in multi-party teleconferencing systems create realistic communication environments with increased immersion compared to monaural systems. Moreover, the introduction of auralization effects can increase even more the immersion but at the expenses of a reduced intelligibility. In this paper we study the influence of some specific auralization processing details for a trade-off between realism and intelligibility. Our own spatial multi-party teleconferencing software running on smartphones and tablets has been developed for carrying out different subjective experiments. By means of subjective testing with a jury, it has been evaluated the influence in immersion, intelligibility, and user preferences in relation with early echoes, late reverberation, and the introduction of simple near-field HRTF processing when audio sources are very close to the user. Results provide interesting guidelines for developing teleconference systems with more subtle auralization and HRTF effects.
Convention Paper 9305

12:00

P12-7 Subjective Assessment of Commercial Sound Enhancement System—*Krzysztof Brawata, Pawel Malecki, Adam Pilch, Tadeusz Kamisinski*, AGH University of Science and Technology, Krakow, Poland

Sound enhancement systems are becoming more and more popular even in very sophisticated concert halls. Especially in places with some acoustics deficiencies, musicians and concert-goers have accepted that kind of solution as very natural sounding and of great possibilities to easily obtain variable acoustics in rooms. On the market, there are some sound enhancement commercial systems, with similar acoustical parameters. Choosing the best one, for defined application, is possible only on the basis of properly designed listening tests. In the paper subjective listening tests of two sound enhancement systems installed in the same room are presented. On the basis of listeners' evaluation, the quality of stage acoustics, naturalness, and spaciousness of sound created by systems were analyzed.
Convention Paper 9306

Tutorial 10 **Saturday, May 9**
09:00 – 11:00 **Room Opera**

MODERN DIGITAL TO ANALOGUE CONVERSION: AUDIO ALCHEMY USING SIGNAL PROCESSING

Presenter: **Jamie A. S. Angus**, University of Salford, Salford, Greater Manchester, UK

Almost all modern digital to analogue convertors (DACs) use over-sampled multibit convertors with noise-shaping to achieve their high performance. Oversampling and noise-shaping allow one to use a DAC with a small number of levels, which is easier to manufacture. Unfortunately traditional noise shaping does nothing to reduce the effect of component tolerances in the DAC because the analogue output cannot be fed back to the input. However, modern DACs do manage to noise shape the output from the DAC without any feedback. This piece of audio alchemy is critical to the exceptional performance of modern digital-to-analogue convertors. Using audio examples, this tutorial will explain how this alchemy is achieved. It will review the problems of component tolerance in DACs and show how they compromise performance. Then noise-shaping, and how it can be applied, without magic or knowing the actual converted output, to a practical DAC will be explained.

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

Workshop 10 **Saturday, May 9**
09:00 – 11:00 **Room Balowa AB**

WORLD CLASS FILM SOUND MIXERS

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

Panelists: *Kacper Habisiak*, Dreamsound, Warsaw, Poland
Marcin Kasinski, Dreamsound, Warsaw, Poland
Filip Krzemien, Dreamsound, Warsaw, Poland

Three of Poland's leading cinema sound mixers discuss their craft—for both the professional doing Sound for Picture as well as students thinking about this for a career.

This session is presented in association with the AES Technical Committee on Sound for Digital Cinema and Television

Student Event and Career Development
STUDENT DESIGN EXHIBITION
Saturday, May 9, 09:00 – 11:00
Foyer

At the Student Design Exhibition, participants of the Student Design Competition will have the opportunity to show off their work to an audience of audio students, educators, and professionals. Anyone attending the Convention is invited to come along.

Apart from being an educational event the Student Design Exhibition also is an invaluable career-building event and a great place for companies to identify prospective employees.

The judges of the Student Design Competition will also roam the Exhibition floors and get in touch with the exhibiting students to make their final decision on who is going to be awarded for their outstanding projects.

Students with both audio and non-audio backgrounds are encouraged to submit their hardware and software projects. Few restrictions are placed on the nature of the submissions but designs must focus on audio applications. Examples include loudspeaker design, DSP plug-ins, analogue hardware, signal analysis tools, mobile applications, and sound synthesis devices.

Products should represent new, original ideas implemented in working-model prototypes.

Saturday, May 9 **09:00** **Room 609A**

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

Spatial Audio Demo 12 **Saturday, May 9**
09:30 – 10:30 **Room Saski**

PSYCHOACOUSTICS OF 3D SOUND RECORDING

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

This tutorial will explain psychoacoustic principles that need to be considered when making 3D sound recordings and demonstrate a number of practical 3D recordings made using various microphone configurations to show how the principles are applied into practice. The topics to be addressed include: vertical interchannel time and level relationship, the effect of decorrelation/mic spacing in vertical stereo perception, the role of spectral cue for vertical localization and image spread, and effective microphone configurations for the capturing of height channel ambience. 3D recordings to be demonstrated will be of various types of music: orchestra, chamber choir, string ensemble, organ, piano, jazz funk, etc.

Project Studio Expo 7
Saturday, May 9 **10:00 – 10:45**
PSE Stage

UNDERSTANDING MONITORS

Presenter: **Jens Schönemann-Paul**, Dynaudio, Denmark

All you ever wanted to know about monitors: how loudspeakers work; the different types of driver used; 2-ways, 3-ways, etc., active and passive, ports and sealed; mono, stereo, and multichannel considerations; subs; hifi versus pro; what does "reference" mean; why engineers use different types and sizes of monitor; and power and size of monitors in relation to room size.

Spatial Audio Demo 13 **Saturday, May 9**
10:30 – 11:30 **Room Saski**

PSYCHOACOUSTICS OF 3D SOUND RECORDING

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

This tutorial will explain psychoacoustic principles that need to be considered when making 3D sound recordings and demonstrate a number of practical 3D recordings made using various microphone configurations to show how the principles are applied into practice. The topics to be addressed include: vertical interchannel time and level relationship, the effect ofdecorrelation/mic spacing in vertical stereo perception, the role of spectral cue for vertical localization and image spread, and effective microphone configurations for the capturing of height channel ambience. 3D recordings to be demonstrated will be of various types of music: orchestra, chamber choir, string ensemble, organ, piano, jazz funk, etc.

Project Studio Expo 8
Saturday, May 9 **11:00 – 11:45**
PSE Stage

OPTIMIZING MONITORS FOR ROOMS

Presenter: **Chris Anet**, Genelec, Finland

The majority of the audible problems in monitoring are caused by the complex interaction between monitors and room. If a monitor produces a flat response in anechoic conditions, once it is placed in a listening room the response changes because of room boundary loading, reflections, reverberation time characteristics, etc. So, precise adjustment of the monitor's response is needed so that any detrimental monitor-room interaction is minimized to achieve a flat and uncolored frequency response at the listening position.

Saturday, May 9 **11:00** **Room 609B**

Technical Committee Meeting on Recording Technology and Practices

Tutorial 11 **Saturday, May 9**
11:15 – 12:45 **Room Opera**

TIME REVERSAL IN SOUND RECORDINGS

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

A survey of the many moments in pop music in which sounds have been presented backwards—ending where they started, starting where they ended, and un-decaying in between—reveals a long history of reveling in recorded reversals. While such sounds can't happen in nature, they are a common studio creation, possessing strong musical value and delivering intriguing technical and perceptual advantages.

Tutorial 12 **Saturday, May 9**
11:30 – 13:30 **Room Balowa AB**

MAIN MICROPHONE TECHNIQUES FOR 2.0, 5.1, AND 3-D AUDIO

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

This tutorial will explain step-by-step, using many practical examples, what a suitable stereophonic microphone array can look like. With 2.0 stereo setups as the starting point, multichannel, and 3-D arrays will also be introduced.

Many factors influence the choice of a stereophonic microphone setup, but the relevance of these factors can vary greatly depending on the application, such that there is never one single “correct” setup. Knowledge of various options gives a Tonmeister the ability to make optimal choices.

With the participation of Mike Williams.

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

Project Studio Expo 9
Saturday, May 9 **12:00 – 12:45**
PSE Stage

IMPLEMENTING PRACTICAL ACOUSTICS

Presenter: **Donato Masci**, Studio Sound Service, Italy

Acoustics for your room. Acoustician Donato takes a “typical” ordinary room, describes it, explains how he appraises it, discuses isolation issues, and then identifies the biggest acoustic problems and explains how to address them in stages of improvement all the way to doing the job “properly.”

Saturday, May 9 **12:00** **Room 609B**

Technical Committee Meeting on Spatial Audio

Tutorial 13 **Saturday, May 9**
12:15 – 13:45 **Room Belweder**

LOUDNESS 101—LUFS IS IN THE AIR

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

This session will bring participants up to speed regarding the latest 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! Be ready to sweat!

Saturday, May 9 **12:30** **Room 609A**

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

Saturday, May 9 **13:00** **Room 609B**

Technical Committee Meeting on Perception and Subjective Evaluation of Audio Signals

Spatial Audio Demo 14 **Saturday, May 9**
13:30 – 15:00 **Room Saski**

RECORDING MUSIC IN IMMERSIVE AUDIO

Presenters: **Morten Lindberg**, 2L (Lindberg Lyd AS), Oslo, Norway; Westerdals - Oslo School of Arts, Communication and Technology
Daniel Shores, Sono Luminus, Boyce, VA, USA; Shenandoah Conservatory Music Production and Recording Technology

In this Demo we will be playing music recorded in 9.1 immersive audio and discussing microphone techniques and production workflow.

Workshop 11 **Saturday, May 9**
14:00 – 15:30 **Room Balowa AB**

WHISPERS AND SCREAMS OF THE XII. MUSE— A WORD ABOUT AUDIO IN GAMES

Presenters: **Daniel Kleczynski**
Maciej Miasik

This talk reviews the development of audio in games, giving a perspective against audio development in other media by reviewing the following topics: • Video games as a discipline of art; • The influence of pro-audio on game development; • Aesthetic, technical, and logical aspects of game sound design; • Differences between game audio and pro-audio; • The future of game audio.

This is designed to be informative to anyone working in audio engineering.

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

Student Event and Career Development
EDUCATION/CAREER FAIR
Saturday, May 9, 14:00 – 15:30
Foyer

One of the best reasons to attend AES conventions is the opportunity to make meaningful connections in the world of audio. The Education & Career Fair offers a place for students to get in touch with representatives from audio schools and institutions, and job seekers to talk to companies searching new talent.

Looking for the best and brightest minds in the audio world? No place will have more of them assembled than the 138th Convention of the Audio Engineering Society.

Academic Institutions offering courses in the field of audio will be represented and information on each school’s respective programs will be displayed.

All attendees of the Convention—students, educators, and professionals alike—are invited to find out more about job and internship opportunities in the audio industry and educational programs. Bring your business cards and resumes!

Companies and educational institutions are invited to participate and can register for an exhibition table free of charge.

Project Studio Expo 10
Saturday, May 9 **14:00 – 14:45**
PSE Stage

INTERFACING THE DAW TO THE HARDWARE

Presenter: **Dave Hill**, Crane Song, USA

It's a hardware world that has to work with a DAW environment. Dave discusses the specifics of working out of the box, dispelling some of the myths and explaining what really matters and what matters less.

Session P13 **Saturday, May 9**
14:30 – 18:00 **Room Belweder**

PERCEPTION—PART 2

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

14:30

P13-1 **Elicitation of the Differences between Real and Reproduced Audio**—*Jon Francombe, Tim Brookes, Russell Mason*, University of Surrey. Guildford, Surrey, UK

To improve the experience of listening to reproduced audio, it is beneficial to determine the differences between listening to a live performance and a recording. An experiment was performed in which three live performances (a jazz duet, a jazz-rock quintet, and a brass quintet) were captured and simultaneously replayed over a nine-channel with-height surround sound system. Experienced and inexperienced listeners moved freely between the live performance and the reproduction and described the difference in listening experience. In subsequent group discussions, the experienced listeners produced twenty-nine categories using some terms that are not commonly found in the current spatial audio literature. The inexperienced listeners produced five categories that overlapped with the experienced group terms but that were not as detailed.
Convention Paper 9307

15:00

P13-2 **Towards Unification of Methods for Speech, Audio, Picture, and Multimedia Quality Assessment**—
Slawomir Zielinski,[1] *Francis Rumsey*,[2] *Søren Bech*[3,4]
[1]Bialystok University of Technology, Bialystok, Poland
[2]Logophon Ltd., Oxfordshire, UK
[3]Bang & Olufsen a/s, Struer, Denmark
[4]Aalborg University, Aalborg, Denmark

The paper addresses the need to develop unified methods for subjective and objective quality assessment across speech, audio, picture, and multimedia applications. Commonalities and differences between the currently used standards are overviewed. Examples of the already undertaken research attempting to "bridge the gap" between the quality assessment methods used in various disciplines are indicated. Prospective challenges faced by researchers in the unification process are outlined. They include development of unified scales, defining unified anchors, integration of objective models, maintaining "backward comparability," and undertaking joint standardization efforts across industry sectors.
Convention Paper 9308

15:30

P13-3 **An Investigation of the Relationship between Listener Envelopment and Room Acoustic Parameters– The Influence of Varied Direct Sound Levels and Onset Times of Late Reverberation on Listener Envelopment**—*Mai Ishida, Toru Kamekawa, Atsushi Marui*, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

It is proposed that the late reflection energy from 80 ms after the direct sound contributes to listener envelopment (LEV). According to previous research, 80 ms is not necessarily the most suitable onset-time of late reverberation for predicting LEV. In addition to this, LEV tends to increase if C80 decreases. However, it is possible for LEV to be low for very low early energy with relatively higher late energy. In this study the LEV for stimuli of varied direct sound levels and onset-times of late reverberation was investigated. As a result, it is suggested that as the direct sound level increases, the LEV increases. Additionally, we confirmed that LEV increased as C10 increased.
Convention Paper 9309

16:00

P13-4 **The Development of a Sound Wheel for Reproduced Sound**—*Torben H. Pedersen, Nick Zacharov*, DELTA SenseLab, Hørsholm, Denmark

Sound quality is an important aspect in many sound reproduction applications. In recent years sensory evaluation techniques have been gaining popularity for the detailed perceptual assessment of device sound characteristics. From the literature, hundreds of descriptors can be found to describe the nature of sound quality and this often becomes the focus of debate among researchers, rather than the product development itself. In an effort to shift the focus back to the areas of importance, i.e., the product, this study seeks to define a common terminology, a lexicon, for the characterization of sound quality in loudspeakers, headphones, or other sound reproduction systems. The study summarized the gathering of descriptors for sound character from the literature and then experimental leading to a structure protocol of perceptual sound quality attributes for this domain of application. A structured sound wheel is presented comprising of different layers of attributes. For each attribute, definitions have been developed with associated sound samples for training. The paper presents the on-going development process, including validation of attributes and their definitions.
Convention Paper 9310

16:30

P13-5 **How Much Is the Use of a Rating Scale by a Listener Influenced by Anchors and by the Listener's Experience?** —*Nadja Schinkel-Bielefeld*,[1,2] *Anna Katharina Leschanowsky*[1]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]International Audio Laboratories, Erlangen, Germany

It has been postulated that anchors in multi-stimulus listening tests for audio quality evaluation should have an item-independent quality, as listeners will likely shift their rating scale if the quality of the anchor varies. However, expert listeners have a very stable internal rating scale, which can be seen from the repeatability of their results when performing the same test multiple times. So they may stick to their usual scale even if the anchor varies. We find that listeners do not shift their rating scale by the full amount the anchor is shifted but only up to 60% of that. Nevertheless this makes quantitative comparisons between different test results difficult even if the anchor varies only by 5 Mushra points.
Convention Paper 9311

17:00

P13-6 **Quantifying Auditory Perception: Dimensions**

of Pleasantness and Unpleasantness—*Judith Liebetrau,*[1] *Thomas Sporer,*[1] *Marius Becker,*[2] *Thanh Phong Duong,*[2] *Andreas Ebert,*[2] *Martin Härtig,*[2] *Phillip Heidrich,*[2] *Jakob Kirner,*[2] *Oliver Rehling,*[2] *Dominik Vöst,*[2] *Roberto Walter,*[2] *Michael Zierenner,*[2] *Tobian Clauß*[1]
[1]Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany
[2]Technische Universität Ilmenau, Ilmenau, Germany

Psychoacoustic attributes like roughness, sharpness, tonality, and fluctuation strength are often used to explain and calculate major properties of sound. While in general these properties are well understood, the concept of "pleasantness" depends on several factors. Investigating the underlying dimensions of pleasantness was the goal of the presented studies. The perception of psychoacoustic attributes was assessed for 22 different audio stimuli by more than 15 listeners. In addition the perception of pleasantness and unpleasantness for these items was evaluated. All tests were conducted in laboratory as well as home environment. The results showed that a link between psychoacoustic attributes and concept of pleasantness could be established. Surprisingly, the relation between the single attributes and pleasantness changed dependent on the applied analysis method.
Convention Paper 9312

17:30

P13-7 Audio Quality Moderates Localization Accuracy: Two Distinct Perceptual Effects?—*PerMagnus Lindborg, Nicholas A. Kwan*, Nanyang Technological University, Singapore

Audio quality is known to cross-modally influence reaction speed, sense of presence, and visual quality. We designed an experiment to test the effect of audio quality on source localization. Stimuli with different MP3 compression rates, as a proxy for audio quality, were generated from drum samples. Participants (n = 18) estimated the position of a snare drum target while compression rate, masker, and target position were systematically manipulated in a full-factorial repeated-measures experiment design. Analysis of variance revealed that location accuracy was better in wide target positions than in narrow, with a medium effect size; and that the effect of target position was moderated by compression rate in different directions for wide and narrow targets. The results suggest that there might be two perceptual effects at play: one, whereby increased audio quality causes a widening of the soundstage, possibly via a SMARC-like mechanism, and two, whereby it enables higher localization accuracy. In the narrow target positions in this experiment, the two effects acted in opposite directions and largely cancelled each other out. In the wide target presentations, their effects were compounded and led to significant correlations between compression rate and localization error.
Convention Paper 9313

Session P14 **Saturday, May 9**
14:30 – 17:30 **Room Królewski**

TRANSDUCERS—PART 2

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

14:30

P14-1 High Power Efficiency and Broad Flat Radiation Bandwidth of a Parametric Array PMUT Loudspeaker—*Kyounghun Been,*[1] *Younghwan Hwang,*[1] *Yub Je,*[2] *Haksue Lee,*[2] *Wonkyu Moon*[1]
[1]POSTECH, Gyeongsangbuk-do, Korea
[2]Agency for Defense Development, Gyeongsangnam-do, Korea

Parametric Array loudspeakers can generate a sound beam using nonlinear acoustic interactions, widely known as "Parametric Array," that can enable private listening in a public area. Parametric array loudspeakers can be applied to many applications, such as information technology devices, that require a high power efficiency and wide bandwidth. In a previous study, a piezoelectric micromachined ultrasonic transducer (PMUT) is shown to be an efficient unit for a parametric array loudspeaker. In this paper we will describe realization of a parametric array loudspeaker with high power efficiency (up to 71%) and wide flat radiation bandwidth (19.5 kHz, difference frequency wave with equalization), which consists of an array of PMUTs with two resonance frequencies (f1 = 100 kHz, f2 = 110 kHz) and use of "out-of-phase" driving techniques.
Convention Paper 9314

15:00

P14-2 Slit-Firing Sound Plate design with Slim Elliptical Speaker—*Gyeong-Tae Lee, Jong-Bae Kim, Seong-Ha Son*, Samsung Electronics Co. Ltd., Suwon, Gyeonggi-do, Korea

Slim design has emerged recently as a new form factor for the speaker system of an electronic device. However, this form factor is disadvantageous to sound performance because of the narrow space for a speaker unit. In this paper, to overcome this drawback, we designed Slit-Firing Sound Plate, which is a slim speaker system using diffraction through a slit, and developed a slim elliptical speaker that is optimized for the boundary conditions for the slit. After building and tuning a prototype, the sound performance of the prototype was assessed by measuring and examining frequency response, transient response, and directivity beam pattern. As a result, proposed novel design shows high performance that makes it suitable for the application of slim electronic devices.
Convention Paper 9315

15:30

P14-3 Improvements in Elimination of Loudspeaker Distortion in Acoustic Measurements—*Finn T. Agerkvist,*[1] *Antoni Torras-Rosell,*[2] *Richard McWalter*[1]
[1]Technical University of Denmark, Kgs. Lyngby, Denmark
[2]Danish National Metrology Institute, Lyngby, Denmark

This paper investigates the influence of nonlinear components that contaminate the linear response of acoustic transducers and presents improved methods for eliminating the influence of nonlinearities in acoustic measurements. The method is evaluated with pure sinusoidal signals as well as swept sine signals and is tested on models of memoryless nonlinear systems as well as nonlinear loudspeakers. The method is shown to give a clear benefit over existing methods. Two techniques that improve the signal to noise ratio are demonstrated: the first uses more measurement level than the number of orders to be separated, whereas the other one is based on standard Tikhonov regularization. Both methods are shown to significantly improve the signal to noise ratio.
Convention Paper 9316

16:00

P14-4 Flux Modulation in the Electrodynamic Loudspeaker—*Morten Halvorsen,*[1] *Carsten Tinggaard,*[1] *Finn T. Agerkvist*[2]
[1]PointSource Acoustics, Roskilde, Denmark
[2]Technical University of Denmark, Kgs. Lyngby, Denmark

This paper discusses the effect of flux modulation in the electrodynamic loudspeaker with main focus on the effect on the force factor. A measurement setup to measure the AC flux modulation with static voice coil is explained and the measurements show good consistency with FEA simulations. Measurements of the generated AC flux modulation shows that eddy currents are the main source to magnetic losses in form of phase lag and amplitude changes. Use of a copper cap shows a decrease in flux modulation amplitude at the expense of increased power losses. Finally, simulations show that there is a high dependency between the generated AC flux modulation from the voice coil and the AC force factor change.
Convention Paper 9317

16:30

P14-5 Validation of Power Requirement Model for Active Loudspeakers—*Henrik Schneider, Anders N. Madsen, Ruben Bjerregaard, Arnold Knott, Michael A. E. Andersen*, Technical University of Denmark, Kgs. Lyngby, Denmark

The actual power requirement of an active loudspeaker during playback of music has not received much attention in the literature. This is probably because no single and simple solution exists and because a complete system knowledge from input voltage to output sound pressure level is required. There are, however, many advantages that could be harvested from such knowledge like size, cost, and efficiency improvements. In this paper a recently proposed power requirement model for active loudspeakers is experimentally validated and the model is expanded to include the closed and vented type enclosures in addition to the main loudspeaker non-linearities.
Convention Paper 9318

17:00

P14-6 Subwoofers in Rooms: Stereophonic Reproduction—
Juha Backman, Microsoft, Espoo, Finland

A study based on computational model of interaural level and time differences at the lowest audio frequencies, often reproduced through subwoofers, is presented. This work studies whether interaural differences can exist, and if they do, what kind of relationship there is between the loudspeaker direction and the interaural differences when monophonic and stereophonic subwoofer arrangements are considered. The calculations are made for both simple amplitude panned signals and for simulated microphone signals. The results indicate that strong narrow-band differences can exist, especially near room eigenfrequencies when the listener is close to nodes of the room modes and that the modes of the recording room can have an effect on the sound field of the listening room. In addition to the computational results an analysis of interchannel level differences in recordings is presented, confirming the computational model.
Convention Paper 9319

Saturday, May 9 14:30 Room 609A

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

Project Studio Expo 11
Saturday, May 9 15:00 – 15:45
PSE Stage

CONVERSION: THE INS AND THE OUTS

Presenter: **Michal Jurewicz**, Mytek, USA

A look at the balance of working analogue alongside digital with particular emphasis on the role of conversion. Michal explains the technicalities, realities, and myths concerning conversion and gives pointers on what is important.

Spatial Audio Demo 15 Saturday, May 9
15:30 – 17:00 Room Saski

MAKING 3D RECORDINGS OF CLASSICAL MUSIC

Presenter: **Malgorzata Albinska-Frank**, Tonstudio arton, Das Tonstudio für klassische Musik, Basel, Switzerland

This session is dedicated to all who would like to start to record classical music using 3D techniques: room acoustics, microphone techniques, mixing tools, and playback systems for working with 3D. Is this technique also useful for surround and stereo? Based on listening to 3D recordings of classical music, the technical and aesthetical aspects of 3D recording production workflow will be presented and discussed.

Workshop 12 Saturday, May 9
15:45 – 16:45 Room Opera

MOBILE PLATFORMS, HEARING LOSS AND LOUDNESS (IPOD EXPOSURE)

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

Millions of people are estimated to be at the risk of developing early hearing loss (HL) as a result of listening to so-called personal media players (PMPs). Legislation now mandated by the European Commission will have a profound influence on the audio delivery to mobile platforms, and therefore impact the broadcast industry at large. This presentation summarizes the current situation, and it discusses a next phase of safety standards. New methods would not only protect hearing better, but also help defuse the loudness wars in music. Maybe, after all there will be some music heritage worth listening to from our time.

Student Event and Career Development
RECORDING COMPETITION—PART 2
Saturday, May 9, 15:45 – 17:45
Room Balowa AB

Presenters: **Simon-Claudius Wystrach, Steven Van Dyne Brecht De Man, Zach Bloomstein**

The Student Recording Competition is a highlight at each convention. A panel of renowned industry professionals will critique and judge the best productions submitted by students from all around the world and make the final decision on the awards they will receive.

Finalists will present their work and detail production intentions and key decisions in front of the Convention audience, after which the judges will ask questions and give their feedback.

The Student Recording Competition Finales are an excellent opportunity for anyone interested to hear the work of the next generation of audio engineers, listen to the comments of distinguished professionals, and pick up new tricks for their own productions. Come along to this great event, learn about what your fellow students are up to and get connected with your international colleagues!

15:45 Category 2, **Traditional Studio Recording**—*Judges:* Richard King; Andrzej Lipinski; Mandy Parnell; Darcy Proper

16:45 Category 3, **Modern Studio Recording & Electronic Music**—*Judges:* Alex Case; Shawn Everett; Ronald Prent; Marek Walaszek

Project Studio Expo 12
Saturday, May 9 15:45 – 16:15
PSE Stage

USING HARDWARE WITH YOUR DAW

Presenter: **Marek Walaszek**, Bettermaker, Poland

Working analogue alongside your DAW can present issues. Marek discusses the specifics of working out of the box, the problems that have to be addressed, ways of getting around them, and tips and tricks that users will appreciate.

Session EB1 **Saturday, May 9**
16:00 – 18:00 **Foyer**

ENGINEERING BRIEFS—POSTERS

16:00

EB1-1 Preferred Sound Level for Concert Listeners and Correlations between Sound Quality Dimensions— *Avo-Rein Tereping*, Tallinn University, Institute of Psychology, Tallinn, Estonia

Increasing loudness at concert performances is not caused by listeners' preferences, but by the opinions of sound engineers and/or producers. The average sound level at public concerts ranges up to 100–105 decibels. Loudness preferences have been examined for listening with earphones but not in the open air. This e-brief describes research on preferred sound quality dimensions at Nordea Concert Hall in Tallinn with live music samples. The experiments revealed that preferred loudness don't differ across age groups or between women and men. The preferred loudness were found 85–87 dB. Fidelity was the most important sound quality parameter influencing to overall pleasantness. No correlation were found between loudness and overall pleasantness.
Engineering Brief 183

16:00

EB-1-2 Evaluation of a Novel Approach to Virtual Bass Synthesis Strategy—*Piotr Hoffmann, Bozena Kostek*, Gdansk University of Technology, Gdansk, Poland

The aim of this paper is to present a novel approach to the Virtual Bass Synthesis (VBS) strategy applied to portable computers. The developed algorithms involve intelligent, rule-based settings of bass synthesis parameters with regard to music genre of an audio excerpt and the type of a portable device in use. The Smart VBS algorithm performs the synthesis based on a nonlinear device (NLD) with artificial controlling synthesis system according to music genre. Classification of musical genres is performed using the k-Nearest Neighbor algorithm and the extracted MPEG 7-based feature vectors optimized by the Principal Component Analysis method. To confirm the relationship between the presented excerpt of music from a variety of music genres and the listener's preferences, subjective tests using the Mushra method are performed. On the basis of the listeners' opinions statistical tests are carried out and show that listeners in most cases prefer the SVBS strategy developed by the authors in comparison to either an audio excerpt with the bass boost algorithm applied and unprocessed audio file. Furthermore, the listeners indicated that perception of the proposed SVBS strategy is similar for different types of portable devices.
Engineering Brief 184

16:00

EB-1-3 Analogue Hearts, Digital Minds? An Investigation into Perceptions of the Audio Quality of Vinyl —*Michael Uwins*, De Montfort University, Leicester, UK; University of Huddersfield, Huddersfield, UK

This study investigates the vinyl revival, with particular focus given to the listener's perception of audio quality. A new album was produced using known source material. Subjects then participated in a series of double-blind listening tests, comparing vinyl to established digital formats. Subsequent usability tests required subjects not only to re-appraise the audio, but also to interact with the physical media and playback equipment. Digital vinyl systems were used in order to investigate the influence of non-auditory factors on their perception of sound quality. Both qualitative and quantitative data was also gathered from subjects of the usability tests, with the correlation (or contradiction) between the results being analyzed. The study concludes that sound quality is not the sole defining factor and that listener preferences are profoundly influenced by other, non-auditory attributes and that such factors are as much a part of the vinyl experience as the music etched into the grooves.
Engineering Brief 185

16:00

EB-1-4 Toward the Development of a Universal Listening Test Interface Generator in Max—*Christopher Gribben, Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

This engineering brief describes HULTI-GEN (Huddersfield Universal Listening Test Interface Generator), a Cycling '74 Max-based tool. HULTI-GEN is a user-customizable environment that takes user-defined parameters (e.g., the number of trials, stimuli, and scale settings) and automatically constructs an interface for comparing auditory stimuli, while also randomizing the stimuli and trial order. To assist the user, templates based on ITU-R recommended methods have been included. As the recommended methods are often adjusted for different test requirements, HULTI-GEN also supports flexible editing of these presets. Furthermore, some existing techniques have been summarized within this brief, including their restrictions and how they might be altered through using HULTI-GEN. A finalized version of HULTI-GEN is to be made freely available online at: http://www.hud.ac.uk/research/researchcentres /mtprg/projects/apl/
Engineering Brief 187

16:00

EB-1-5 Multidimensional Ability Evaluation of Participants of Listening Tests: Part II—*Tomasz Dziedzic, Piotr Kleczkowski*, AGH University of Science and Technology, Krakow, Poland

The problem of selecting appropriate participants for a listening test was addressed by the authors' e-brief presented during the 136th AES Convention. After the Convention, the project was continued and the application was further developed. The next stage consisted in the comparison of resolutions in the time and frequency domains for particular listeners. The resolution in time was investigated by a task for gap detection in noise. The resolution in frequency was tested by two tasks: a two-tone detection and distinguishing pitches of two tones. The results showed that there was no correlation between these resolutions (neither negative nor positive), but considerable frequency dependent differences were found. A new version of the test application was developed for this stage.
Engineering Brief 188

16:00

EB-1-6 Database of Single-Channel and Binaural Room Impulse Responses of a 64-Channel Loudspeaker Array—*Vera Erbes*,[1] *Matthias Geier*,[1] *Stefan Weinzierl*,[2] *Sascha Spors*[1]
[1]University of Rostock, Rostock, Germany
[2]Technical University of Berlin, Berlin, Germany

A freely available database of measured single-channel and binaural room impulse responses (RIRs and BRIRs) of a 64-channel loudspeaker array of rectangular shape under varying room acoustical conditions is presented. The RIRs have been measured at three receiver positions for four different absorber configurations. Corresponding BRIRs for head-orientations in the range of ±80° in 2° steps with a KEMAR manikin have been captured for a subset of seven combinations of position and absorber configurations. The data is provided in the Spatially Oriented Format for Acoustics (SOFA). It can be used to study the influence of the listening room on multichannel audio reproduction. As an application RIRs for the synthesis of a sound field by Wave Field Synthesis are shown.
Engineering Brief 189

16:00

EB-1-7 HAART: A New Impulse Response Toolbox for Spatial Audio Research—*Dale Johnson, Alex Harker, Hyunkook Lee*, University of Huddersfield - Huddersfield, UK

This engineering brief describes a new, open source code library named HAART (Huddersfield Acoustical Analysis Research Toolbox). HAART simplifies the measurement and analysis of multi-channel impulse responses (IRs). For the purposes of this engineering brief the code library is compiled as a set of Max objects that form a prototype program in Max. This program is able to perform the acquisition, manipulation and analysis of IRs using subjective and objective measures described in acoustics literature. HAART is also able to convolve IRs with audio material and, most importantly, able to binaurally synthesize virtual, multichannel speaker arrays over headphones, negating the need for multichannel setups when out in the field. The code library is freely available from: http://www.hud.ac.uk/research/researchcentres/mtprg/projects/apl/
Engineering Brief 190

Saturday, May 9 16:00 Room 609A

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

Project Studio Expo 13
Saturday, May 9 16:30 – 17:15
PSE Stage

MIXING OUT OF THE BOX

Presenter: **Jacek Gawłowski**, JGMasterLab, Warsaw, Poland

Grammy award winner Jacek talks us through his mixing+mastering hybrid approach where both duties are performed simultaneously by one man in the same room combining gear strictly for mixing with dedicated mastering equipment.

Tutorial 14 Saturday, May 9
17:00 – 18:00 Room Opera

OBJECT-BASED BROADCASTING—
THE FUTURE OF BROADCASTING?

Presenter: **Chris Baume**, BBC Research and Development, London, UK; University of Surrey, Guildford, Surrey, UK

The audio research group at BBC R&D has for the past four years been developing the next generation of audio for broadcast. Working with academic partners in the BBC Audio Research Partnership, the group has created novel audience experiences based on new forms of audio content representations. This work has led to the launch of large-scale funded projects involving a significant number of researchers and international industrial partners. This talk will outline a number of public trials around object-based audio and binaural experiences which demonstrates the potential future of broadcast. The talk will focus on the audience experience, but will also detail the technological challenges from a content provider perspective.

Special Event
OPEN HOUSE OF THE TECHNICAL COUNCIL AND THE RICHARD C. HEYSER MEMORIAL LECTURE
Saturday, May 9, 18:15 – 19:30
Room Belweder

Lecturer: **Ilpo Martikainen**

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 138th AES Convention is Ilpo Martikainen. Born in 1947, Martikainen studied applied electronics, digital techniques, and acoustics in Helsinki University of Technology and received M.Sc. and D.Sc (Tech.) h.c. degrees in 1977 and 2008, respectively. Dr. Martikainen worked as Managing Director at KLT Electronics Oy from 1972 to 1977 and founded Genelec Oy in 1978, where he worked as Managing Director until 2005, and as Chairman since 2006. Dr. Martikainen has been involved in tens of speaker designs and his current interests include new transducers. Dr. Martikainen is Fellow of AES and holds several patents. The title of his lecture is “Monitoring Loudspeakers—Can They Be Truthful Quality Control Tools?”

Monitoring speakers are quality control tools, which should give a truthful acoustical presentation of the electrical signal in question. Over the decades this goal has been approached with different means. This presentation looks what the technical means have been since 1940s in different parts of the world and whether the evolution has converged towards some solutions giving sustainable value to the industry. The future trends are also discussed.

Special Event
BANQUET
Saturday, May 9, 20:30 – 23:30
Galaria Freta
Freta 39 Street (“New Town” Market Square)

This year, the Banquet will be held in one of the trendiest venues in Warsaw—in the charming Old Town / New Town district. The Restaurant and Gallery Freta combines culinary artistry with artistic ambiance of an art gallery, full of classic pieces of furniture and modern pieces of art.

The AES banquet will use three rooms, the main hall “Patio” on the ground floor with a glamorous glass roof, together with two adjacent halls. The Arkady Room divided with a stylish bar creates a relaxing climate for conversations, while Gray Hall, an intriguing room with unusual Parisian charm, can be also used by those who want to smoke.

Tasty hot and cold food will be served in a buffet style—so we all can move around and meet people.

How to get there:

From Sofitel hotel it’s a 20 minute walk through the scenic street named Krakowskie Przedmiescie and through the charming Old and New Town districts. It is also possible to get there by bus nr 116, 178, 195 going to Franciszkanska street (four stops), then take a five minute walk down Franciszkanska street to Freta street, nr 39.

Tickets will be available at the Special Events desk.

Session P15 **Sunday, May 10**
09:00 – 12:30 **Room Belweder**

SPATIAL AUDIO—PART 2

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

09:00

P15-1 Analysis on the Timbre Coloration of Wave Field Synthesis Using a Binaural Loudness Model— *Bosun Xie, Haiming Mai, Yang Liu, Xiaoli Zhong*, South China University of Technology, Guangzhou, China

Wave field synthesis (WFS) aims to reconstruct a target sound field within an extending region. An ideal WFS system requires continuous loudspeakers array. Discrete loudspeaker array in practical WFS causes spatial aliasing errors above the Nyquist frequency limit, and thus results in timbre coloration. The present work analyzes the timbre in WFS using Moore's modified binaural loudness model, in which the binaural loudness level spectra is used as a criterion to evaluate the timbre coloration. The results prove that timbre coloration reduces with the increasing distance between field point and active loudspeakers; and reducing the space between adjacent loudspeakers reduces perceivable timbre coloration. A psychoacoustic experiment yields consistent results with those of analysis, and therefore validates the proposed method.
Convention Paper 9320

09:30

P15-2 Physical Properties of Local Wave Field Synthesis Using Linear Loudspeaker Arrays—*Fiete Winter, Sascha Spors*, University of Rostock, Rostock, Germany

Wave Field Synthesis aims at a physically accurate synthesis of a desired sound field inside an extended listening area. Due to limitation of practical loudspeaker setups, the accuracy of this sound field synthesis technique over the entire listening area is limited. Local Wave Field Synthesis narrows the spatial extent down to a local listening area in order to improve the reproduction accuracy inside this limited region. Recently a method has been published, which utilizes focused sources as a distribution of more densely placed virtual secondary sources around the local area. Within this paper an analytical framework is established to analyze the physical properties of this approach for linear loudspeaker setups.
Convention Paper 9321

10:00

P15-3 Pressure-Matching Beamforming Method for Loudspeaker Arrays with Frequency Dependent Selection of Control Points—*Ferdinando Olivieri, Filippo Maria Fazi, Mincheol Shin, Philip Nelson*, ISVR, University of Southampton, Southampton, UK

The Pressure-Matching Method (PMM) is a signal processing technique used to generate the digital filters required by a loudspeaker array to reproduce a desired sound field. System performance depends on the choice of a number of parameters of the numerical algorithm, such as the target field and the regularization factor. If a target sound field is chosen with large amplitude variation between the so-called control points, performance might also depend on the relative distance between these points in relation to a the wavelength of the sound to be reproduced. If this distance is too small, the accuracy of the reproduced field may be reduced at the listener location. A strategy is proposed to improve the PMM that is based on a frequency-dependent selection of the control points that contribute to the PMM cost function. By means of numerical simulations and experiments in anechoic environment, it is shown that the proposed method allows for accurate control of the response of the reproduced field at the listener location.
Convention Paper 9322

10:30

P15-4 Discussion of the Wavefront Sculpture Technology Criteria for Straight Line Arrays—*Frank Schultz,*[1] *Florian Straube,*[2] *Sascha Spors*[1]
[1]University of Rostock, Rostock, Germany
[2]TU Berlin, Berlin, Germany

Wavefront Sculpture Technology introduced line source arrays for large scale sound reinforcement, aiming at the synthesis of highly spatial-aliasing free sound fields for full audio bandwidth. The paper revisits this technology and its criteria for straight arrays using a signal processing model from sound field synthesis. Since the latest array designs exhibit very small driver distances, the sampling condition for grating lobe free electronic beam forming regains special interest. Furthermore, a discussion that extends the initial derivations of the spatial lowpass characteristics of circular and line pistons and line pistons with wavefront curvature applied in subarrays is given.
Convention Paper 9323

11:00

P15-5 Sound Field Synthesis of Virtual Cylindrical Waves Using Circular and Spherical Loudspeaker Arrays—*Nara Hahn, Sascha Spors*, University of Rostock, Rostock, Germany

In sound field synthesis, like near-field compensated higher-order Ambisonics or Wave Field Synthesis, various virtual source models are used to describe a virtual sound scene. In near-field compensated higher-order Ambisonics, the virtual sound field has to be expanded into spherical harmonics. Unlike plane waves or spherical waves, cylindrical waves are not conveniently represented in the spherical harmonics domain. In this paper we tackle this problem and derive closed form driving functions for virtual cylindrical waves. The physical properties of synthesized sound fields are investigated through numerical simulations, where the results are compared with virtual cylindrical waves in wave field synthesis.
Convention Paper 9324

11:30

P15-6 Perceptual Band Allocation (PBA) for the Rendering of Vertical Image Spread with a Vertical 2D Loudspeaker Array—*Hyunkook Lee*, University of Huddersfield, Huddersfield, UK

A series of subjective experiments were conducted to investigate a novel vertical image rendering method named "Perceptual Band Allocation (PBA)," using octave bands of pink noise with a vertical 2D reproduction setup with main and height loudspeaker pairs. The perceived height of each octave band was first measured for the main and height loudspeakers individually. Results suggested a significant difference between monophonic and stereophonic images in the perceived relationship between frequency and height. Six different test conditions have been created aiming for various degrees of vertical image spread, in such a way that each frequency band was mapped to either the main or height loudspeaker layer based on the results from the localization experiment. Multiple comparison tests were conducted to grade the perceived magnitude of vertical image spread. It was generally found that various degrees of vertical image spread could be rendered using

different PBA schemes, but the perceived results did not fully match predicted results based on the localization results. Differences between the main and height loudspeaker layers in the spectral weightings of ear-input signal at certain frequencies was identified as one of the factors that influenced this result.
Convention Paper 9325

12:00

P15-7 Synthesis of Moving Reverberation Using Active Acoustics—Preliminary Report—Jung Wook *(Jonathan) Hong,*[1] *Wieslaw Woszczyk,*[1] *Durand R. Begault,*[2] *David Benson*[1]
[1]McGill University, Montreal, Quebec, Canada
[2]NASA Ames Research Center, Moffett Field, CA, USA

An ambient sound field created artificially using active acoustics (virtual acoustics) attempts to resurrect the perceived naturalness of the original architectural space and the distinct responsiveness to musical sound sources. Moving reverberation is typically associated with coupled volumes contained within a larger architectural space where each volume is activated by a sound source at a different moment in time due to propagation delay. The produced energy has a diverse characteristic rate of decay with which its energy is mixed within the common space causing multiple slopes on the decay. This causes a sensation of a decaying diffused sound that is not homogenized but distinctly appearing in different zones of the space as a shifting acoustic energy. In order to reconstruct the moving reverberation, an active acoustics system was used to render an ambient sound field from measured impulse responses of large architectural space, the Grace Cathedral in San Francisco.
Convention Paper 9326

Session EB2 — Sunday, May, 10
09:00 – 12:00 — Room Królewski

ENGINEERING BRIEFS—LECTURES

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

09:00

EB2-1 Comparing Room Acoustics for the Performance of Wagners‘ Lohengrin—*Winfried Lachenmayr, Gunter Engel*, Mueller-BBM, Munich, Germany

Wagner is *the* example of a composer caring for the entire “production” chain of his works. A specific venue, the Festspielhaus in Bayreuth, was dedicated and built for a handful of his own compositions. But through their popularity Wagner’s operas are nowadays played in theaters and opera houses all over the world with different acoustic conditions. How does this affect the performance and musical perception and what can be observed in recordings or measurements? As an example, the acoustic recordings of performances of “Lohengrin” in (1) a typical smaller opera house in Germany and (2) in the Festspielhaus Bayreuth are compared and analyzed. Differences regarding running- and decay reverberation are addressed.
Engineering Brief 191

09:15

EB-2-2 Progress in Power Amplifiers: Thermal Distortion— *Douglas Self,* The Signal Transfer Company, London, UK

You sometimes read of "thermal distortion" in power amplifiers supposed to be due to cyclic changes in device parameters due to varying heating over that cycle. Until recently I was unimpressed by the likelihood of its existence, reasoning that it must cause rising distortion with falling frequency, and this was not observable in measurements to 0.001% THD. (Changes in distortion due to long-term thermal changes in the output stage quiescent conditions are another matter and clearly do exist.) Improvements in test gear allowing 0.0002% measurements and the use of higher powers discloses that THD can rise with falling frequency. The component concerned is identified, and methods given for both confirming the presence of thermal distortion and reducing it.
Engineering Brief 192

09:30

EB-2-3 Immersive Sound Design Using Particle Systems— *Nuno Fonseca*, ESTG/CIIC, Polytechnic Institute of Leiria, Leiria, Portugal

With the release of major immersive audio formats for cinema, including Auro-3D and Dolby Atmos, a new sound dimension is getting the attention of sound professionals—height. Although traditional panning techniques are still possible, more interesting approaches could be used to better explore space. Particles systems, which are widely used on computer graphics, could present themselves as a very interesting sound design approach. With immersive virtual microphones, capable of supporting many different setups, perfect sound and space coherence can be obtained. By controlling the particle system, instead of individual sound sources, a high number of sounds can be rendered in 3D. A particle system software was created, capable of running highly complex situations with up to several millions of sound sources, which is currently under testing by major Hollywood studios.
Engineering Brief 193

09:45

EB-2-4 Examining Influence of Distance to Microphone on Accuracy of Speech Recognition—*Piotr Bratoszewski, Marcin Szykulski, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

The problem of controlling a machine by the distant-talking speaker without a necessity of handheld or body-worn equipment usage is considered. A laboratory setup is introduced for examination of performance of the developed automatic speech recognition system fed by direct and by distant speech acquired by microphones placed at three different distances from the speaker (0.5 m to 1.5 m). For feature extraction from the voice signal the Mel- Frequency Cepstral Coefficients (MFCC) are used. The experiments are conducted employing the HTK engine (Hidden Markov Toolkit) for the Automatic Speech Recognition (ASR) task. The dictionary of 184 words was employed and WER (Word Error Rate), correctness and accuracy measures were calculated in order to verify and to compare obtained results of speech recognition.
Engineering Brief 194

10:00

EB-2-5 Investigating the Sound Quality Lexicon of Analogue Compression Using Category Analysis—*Malachy Ronan, Nicholas Ward, Robert Sazdov*, University of Limerick, Limerick, Ireland

This study investigates the lexicon used to describe analogue compression. Extant documents comprising 51 reviews of analogue compressors over 15 years are coded ➥

using category analysis. A total of 160 adjectives emerge and are further refined to nine inductive categories: transparency; frequency related; signal distortion; energy; transient shaping; glue; association; spatial dimensions; and character. A final abstraction reveals two primary attributes governing the perceived sound quality of analogue compression: "character" and "transient shaping." Transparent dynamic range compression is found to be less important. This investigation clarifies the lexicon used to describe the sound quality attributes of analogue compression.
Engineering Brief 195

10:15

EB-2-6 Exploratory Microphone Techniques for Three-Dimensional Classical Music Recording—*Will Howie, Richard King*, McGill University, Montreal, Quebec, Canada; The Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada

At McGill University's Redpath Hall, a conventional stereo recording array was augmented with additional microphones in both the horizontal and vertical planes, yielding a 14-channel surround sound recording featuring seven discrete height channels. Based on existing multichannel recording models, microphone placement was designed to prioritize listener envelopment. Preliminary evaluations of the recordings by the authors and fellow researchers from the Graduate Program in Sound Recording at McGill University found that these 3D recordings have an increased sense of envelopment and realism as compared to traditional 5.1 surround sound. The authors have identified several areas to be further investigated through future recordings and listening tests.
Engineering Brief 196

10:30

EB-2-7 Automated (Microcontroller-Based) Impedance Tube—*Alex Faron, Connor McCullough, Diego Ugaz*, University of Miami, Miami, FL, USA

The acquisition of acoustic properties such as absorption and transmission-loss coefficients is necessary for the analysis and synthesis of acoustic materials. Because of its accuracy, the impedance tube is the preferred method of measurement. However, current implementations require a strong technical background in acoustics and considerable time to produce results. The system described in this work will use a microcontroller to automate the measurement process and expedite the delivery of information to a non-expert end-user.
Engineering Brief 197

10:45

EB-2-8 Analysis of DML Sound Reinforcement Systems Behavior in Large Concert Halls—*Dragan Novkovic, Stefan Dimitrijevic*, College of Electrical Engineering and Computer Sciences, Belgrade, Serbia

Sound reinforcement systems based on DML technology, as a sound sources with unique signal generation and radiation characteristics, provides some particular features when it comes to audience coverage and speech intelligibility. Although this technology exists for over a quarter of century, not so many sound reinforcement systems based on this technology exists and, therefore, it is possible to perceive a lack of data on the behavior of this type of system in real-life conditions. This paper has a goal to present the results of impulse response measurements conducted in big concert venue that was alternately excited with conventional and DML sound reinforcement systems.
Engineering Brief 198

11:00

EB-2-9 Production and Reproduction of Program Material for a Variety of Spatial Audio Formats—*Jon Francombe, Tim Brookes, Russell Mason, Rupert Flindt, Philip Coleman, Qingju Liu, Philip J.B. Jackson*, University of Surrey, Guildford, Surrey, UK

For subjective experimentation on 3D audio systems, suitable program material is needed. A large-scale recording session was performed in which four ensembles were recorded with a range of existing microphone techniques (aimed at mono, stereo, 5.0, 9.0, 22.0, ambisonic, and headphone reproduction) and a novel 48-channel circular microphone array. Further material was produced by remixing and augmenting pre-existing multichannel content. To mix and monitor the program items (that included classical, jazz, pop, and experimental music, and excerpts from a sports broadcast and a film soundtrack), a flexible 3D audio reproduction environment was created. Solutions to the following challenges were required: level calibration for different reproduction formats; bass management; and adaptable signal routing from different software and file formats.
Engineering Brief 199

11:15

EB-2-10 Low Frequency Performance of Circular Loudspeaker Arrays—*Filippo Maria Fazi,*[1] *Mincheol Shin,*[1] *Ferdinando Olivieri,*[1] *Simone Fontana*[2]
[1]University of Southampton, Southampton, Hampshire, UK
[2]Huawei European Research Center, Munich, Germany

Compact loudspeaker arrays are widely used for the localized delivery of audio messages and beamforming applications. The optimal directivity performance of these devices is limited to a given frequency limit, whose upper bound is defined by the occurrence of spatial aliasing. The lower bound of this frequency range is caused by the limited capability of the array to generate a directional sound beam when the wavelength of the sound to be reproduced is large in comparison to the size of the array. In this work a theoretical and experimental study is presented of the directivity limitations of circular loudspeaker arrays at low frequencies. The frequency at which the array directivity pattern starts to diverge from the desired one is calculated analytically and put into relation with the dynamic range of the transducers and with the regularization scheme used when designing the beamforming digital filters.
Engineering Brief 200

11:30

EB-2-11 An Exploratory Evaluation of User Interfaces for 3D Audio Mixing—*Steven Gelineck, Dannie Korsgaard*, Aalborg University, Copenhagen, Denmark

The paper presents an exploratory evaluation comparing different versions of a mid-air gesture based interface for mixing 3D audio exploring: (1) how such an interface generally compares to a more traditional physical interface, (2) methods for grabbing/releasing audio channels in mid-air, and (3) representation of sources in separate 3D views vs. in one shared 3D view. Results suggest that while the traditional physical interface is generally intuitive and easy to use, the 3D gesture interface provides an improved understanding of the 3D space and provides a better control of especially moving sources. The shared view provides a better overview and workflow than the separated view and grabbing sounds using hand-gestures causes difficulties.
Engineering Brief 201

11:45

EB-2-12 Differences between Recorded and Emulated Guitar Sounds—*Maciej Majewski*,[1] *Pawel Malecki*[2]

[1]Akademia Górniczo-Hutnicza, Krakow, Poland
[2]AGH University of Science and Technology, Krakow, Poland

Is it possible to create an emulated guitar sound similar to the recorded one? Why not! First of all, the direct signal from the guitar is prepared. After that, using a "reamping" technique the desired sound is recorded. Subsequently the whole audio track is emulated using device called "Kemper." Then the listening tests among people who work in the music industry were performed for subjective comparison of prepared sounds. The comparison of numerical audio parameters is provided using Matlab scripts. The results analysis show the performance of modern emulation techniques in compare to the traditional multitrack recording. Major benefits and losses are discussed.
Engineering Brief 186

Workshop 13 **Sunday, May 10**
09:00 — 11:00 **Room Balowa AB**

NEW CINEMA AUDIO STANDARDS

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

Panelists: *TBA*

Significant changes are occurring to the production and exhibition environment for cinema sound. As the SMPTE works to consolidate the immersive sound formats into one distribution master, improvements to the playback environment continue. TC-SDCTV will look at some of the changes being made to improve the reproduction of immersive sound, as well as updates to the existing technologies.

This session is presented in association with the AES Technical Committee on Sound for Digital Cinema and Television

Sunday, May 10 **09:00** **Room 609B**

Technical Committee Meeting on Transmission and Broadcasting

Sunday, May 10 **09:00** **Room 609A**

AESSC Plenary Meeting

Tutorial 15 **Sunday, May 10**
09:15 – 10:15 **Room Opera**

MIC IT AND RECORD IT!

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

You should not start a recording project without knowing how you want the mix to sound, and how to record the sounds you need to create that mix. Learning to use microphones and mic techniques to capture sounds "in their place in the mix" will not only produce more spacious multi-dimensional mixes needing less electronic processing to achieve that goal, but also create a more efficient workflow that speeds up the mixing process. Come and learn how to approach the recording session from the perspective of the mix, how different microphone technologies can beneficially color and shape the sounds you're recording, how different instrument, vocal, and stereo mic techniques affect the sounds captured, and how the recording room's affect on the sound source and the microphones can be exploited and explored to help you capture sounds you actually need for the mix. For students and professionals alike, understanding these techniques will increase the variety of mix styles and sounds you are able to produce, making you a more versatile audio engineer ready to meet the needs of clients and employers today and tomorrow.

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

Spatial Audio Demo 16 **Sunday, May 9**
09:30 – 11:00 **Room Saski**

MAKING 3D RECORDINGS OF CLASSICAL MUSIC

Presenter: **Malgorzata Albinska-Frank**, Tonstudio arton, Das Tonstudio für klassische Musik, Basel, Switzerland

This session is dedicated to all who would like to start to record classical music using 3D techniques: room acoustics, microphone techniques, mixing tools, and playback systems for working with 3D. Is this technique also useful for surround and stereo? Based on listening to 3D recordings of classical music, the technical and aesthetical aspects of 3D recording production workflow will be presented and discussed.

Session P16 **Sunday, May 10**
10:00 – 12:00 **Foyer**

POSTERS: APPLICATIONS IN AUDIO

10:00

P16-1 Dubbing Studio for 22.2 Multichannel Sound System in NHK Broadcasting Center—*Ikuko Sawaya*,[1] *Kengo Sasaki*,[2] *Shinji Mikami*,[3] *Hiroyuki Okubo*,[1] *Kazuho Ono*[1]

[1]Science & Technology Research Laboratories, Japan Broadcasting Corp., Setagaya, Tokyo, Japan
[2]Broadcast Engineering Department, Japan Broadcasting Corporation, Shibuya, Tokyo, Japan
[3]Engineering Administration Department, Japan Broadcasting Corporation, Shibuya, Tokyo, Japan

8K Super Hi-Vision is planned to be on test broadcasting in 2016 and to launch a full broadcasting service in 2018 in Japan. NHK has developed a program production system for 22.2 multichannel sound for 8K Super Hi-Vision. As part of the development NHK completed the construction of a 22.2 ch dubbing studio in the NHK Broadcasting Center in July 2014. This is the first 22.2 ch dubbing studio in the production field in the world with a loudspeaker configuration that meets the standard Recommendation ITU-R BS.2051. In this paper we discuss the 22.2 ch production system, including its sound mixing system, loudspeaker system for monitoring, and perforated screen for 8K resolution, as well as the room design and the characteristics of the room acoustics in the studio.
Convention Paper 9327

10:00

P16-2 A Floor Acoustic Sensor for Fall Classification—*Emanuele Principi*,[1] *Paolo Olivetti*,[2] *Stefano Squartini*,[1] *Roberto Bonfigli*,[1] *Francesco Piazza*[1]

[1]Università Politecnica delle Marche, Ancona, Italy
[2]Scientific Direction, Italian National Institute of Health and Science on Aging (INRCA), Ancona, Italy

The interest in assistive technologies for supporting people at home is constantly increasing, both in academia and industry. In this context the authors propose a fall classification system based on an innovative acoustic sensor that operates similarly to stethoscopes and captures the acoustic waves transmitted through the floor. The sensor is designed ➥

to minimize the impact of aerial sounds in recordings, thus allowing a more focused acoustic description of fall events. In this preliminary work, the audio signals acquired by means of the sensor are processed by a fall recognition algorithm based on Mel-Frequency Cepstral Coefficients, Supervectors, and Support Vector Machines to discriminate among different types of fall events. The performance of the algorithm has been evaluated against a specific audio corpus comprising falls of persons and of common objects. The results show the effectiveness of the approach.
Convention Paper 9329

10:00

P16-3 Active Field Control in the Teatr Wielki—Opera Narodowa—*Takayuki Watanabe,*[1] *Hideo Miyazaki,*[1] *Shinichi Sawara,*[1] *Masahiro Ikeda,*[1] *Ron Bakker*[2]
[1]Yamaha Corp., Hamamatsu, Shizuoka, Japan
[2]Yamaha Commercial Audio Systems Europe, Rellingen, Germany

This opera house of 1,828 seats boasts one of Europe's largest stages and is highly reputed for its repertoire and acoustics. However, it presented a number of issues including poor communication between the singers and the orchestra pit, insufficient loudness of the upstage singers for the audience, a lack of reverberation when the house was occupied, and insufficient loudness at the seats under the balconies. For these reasons Active Field Control System (AFC) was adopted as a means to improve the acoustics while preserving the historic architecture of the opera house. This paper presents an overview of that system and the benefits achieved by its introduction.
Convention Paper 9330

10:00

P16-4 An Environment for Submillisecond-Latency Audio and Sensor Processing on BeagleBone Black—*Andrew McPherson,*[1] *Victor Zappi*[2]
[1]Queen Mary University of London, London, UK
[2]University of British Columbia, Vancouver, BC, Canada

This paper presents a new environment for ultra-low-latency processing of audio and sensor data on embedded hardware. The platform, which is targeted at digital musical instruments and audio effects, is based on the low-cost BeagleBone Black single-board computer. A custom expansion board features stereo audio and 8 channels each of 16-bit ADC and 16-bit DAC for sensors and actuators. In contrast to typical embedded Linux approaches, the platform uses the Xenomai real-time kernel extensions to achieve latency as low as 80 microseconds, making the platform suitable for the most demanding of low-latency audio tasks. The paper presents the hardware, software, evaluation, and applications of the system.
Convention Paper 9331

10:00

P16-5 Commonwealth Games 2014 Host Broadcaster Training Initiative–A Game Changer?—*Patrick Quinn, David Moore*, Glasgow Caledonian University, Glasgow, Lanarkshire, UK

Glasgow Commonwealth Games 2014 provided an ideal platform for over 200 students to gain work experience in sports broadcasting as part of the Host Broadcaster Training Initiative. Organized by SVGTV and Creative Loop with the intention of attracting students into this growing area of broadcasting, the successful initiative has encouraged many of the students involved, including Audio Technology students from Glasgow Caledonian University, to consider and subsequently pursue careers in broadcasting. In addition as a legacy from the initiative a new course is planned at Glasgow Caledonian University in Broadcasting Technology.
Convention Paper 9332

10:00

P16-6 Influence of Noise on the Effectiveness of Speaker Identification in the Acoustics of Crime—*Tomasz Smutnicki, Stefan Brachmanski*, Wroclaw University of Technology, Wroclaw, Poland

One of the main elements of the research in acoustics of crime is to compare the evidential recording with the comparative adequate pattern. Unfortunately, the evidential recording usually has poor quality and contains relatively high level of noise, which results from the way of its acquiring, namely eavesdropping or record of automatic monitoring. The signal quality and the noise to signal ratio have an impact on the value of the extracted voice metrics. In this paper we analyze factors that may have an impact on formants value in the human voice. Based on Six Sigma methodology we also designed and performed an experiment that allowed us to determine the extent in which various factors influence on the resulting parameters.
Convention Paper 9333

10:00

P16-7 Acoustic Profile of Identified Speaker in Forensics—*Krystian Kapala, Stefan Brachmanski*, Wroclaw University of Technology, Wroclaw, Poland

Speaker identification is deemed to be one of the basic tasks in audio forensics. Delivering a categorical opinion is often difficult due to insufficient quality of the recorded material, simulation or modulation of speaker's voice. Hence, a wide-ranging approach to the identification process is used, including both subjective and objective methods. With their help, it becomes possible to obtain a broad spectrum of speech characteristics ranging from low-level features relating to physical construction of the vocal tract to advanced ones concerning various ways of expressing oneself and articulation, acquired during socialization process. This paper describes an experiment undertaken to create acoustic profiles of a chosen group of speakers based on the features mentioned above.
Convention Paper 9334

10:00

P16-8 An Implementation Of Beamforming Algorithm On FPGA Platform with Digital Microphone Array—*Iva Salom,*[1] *Vladimir Celebic,*[1]*Milan Milanovic,*[1] *Dejan Todorovic,*[2] *Jurij Prezelj*[3]
[1]Institute Mihajlo Pupin, University of Belgrade, Belgrade, Serbia
[2]Dirigent Acoustics Ltd., Belgrade, Serbia
[3]University of Ljubljana, Ljubljana, Slovenia

The goal of the project described in this paper was to design an acoustic system for localization of the dominant noise source by implementation of the conventional delay-and-sum beamforming algorithm on FPGA platform with a sound receiver system based on digital MEMS microphone array. The system consists of a platform for acoustic signal acquisition and data processing (microphone array, interface, and central block), and a platform for monitoring and control (a computer with a user application). Such configuration provides the execution of the beamforming algorithm in real time. Additionally, FPGAs are bringing many benefits in terms of safety, reliability, rapidity, and power consumption. The platform was tested and verified with

various microphone array configurations and results are presented in the paper.
Convention Paper 9335

10:00

P16-9 Measuring and Analyzing Audio Levels in Film, Commercials, and Movie Trailers Using Leq(A) Values and the LUFS Loudness Model—*Bozena Kostek, Kamila Milarska, Aleksander Zakrzewski*, Gdansk University of Technology, Gdansk, Poland

The purpose of this paper is to describe the measurement of loudness levels in movies, movie trailers, and commercials displayed before feature films at movie theaters. In the initial section, the paper discusses the issues related to measurement of loudness levels, provides recommendations regarding permissible loudness levels during movie screenings, and mentions the applied units of measurement. The following section of the paper describes the actual measurements, measuring equipment, as well as analysis of the results of the measurements. The summary provides conclusions about the measured loudness levels at movie theaters, for DVD and Blu-ray discs, and for YouTube videos.
Convention Paper 9336
This paper was presented by Aleksander Zakrzewski

Student Event and Career Development
STUDENT DELEGATE ASSEMBLY MEETING—PART 2
Sunday, May 10, 10:30 – 12:00
Room Opera

Presenters: **Simon-Claudius Wystrach**
Steven Van Dyne
Brecht De Man
Zach Bloomstein

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.

Tutorial 16 Sunday, May 10
11:30 – 13:30 Room Balowa AB

CREATING MUSIC FOR FILM AND TV—A MASTER CLASS

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

Presenters: *Maciej Zielinski*
TBA

The composition of music for film and television is a detailed process with its own challenges and special recording techniques. Here we present two of Poland's best.

This session is presented in association with the AES Technical Committee on Sound for Digital Cinema and Television

Session P17 Sunday, May 10
13:00 – 15:00 Foyer

POSTERS: RECORDING AND PRODUCTION

13:00

P17-1 Tom–Tom Drumheads Miking Analysis—*Andrés Felipe Quiroga, Juan David Garcia, Dario Páez*, Universidad de San Buenaventura, Bogotá, Colombia

Different drum recording techniques have been developed through time, from stereo to close miking techniques. This is relevant, since the techniques and the characteristics of the instrument will define its sound within the final mix. A study was designed that gives experienced and non-experienced recording engineers tools and specific characteristics of tom–tom close miking techniques with different drumheads, microphones, and capture positions. Results indicate the behaviors of the different drumheads and capture positions with the different microphones. The first frequency band of resonance (attack) shows the highest decay level compared to the second band of resonance (tone), and the edge position presented the lower decay level on the second band of resonance, showing its resonant behavior on the envelope.
Convention Paper 9338

13:00

P17-2 Concept of Film Sound Restoration by Adapting to Contemporary Cinema Theatre—*Joanna Napieralska*, Frederic Chopin University of Music, Warsaw, Poland

This paper presents an individual approach to restoration of Polish film sound based on the author's own works. It answers the following question: under what conditions, and by the use of which techniques, may the restoration of archive film sound provide the viewer with a cleaner reproduction of the original sound while maintaining the standard expected by a modern cinema-going audience. At its basic level, the sound restoration routine comprises the following: transfer from the magnetic tape, syncing, cleaning of low/high frequency noises, repair of material impairments and reprinting effect, and mastering for broadcasting, cinema, DVD/Blu-ray, and internet formats. The reconstruction discussed modifies the sound quality and sometimes the contents. However, it can be performed only under certain legal restrictions.
Convention Paper 9339
[The author presented a demonstration following this poster session in Room Saski at 16:15]

13:00

P17-3 Deep Sound Design: Procedural Implementations Based on General Audiovisual Production Pipeline Integration—*José Roberto Cabezas Hernández*, Universidad Nacional Autónoma de México, Mexico City, Mexico

This work is an exploration for the integration of data available on the visual post-production pipeline for the development of procedural sound design and composition techniques, by implementing different methods to allow access of different file formats for scene and shot reconstruction. The main purpose in an audiovisual creation context is to investigate stronger and inmost image-sound cognitive perceptions and relationships generated by data usage and analysis; also, reducing automation time by directly linking data to parameters for developing a creative editing, mixing, design, and compositional workflow based on a shot by shot manipulation.
Convention Paper 9340

13:00

P17-4 An Investigation into the Efficacy of Methods Commonly Employed by Mix Engineers to Reduce Frequency Masking in the Mixing of Multitrack Musical Recordings—*Jonathan Wakefield, Christopher Dewey*, University of Huddersfield, Huddersfield, UK

Studio engineers use a variety of techniques to reduce frequency masking between instruments when mixing multitrack musical recordings. This study evaluates the efficacy

of three techniques, namely mirrored equalization, frequency spectrum sharing, and stereo panning against their variations to confirm the veracity of accepted practice. Mirrored equalization involves boosting one instrument and cutting the other at the same frequency. Frequency spectrum sharing involves low pass filtering one instrument and high pass filtering the other. Panning involves placing two competing instruments at different pan positions. Test subjects used eight tools comprising a single unlabeled slider to reduce frequency masking in several two instrument scenarios. Satisfaction values were recorded. Results indicate subjects preferred using tools that panned both audio tracks.
Convention Paper 9341

13:00

P17-5 An Interactive Multimedia Experience: A Case Study— *Andrew J. Horsburgh*, Southampton Solent University, Southampton, UK

Accurate representation of three dimensional spaces, both real and virtual, within an environment is a matter of concern for researchers and content producers in the media industry; it is expected that truly immersive experiences will become more desirable outside of research labs and bespoke facilities. This paper presents a case study examining the implementation between visual and audible elements to form a singular experience of immersion, AIME, at Solent University. The computer-based system uses a time-code generator that allows for seamless integration between audio workstations, visual playback, and external lighting. The prototype system uses 2nd order ambisonic audio reproduction, three large panel displays for vision, and an external lighting rig running from time code.
Convention Paper 9342

13:00

P17-6 Evaluation of an Algorithm for the Automatic Detection of Salient Frequencies in Individual Tracks of Multitrack Musical Recordings— *Jonathan Wakefield, Christopher Dewey*, University of Huddersfield, Huddersfield, UK

This paper evaluates the performance of a salient frequency detection algorithm. The algorithm calculates each FFT bin maximum as the maximum value of that bin across an audio region and identifies the FFT bin maximum peaks with the highest five deemed to be the most salient frequencies. To determine the algorithm's efficacy test subjects were asked to identify the salient frequencies in eighteen audio tracks. These results were compared against the algorithm's results. The algorithm was successful with electric guitars but struggled with other instruments and in detecting secondary salient frequencies. In a second experiment subjects equalised the same audio tracks using the detected peaks as fixed centre frequencies. Subjects were more satisfied than expected when using these frequencies.
Convention Paper 9343

13:00

P17-7 The Sonic Vernacular: Considering Communicative Timbral Gestures in Modern Music Production— *Leah Kardos*, Kingston University London, Kingston Upon Thames, Surrey, UK

Over the course of audio recording history, we have seen the activity of sound recording widen in scope "from a technical matter to a conceptual and artistic one" (Moorefield 2010) and the producer's role evolving from technician to "auteur." For recording practitioners engaged in artistic and commercial industry and discourse, fluency in contemporary and historic sound languages is advantageous This paper seeks to find the best, most practically useful method to describe these characteristics in practice, identify a clear and suitable way to talk about and analyze these uses of communicative timbral gestures, as heard in modern music productions."
Convention Paper 9344

13:00

P17-8 Auto Panning In-Ear Monitors for Live Performers— *Tom Webb, Andrew J. Horsburgh*, Southampton Solent University, Southampton, UK

In a live musical performance, accurate stage monitoring is a vital element to achieve the optimal performance. Current stage monitoring uses traditional musician-facing loudspeakers. Problems can be surmised as excessive Sound Pressure Level (SPL), the inability to hear themselves, acoustic feedback, and general stage untidiness /space requirements. In-ear monitors (IEM's) can offer a solution to these problems when the IEM system has been properly designed [7]. One crucial issue with IEMs is the sense of isolation and disconnection from stage noise and crowd. To overcome this issue, an auto-panning system that adjusts spatial placement of audio channels within the performers stage mix has been designed and built.
Convention Paper 9345

13:00

P17-9 An Investigation into Plausibility in the Mixing of Foley Sounds in Film and Television—*Braham Hughes, Jonathan Wakefield*, University of Huddersfield, Huddersfield, UK

This paper describes an experiment that tested the plausibility of a selection of post-production audio mixes of Foley for a short film. The mixes differed in the implementation of four primary audio mixing parameters: panning, level, equalization, and the control of reverberation effects. The experiments presented test subjects with mixes in which one of the four primary parameters was altered while the rest remained at levels deemed to conform to an "industry standard" reference mix that had been verified by an expert industry practitioner. Results show that there is a statistically significant affect on plausibility of using even slight dynamic variation of pan, level, and equalization control to enhance the perception of realism of Foley that move in a scene.
Convention Paper 9346

13:00

P17-10 A Semantically Motivated Gestural Interface for the Control of a Dynamic Range Compressor— *Thomas Wilson, Steven Fenton, Matthew Stephenson*, University of Huddersfield, Huddersfield, West Yorkshire, UK

This paper presents a simplified 2D gesture based approach to modifying dynamics within a musical signal. Despite the growth in gesture-controlled audio seen over recent years, it has primarily been limited to the upper workflow/navigation level. This has been compounded by the Skeuomorphic design approaches of graphical user interfaces (GUI). This design approach, although representative of the original piece of audio equipment, often lowers workflow and hinders the simultaneous control of parameters. Following a large scale gesture elicitation exercise utilizing a common 2D touch pad and analysis of semantic audio control parameters, a set of reduced multi-modal parameters are proposed that offers both workflow efficiency and a much simplified method of control for dynamic range compression.
Convention Paper 9347

13:00

P17-11 Natural Sound Recording of an Orchestra with Three-Dimensional Sound—*Kimio Hamasaki,*[1] *Wilfried Van Baelen*[2]
[1]Artsridge LLC, Chiba, Japan
[2]Auro Technologies N.V., Mol, Belgium

This paper introduces the microphone techniques for recording an orchestra with three-dimensional multichannel sound and discusses the spatial impression provided by the recorded sound of an orchestra. Listeners in a concert hall simultaneously hear both a direct sound arriving from each musical instrument and an indirect sound reflected from the walls and the ceiling. Concerning a direct sound, existing microphone techniques can be used for three-dimensional multichannel sound with necessary modification, but new microphone techniques should be developed for an indirect sound. This paper will propose the microphone technique consisting of a main microphone array and an ambience microphone array, which will enable us to control spatial impressions easily and realize the stable sound source localization.
Convention Paper 9348

Session P18 **Sunday, May 10**
14:00 – 17:00 **Room Belweder**

SEMANTIC AUDIO

Chair: **Pedro Pestana**, Catholic University of Oporto, Oporto, Portugal

14:00

P18-1 Music Onset Detection Using a Bidirectional Mismatch Procedure Based on Smoothly Varying-Q Transform—*Li Luo, Guido H. Bruck, Peter Jung*, University of Duisburg-Essen, Duisburg, Germany

This paper describes a novel onset detector for music signal based on the smoothly varying-Q transform, where the Q-factors vary following a linear function of the center frequencies. The smoothly varying-Q factors allow the time-frequency representation to coincide with the auditory critical-band scale. As the analysis basis of the input signal, the time-frequency image generated by smoothly varying-Q transform indicates the frequency evolution. On the detection stage, a bidirectional mismatch procedure is designed to estimate the discrepancies of frequency partials between currently processed frame and its bidirectional neighboring frames. An onset strength signal is obtained by measuring the mismatch error between the neighboring frames. The evaluation of the proposed algorithm is performed on a fully onset annotated music database and the results show that the proposed algorithm can achieve high detection accuracy and satisfied results.
Convention Paper 9349

14:30

P18-2 A Real-Time System for Measuring Sound Goodness in Instrumental Sounds—Oriol Romani Picas,[1] *Hector Parra Rodriguez*,[1] *Dara Dabiri*,[1] *Hiroshi Tokuda*,[2] *Wataru Hariya*,[2] *Koji Oishi*,[2] *Xavier Serra*[1]
[1]Universitat Pompeu Fabra, Barcelona, Spain
[2]KORG Inc., Tokyo, Japan

This paper presents a system that complements the tuner functionality by evaluating the sound quality of a music performer in real-time. It consists of a software tool that computes a score of how well single notes are played with respect to a collection of reference sounds. To develop such a tool we first record a collection of single notes played by professional performers. Then, the collection is annotated by music teachers in terms of the performance quality of each individual sample. From the recorded samples, several audio features are extracted and a machine learning method is used to find the features that best described performance quality according to musician's annotations. An evaluation is carried out to assess the correlation between systems' predictions and musicians' criteria. Results show that the system can reasonably predict musicians' annotations of performance quality.
Convention Paper 9350

15:00

P18-3 *Timbre Solfege*: Development of Auditory Cues for the Identification of Spectral Characteristics of Sound—*Teresa Rosciszewska, Andrzej Miskiewicz*, Fryderyk Chopin University of Music, Warsaw, Poland

This paper is concerned with listening exercises conducted during a technical ear training course called *Timbre Solfege*, taught to the students of sound engineering at the Fryderyk Chopin University of Music in Warsaw. Discussed are auditory cues used for identification of the characteristics of timbre produced by varying the sound frequency bandwidth and by boosting of selective frequency bands with the use of a spectrum equalizer. The students' ability of identifying those modifications of the spectrum envelope has been assessed in a variety of progress tests. Results of the tests show that systematic training during the *Timbre Solfege* course considerably improves memory for timbre and develops the ability of associating the perceived characteristics of timbre with the spectral properties of sounds.
Convention Paper 9351

15:30

P18-4 Automatic Vocal Percussion Transcription Aimed at Mobile Music Production—*Héctor A. Sánchez-Hevia, Cosme Llerena-Aguilar, Guillermo Ramos-Auñón, Roberto Gil-Pita*, University of Alcalá, Alcalá de Henares, Madrid, Spain

In this paper we present an automatic vocal percussion transcription system aimed to be an alternative to touch-screen input for drum and percussion programming. The objective of the system is to simplify the workflow of the user by letting him create percussive tracks made up of different samples triggered by his own voice without the need of any demanding skill by creating a system tailored to his specific needs. The system consists of four stages: event detection, feature extraction, and classification. We are employing small user-generated databases to adapt to particular vocalizations while avoiding overfitting and maintaining computational complexity as low as possible.
Convention Paper 9352

16:00

P18-5 Training-Based Semantic Descriptors Modeling for Violin Quality Sound Characterization—*Massimiliano Zanoni,*[1] *Francesco Setragno*,[1] *Fabio Antonacci*,[1] *Augusto Sarti*,[1] *György Fazekas*,[2] *Mark B. Sandler*[2]
[1]Politecnico di Milano, Milan, Italy
[2]Queen Mary University of London, London, UK

Violin makers and musicians describe the timbral qualities of violins using semantic terms coming from natural language. In this study we use regression techniques of machine intelligence and audio features to model in a training-based fashion a set of high-level (semantic) descriptors for the automatic annotation of musical instruments. The most relevant semantic descriptors are collected through interviews to ➥

violin makers. These descriptors are then correlated with objective features extracted from a set of violins from the historical and contemporary collections of the Museo del Violino and of the International School of Luthiery both in Cremona. As sound description can vary throughout a performance, our approach also enables the modeling of time-varying (evolutive) semantic annotations.
Convention Paper 9353

16:30

P18-6 Audibility of Lossy Compressed Musical Instrument Tones—*Agata Rogowska*, Warsaw University of Technology, Warsaw, Poland

The aim of the conducted study was to evaluate differences in the audibility of different instruments by three commonly used lossy codecs. Seven instrument tones were compressed using MP3-LAME, Vorbis, and Opus to determine how the detection of compressed sounds varies with bit rate, instrument, and compression formats. Audibility of lossy compression was examined on six naïve subjects during 60 hours of listening. At the bit rate of 32 kbps the compressed signals were easily discriminable with significant differences between subjects. With magnifying the bit rate audibility decreased, the signal becoming inaudible at 64–96 kbps. Discrimination varied significantly from instrument to instrument.
Convention Paper 9232

Workshop 14 **Sunday, May 10**
14:00 — 15:30 **Room Królewski**

AUDIO RESTORATION: FROM ANALOGUE TRANSFER TO DIGITAL SIGNAL ENHANCEMENT

Presenters: **Andrzej Czyzewski**
Dietrich Schüller
Nadja Wallaszkovits

The workshop will lead through the complete workflow chain of high quality audio restoration—from analogue to digital. Starting with the introduction to guidelines and best practices to preservation and restoration of analogue audio materials, practical examples of restoration and digitization procedures are outlined and presented. Finally signal enhancement in the digital domain is discussed and demonstrated on the basis of an experimental automated restoration process.

Workshop 15 **Sunday, May 10**
14:00 — 16:00 **Room Balowa AB**

MIX KITCHEN WITH MAREK "MARO MUSIC" WALASZEK

Chair: **Marek Walaszek**, Addicted to Music Media Group, Jozefow, Poland

Panelists: *Bilon*
Junior Stress
Szwed SWD

Mix Kitchen is a continuation of mixing tutorial presented at AES Budapest, AES Krakow, and at the AES student meet up in Graz. I will show mixing techniques based on modern recordings and interview the artist behind the music.

Workshop 16 **Sunday, May 10**
14:00 — 15:30 **Room Opera**

AUDIO DEVELOPMENT IN DYING LIGHT

Presenter: **Paula Karbowniczek**, Techland, Wroclaw, Poland

Audio development is a crucial process that facilitates creation of engaging and atmospheric gameplay. This workshop is an in-depth look at audio design and implementation, focused on such related issues as creation of sound assets, composition of immersive soundscapes, or synchronization of music and action. It will also cover specific challenges that had to be overcome at various stages of developing *Dying Light* – the award-winning survival horror game set in a city devastated by a zombie epidemic.

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

* * *

Upcoming AES Conventions and Conferences

58th International Conference
"Music-Induced Hearing Disorders"
Aalborg, Denmark
2015 June 28–30
•
59th International Conference
"Live Sound Reinforcement"
Montreal, Canada, McGill University
2015 July 15–17
•
139th CONVENTION
New York, NY, USA
2015 October 29–November 1
•
60th International Conference
"Dereverberation and Reverberation of Audio Music and Speech"
Leuven, Belgium
2016 February 3–5

For the latest information on AES conventions and conferences, visit the AES Web site at www.aes.org