AES 124th CONVENTION PROGRAM

May 17 – 20, 2008

RAI International Exhibition & Congress Centre, Amsterdam

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

Forty-two 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 published in a timely manner in 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 124th 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 124th AES Convention Student Paper Award is:
Modeling Frequency-Dependent Boundaries as Digital Impedance Filters in FDTD and K-DWM Room Acoustic Simulations
—*Konrad Kowalczyk* (Presenting Author) and *Maarten van Walstijn*, Convention Paper 7430

To be presented on Monday, May 19 in Session P18 —Room and Architectural Acoustics and Sound Reinforcement

* * * * *

Friday, May 16 13:30 Room H
Standards Committee Meeting on SC-02-02 Digital Input/Output Interfacing

Session P1 Saturday, May 17
09:00 – 12:00 Room C-D

LOUDSPEAKERS, PART 1

Chair: **Stanley Lipshitz,** University of Waterloo, Waterloo, Ontario, Canada

09:00

P1-1 Audio Capacitors. Myth or Reality?—*Philip Duncan,*[1] *Paul Dodds,*[2] *Nigel Williams*[2]
[1]University of Salford, Salford, Greater Manchester, UK
[2]ICW, Ltd., Wrexham, Wales, UK

This paper gives an account of work carried out to assess the effects of metallized film polypropylene crossover capacitors on key sonic attributes of reproduced sound. The capacitors under investigation were found to be mechanically resonant within the audio frequency band, and results obtained from subjective listening tests have shown this to have a measurable effect on audio delivery. The listening test methodology employed in this study evolved from initial ABX type tests with set program material to the final A/B tests where trained test subjects used program material that they were familiar with. The main findings were that capacitors used in crossover circuitry can exhibit mechanical resonance, and that maximizing the listener's control over the listening situation and minimizing stress to the listener were necessary to obtain meaningful subjective test results.
Convention Paper 7314

09:30

P1-2 Perceptual Study and Auditory Analysis on Digital Crossover Filters—*Henri Korhola, Matti Karjalainen*, Helsinki University of Technology, Espoo, Finland

Digital crossover filters offer interesting possibilities for sound reproduction, but there does not exist many publications on how they behave perceptually. In this paper phase and magnitude errors in digital implementations of linear phase FIR as well as Linkwitz-Riley crossover filters are studied perceptually and by auditory analysis. In a headphone simulation listening experiment we explored the just noticeable level of degradation due to crossover filter artifacts. In a real loudspeaker experiment we explored rough guidelines for "safe" filter orders of linear-phase FIR crossover filters, which would not produce audible errors. Possibilities to predict the perceived errors were then explored using auditory analysis, including also third-octave magnitude spectrum and group delay as simple auditory correlates. Linear-phase FIR crossovers were found to produce different kinds of phase errors than Linkwitz-Riley crossovers. The auditory analysis can qualitatively explain the perceptibility degradation.
Convention Paper 7315

10:00

P1-3 **The Air Spring Effect of Flat Panel Loudspeakers**—*Daniel Beer*,[1] *Michaela Schuster*,[2] *Michael Jahr*,[2] *Alexander Reich*[1]
[1]Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany
[2]Technical University Ilmenau, Ilmenau, Germany

Flat panel loudspeakers are characterized by their low manufactured depth. Compared with conventional loudspeakers the space saving integration in existing surroundings is an advantage. From the acoustics point of view disadvantages come along with the low manufactured depth that influence the reproduction in the lower and middle frequency range. Based on measurements and FEM-simulations the reasons for this behavior were analyzed. Supplementary methods for solving this problem have been considered that are derived from conventional loudspeaker technologies.
Convention Paper 7316

10:30

P1-4 **The Inertial Air Load of a Loudspeaker Diaphragm**—*John Vanderkooy*, University of Waterloo, Waterloo, Ontario, Canada, and B&W Group Ltd., Steyning, West Sussex, UK

A typical bass loudspeaker driver has an inertial air load, which is about 30% of its actual cone mass. This air mass is often poorly understood, but it is significant in defining the resonance frequency; and the purpose of this paper is to understand the concept, clarify important aspects, and present corroborative measurements. The immediate surroundings of the diaphragm determine the low-frequency air load, and measurements on a test driver with different mountings arrangements are made and assessed, including measurements in vacuum. A loudspeaker box presents its own complications. Simulations are used to show how the air load depends on baffle size. In general, the air load may not be accurately represented by the usual approximations that apply to a piston in an infinite baffle or to a freely oscillating disk, but they do give a rough estimate.
Convention Paper 7317

11:00

P1-5 **Horn Loudspeaker Nonlinearty Comparison and Linearization Using Volterra Series**—*Delphine Bard*, University of Lund, Lund, Sweden

The characterization of a weakly nonlinear electroacoustic device with usual methods of measurement (THD, intermodulation) does not illustrate the nonlinearities themselves, but only some of their effects. Device linearization can be achieved by applying the inverse nonlinearity upstream of the device, under the condition that the nonlinearity law is known in detail. This paper presents nonlinearities behavior comparison of horn loudspeakers of different frequency ranges using an experimental method of weak nonlinearity characterization and compensation, based on a representation of the nonlinearity by Volterra series using multitone excitations.
Convention Paper 7318

11:30

P1-6 **Audibility of Phase Response Differences in a Stereo Playback System. Part I: Headphone Reproduction of Wide-Band Stimuli**—*Geoff Martin, Sylvain Choisel*, Bang & Olufsen A/S, Struer, Denmark

The audibility of phase distortion in sound reproduction systems has been the subject of many studies. However, it remains a topic of controversy, in particular in the field of loudspeaker or headphone equalization. Most studies lead to the conclusion that, although phase distortion may be audible for specific stimuli, in realistic listening situations in a room, they will go largely unnoticed. These studies, however, have focused on monophonic phase distortion; a severe limitation, since ignoring phase response in equalization can result in different phase distortion in different channels. It is the purpose of the present study to investigate the audibility of stereophonic phase mismatch in the specific case of headphone reproduction. In addition, the implications on microphone design and production are discussed.
Convention Paper 7319

Workshop 1 **Saturday, May 17**
09:00 – 10:30 **Room L**

PRODUCTION OF SURROUND SOUND WITH HDTV

Chairs: **Kimio Hamasaki**, NHK Science and Technical Research Laboratories, Tokyo, Japan
Christian Hugonnet, Audio Engineering Consultant, Paris, France

Panelists: *Florian Camerer*, ORF - Austrian TV
Toru Kamekawa, Tokyo University of Fine Arts and Music, Tokyo, Japan

HDTV is currently broadcasted in many countries over the world. 5.1 channel surround sound has made great progress in digital broadcasting. Many workshops have discussed the 5.1 channel surround sound at previous AES conventions; however, 5.1 channel sound production and broadcasting accompanying HDTV has not been discussed as often at previous conventions. This workshop will focus on the 5.1 surround sound production and broadcasting for HDTV and discuss various issues such as sound recording, sound design, postproduction, and aesthetic approaches in terms of 5.1 channel sound with large and high-resolution pictures.

Tutorial 1 **Saturday, May 17**
09:00 – 12:00 **Room N101**

PERCEPTUAL AUDIO EVALUATION

Presenters: **Søren Bech**
Nick Zacharov

The aim of this tutorial is to provide an overview of perceptual evaluation of audio through listening tests, based on good practices in the audio and affiliated industries. The tutorial is geared to anyone interested in the evaluation of audio quality and will provide a highly condensed overview of all aspects of performing listening tests in a robust manner.

Topics will include:

(1) Definition of a suitable research question and associated hypothesis;
(2) Definition of the question to be answered by the subject;
(3) Scaling of the subjective response;
(4) Control of experimental variables such as choice of signal, reproduction system, listening room, and selection of test subjects;
(5) Statistical planning of the experiments; and
(6) Statistical analysis of the subjective responses.

The tutorial will include both theory and practical examples including discussion of the recommendations of relevant international standards (IEC, ITU, ISO). The presentation will be made available to attendees and an extended version will be available in the form of the text *Perceptual Audio Evaluation*, authored by Søren Bech and Nick Zacharov.

Tutorial 2 — Saturday, May 17
09:00 – 11:00 — Room B

DIGITAL AUDIO SIGNALS, FILTERS, AND EQUALIZERS

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

In this tutorial we will first present the principles of sampling and the basic form and function of digital filters, explain the link between impulse response and frequency response, and how it leads to different forms of filtering. Next we will relate that to the similarities and differences between the designs of analog filters and digital filters. Finally, we will discuss some of the filter structures that might be used for audio equalizers.

Saturday, May 17 — 09:00 — Room H
Standards Committee Meeting on SC-02-08 Audio-File Transfer and Exchange

Saturday, May 17 — 09:00 — Room K
Technical Committee Meeting on Audio for Games

Session P2 — Saturday, May 17
09:30 – 12:00 — Room E-F

AUDIO NETWORKING

Chair: **Sascha Spors**, Technische Universität Berlin, Berlin, Germany

09:30

P2-1 EBU Tech.doc. 3326 for Interoperability between Audio over IP Units—*Lars Jonsson*,[1] *Mathias Coinchon*[2]
[1]Swedish Radio, Stockholm, Sweden
[2]European Broadcasting Union, Geneva, Switzerland

Audio over IP end units are now common in radio and TV operations for streaming programs over IP networks. The units are used to create contribution circuits from remote sites or local offices into main studio centers. The IP networks used are usually well managed corporate networks with good Quality of Service (QoS) and usually high bandwidth. Due to its availability, the Internet is also increasingly used for various cases of radio and television contribution, especially over longer distances. However, the use of high bit rates and reliable contribution transmissions over the Internet cannot be guaranteed. Correspondents have the choice in their equipment to use either ISDN or the Internet to deliver their reports. More than 20 manufacturers now provide equipment for audio over IP applications. The EBU has issued and verified a standard, EBU TECH 3326-2007, which allows for interoperability between previously not compatible Audio over IP codecs. A plug-test between nine manufacturers held in February 2008 proved that earlier incompatible units now can connect according to the new standard.
Convention Paper 7322
[This paper was presented by Mathias Coinchon]

10:00

P2-2 Audio Fingerprint and its Applications to Peer-to-Peer Systems—*Antonello D'Aguanno, Goffredo Haus*, Università degli Studi di Milano, Milano, Italy

In this paper we want to analyze the applicability of audio-fingerprint technology to peer-to-peer systems. Audio-fingerprint is a technology commonly applied to scopes like audio identification or digital rights management. Peer-to-peer is a common Internet paradigm to share various digitalized contents. We propose an improvement for typical peer-to-peer architectures (query flooding, centralized directory, hybrid architecture) that permits the application of audio fingerprint technology to these systems.
Convention Paper 7321

10:30

P2-3 Audible ICMP Echo Responses for Monitoring Ultra Low Delayed Audio Streams—*Alexander Carôt*,[1] *Alain B. Renaud*,[2] *Christian Werner*[1]
[1]University of Lübeck, Lübeck, Germany
[2]Queen's University Belfast, Belfast, Northern Ireland, UK

Playing live music on the Internet is very demanding in terms of delay, loss, or jitter and hence requires extremely reliable network conditions. Jitter is the most problematic factor because it has a direct influence on the required network buffer sizes for receiving low delay audio streams. Therefore, measuring the amount of jitter is a very complex task due to the multi-hop architecture of the Internet. So far it has been impossible to know at which hop these delay variances appear. The authors propose a solution that is able to generate an audible impression of the jitter problem for each hop.
Convention Paper 7320

11:00

P2-4 A Grid-Based Approach to the Remote Control and Recall of the Properties of IEEE1394 Audio Devices—*Philip Foulkes, Richard Foss*, Rhodes University, Grahamstown, South Africa

Typically, the configuration of audio hardware and software is not integrated. This paper discusses a software system that has been developed to remotely control and recall the properties of IEEE1394 (FireWire) audio devices via a series of graphical routing matrices. The software presents sound engineers with a graphical routing matrix that shows, along its axes, the available FireWire audio devices on a FireWire network. Inter device connection management may be performed by selecting the cross points on the grid, and intra device control may be performed via device editors that are displayed via the axes of the matrix. The software application may be hosted by a compatible Digital Audio Workstation (DAW) application to allow for the storing and recalling of the various properties associated with the devices.
Convention Paper 7323

11:30

P2-5 Can the Public Internet Be Used for Broadcast?— *Simon Daniels*, Audio Processing Technology, Belfast, Northern Ireland, UK

This paper will look at a number of examples of remote broadcasts over contended IP links and examine the key points in their success. We will talk about issues such as jitter and latency and considerations regarding essential features on IP codec equipment. The experiences of major European broadcasters trialing audio over the Public Internet will form the basis of a discussion of the pitfalls and possibilities associated with using the public Internet for essential broadcast links.
Convention Paper 7324

Historical Event
HISTORICAL CAFÉ
Saturday/Sunday/Monday 10:00 – 17:00
Tuesday 10:00 – 13:00
Room N

In the historical Café there will be on display a selection of vintage CD and compact cassette equipment from the early days; a book selection of vintage acoustic matters; a Slide Rule used for electro acoustical calculations; samples of triode applications (100 years triode use)

Saturday, May 17 10:00 Room K
Technical Committee Meeting on Coding of Audio Signals

Workshop 2 Saturday, May 17
10:30 – 12:00 Room L

USE OF RF 64: A FILE FORMAT FOR SURROUND SOUND

Chair: **Lars Jonsson**, Swedish Radio

Panelists: *Axel Holzinger*, D.A.V.I.D. GmbH, Munich, Germany
Erik Lundbeck, Swedish Television Stockholm, Stockholm, Sweden
Heinz-Peter Reykers, WDR Cologne, Cologne, Germany
Mark Yonge, AES Standards, UK

In this workshop experiences and utilization by manufacturers and users of the EBU RF 64 standard file format for surround and stereo in TV and radio production will be reviewed. AES Standard Working Group SC-02-08 have begun working on the adoption of this format as an AES standard. The format began to be used with radio automation systems in Germany and in editors. It has now found applications also for file-based TV 5.1 audio operations in Sweden and elsewhere.

Workshop 3 Saturday, May 17
10:30 – 12:00 Room O

RECORDING STUDIO DESIGN

Chair: **Fritz Fey**

Editor of *Studio Magazin*, Fritz Fey has successfully developed its activities to studio design (Studioplan). He will present various types of studio concepts, from musician personal rooms to more sophisticated installations, with testimonials of users.

Being presented will be some VIPs (very important points) on studio design from a practical point of view that are REALLY helpful for participants. There are different aspects that come across besides the usually well-bundled package of formula and theoretical know how, like basic geometry, ergonomics, economics, stereo + surround, working atmosphere, multi-use, loudspeaker environment, and many others. This workshop will show that studio design is not just a matter of absorbers and reverb time issues.

Live Sound Seminar 1 Saturday, May 17
10:30 – 12:00 Forum

DEMYSTIFYING AUDIO

Presenter: **Ralf Zuleeg**, d & b

This session describes the principles of acoustics from the wavelength to the properties of a good loudspeaker. The session is spiced with demonstrations and experiments to make it easy to digest the information.

Saturday, May 17 11:00 Room H
Standards Committee Meeting on SC-05-05 Grounding and EMC Practices

Saturday, May 17 11:00 Room K
Technical Committee Meeting on Hearing and Hearing Loss Prevention

Special Event Saturday, May 17
12:00 – 13:30 Room B

AWARDS PRESENTATION AND KEYNOTE ADDRESS

Opening Remarks:
- Executive Director Roger Furness
- President Bob Moses
- Convention Chair Peter Swarte

Program:
- AES Awards Presentation
- Introduction of Keynote Speaker by Convention Chair Peter Swarte
- Keynote Address by Tammo Houtgast

Awards Presentation
Please join us as the AES presents special awards to those who have made outstanding contributions to the

Society in such areas of research, scholarship, and publications, as well as other accomplishments that have contributed to the enhancement of our industry.

Keynote Speaker

This year's Keynote Speaker is Tammo Houtgast. Houtgast was trained as a physical engineer and worked at the Human Factors Laboratory of the Dutch Organization for Applied Physical Research. In 1974 he received his PhD with a thesis on "Lateral Suppression in Hearing." In the field of speech communication, he was involved in the development of the STI (Speech Transmission Index), i.e. modeling and predicting the effect of any speech transmission channel on speech intelligibility. Since 1994 he has been affiliated with the ENT department of the VU University Medical Center in Amsterdam, as professor of experimental audiology, working on speech reception by hearing impaired persons. His main interests are the relation between types of hearing impairment and the reduced intelligibility of speech in noise; signal-processing strategies for hearing aids to (partly) compensate for that impairment. He retired in 2006, but continues the scientific coordination of an extensive EU-funded project HearCom, aimed at improving the participation of hearing impaired persons in our modern communication society.

Houtgast's keynote address is entitled: "Speech Reception in Noise: a Review from a Personal Perspective."

It will be described how the concept of the Articulation Index, as developed over fifty years ago by French and Steinberg and others, led to the present models for quantifying the effect of noise on the intelligibility of speech. Among these models, the Speech Transmission Index (STI) is of some importance since, besides the effect of interfering noise, it also accounts for the effect of room acoustics (reverberation). In addition to its use for design purposes, the STI has proven a valuable tool for assessing the quality of speech transmission in actual situations. Besides the properties of the sound transmission path from the source to the listener, also the hearing characteristics of the listener are to be considered. Many hearing impaired and older persons experience great difficulties in understanding speech in noisy situations. The present insight in this field will be briefly reviewed.

Live Sound Seminar 2 **Saturday, May 17**
12:15 – 14:30 **Forum**

LINE ARRAYS

Presenter: **Ralf Zuleeg**, d & b

This is a session about the principles of line arrays and describes in a non-mathematical way their behavior. This unique style of explaining the physics makes it easy to adapt the theory to the real world. Demonstrations will round this session up.

Saturday, May 17 **13:00** **Room H**
Standards Committee Meeting on SC-05-02 Audio Connectors

Workshop 4 **Saturday, May 17**
13:30 – 15:30 **Room O**

MULTISENSORY INTERACTION: DOES INTERACTION IMPROVE PRESENCE?

Chair: **Renato S. Pellegrini**, sonic emotion ag, Oberglatt, Switzerland

This workshop investigates the impact of interactivity in audio environments on improving the feeling of presence within those environments. This workshop extends a series of related workshops on perception and interactivity presented at recent AES Conventions. The workshop considers the technical, creative, and psychoacoustic constraints when building interactive environments. A selection of panelists, with varying perspectives on the meaning of interactive audio, will discuss the importance of presence within their own application areas. These include gaming, training, and education. The workshop has an open format and encourages discussion between panelists and the audience relating to the meaning of interactivity and presence. The workshop is therefore suitable to experts and newcomers alike and offers a chance to share different perspectives of this topic of growing importance.

Workshop 5 **Saturday, May 17**
13:30 – 15:30 **Room L**

SPATIAL AUDIO AND SOUND DESIGN IN SPORTS

Chair: **Gerhard Stoll**, IRT, Munich, Germany

Panelists: *Dennis Baxter*, Sound for the Olympics
Andrea Borgato, Dolby
Hans Huber, IRT
Ales Koman, Slovenia TV

Sports productions are a great platform for the combination of HDTV and multichannel audio, i.e. the combination of "best picture and best sound." Since audio has a very high impact on the emotions of the viewer, in particular in the case of very good pictures, more and more attention should be paid to a stimulating sound design, which conveys a major part of the tension in a sport event. The forthcoming Euro 2008, the Beijing 2008 Summer Olympics, and the Vancouver 2010 Winter Olympics will be important examples for bringing exciting sound together with crisp pictures to the home. The applied techniques range from special microphones and according placement to the use of additional sound effects with samplers, etc., as well as the specific issues of crowd pickup in surround sound. The panelists will explain and discuss their ideas to provide new tools for the production and design of exciting sound in sports.

Tutorial 3 **Saturday, May 17**
13:30 – 15:00 **Room N101**

ABSORBERS AND DIFFUSERS

Presenter: **Trevor Cox**, University of Salford, Salford, Greater Manchester, UK

Absorbers and diffusers are two of the main design tools for altering the acoustic conditions of rooms such as studios and auditoria, semi-enclosed spaces such as stadia and the outdoor environment. Absorbers also have a crucial role within loudspeaker enclosures and in the reduction of machinery noise. This tutorial will describe state-of-the-art designs for diffusers and absorbers, as well as touching on techniques for measuring and modeling these treatments. Surface diffusion is a relatively young subject area, and case studies will be drawn upon to show when are where diffusers can be used to improve conditions. Absorption is a more established technology, and yet new materials and techniques continue to be developed today, for instance driven by the need for more sustainable buildings.

Session P3 **Saturday, May 17**
14:00 – 18:00 **Room C-D**

SPATIAL AUDIO PERCEPTION AND PROCESSING, PART 1

Chair: **Francis Rumsey**, University of Surrey, Guildford, Surrey, UK

14:00

P3-1 Objective and Subjective Evaluation of Urban Acoustic Modeling and Auralization—*Yuliya Smyrnova, Yan Meng, Jian Kang*, University of Sheffield, Western Bank, Sheffield, UK

This paper presents the results of objective and subjective evaluation of a simulation and auralization system based on model CRR—combined ray-tracing and radiosity. Auralization of an urban square has been carried out with various boundary reflection patterns (purely specular, purely diffuse, and a mix of specular and diffuse) using two audio stimuli. The subjective evaluation results reveal a strong impact of sound sources and reflection pattern. Despite similarities in objective measures, there are noticeable differences in subjective attributes between signals based on simulated and measured impulse responses, but current auralization algorithms are still adequate in simulating real urban environments.
Convention Paper 7325

14:30

P3-2 Virtual vs. Actual Multichannel Acoustical Recording—*Gavin Kearney,*[1] *Jeff Levison*[2]
[1]Trinity College, Dublin, Ireland
[2]Euphonix, Inc., Palo Alto, CA, USA

We present a comparison of live recordings of a choral ensemble versus dry recordings of the same players, with the acoustic environment reconstructed from impulse responses of the original reverberant performance space. Binaural measurements are used to objectively classify the recordings, and the perceptual attributes are investigated through a series of subjective listening tests. It is shown that the differences between dry recordings convolved with linear time-invariant (LTI) impulse responses and actual acoustical recordings can be perceived by a panel of expert listeners.
Convention Paper 7326

15:00

P3-3 Virtual Sources and Moving Targets—*Glenn Dickins, David Cooper, David McGrath*, Dolby Laboratories, Sydney, NSW, Australia

This paper presents an analysis of the effects of listener mobility on the stability of virtual source images created by a pair of loudspeakers. A spherical head is used to generate analytic head related transfer functions from which we create a simple perceptual localization model for the forward half of the horizontal plane. This model is then used to investigate changes in perceived source localization as the listener moves. The analysis demonstrates that even with this simple model, and the assumption of small listener movements, the source image becomes unstable at a relatively low frequency. Given that for such low frequencies the spherical head model is a reasonable approximation of measured HRTFs, this work suggests that individualized HRTF and pinnae functions are of little benefit when designing a virtualizer system that allows for some listener mobility.
Convention Paper 7327

15:30

P3-4 On the Use of Directional Loudspeakers to Create a Sound Source Close to the Listener—*Aki Härmä, Steven van de Par, Werner de Bruijn*, Philips Research Laboratories, Eindhoven, The Netherlands

It is sometimes desired to create an illusion that a sound source appears closer to the listener than the nearest loudspeaker location. By using highly directional loudspeakers one may manipulate the relation between direct and reverberant energy and therefore change the distance cues to make the sound source appear very close to the listener. In this paper we present a method combining highly directional sound with surround audio reproduction to produce controllable distance effects between the listener location and the nearest loudspeakers.
Convention Paper 7328

16:00

P3-5 Directional Analysis of Sound Field with Linear Microphone Array and Applications in Sound Reproduction—*Jukka Ahonen,*[1] *Ville Pulkki,*[1] *Fabian Küch,*[2] *Markus Kallinger,*[2] *Richard Schultz-Amling*[2]
[1]Helsinki University of Technology, Espoo, Finland
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

The use of a linear microphone array composed of two closely spaced omnidirectional microphones as input to teleconference application of Directional Audio Coding (DirAC) is presented. DirAC is a method for spatial sound processing, where the direction of the arrival of sound and diffuseness are analyzed and used for different purposes in reproduction. Two-dimensional plane arrays have been used so far to generate input signals for DirAC, in which case it is possible to measure directly a two-dimensional sound field. In this paper a one-dimensional linear array is used to provide input signals for one-dimensional direction and diffuseness analysis in DirAC. Listening tests are conducted to evaluate the intelligibility of speech with simultaneous talkers when the linear array is used in teleconference applications.
Convention Paper 7329

16:30

P3-6 The SoundScape Renderer: A Unified Spatial Audio Reproduction Framework for Arbitrary Rendering Methods—*Matthias Geier, Jens Ahrens, Sascha Spors*, Technische Universität Berlin, Berlin, Germany

The *SoundScape Renderer* is a versatile software framework for real-time spatial audio ren-

dering. The modular system architecture allows the use of arbitrary rendering methods. Three rendering modules are currently implemented: Wave Field Synthesis, Vector Base Amplitude Panning, and Binaural Rendering. After a description of the software architecture, the implementation of the available rendering methods is explained and the graphical user interface is shown as well as the network interface for the remote control of the virtual audio scene. Finally, the Audio Scene Description Format, a system-independent storage file format, is briefly presented.
Convention Paper 7330

17:00

P3-7 Initial Investigation of Signal Capture Techniques for Objective Measurement of Spatial Impression Considering Head Movement—*Chungeun Kim, Russell Mason, Tim Brookes*, University of Surrey, Guildford, Surrey, UK

In a previous study it was discovered that listeners normally make head movements attempting to evaluate source width and envelopment as well as source location. To accommodate this finding in the development of an objective measurement model for spatial impression, two capturing models were introduced and designed in this research, based on binaural technique: 1) rotating Head And Torso Simulator (HATS), and 2) a sphere with multiple microphones. As an initial study, measurements of interaural time difference, level difference and cross-correlation made with the HATS were compared with those made with a sphere containing two microphones. The magnitude of the differences was judged in a perceptually relevant manner by comparing them with the just-noticeable differences of these parameters.
Convention Paper 7331

17:30

P3-8 A Second Order Differential Microphone Technique for Spatially Encoding Virtual Room Acoustics— *Alexander Southern, Damian Murphy*, University of York, Heslington, York, UK

Room acoustics modeling using a numerical simulation technique known as the Digital Waveguide Mesh (DWM) has previously been presented as a suitable method for measuring spatial Room Impulse Responses (RIR) of virtual enclosed spaces. In this paper a new method for capturing the DWM modeled soundfield using an array of spatially distributed pressure-sensitive receivers is presented. The polar response of the formed 2nd order virtual microphone is measured and compared to the theoretical polar response. This approach is proven to be capable of decomposing the modeled soundfield into second order spherical harmonic components that are typically associated with 2nd order Ambisonics.
Convention Paper 7332

Session P4 **Saturday, May 17**
14:00 – 18:00 **Room E-F**

LOW BIT-RATE AUDIO CODING

Chair: **Karlheinz Brandenburg**, Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany

14:00

P4-1 Time-Varying Transform for High Quality Audio Communication Codecs—*Pierrick Philippe*,[1] *David Virette*,[2] *Balázs Kövesi*[2]
[1]France Télécom R&D, Cesson Sevigne, France
[2]France Télécom R&D, Lannion, France

High quality audio communication is a current challenge addressed by the standardization committees. In this context, ITU and MPEG recently issued standards for high quality coding of both speech and music contents. Transform coding is used and allows quality commensurate with bit rates regardless of the audio content. Up to now, only constant transform sizes were used in these coding schemes since time varying transform needed look-ahead for perfect reconstruction, hence adding further delay. In this paper we demonstrate how variable transform sizes can be used without affecting the coding delay. Based on filter bank theory, a framework avoiding look-ahead is presented. The quality improvement offered by the proposed solution is illustrated in the context of MPEG4 enhanced low delay AAC.
Convention Paper 7333

14:30

P4-2 Differential Graph-Based Coding of Spikes in a Biologically-Inspired Universal Audio Coder — *Ramin Pichevar, Hossein Najaf-Zadeh, Louis Thibault, Hassan Lahdili*, Communications Research Centre, Ottawa, Ontario, Canada

In a previous work we showed that it is possible to code audio materials using a biologically-inspired universal audio coder based on matching pursuit. The best atoms/kernels chosen by matching pursuit are represented by spikes to reflect the biologically-inspired nature of the algorithm. In that work, each spike or atom was defined by parameters such as timing, channel frequency, amplitude, chirp factor, etc., that were encoded independently. However, encoding each atom/spike as a separate entity is very bit consuming. In the present paper, we propose algorithms to encode only the difference between parameters associated with spikes. Hence, we assume that each spike/atom is a node in a graph and choose the sequence of spikes that will minimize the differential encoding costs. Methods based on minimum spanning tree and traveling salesman are proposed and compared for the graph-based optimization of the code.
Convention Paper 7334

15:00

P4-3 Unraveling the Relationship between Basic Audio Quality and Fidelity Attributes in Low Bit-Rate Multichannel Audio Codecs—*Paulo Marins, Francis Rumsey, Slawomir Zielinski*, University of Surrey, Guildford, Surrey, UK

Prior to this study the evaluation of multichannel audio codecs has been done mainly according to the ITU-R standards BS.1116 and BS.1534. Basic audio quality is the only perceptual attribute assessed in the majority of these tests. This approach, although efficient for measuring the overall quality of several codecs at once, does not provide reasons why a particular codec is rated as better as or worse than another. In this paper fidelity attributes were included; these were based on the attributes suggested in the ITU-R standards but have not been used explicitly in codec evaluation up to now. In this experiment the perceptual importance of these attributes and their contribution to the basic audio quality of low bit-rate surround sound codecs were investigated.
Convention Paper 7335

15:30

P4-4 A New Perceptual Model for Audio Coding Based on Spectro-Temporal Masking—
Steven van de Par,[1] *Jeroen Koppens*,[2] *Armin Kohlrausch*,[1,3] *Werner Oomen*[2]
[1]Philips Research Europe, Eindhoven, The Netherlands;
[2]Philips Applied Technologies, Eindhoven, The Netherlands;
[3]Eindhoven University of Technology, Eindhoven, The Netherlands;

In psychoacoustics, considerable advances have been made recently in developing computational models that can predict the discriminability of two sounds taking into account spectro-temporal masking effects. These models operate as artificial observers by making predictions about the discriminability of arbitrary signals [e.g., Dau et al., J. Acoust. Soc. Am. 99, Vol. 36(15), 1996]. Therefore, such models can be applied in the context of a perceptual audio coder. A drawback, however, is the computational complexity of such advanced models, especially because the model needs to evaluate each quantization option separately. In this paper a model is introduced and evaluated that is a computationally lighter version of the Dau model but maintains its essential spectro-temporal masking predictions. Listening test results in a transform coder setting show that the proposed model outperforms a conventional purely spectral masking model and the original model proposed by Dau.
Convention Paper 7336

16:00

P4-5 Delayless Mixing—On the Benefits of MPEG-4 AAC-ELD in High Quality Communication Systems—*Markus Schnell,*[1] *Markus Schmidt*,[1] *Per Ekstrand*,[2] *Tobias Albert,*[1] *Daniel Przioda,*[1] *Manfred Lutzky,*[1] *Ralf Geiger*,[1] *Vesa Ruoppila*,[3] *Fredrik Henn*,[2] *Erlend Tårnes*[4]
[1]Fraunhofer IIS, Erlangen, Germany
[2]Dolby Sweden, Stockholm, Sweden
[3]Dolby Germany, Nuremberg, Germany
[4]Tandberg, Oslo, Norway

Tele- and video conferencing systems for modern business communication are managed by central hubs, so-called multipoint control units (MCU). One major task of these units is the mixing of audio streams from the participating sites. This is traditionally done by decoding the streams, mixing in time domain, and then re-encoding of the mixed signals. This requires additional processing power, leads to increased delay, and degraded audio quality. The paper demonstrates how the recently standardized MPEG-4 Enhanced Low Delay AAC (AAC-ELD) codec offers a solution to these problems by efficient and delayless mixing in the transform domain of the codec.
Convention Paper 7337

16:30

P4-6 Low-Power MPEG-4 HE-AAC Version-2 Encoder— *Chi-Min Liu, Han-Wen Hsu, Chung-Han Yang, Wen-Chieh Lee*, National Chiao Tung University, Hsinchu, Taiwan

In MPEG-4 HE-AAC version-2 encoder, the analysis/synthesis complex-exponential modulation filter banks are used in spectral band replication (SBR) and parametric stereo (PS) coding. Due to the aliasing interference, the complex banks instead of real banks are adopted in the SBR and PS coding. However, the additional overhead from the complex values in the CEMFB and the subsequent processing have led to high operational overhead. Our previous work has designed the SBR encoders based on the real-domain cosine modulation filter banks; we proposed a complexification-based approach for the SBR coding. This paper extends the work into PS coding. An approximate method for parameters estimation is proposed to save operational overhead with only one CEMFB-analysis channel. Also, a phase-adjustment down-mixing method is proposed to reduce energy vanish effects.
Convention Paper 7338

17:00

P4-7 Low Complexity Bit Allocation Algorithms for MP3/AAC Encoding—*S Nithin*,[1] *Kumaraswamy Suresh,*[2] *T. V. Sreenivas*[2]
[1]National Institute of Technology, Surathkal, India
[2]Indian Institute of Science, Bangalore, India

We have developed two reduced complexity bit-allocation algorithms for MP3/AAC based audio encoding, which can be useful at low bit-rates. One algorithm derives optimum bit-allocation using constrained optimization of weighted noise-to-mask ratio and the second algorithm uses decoupled iterations for distortion control and rate control, with convergence criteria. MUSHRA based evaluation indicated that the new algorithm would be comparable to AAC but requiring only about 1/10th the complexity.
Convention Paper 7339
[Paper presented by TV Sreenivas]

17:30

P4-8 Linear Filtering in MDCT Domain—*Kumaraswamy Suresh, T. V. Sreenivas*, Indian Institute of Science, Bangalore, India

In this paper expressions for convolution multiplication properties of MDCT are derived starting from equivalent DFT representations. Using these expressions, methods for implementing linear filtering through block convolution in the MDCT domain are presented. The implementation is exact for symmetric filters and approximate for non-symmetric filters in the case of rectangular window-based MDCT. For a general MDCT window function, the filtering is done on the windowed segments and hence the convolution is approximate for symmetric as well as non-symmetric filters. This approximation error is shown to be perceptually insignificant for symmetric impulse response filters. Moreover, the inherent 50% overlap between adjacent frames used in MDCT computation does reduce this approximation error similar to smoothing of other block processing errors. The presented techniques are useful for compressed domain processing of audio signals.
Convention Paper 7340
[Paper presented by TV Sreenivas]

Session P5 Saturday, May 17
14:00 – 15:30 Topaz Lounge

POSTERS: MICROPHONES AND LOUDSPEAKERS

14:00

P5-1 A Study of Electrostatic Forces in Single-Acting Condenser Digital Transducer—*Libor Husník*, Czech Technical University in Prague, Prague, Czech Republic

One of the possibilities to design a transducer with the direct digital-to-analog conversion, sometimes called a digital loudspeaker, is the miniature condenser transducer manufactured on a silicon chip. Only recently has this micro technology been made available commercially, which can open further application possibilities. This paper is aimed at the study in which the back electrode of the electrostatic transducer is partitioned into sections having total areas proportional to powers of 2. Since electrostatic force acting on the membrane is affected by the distribution of bit groups, which cannot be even, said electrostatic force will not be a linear function of the signal voltage. Correcting coefficients for some arrangements are searched for.
Convention Paper 7341

14:00

P5-2 Ultra-Thin Micro-Loudspeaker Using Oblique Magnetic Circuit—*Toshiyuki Matsumura,*[1] *Shuji Saiki,*[1] *Sawako Usuki,*[1] *Koji Sano*[2]
[1]Matsushita Electric Industrial Co., Ltd., Kadoma City, Osaka, Japan
[2]Panasonic Electronic Devices Co., Ltd., Matsusaka City, Japan

More and more functions are installed to a mobile phone, but the size of the handset has become smaller. Devices installed in the mobile phone have been required to be downsized or thinner. Micro-loudspeakers installed to the mobile phone are required to be thinner. They are required to become both thinner and reproduce high quality sound. However, it has been very difficult to make thinner micro-loudspeakers without deteriorating the acoustic performance because the structure of conventional dynamic micro-loudspeaker is not suitable to make it thinner. We have succeeded in developing an ultra-thin micro-loudspeaker using Oblique Magnetic Circuit, which is 1.5 mm thick (45% thinner than conventional dynamic micro-loudspeakers) without deteriorating the acoustic performance.
Convention Paper 7342

14:00

P5-3 A Novel Glass Laminated Structure for Flat Panel Loudspeakers—*Olivier Mal,*[1] *Marek Novotny,*[1] *Bart Verbeeren*,[1] *Neil Harris*[2]
[1]AGC Research & Development Centre, Jumet, Belgium
[2]New Transducers Ltd. (NXT), Cambourne, UK

A new, patented "sandwich structure" has been developed for various audio applications, in which thin glass sheets are laminated with a special PVB (Polyvinyl Butyral) film to eliminate typical acoustical weaknesses of monolithic glass and standard laminated solutions. The glass improvements include suppression of ringing of the audio signal and a much more flexible and lightweight glass structure. It results in flatter frequency response (both on-axis and 180° power response) and better transport of vibrations in the glass surface. In addition, better acoustical sensitivity and mechanical resistance are achieved. In this paper, after defining the structure of the developed laminated glass solution, we compare its performances to previously tried monolithic and laminated glass solutions. We also emphasize the key factors influencing the final acoustical properties. Finally, we introduce potential application fields for the developed structure.
Convention Paper 7343

14:00

P5-4 A Digitally Direct Driven Dynamic-Type Loudspeaker—*Ryota Saito, Akira Yasuda, Kazushige Kuroki, Tomohiro Tsuchiya, Naoto Shinkawa*, Hosei University, Koganei, Tokyo, Japan

If a loudspeaker can be driven digitally, all processes from the input to the output can be performed digitally without the use of analog components such as power amplifiers; and a small, light, and high-quality speaker system can be realized. In this paper we propose the basic principle behind Digital-Speaker, and a digitally driven dynamic-type loudspeaker provided with multiple voice coils employing multibit delta-sigma modulation. The piezoelectric-speaker used in our previous study is replaced by the voice coil. The prototype is implemented along with a FPGA, CMOS dri-

vers, and a dynamic-type loudspeaker. The THD and SPL are approximately 0.1% and 104 dB, respectively, and the output power is 1 W even when the power supply voltage is 1 V.
Convention Paper 7344

14:00

P5-5 Accelerated Power Test Analysis Based on Loudspeaker Life Distribution—*Xu Wang, Yong Shen, Zhicheng Wu*, Nanjing University, Nanjing, China

For the loudspeaker manufacturers, the long time spent on power tests made by relative standards or buyers has deeply influenced the period of product design and development. The authors apply the theory of reliability to cut the duration of loudspeaker power tests. On the basis of experiment data, a model of loudspeaker life distribution is propounded, from which an accelerated factor of the loudspeaker power test is derived, and then the characteristics of the loudspeaker under normal working conditions can be estimated. The method can be conveniently used on relative power tests and shorten the duration of the tests effectively.
Convention Paper 7345

14:00

P5-6 Perception and Physical Behavior of Loudspeaker Nonlinearities at Bass Frequencies in Closed vs. Reflex Enclosures—*Jukka Rauhala, Jukka Ahonen, Miikka Tikander, Matti Karjalainen*, Helsinki University of Technology, Espoo, Finland

This paper examines loudspeaker nonlinearities at bass frequencies in closed and reflex enclosures using signal analysis and perceptual evaluation methods. The nonlinearities are investigated by driving the loudspeakers to be compared with sinusoidal and musical test tones. The produced responses are evaluated in terms of diaphragm displacement, harmonic distortion, and bandwise distortion. In addition, a listening experiment is conducted in order to determine how the nonlinearities are perceived in both reflex and closed enclosures. The results show that with signals that have energy close to the tuning frequency of the reflex port produce more distortion with the closed enclosure. On the other hand, acoustic bass test tone behaved in an opposite way causing more distortion with the reflex enclosure. These phenomena were verified with the listening tests.
Convention Paper 7346

Workshop 6 **Saturday, May 17**
14:00 – 16:00 **Room B**

TRENDS OF STORAGE TECHNOLOGIES FOR AUDIO RECORDING AND MASTERING

Chair: **Kimio Hamasaki**, NHK Science and Technical Research Laboratories, Tokyo, Japan

Panelists: *Noboru Harada*, NTT Communication Science Laboratories, Tokyo, Japan
Koichi Kitamura, Recording Industry Association of Japan, Tokyo, Japan
Kunimaro Tanaka, Teikyo Heisei University, Tokyo, Japan
Junichi Yoshio, Pioneer Corporation, Tokyo, Japan

Various technologies in terms of storage are currently used for audio recording, mastering, and archiving. Necessary data rate is also increasing due to the complexity of audio production —such as multichannel sound production. This workshop reviews the current situation of technologies regarding the storage for audio recording, mastering, and archiving and will discuss the future of storage and archives. Panels from both media industries and production will discuss the requirements for the next-generation storage and next-generation audio recording and archiving systems.

Historical Event
SHANNON, BEETHOVEN, AND THE COMPACT DISC
Saturday, May 17, 14:00 – 15:00
Room N

Presenter: **Kees Schouhamer Immink**

An audio compact disc (CD) holds up to 74 minutes, 33 seconds of sound, just enough for a complete mono recording of Ludwig van Beethoven's Ninth Symphony ("Alle Menschen werden Brüder") at probably the slowest pace it has ever been played, during the Bayreuther Festspiele in 1951 and conducted by Wilhelm Furtwängler. Each second of music requires about 1.5 million bits, which are represented as tiny pits and lands ranging from 0.9 to 3.3 micrometers in length. More than 19 billion channel bits are recorded as a spiral track of alternating pits and lands over a distance of 5.38 kilometers (3.34 miles), which are scanned at walking speed, 4.27 km per hour. This year it is 25 years ago that Philips and Sony introduced the CD. We will discuss the various crucial technical decisions made that would determine the technical success or failure of the new medium.

Special Event
WAVE FIELD SYNTHESIS DEMONSTRATION
Saturday, May 17, 14:00 – 17:00
Sunday, May 18, 10:00 – 17:00
Monday, May 19, 10:00 – 17:00
Tuesday, May 20, 10:00 – 15:00
Room G

In 1988, Guus Berkhout published his paper "A Holographic Approach to Acoustic Control" [J. Audio Eng. Soc. 36, pp. 977 - 995] where he introduced Wave Field Synthesis as a new concept for sound reproduction without "sweet spot" limitations. Now, 20 years later, WFS is recognized as a favorite technique for spatial sound reproduction with a high potential of applications. Research is done at many institutes to match the perception of critical listeners with the technical state-of-the-art. The production of content dedicated to WFS performance is slowly but steadily increasing.

In this special event, "20 Years WFS," a series of examples of such content will be demonstrated. Old friends that took part in the development of WFS from the beginning as well as later converts will cooperate in a program where traditional and electronic music, acoustic landscapes, and movie fragments illustrate the possibilities of Wave Field Synthesis. The demos are realized with the portable WFS system of the organization for spatial rendering of electronic music "The Game of Life."

Moreover, this special event will be a platform where people can discuss past, present, and—most important—the future of Wave Field Synthesis.

Saturday, May 17 14:00 Room K
Technical Committee Meeting on High Resolution Audio

Live Sound Seminar 3 Saturday, May 17
15:00 – 18:00 Forum

MIXING CONSOLES

Presenter: **Ruben van der Goor**

The digital mixing console as the center of a PA audio system. Under this headline theoretical and practical aspects of system EQs, room alignment, sound creation, signal processing, and applications of VCM technology will be explained as well as analog and digital pros and cons with live demonstrations.

Saturday, May 17 15:00 Room H
Standards Committee Meeting on SC-04-03 Loudspeker Modeling and Measurement

Saturday, May 17 15:00 Room K
Technical Committee Meeting on Human Factors in Audio Systems

Tutorial 4 Saturday, May 17
15:30 – 18:00 Room N101

INTELLIGENT AUDIO SYSTEMS: A REVIEW OF THE FOUNDATIONS AND APPLICATIONS OF SEMANTIC AUDIO ANALYSIS AND MUSIC INFORMATION RETRIEVAL

Presenters: **Perfecto Herrera-Boyer**, Universitat Pompeu Fabra
Jay LeBoeuf, Imagine Research

This tutorial will target students, researchers, and industry audio engineers who are unfamiliar with the field of Music Information Retrieval (MIR). We will demonstrate the myriad of exciting technologies enabled by the fusion of basic signal processing techniques with machine learning. The presentation will be a high-level, applied, multimedia-rich, overview of the building blocks of MIR systems. Our goal is to make highly-interdisciplinary technologies and dauntingly-complex algorithms approachable. In the spirit of modern cooking shows, we will perform numerous demonstrations, including on-the-fly coding of basic "intelligent audio" systems, prepared system demonstrations, and prepared audio examples demonstrating more complex systems.

Session P6 Saturday, May 17
16:00 – 17:30 Topaz Lounge

POSTERS: MOBILE PHONE AUDIO

16:00

P6-1 Enhancements to the SBC CODEC for Voice Communication in Bluetooth Devices—*Laurent Pilati*,[1] *Mohammad Zad-Issa*[2]
[1]Broadcom Corp., Sophia Antipolis, France
[2]Broadcom Corp., Irvine, CA, USA

The Bluetooth Audio Distribution profile has uses low complexity sub-band coder (SBC) as its mandatory audio compression codec. More recently, SBC has been selected for Bluetooth wideband voice communication. Since SBC was first designed with audio compression, it does not incorporate the features that speech coders commonly use. The use of voice activity detection and comfort noise generation to reduce bandwidth usage and power consumption is an example. In this paper we investigated extensions for SBC that would make it better suited for voice compression in the Bluetooth framework. The proposed enhancements were evaluated on the basis of their impact on voice quality, their implementation requirements, and their bandwidth power savings.
Convention Paper 7347

16:00

P6-2 Efficiently Shuffling Large Sets of Clips—*Ulrich Herrmann*, austriamicrosystems, Graz, Austria

A method for randomly shuffling through large sets of video or audio clips is presented in this paper. Many up-to-date devices have only a rather limited capability of shuffling only up to 200 or 256 songs. This algorithm presents a way of shuffling even large sets with an almost unlimited number of items. It also provides the ability to traverse back and forth with little processing power on today's micro controllers. All this is done with few bytes of code and almost no RAM.
Convention Paper 7348

16:00

P6-3 Hardware/Software Co-Design of Multi-Format Audio Decoder—*ChangYong Son, KangEun Lee, DoHyung Kim, Soojung Ryu, Shihwa Lee*, Samsung Advanced Institute of Technology, Suwon, Korea

This paper presents a hardware/software co-design method for the implementation of a multi-format audio decoder with ultra low power, small chip size, and high flexibility, which are the most critical factors in embedded devices. This approach can provide both flexibility and low power with high performance in such a way that hardware implementation has been focused on the commonly used critical blocks of multiple audio decoders having intensive computations. Hardware blocks are well modularized to allow easy and rapid architecture exploration of several digital audio standards. The proposed system can decode an MP3 bitstream using only about 4 MHz clock frequency and AAC bitstream using only about 7 MHz clock frequency on average at the sampling rate of 48 kHz and the target bit rate of 128 kbps/stereo.
Convention Paper 7349

16:00

P6-4 Audio Enhancement for Portable Device-Based Speech Applications—*Rory Turnbull, Peter Hughes, Steve Hoare*, BT Group CTO, Ipswich, Suffolk, UK

Portable devices with audio capabilities necessitate the use of small transducers, often with poor frequency responses. This can be a limiting factor in the perception of the speech quality of VoIP

services hosted on such a device. This paper seeks to investigate the problem and provide practical solutions through the use of appropriate enhancement technologies. The paper covers the use of equalization, dynamic range compression, and psychoacoustic bass enhancement as possible methods for improving intelligibility. Subjective tests are used to evaluate the enhancements prior to making practical recommendations.
Convention Paper 7350
[Paper was presented by Peter Hughes]

16:00

P6-5 An Efficient, Low-Noise Filter Architecture for Bass Processing on a Processor Core—*Peter Eastty, Nathan Bentall, Duncan Stott*, Oxford Digital Limited, Stonesfield, Oxfordshire, UK

Bass Enhancement is becoming popular in many forms of consumer devices. Whatever technique is used on whatever processor, the low frequency filtering involved is frequently the major determinant of system signal-to-noise ratio. The architecture described combines an efficient, cascaded, low-pass FIR filter and a poly-phase adaptation of standard low frequency IIR filtering. The resulting circuit achieves a 20 to 30 dB improvement in signal to noise ratio at the cost of only 12 instructions per sample. The technique may be applied to any bass processing using fixed or floating point processors. Complete design tables for the cascaded FIR filters are given as are noise spectrum plots of the results.
Convention Paper 7351

16:00

P6-6 Implementation of Dynamic Voltage and Frequency Scaling on Portable Audio Players—*Dahyanto Harliono, Woon-Seng Gan*, Nanyang Technological University, Singapore

Current portable computing devices demand not only higher performance but also lower power consumption. For the same reason, this research aims to build a framework that enables a rapid design of energy-efficient embedded systems. Specifically, this research is focused on a dynamic voltage scaling algorithm, which has been found effective in saving power consumption. We developed a method of scaling voltage and frequency dynamically on the latest embedded processor, jointly designed by Analog Devices and Intel. The rationale behind this method is to avoid the processor being idle in high operating frequency and voltage. Instead, the processor can save power by running its task at a lower frequency and voltage, and completing it just before the real-time deadline. Furthermore, our method can also be implemented in other embedded processors with voltage-frequency scaling features.
Convention Paper 7352

Workshop 7 — Saturday, May 17
16:00 – 18:00 — Room B

LOUDNESS—A CHANGE OF PARADIGMS?

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

Panelists: *Andrew Mason*, BBC R&D
Jean-Paul Moerman, VRT – Flemish Broadcasting
Gerhard Spikofski, IRT Munich
Alessandro Travaglini, Fox Italia

Canada and Fox TV in Italy have proven that the switch to loudness measurement instead of peak measurement is a possible route—and one which serves to solve the most frequent complaint by consumers. Are we at the beginning of a new practice in audio levels in Broadcasting? Standards are also in place (ITU-R BS.1770), metering tools incorporating it appear, etc. The panel will discuss the state of things, the transition workflow, the tools available, different possible solutions, and practical experiences, also with different methods of measurement.

Historical Event
ELECTRO ACOUSTIC SLIDE RULE
Saturday, May 17, 16:00 – 17:00
Room N

Presenter: **Peter Swarte**

In the transition of the seventies to the eighties in the previous century, the programmable calculators could be applied for the design of a sound reinforcement system. The disadvantage was that you were not in the position to make quick changes in the parameters, e.g., the reverberation time of the room, the directivity factor of the sources, and the speech intelligibility based upon the development work of Peutz as he had published in a paper during the AES convention in 1971, New York. A slide rule reduces these disadvantages in a dramatic way. By sliding you were able to make an estimation of the results (costs) and the margins you worked in. Peter Swarte will demonstrate some examples of real situations with the help of this slide rule. The results are still valid and based upon the statistical approach of sound in (semi) enclosed rooms. Furthermore, the relationship between several parameters like acoustic power, directivity, diffuse sound, direct sound, absorption, and reverberation are part of the demonstration.

Saturday, May 17 — 16:00 — Room K
Technical Committee Meeting on Studio Practices and Production

Student Event/Career Development
OPENING AND STUDENT DELEGATE ASSEMBLY MEETING – PART 1
Saturday, May 17, 16:30 – 18:00
Room O

Chair: **Suzana Jakovic**

Vice Chair: **Misato Yamada**

The first Student Delegate Assembly (SDA) meeting is the official opening of the convention's student program and a great opportunity to meet with fellow students from all corners of the world. This opening meeting of the Student Delegate Assembly will introduce new events and election proceedings, announce candidates for the coming year's election for the Europe/International Regions, announce the finalists in the recording competition categories, hand out the judges' sheets to the nonfinalists, and announce any upcoming events of the convention. Students and student sections will be given the opportunity to introduce themselves and their activities, in order to stimulate international contacts.

All students and educators are invited to participate in this meeting. Election results and Recording Competition and Poster Awards will be given at the Student Delegate Assembly Meeting–2 on Tuesday, May 20, at 14:30.

Saturday, May 17 17:00 Room H
Standards Committee Meeting on SC-04-01 Acousics and Sound Source Modeling

Saturday, May 17 17:00 Room K
Technical Committee Meeting on Signal Processing

Special Event
MIXER PARTY
Saturday, May 17, 18:00 – 19:30
See Flyer for Details

A mixer party will be held on Saturday evening to enable convention attendees to meet in a social atmosphere after the opening day's activities to catch up with friends and colleagues from the industry. There will be a cash bar and snacks.

Student Event/Career Development
STUDENT PARTY
Saturday, May 17, 20:00 –22:00
De Valk

To welcome you to Amsterdam the Dutch student section has organized a party in the evening on the first day of the convention (Saturday, May 17). The party will take place in the venue De Valk at the IJplein 3 (at the back of Amsterdam central station, across the water). At the party there will be audio colleagues from all over the world, drinks, and music. Music is provided by the band Dear Watson and later in the evening DJ's will heat up the dance floor. Of course everybody is welcome to this party—you don't have to be a student to party like a student.

Contribution toward expenses: 5 EUROS. Tickets are sold at SDA-1.

The Student party is sponsored by SAE Institute.

Special Event
BOAT TRIP
Saturday, May 17, 20:30 – 22:00

Join us for the Grachten Tour boat trip through the Amsterdam canals. The tour Ishows parts of Amsterdam that cannot be seen "on land." What better way to explore Amsterdam's ancient city center than by going on a tour through the city's canals? This is definitely an experience not to be missed during a visit to Amsterdam. Tickets will be available at the Special Events desk.

Session P7 Sunday, May 18
09:00 – 11:30 Room C-D

LOUDSPEAKERS, PART 2

Chair: **Ronald Aarts**, Philips Research, Eindhoven, The Netherlands

09:00

P7-1 Low-Frequency Extension of Gated Loudspeaker Measurements—*Juha Backman*, Nokia Corporation, Espoo, Finland

The free-field response of a loudspeaker system can be approximated through a gated measurement, made in a sufficiently large space. The frequency resolution is nominally determined by the time gap between the direct sound and the first reflection, but the actual low-frequency accuracy of gated measurements is reduced also by the group delay of the loudspeaker itself. The group delay at low frequencies may cause a large fraction of the energy sound radiation to be cut off, underestimating the low-frequency response. A method is presented to estimate the approximate low-frequency response from the impedance measurement of the loudspeaker and to use the response to pre-process the acoustical measurement to improve the accuracy of the gated measurement.
Convention Paper 7353

09:30

P7-2 Measurement and Fourier-Bessel Analysis of Loudspeaker Radiation Patterns Using a Spherical Array of Microphones—*Filippo M. Fazi,*[1] *Vincent Brunel,*[1] *Philip A. Nelson*,[1] *Lars Hörchens*,[2] *Jeongil Seo*[3]
[1]University of Southampton, Highfield, Southampton, UK
[2]Delft University of Technology, Delft, The Netherlands
[3]Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea

Loudspeakers are widely used in three-dimensional sound field reconstruction systems, but their spatial directivity features are relatively little-known. In this paper a hemispherical array of 40 microphones was designed and built in order to measure the pressure field radiated by different commercially available loudspeakers. The spatial samples of the acoustic pressure were processed in order to estimate the truncated Fourier-Bessel expansion of the sound field, which allows the reconstruction of the 3-D radiation pattern. An analysis of the errors involved in the estimation was also performed with a numerical model of the array.
Convention Paper 7354

10:00

P7-3 Turbulent and Viscous Air Friction in the Mid-High Frequency Loudspeaker—*Ivan Djurek*,[1] *Antonio Petosic*,[1] *Danijel Djurek*[2]
[1]University of Zagreb, Zagreb, Croatia
[2]Alessandro Volta Applied Ceramics (AVAC), Zagreb, Croatia

Mid-high frequency loudspeaker with resonant frequency f = 982 Hz has been studied in atmospheres of air, He^4, D_2 and H_2 at pressures ranging 0-1 bar. The measurements of viscous and turbulent contributions to the friction entering Q-factor showed significant difference as compared to a low frequency loudspeaker. The resonant frequency in air is considerably lower in an evacu-

ated space than at 1 bar, and this differs from the low frequency loudspeaker, when the opposite is true. Measurements showed that imaginary part of viscous friction in Navier-Stokes equation is dominant, while contribution of the real part to the friction term is less significant, and Navier-Stokes equation reduces to the Stokes form [. . .], when imaginary part of the viscous force reduces effective vibration mass, which in turn enables operation of the loudspeaker at high frequency. The data were interpreted in terms of Greenspan theory of the piston radiator.
Convention Paper 7355

10:30

P7-4 Modeling of an Electrodynamic Loudspeaker including Membrane Viscoelasticity—*Antonio Petosic,*[1] *Ivan Djurek*,[1] *Danijel Djurek*[2]
[1]University of Zagreb, Zagreb, Croatia
[2]Alessandro Volta Applied Ceramics (AVAC), Zagreb, Croatia

The model is proposed based upon viscoelastic properties of the loudspeaker membrane, and properties considered include stress-strain hysteresis, creeping effect, initial stress effect, and appearance of the temperature fluctuations on the membrane surface. The creeping displacement response dependent on the step-like excitation current has been measured on different loudspeaker configurations, and listed effects were analyzed in terms of the *N*-order Bennewitz-Rötger differential equation, commonly used for description of the system of vibrating viscoelastic body. The main parameter in this equation is inverse stress parameter which connects friction and restoring term in the loudspeaker vibrating system.
Convention Paper 7356

11:00

P7-5 On a Novel Concept of Membrane Suspension in an Electrodynamic Loudspeaker—*Danijel Djurek*,[1] *Ivan Djurek,*[2] *Antonio Petosic*[2]
[1]Alessandro Volta Applied Ceramics (AVAC), Zagreb, Croatia
[2]University of Zagreb, Zagreb, Croatia

A laboratory model of an electrodynamic loudspeaker has been realized with the membrane suspended on a hollow elastic torus positioned in the bottom of the membrane, close to the voice coil. This geometry removes the torque in the membrane coming from the maximum possible distance of the suspension on the outer rim from the voice coil. The suppressed torque results in the suppression of the Bessel vibration modes, which generate stochastic deformation tilts on the membrane surface. Such tilts contribute to the intrinsic friction of the membrane, and their absence results in minor viscoelastic losses. Lateral rigidity of the torus is sufficient for operation of the loudspeaker without centric fixation.
Convention Paper 7357

Session P8 Sunday, May 18
09:00 – 12:30 Room E-F

WAVE FIELD SYNTHESIS

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

09:00

P8-1 The Theory of Wave Field Synthesis Revisited—*Sascha Spors*,[1] *Rudolph Rabenstein*,[2] *Jens Ahrens*[1]
[1]Technische Universität Berlin, Berlin, Germany
[2]University of Erlangen-Nuremberg, Erlangen, Germany

Wave field synthesis is a spatial sound field reproduction technique aiming at authentic reproduction of auditory scenes. Its theoretical foundation was developed almost 20 years ago and has been improved considerably since then. Most of the original work on wave field synthesis is restricted to the reproduction in a planar listening area using linear loudspeaker arrays. Extensions like arbitrarily shaped distributions of secondary sources and three-dimensional reproduction in a listening volume have not been discussed in a unified framework so far. This paper revisits the theory of wave field synthesis and presents a unified theoretical framework covering arbitrarily shaped loudspeaker arrays for two- and three-dimensional reproduction. The paper additionally gives an overview on the artifacts resulting in practical setups and briefly discusses some extensions to the traditional concepts of WFS.
Convention Paper 7358

09:30

P8-2 A Finite Difference Time-Domain Approach to Analyzing Room Effects on Wave Field Synthesis Reproduction—*Robert Oldfield, Ian Drumm, Jos Hirst*, University of Salford, Salford, Greater Manchester, UK

Probably the largest pit-fall to accurate audio reproduction using wave field synthesis (WFS) is the listening space. The WFS theory assumes free field, source free conditions that are seldom the case for practical sound reproduction. There is consequently a need to determine what effect the reproduction room has upon the synthesized sound field. This paper presents a finite difference time-domain (FDTD) approach to predicting the sound field in a room with arbitrary geometry and frequency dependent absorbing boundaries. A significant benefit to using FDTD is that the WFS system can be modeled both as part of the room and also in free-field conditions; therefore distortion of the sound field from the acoustics of the reproduction room can be quantified.
Convention Paper 7359

10:00

P8-3 Wave Field Synthesis Evaluation Using the Minimum Audible Angle in a Concert Hall—*Georgios Marentakis*,[1] *Etienne Corteel*,[2] *Stephen McAdams*[1]
[1]McGill University, Montreal, Quebec, Canada
[2]sonic emotion ag, Oberglatt, Switzerland

Localization accuracy with Wave Field Synthesis (WFS) was estimated in a variable-acoustics concert hall. Contrary to previous studies, we employed a Minimum Audible Angle (MAA) paradigm as a measure of localization performance. The MAA was estimated for three different listening positions, three orientations of the listeners (0, 60, 90 degrees) and two acoustical conditions. WFS was found to produce satisfying localization cues that depend little on the reverberation time of the room and only weakly on the position of the listener.
Convention Paper 7360

10:30

P8-4 Objective and Subjective Analysis of Localization Accuracy in Wave Field Synthesis—*Joseph Sanson*,[1] *Etienne Corteel*,[2] *Olivier Warusfel*[1]
[1]IRCAM, Paris, France
[2]sonic emotion ag, Oberglatt, Switzerland

This paper analyses localization inaccuracies in the synthesis of virtual sound sources using Wave Field Synthesis (WFS), particularly at high frequencies. Objective and perceptual analyses are conducted through a binaural simulation of the actual sound field reproduced at the listener's ears. The simulation consists in summing the respective contribution of each array transducer after filtering it with the appropriate HRTF according to the considered listener's position. High-pass filtered white noises are used as a critical signal to investigate the impact of aliasing on localization accuracy. Objective and perceptual observations show that localization accuracy may degrade for off-centered listening positions, which can be mainly attributed to a mismatch in the elicited Interaural Level Differences (ILD) above the aliasing frequency.
Convention Paper 7361

11:00

P8-5 Wave Field Synthesis with Increased Aliasing Frequency—*Etienne Corteel, Renato Pellegrini, Clemens Kuhn-Rahloff*, sonic emotion ag, Obergltt (Zurich), Switzerland

Wave Field Synthesis (WFS) is a sound reproduction technique that enables the synthesis of target sound fields without any assumption on the listening position. Spatial aliasing is one of the remaining artifacts of WFS that limits the exact synthesis below a corner frequency referred to as spatial aliasing frequency. This paper presents a new technique that enables an increase of the spatial aliasing frequency of WFS assuming a preferred listening area. The presented technique is fully scalable and may be adapted to any listening zone shape or location. Applications in the domain of simulation environments and home entertainment are discussed.
Convention Paper 7362

11:30

P8-6 Reproduction of Moving Virtual Sound Sources with Special Attention on the Doppler Effect—*Jens Ahrens, Sascha Spors*, Technische Universität Berlin, Berlin, Germany

In this paper we outline a basic framework for the reproduction of the wave field of moving virtual sound sources. Conventional implementations usually reproduce moving virtual sources as a sequence of stationary positions. This process leads to various artifacts as reported in the literature. On the example of wave field synthesis, we show that the explicit consideration of the physical properties of the wave field of moving sources avoids these artifacts and allows for the accurate reproduction of the Doppler Effect. However, numerical simulations suggest that the artifacts inherent to the reproduction system can lead to a heavy degradation of the reproduction quality.
Convention Paper 7363

12:00

P8-7 A Graphical Tool Set for Analyzing Wave Field Synthesis Algorithms—*Thomas Korn*, Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany

Current Wave Field Synthesis (WFS) rendering realizations consist of large structures of audio signal processing components (filters, delays, amplitude weighting) that are controlled by complex algorithms based on the virtual source's properties. This paper proposes a set of tools that is used to analyze the underlying WFS coefficient calculation algorithms visually by mapping characteristic measures dependent on the source's and listener's position. These measures are derived from the reproduction system's idealized transfer function and parametric impulse response description. They reveal functional aspects of the algorithm's behavior. The measures aim at supporting an intuitive understanding of the perception of virtual sound events in a Wave Field Synthesis system, but also they facilitate the basic algorithm development process.
Convention Paper 7364

Workshop 8 — Sunday, May 18
09:00 – 11:00 — Room L

SURROUND SOUND IN RADIO

Chair: **Bosse Ternstrom**, Swedish Radio

Panelists: *Karl Petermichl*, ORF Radio
Heinz-Peter Reykers, WDR Köln
Gerhard Stoll, IRT – Munich, Germany

Multichannel audio in radio is steadily catching up and sometimes surpassing the TV realm with regards to number of productions, diversity of genres, and dedication of the people involved. Although there shouldn't be a direct competition, the inherent greater freedom of radio concerning the dramaturgical creativity in genres like radio drama or experimental and electro-acoustic music has led to a promising output of surround sound programs in those stations that pioneer this new format. The panel consists of what you already can call “veterans” of surround sound in radio, and they will discuss the state of the art, current issues, and solutions like file formats, LFE, signal distribution, etc., as well as showcase some of their respective work.

➥

Workshop 9 **Sunday, May 18**
09:00 – 11:30 **Room O**

UNEVEN BASS REPRODUCTION IN AUTOMOBILES

Chair: **Tom Nousaine**, Listening Technology

Panelists: *David Calristrom*, Chrysler Automotive
David Clark, Alpine, USA
Robert Klacza, Chrylser Automotive
Alfred Svobodnik, Harman International
Mark Ziemba, Panasonic-US

In the evaluation of 800 OEM Autosound systems Tom Nousaine, Listening Technology, has found that a majority of systems suffer from spectral uniformity problems at low frequencies. This simply means that bass sounds are uneven. For example, on a program with an acoustic bass solo, some notes almost disappear while others may seem unusually loud. It is commonly felt that this is a simple equalization issue. Thus, it is puzzling that this problem continues to exist when modern car electronics have significant active sound control capabilities. The panel will discuss the issue, causes, and possible solutions.

Tutorial 5 **Sunday, May 18**
09:00 – 11:00 **Room N101**

ELECTROACOUSTIC MEASUREMENTS, PART 1

Presenter: **Christopher J. Struck**, CJS Labs, San Francisco, CA, USA

This tutorial focuses on the fundamentals of electroacoustic measurements, including principles of acoustics, instrumentation, and data interpretation as well as practical information on how to perform appropriate tests. The basic components of a measuring system are examined followed by a review of basic acoustics, signals, sound sources, and sound fields. Psychoacoustics and hearing are examined with respect to objective measurements. Measurement transducers are explained in terms of selection and application. Finally, an overview of frequency analysis and the FFT is presented. This tutorial is intended to enable the participants to perform accurate audio and electroacoustic tests and provide them with the necessary tools to understand and correctly interpret the results.

Sunday, May 18 **09:00** **Room H**
Standards Committee Meeting on SC-02-01 Digital Audio Measurement Techniques

Sunday, May 18 **09:00** **Room K**
Technical Committee Meeting on Network Audio Systems

Session P9 **Sunday, May 18**
09:30 – 11:00 **Topaz Lounge**

POSTERS: SPATIAL AUDIO PERCEPTION AND PROCESSING

09:30

P9-1 **Audio-Visual Processing Tools for Auditory Scene Synthesis**—*Gavin Kearney, Rozenn Dahyot, Frank Boland*, Trinity College Dublin, Dublin, Ireland

We present an integrated set of audio-visual tracking and synthesis tools to aid matching of the audio to the video position in both horizontal and periphonic sound reinforcement systems. Compensation for screen size and loudspeaker layout for high definition formats is incorporated and the spatial localization of the source is rendered using advanced spatialization techniques. A subjective comparison of several original and enhanced film sequences using the Vector Base Amplitude Panning (VBAP) method is presented. The results show that the encoding of non-contradictory audio-visual spatial information, for presentation on different loudspeaker layouts significantly improves the naturalness of the listening/viewing experience.
Convention Paper 7365

09:30

P9-2 **Encoding Higher Order Ambisonics with AAC**—*Erik Hellerud,*[1] *Ian Burnett,*[2] *Audun Solvang,*[1] *U. Peter Svensson*[1]
[1]Norwegian University of Science and Technology, Trondheim, Norway
[2]University of Wollongong, Wollongong, New South Wales, Australia;

In this paper we explore a simple method for reducing the bit rate needed for transmitting and storing Higher Order Ambisonics (HOA). The HOA B-format signals are simply encoded using Advanced Audio Coding (AAC) as if they were individual mono signals. Wave field simulations show that by allocating more bits to the lower order signals than the higher the resulting error is very low in the sweet spot but increases as function of distance from the center. Encoding the higher order signals with a low bit rate does not lead to a reduced audio quality. The spatial information is improved when higher-order channels are included, even if these are encoded with a low bit rate.
Convention Paper 7366

09:30

P9-3 **Virtualized Listening Tests for Loudspeakers** —*Timo Hiekkanen,*[1] *Aki Mäkivirta,*[2] *Matti Karjalainen*[1]
[1]Helsinki University of Technology, Espoo, Finland
[2]Genelec Oy, Iisalmi, Finland

The precise location of a loudspeaker in a listening room is known to affect loudspeaker preference ratings. When multiple loudspeakers are compared the evaluation is limited by the poor human auditory memory. To overcome these problems, a method to evaluate and compare loudspeakers using headphones is proposed. The method utilizes personal head-related transfer functions in rendering the sound field recorded in a standard listening room with an artificial head. Equalization of circumaural headphones and the artificial head are investigated. Formal listening tests are conducted to examine differences between the proposed binaural method and real loudspeakers in a standard listening room. Listening tests show that the virtualized loudspeakers can be nearly imperceptible from reality in many but not in all cases.
Convention Paper 7367

09:30

P9-4 Binaural Rendering in MDCT Domain for Multi-Object Audio Coding—*Shinya Iizuka, Kei Kikuiri, Nobuhiko Naka*, NTT DoCoMo, Inc., Yokosuka, Kanagawa, Japan

We propose a binaural rendering method in Modified Discrete Cosine Transform (MDCT) domain. It has good compatibility with audio codecs because a number of audio codecs utilize an MDCT filter bank for time-frequency transform. The proposal maps MDCT coefficients to the real part of the Modulated Complex Lapped Transform (MCLT) coefficients and processes the amplitudes and phases according to the binaural information. The inverse MCLT is applied to the coefficients with a synthesis window function, which is derived from the perfect reconstruction condition for the phase shifted signal under the assumption of linear phase property. The proposed method is applicable to the Binaural Cue Coding Type I and offers equivalent subjective quality to the original binaural signal.
Convention Paper 7368

09:30

P9-5 Room-Dependent Preference of Virtual Surround Sound—*Frederick Scott, Agnieszka Roginska*, New York University, New York, NY, USA

A common method for simulating surround sound over headphones, so-called virtual surround sound, is the convolution of content information with binaural cues. Often, room information is included. This paper examines if using HRTFs with room impulse responses customized to the room the listener is in enhances the listening experience. Perceptual experiments were conducted to evaluate whether or not listeners prefer a room accurate rendering versus a room that is dissimilar to the one a listener is seated in. A preference test was conducted using music as the test material.
Convention Paper 7369

09:30

P9-6 Quantization of 2-D Higher Order Ambisonics Wave Fields—*Audun Solvang, U. Peter Svensson, Erik Hellerud*, Norwegian University of Science and Technology, Trondheim, Norway

The spatial distribution of the quantization noise for a 2-D Higher Order Ambisonics (HOA) signal is investigated analytically. Uniformly distributed loudspeakers radiating plane waves in a nonreverberant environment and frequency domain quantization are presumed. It is found that employing the same quantization interval for all orders leads to uniformly distributed quantization noise in space. Assigning a larger quantization interval (i.e., fewer bits) to higher orders leads to a radially increasing quantization noise. Matching the quantization error to the reproduction error at the near perfect reconstruction boundary suggests that as little as four bits per sample can be used for quantization. Furthermore, high-pass filtering the HOA components opens up for employing as little as three bits per sample. This quantization strategy seems very promising for reducing the rate of HOA.
Convention Paper 7370

09:30

P9-7 A Binaural Auditory Model for the Evaluation of Reproduced Stereophonic Sound—
Marko Takanen,[1] *Gaëtan Lorho*[2]
[1]Helsinki University of Technology, Espoo, Finland
[2]Nokia Corporation, Helsinki, Finland

Binaural cues describing the differences in phase and power between signals at the two ears enable our auditory system to localize sound sources and segregate spatially multiple auditory events. Recent publications on binaural auditory models have shown how the interaural coherence can be utilized to estimate these cues and therefore model the localization ability of our auditory system. This approach is exploited in this paper to estimate the binaural cues at different frequency bands and identify the spatial location of sound sources from recorded broadband signals. We illustrate the application of a binaural auditory model to evaluate sound reproduced by a stereophonic loudspeaker setup in terms of source localization and specific loudness.
Convention Paper 7371

09:30

P9-8 An Augmented Reality Audio Mixer and Equalizer—*Ville Riikonen, Miikka Tikander, Matti Karjalainen*, Helsinki University of Technology, Espoo, Finland

In Augmented Reality Audio (ARA) applications the real sound environment of the user is extended with virtual objects. The real environment is reproduced as a pseudo-acoustic world via a special ARA headset that consists of binaural microphones and headphones. However, the headset causes coloration to the pseudo-acoustic representation. In order to make the headset acoustically transparent, equalization is needed. Digital equalization easily causes unacceptable delays. This paper presents a novel ARA mixer with real-time analog equalization to correct the coloration caused by the leakage through the headset and changed resonances in the closed ear canal.
Convention Paper 7372

09:30

P9-4 Sub-Band Adaptive Crosstalk Cancellation: A Novel Approach for Immersive Audio—
Stefania Cecchi,[1] *Lorenzo Palestini,*[1] *Paolo Peretti,*[1] *Francesco Piazza*,[1] *Ferruccio Bettarelli*[2]
[1]Universita Politecnica delle Marche, Ancona, Italy
[2]Leaff Engineering, Porto Potenza Picena (MC), Italy

In the field of immersive audio, crosstalk canceller is required when a virtual sound is rendered over two loudspeakers. In the last decade several adaptive algorithms have been proposed: nowadays the least square (LMS) algorithm seems to be the best compromise ➥

between simplicity and robustness although its convergence is weakened for colored inputs. In this paper a new approach for crosstalk cancellation based on a sub-band adaptive algorithm will be derived. The effectiveness of this algorithm, considering colored input, will be presented in terms of matrix inversion quality and fast convergence rate comparing it with the conventional LMS algorithm.
Convention Paper 7373

Workshop 10 — Sunday, May 18
09:30 – 11:00 — Room B

PROJECT MANAGEMENT IN PRODUCT DEVELOPMENT

Chair: **Jeroen Langevoort**, NXP Semiconductors, Nijmegen, The Netherlands

As car radio systems become more and more complex proper project management is key to deliver the products on time and in full to the market.

The development project of a car radio system with ROM-codes and IC's, like the LeafDice and Dirana 2, will be taken as reference. In the workshop attention will be paid to the requirements definition, risk management, and design processes within NXP. Special requirements on quality that are common in the automotive industry will be briefly touched on, as these are very different from those of the consumer industry. Among the deliverables of such projects is an extensive communication package for the OEM to enable them to use the system in a proper way. This includes application board and application notes.

Live Sound Seminar 4 — Sunday, May 18
09:30 – 13:00 — Forum

CABLES AND CONNECTORS

Presenters: **Wolfgang Feß**, Klotz-a.i.s
Norbert Nachbauer, Neutrik, Schaan, Liechtenstein

This session describes in a wide range the structures, materials, and application ranges of cables used in PA systems (copper - balanced, unbalanced; fiber optic - single mode, multi mode, a.o.). Parameters that have determining influence on the characteristic impedance of copper cables will be pointed out as well as an overview being shown about typical audio, video, and data applications and the required cables (analog, digital, data, antenna, etc.). Audible effects will be demonstrated.

Student Event/Career Development
CAREER/JOB FAIR
Sunday, May 18, 09:30 – 11:30
Topaz Lounge

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

Sunday, May 18 — 10:00 — Room K
Technical Committee Meeting on Audio for Telecommunications

Special Event
GEOFF EMERICK/SGT. PEPPER
Sunday, May 18, 11:00 – 12:30
Room L

Marking the 40th Anniversary of the release of *Sgt. Pepper's Lonely Hearts Club Band*, Geoff Emerick, the Beatles engineer on the original recording was commissioned by the BBC to re-record the entire album on the original vintage equipment using contemporary musicians for a unique TV program.

Celebrating its own anniversary, the APRS is proud to present for a select AES audience, this unique project featuring recorded performances by young UK and US artists including the Kaiser Chiefs, The Fray, Travis, Razorlight, the Stereophonics, the Magic Numbers, and a few more—and one older Canadian, Bryan Adams.

These vibrant, fresh talents recorded the original arrangements and orchestrations of the Sgt. Pepper album using the original microphones, desks, and hard-learned techniques directed and mixed in mono by the Beatles own engineering maestro, Geoff Emerick.

Hear how it was done, how it should be done, and how many of the new artists want to do it in the future. Geoff will be available to answer a few questions about the recording of each track and, of course, more general questions regarding the recording processes and the innovative contribution he and other Abbey Road wizards made to the best ever album.

APRS, The Association of Professional Recording Services, promotes the highest standards of professionalism and quality within the audio industry. Its members are recording studios, postproduction houses, mastering, replication, pressing, and duplicating facilities, and providers of education and training, as well as record producers, audio engineers, manufacturers, suppliers, and consultants. Its primary aim is to develop and maintain excellence at all levels within the UK's audio industry.

Exhibitor Seminar 1 — Sunday, May 18
11:00 – 12:00 — Room P

SENNHEISER ELECTRONIC CORPORATION

Presenters: **Sebastian Schmitz**
Gregor Zielinsky

MKH 800 Twin with Remote Controllable Directivity

MKH 800 Twin The MKH 800 Twin is a Doublecapsule Microphone with remote-controllable pattern for stereo and surround recording. Each of the capsules can be recorded separately as they have separate outputs. This enables you to seamlessly adjust the pickup pattern at post-production stage. The presentation shows several examples of how to use this technique in stereo and surround examples.

Sunday, May 18 — 11:00 — Room H
Standards Committee Meeting on SC-04-04 Microphone Measurement and Characterization

Sunday, May 18 — 11:00 — Room K
Technical Committee Meeting on Archiving, Restoration, and Digital Libraries

Workshop 11 **Sunday, May 18**
11:30 – 13:30 **Room B**

CREATING A MUSIC DOWNLOAD WEBSITE

Chair: **Vicki Melchior**, Consultant, San Anselmo, CA, USA

Panelists: *Brian Armour*, CC Technology, Glasgow, Scotland
Philip Hobbs, Linn Audio
Bosse Ternstrom, Swedish Radio

For labels, bands, and artists facing the decline of physical media, the potential now exists to release new and prior catalogs directly through digital downloads. Downloads can readily be targeted to specialized markets as well as to the broader online community, and at the same time may offer selective digital formats—for example, high resolution or lossless standard resolution, multichannel, surround algorithms such as ambisonics, surround with designated height channels, etc. In this workshop a number of individuals with recent web development experience will describe the challenges and methodology of creating a music download website. The discussion should provide insight for those getting started in this area and will include: what is involved with setting up for web delivery from a producer's viewpoint, overview of server hardware and network requirements, programming and software development, website issues and licensing, business models and expected volume, choosing download formats, reaching target audiences, and economies of scale including when it's better to join a much larger website.

Tutorial 6 **Sunday, May 18**
11:30 – 13:00 **Room N101**

YESTERDAY'S FX TODAY

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

With affordable digital audio tools being continuously invented, refined, improved, extended, and upgraded, we are lucky to be a part of the audio industry at this moment. We have no excuse not to create original, beautiful art. What we do with today's ability to do anything can be informed by the creative and technical achievements expressed in touchstone recordings decades ago. This tutorial takes a close look at some iconic moments of signal processing innovation in recorded music history, undoing, isolating, and analyzing the effects for our edification.

Student Event/Career Development
EDUCATION FAIR
Sunday, May 18, 11:30 – 13:30
Topaz Lounge

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

Workshop 12 **Sunday, May 18**
12:00 – 13:00 **Room O**

AUTOMATED DIRECTIONAL LIVE SOUND REINFORCEMENT

Chair: **Dave Haydon**, Out Board Electronics, UK

Panelists: *Tom Strebel*, Audiopool, Switzerland
Cees Wagenaar, Peutz, The Netherlads
Robin Whittaker, Out Board Electronics, UK
Karsten Wolstadt, Danish Royal Theatre, Denmark

Directional amplification, also referred to as Source Oriented Reinforcement (SOR), describes a practical technique to deliver amplified sound to a large listening area with even coverage while providing directional information to reinforce visual cues and create a realistic and noncontradictory auditory panorama. Audio demonstrations of the fundamental psychoacoustic techniques employed in an SOR design will be presented and limits discussed. The panel of presenters will outline the history of SOR from Steinke & Fels pioneering work with their Delta Stereophony System in the mid 1970s (later licensed to AKG) to Out Boards current-day TiMax Audio Imaging Delay Matrix including the very latest ground breaking technology employed to enable control of precedence by radar tracking the actors on the stage. Descriptions of venues and productions that have employed SOR include the new Copenhagen Royal Theatre Drama House, Einsiedeln Welttheater 2007, the opera Aïda in Rotterdam Ahoy and Basel Tattoo among others. TiMax, a DSP-based audio control system, was launched in 1997 and has since received a number of industry awards for innovation. TiMax has become synonymous with the techniques for localization of sound in theaters and auditoria, and was one of the first commercially available products offering dynamic control of time delay and level for recreating realism in sound reinforcement.

Human ability to determine the direction from which sound arrives is due to binaural hearing, the brain being able to detect differences between sounds received by the two ears from the same source and thus to determine angular direction. Alan Blumlein 1931

If two successive sounds are heard as fused, the location of the total sound is determined largely by the location of the first arriving sound. This, among other things, stops you getting confused when a sound comes at you from two speakers or audio sources at once. Helmut Haas 1946

Tutorial 7 **Sunday, May 18**
12:30 – 14:30 **Room L**

SURROUND SOUND CONTRIBUTION METHODS

Presenters: **John McClintock**, APT UK (Chair)
Karl Petermichl, ORF Radio
Heinz-Peter Reykers, WDR
Geir Skaden, Neural Audio

Surround contribution over networks are coming into use in, e.g., Austria, Germany, by the EBU world wide and in the U.S. The intention is to describe these practical implementations as case studies. ➥

Session P10 **Sunday, May 18**
13:00 – 17:30 **Room C-D**

SPATIAL AUDIO PERCEPTION AND PROCESSING, PART 2

Chair: **Renato Pelligrini**, sonic emotion ag, Obergltt (Zurich), Switzerland

13:00

P10-1 Analysis and Adjustment of Planar Microphone Arrays for Application in Directional Audio Coding—*Markus Kallinger,*[1] *Fabian Kuech,*[1] *Richard Schultz-Amling,*[1] *Giovanni Del Galdo,*[1] *Jukka Ahonen,*[2] *Ville Pulkki*[2]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]Helsinki University of Technology, Espoo, Finland

Directional Audio Coding (DirAC) is a well-established and efficient way to capture and reproduce a spatial sound event. In a recording room, DirAC requires four spatially coincident microphones to estimate the desired parameters, i.e., direction-of-arrival and diffuseness of sound: one omnidirectional and three figure-of-eight microphones pointing along the axes of a three-dimensional Cartesian coordinate system. In most consumer applications only two dimensional scenes need to be reproduced, implying that only two figure-of-eight microphones are required. Furthermore, instead of directional microphones, arrays of omnidirectional microphones are considered for economic reasons. Therefore, we investigate various two-dimensional microphone configurations with respect to their usability for DirAC. We derive theoretical limits for the correct estimation of both direction-of-arrival and diffuseness for the most suitable planar arrays. Furthermore, we suggest a way to equalize the systematic bias for the direction-of-arrival estimation, introduced by the discrete planar arrays.
Convention Paper 7374

13:30

P10-2 Planar Microphone Array Processing for the Analysis and Reproduction of Spatial Audio Using Directional Audio Coding—*Richard Schultz-Amling,*[1] *Fabian Kuech,*[1] *Markus Kallinger,*[1] *Giovanni Del Galdo,*[1] *Jukka Ahonen,*[2] *Ville Pulkki*[2]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]Helsinki University of Technology, Espoo, Finland

Recording and reproduction of spatial audio becomes more and more important, as multichannel audio applications gain increasing attention. Directional Audio Coding (DirAC) represents a well proven approach for the analysis and reproduction of spatial sound. In the analysis part, the direction-of-arrival and the diffuseness of the sound field is estimated in subbands using B-format signals, which can be created with 3-D omnidirectional microphone arrays. However, 3-D microphone configurations are not practical in consumer applications, e.g., due to physical design constraints. In this paper we propose a new approach that allows for an approximation of the required B-format signals but is based on a planar microphone configuration only. Comparisons with the standard DirAC approach confirm that the proposed method is able to correctly estimate the desired parameters within a wide range of frequency and the spatial resolution matches the human perception.
Convention Paper 7375

14:00

P10-3 User-Dependent Optimization of Wave Field Synthesis Reproduction for Directive Sound Fields—*Frank Melchior,*[1,2] *Christoph Sladeczek,*[1,3] *Diemer de Vries,*[3] *Bernd Fröhlich*[2]
[1]Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany
[2]Delft University of Technology, Delft, The Netherlands;
[3]Bauhaus Universität Weimar, Germany

The use of wave field synthesis (WFS) enables the correct localization of sources over a large listening area. This works well for simulated point sources outside the listening area. The perception of focused sources is only correct for a subspace of the listening area. The subspace depends on the selected set of loudspeakers used for the reproduction of the focused source. If the position of the listener is known the selection of loudspeakers as well as the signal processing can be optimized. By the use of continuous tracking of the listener this adaptation can be done in real time. The same data can be used to simulate a specific directivity of a source and optimize a corresponding room simulation for the tracked listener. We present a wave field synthesis system for the simulation of directive focused sources including room simulation, which is continuously optimized for the position of a tracked listener. Our observations confirm that this approach significantly improves the localization and sound quality of focused sources located inside the listening area.
Convention Paper 7376

14:30

P10-4 Spatial Audio Object Coding (SAOC)—The Upcoming MPEG Standard on Parametric Object Based Audio Coding—*Jonas Engdegård,*[1] *Barbara Resch,*[1] *Cornelia Falch,*[2] *Oliver Hellmuth,*[2] *Johannes Hilpert,*[2] *Andreas Hoelzer,*[2] *Leonid Terentiev,*[2] *Jeroen Breebaart,*[3] *Jeroen Koppens,*[4] *Erik Schuijers,*[4] *Werner Oomen*[4]
[1]Dolby Sweden, Stockholm, Sweden
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[3]Philips Research Laboratories, Eindhoven, The Netherlands;
[4]Philips Applied Technologies, Eindhoven, The Netherlands

Following the recent trend of employing parametric enhancement tools for increasing coding or spatial rendering efficiency, Spatial Audio Object Coding (SAOC) is one of the recent additions to the standardization activities in the MPEG audio group. SAOC is a technique for

efficient coding and flexible, user-controllable rendering of multiple audio objects based on transmission of a mono or stereo downmix of the object signals. The SAOC system extends the MPEG Surround standard by re-using its spatial rendering capabilities. This paper will describe the chosen reference model architecture, the association between the different operational modes and applications, and the current status of the standardization process.
Convention Paper 7377

15:00

P10-5 Focusing of Virtual Sound Sources in Higher Order Ambisonics—*Jens Ahrens, Sascha Spors*, Technische Universität Berlin, Berlin, Germany

Higher order Ambisonics (HOA) is an approach to the physical (re-)synthesis of a given wave field. It is based on the orthogonal expansion of the involved wave fields formulated for interior problems. This implies that HOA is per se only capable of recreating the wave field generated by events outside the listening area. When a virtual source is intended to be reproduced inside the listening area, strong artifacts arise in certain listening positions. These artifacts can be significantly reduced when a wave field with a focus point is reproduced instead of a virtual source. However, the reproduced wave field only coincides with that of the virtual source in one half-space defined by the location and nominal orientation of the focus point. The wave field in the other half-space converges toward the focus point.
Convention Paper 7378

15:30

P10-6 Listener Envelopment—What Has Been Done and What Future Research Is Needed?—*Dan Nyberg, Jan Berg*, Luleå University of Technology, Piteå, Sweden

In concert hall acoustics, the perceived spatial impression and/or spaciousness are characterized by the two attributes apparent source width (ASW) and listener envelopment (LEV). For LEV there are no clear consensus across the results of previous work. This paper aims to discuss the research performed on LEV and how these research results are confirming or contradicting each other. There is a consensus on the arrival angle of the later sound energy and its influence on LEV, whereas there is no clear agreement on the delay time and frequency content of the late reflections.
Convention Paper 7379

16:00

P10-7 Obtaining a Highly Directive Center Channel from Coincident Stereo Microphone Signals—*Christof Faller*, Illusonic LLC, Lausanne, Switzerland

Time-frequency based postprocessing applied to a coincident stereo recording is proposed to generate an audio signal with a highly directive directional response pointing straight forward. Assuming an ideal coincident stereo microphone, the directional response of this center channel is effectively time and frequency invariant. Further, the look direction can be steered to left and right front directions. The technique is based on the insight that the signal that predicts left from right, modified by limiting the magnitude of the frequency domain prediction gains, has a center forward directional response. The center channel is generated using both, a left-right and a right-left magnitude-limited-predictor signal. Applications of the proposed scheme are use of stereo microphones as “digital steerable shot gun microphones” and center channel generation for music recording.
Convention Paper 7380

16:30

P10-8 Spatial Sound in the Use of Multimodal Interfaces for the Acquisition of Motor Skills—*Pablo F. Hoffmann*, Aalborg University, Aalborg, Denmark

This paper discusses the potential effectiveness of spatial sound in the use of multimodal interfaces and virtual environment technologies for the acquisition of motor skills. Because skills are generally of multimodal nature, spatial sound is discussed in terms of the role that may play in facilitating skill acquisition by complementing, or substituting, other sensory modalities. An overview of related research areas on audiovisual and audiotactile interaction is given in connection to the potential benefits of spatial sound as a means to improve the perceptual quality of the interfaces as well as to convey information that may prove critical for the transfer of motor skills.
Convention Paper 7381

17:00

P10-9 Evaluating the Sensation of Envelopment Arising from 5-Channel Surround Sound Recordings—*Sunish George,*[1] *Slawomir Zielinski,*[1] *Francis Rumsey*,[1] *Søren Bech*[2]
[1]University of Surrey, Guildford, Surrey, UK
[2]Bang & Olufsen a/s, Struer, Denmark

This paper discusses a series of listening tests conducted in the UK and Denmark to evaluate the perceived envelopment of surround audio recordings. The listening tests were designed to overcome some drawbacks (such as range equalization bias) present in the scores of a listening test based on ITU-R. BS. 1534-1 Recommendation (MUSHRA). In this method the listeners were asked to evaluate the envelopment of 5-channel surround sound recordings using a 100-point continuous scale. In order to calibrate the scale, two anchor recordings were used to define points 15 and 85 on the scale. The anchor recordings were selected by means of a formal listening test and interviews with the listeners. According to the obtained results, the proposed method provides repeatable results.
Convention Paper 7382

➥

Session P11 **Sunday, May 18**
13:00 – 17:30 **Room E-F**

ANALYSIS AND SYNTHESIS OF SOUND

Chair: **Olivier Warusfel**, IRCAM, Paris, France

13:00

P11-1 An Improved Pattern-Matching Method for Piano Multipitch Detection—*Luis Ortiz-Berenguer, Francisco J. Casajus-Quiros, Elena Blanco-Martin*, Technical University of Madrid, Madrid, Spain

A previous method presented by the authors carried out multipitch piano sound identification by using a pattern-matching process. In that method, the identification required, besides the matching-metric calculation, both a spectral predetection process and a validation step. Predetection allowed selection of a subset out of the eighty-eight patterns, whereas the validation verified whether the detected note were actually in the analyzed spectrum. Both highly increased the true-positive detections ratio, but they imposed restrictions to the identification of complex real sounds (e.g., two-hands playing). This paper presents an improvement in the method that allows getting rid of both, predetection and validation, by using a modified matching-metric algorithm. This work has been supported by the Spanish National Project TEC2006-13067-C03-01/TCM.
Convention Paper 7383

13:30

P11-2 Polyphonic Piano Transcription Based on Spectral Separation—*Julio Jose Carabias-Orti, Pedro Vera-Candeas, Nicolas Ruiz-Reyes, Raul Mata-Campos, Francisco Jesus Cañadas-Quesada*, University of Jaén, Linares, Jaén, Spain

We propose a discriminative model for polyphonic piano transcription. Spectral features are obtained individually for each note. To solve the overlapping partial problem, we apply spectral separation by estimating the spectral envelope for each note. For classifying purposes, support vector machines (SVM) are trained on the spectral energy inferred from these spectral features. We apply a scheme of one-versus-all (OVA) SVM classifiers to discriminate frame-level note instances. To decrease the high frequency notes residual energy due to the downward notes shared partials, a method to cancel the interferences from the downward notes to the upward notes has been developed. The classifier output is filtered with a hidden Markov model. Our approach has been tested with synthesized and real piano recordings obtaining very promising results.
Convention Paper 7384

14:00

P11-3 Toward a Real-Time Implementation of a Physical Modeling Based Percussion Synthesizer—*Katarzyna Chuchacz, Roger Woods, Sile O'Modhrain*, Queen's University Belfast, Belfast, Northern Ireland, UK

This paper presents work carried out with the objective of designing a novel percussion synthesizer based on a physical model of a plate-based percussion instrument. The algorithm has been implemented in real-time for the first time, on a Field Programmable Gate Array (FPGA) chip allowing a number of parameters such as excitation value, stroke location, and plate stiffness, to be changed in real-time. This presents the player with a number of new modes of playability but requires the definition and design of a flexible interface that gives the extensive access to the sound world of the synthesis model. Details of the hardware implementation architecture are put forward as well as fixed point/floating point computation aspects that impact the instrument's playability.
Convention Paper 7385

14:30

P11-4 Dual Noise Suppression in Hearing Aids—*Anton Schlesinger, Marinus M. Boone*, Delft University of Technology, Delft, The Netherlands

A combined processing scheme for the enhancement of speech intelligibility in hearing aids is presented. The approach utilizes an optimized beam-forming method in connection with a biologically inspired processing model of modulation perception and binaural interaction.
Convention Paper 7386

15:00

P11-5 Automatic Sound Recognition for Security Purposes—*Pawel Zwan*, Gdansk University of Technology, Gdansk, Poland

In the paper an automatic sound recognition system is presented. It forms a part of a larger security system developed in order to monitor outdoor conditions for non-typical audio-visual events. The analyzed audio signal is being recorded from a microphone mounted outdoor, thus non-stationary noise of a significant energy may be present in it. In the paper an especially designed algorithm for an outdoor noise reduction is presented, non-typical events in audio stream are automatically detected and parameterized. Parameter values of various audio events are analyzed and sounds are automatically recognized. The automatic recognition accuracy obtained for various feature vectors and some chosen recognition systems is compared. The conclusions are derived and a future plan of experiments is proposed.
Convention Paper 7387

15:30

P11-6 Multipitch Estimation of Harmonically-Related Event-Notes by Improving Harmonic Matching Pursuit Decomposition—*Francisco Jesus Canadas-Quesada, Pedro Vera-Candeas, Nicolas Ruiz-Reyes, Raul Mata-Campos, Julio Jose Carabias-Orti*, University of Jaén, Linares, Jaén, Spain

In this paper we propose a note detection approach based on harmonic matching pursuit (HMP) and specifically designed to detect simul-

taneous notes. However, HMP is not able to decompose harmonic sounds in different harmonic atoms when their fundamental frequencies are harmonically-related. To solve this problem, we propose an algorithm, called atomic spectral smoothness (SS), which works over the harmonic atoms obtained by HMP. This algorithm is based on the spectral smoothness principle that supposes that the spectral envelope of a harmonic sound usually forms smooth contours. Our proposal shows promising results for polyphonic musical signals with two harmonically-related note-events.
Convention Paper 7388

16:00

P11-7 Amplitude Modification Algorithms within the Framework of Physical Modeling and of Haptic Gestural Interaction—*Alexandros Kontogeorgakopoulos, Claude Cadoz*, Institute National Polytechnique de Grenoble, Grenoble, France

Every underlying technique that has been used for the realization of audio effects since the beginning of electronic and computer music, introduced different types of sound modifications and proposed new ways of control. The advent of digital signal processing has stimulated the audio processing researchers to a great extent; thus a variety of algorithms were designed to provide novel sound modifications. On the other hand, physical modeling and digital simulation formalisms have been principally used for the merely imitation and emulation of older sound processing systems. The aim of this paper is to propose three physical models conceived to offer sound modifications that mainly alter the amplitude of audio signals. The originality of this case is not the resulted audio modifications but their transposition in the framework of physical modeling and digital simulation, which outlines an alternative control procedure.
Convention Paper 7389

16:30

P11-8 Circular Pitch Space Based Harmonic Change Detection—*Markus Mehnert,*[1] *Gabriel Gatzsche,*[2] *Daniel Arndt,*[1] *Karlheinz Brandenburg*[2]
[1]Technische Universität Ilmenau, Ilmenau, Germany
[2]Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany

This paper introduces a novel method for detecting harmonic boundaries in musical audio signals. These boundaries are important for chord analysis and between two boundaries is just one particular chord. This event-driven analysis of musical audio signals is a better basis for a following chord analysis than the traditionally frame-based-only concept. The method itself works with circular pitch spaces (CPS). The idea behind CPS is the calculation of parameters that summarize high level aspects of the audio signal such as semantic and music theoretical relationships. Using CPSs entails good results in detecting harmonic changes.
Convention Paper 7390

17:00

P11-9 Circular Pitch Space Based Musical Tonality Analysis—*Gabriel Gatzsche,*[1] *Markus Mehnert,*[2] *Daniel Arndt,*[2] *Karlheinz Brandenburg*[1]
[1]Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany
[2]Technische Universität Ilmenau, Ilmenau, Germany

The focus of this paper is to give an overview of existing circular pitch spaces, its special properties and application for semantic audio analysis. Beside this the symmetry model is proposed as a framework to describe the inter-model relationships between different circular pitch spaces. Similar to color spaces in vision musical pitch spaces organize pitches in a way that semantic/cognitive/theoretical/physical relationships between tones become geometrically apparent. Within the last years pitch spaces were mainly the subject of music theory. But they become more and more interesting for semantic analysis of musical audio signals. Pitch spaces can be applied to key and chord recognition, similarity calculation of musical pieces, genre estimation, tension analysis, or harmonic change detection.
Convention Paper 7391

Exhibitor Seminar 2 — Sunday, May 18
13:00 – 14:00 — Room P

FERROTEC

Presenters: **M. Klasco**
B. Moskowitz

The Past, Present and Future of Ferrofluids in the Loudspeaker Industry

Ferrofluids were introduced to the loudspeaker marketplace by Ferrotec (formerly Ferrofluidics) in the early 1970s. Millions of loudspeakers have been produced with ferrofluid in the voice coil assembly to improve performance and speaker life. New advancements in research have expanded the range of ferrofluid products available to the industry.

Sunday, May 18 — 13:00 — Room H
Standards Committee Meeting on SC-03-12 Forensic Audio

Sunday, May 18 — 13:00 — Room K
Technical Committee Meeting on Audio Recording and Mastering Systems

Workshop 13 — Sunday, May 18
13:30 – 16:30 — Room N101

AUDIO AND AUDIO-TACTILE WARNINGS AND ALERTS: FORMING TODAY'S SOUNDSCAPE

Chairs: **Durand Begault**, NASA Ames Research Center, Moffett Field, CA, USA
Ellen Haas, Army Research Lab

Panelists: *Caryl Baldwin*, George Mason University
Judy Edworthy, University of Plymouth
Anne Guillaume, ISMA

Elif Ozcan, Technical University of Delft
Rene VanEgmond, Technical University of Delft

Audio warning and alert signals are omnipresent in the soundscape of modern society, ranging from warnings (sirens), to sounds from human interfaces (pushbutton sounds on devices), to cell phone ring tones. Attention has only been given fairly recently in the human factors research community toward the best means of making caution and warning signals discriminable, audible, and tolerable; some high-stress human interfaces are also concerned with increasing the richness of the semantic content of these alarms. Ironically, while it is well documented that people have ignored critical audio alarms in flight and railroad operations, personalized cell telephone ring tones have an apparently high "hit" rate, while simultaneously alerting (and annoying) others. Discussion topics for this workshop include assessment of current standards and best practices for design of auditory alerts; new approaches to forming the sound design of alarms, including multimodal approaches such as including tactile displays; new approaches to forming the sound design of alarms, and defining what best (or worst!) practices might be. We may obtain as a result an idea of how the soundscape of future decades may sound like (or what we would like it to sound like).

Tutorial 8 **Sunday, May 18**
13:30 – 15:00 **Room O**

CURRENT STANDARDS OF FILE-INTERCHANGE FORMATS

Presenter: **Mark Yonge**, AES Standards Manager

As the broadcasting world is switching to file-based production, it is vital to get to grips with current formats one is likely to encounter as well as to know the different flavors, strengths, and weaknesses. The presenter has an intimate insight into file-formats, not at least through his work as AES standards manager.

Session P12 **Sunday, May 18**
14:00 – 15:30 **Topaz Lounge**

POSTERS: AUDIO ARCHIVING, STORAGE, RESTORATION, AND CONTENT MANAGEMENT AND AUDIO NETWORKING

14:00

P12-1 Drift, Wow, and Flutter Measurement and Reduction in Shrunken Movie Soundtracks—*Przemek Maziewski, Adam Kupryjanow, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

The paper presents the method and algorithms used to determine and reduce drift, wow, and flutter in shrunken movie tapes. The idea behind the algorithms is to use image processing for calculating the local tape shrinkage, which is one of the reasons for drift, wow, and flutter. The shrinkage can be calculated via analyzing the image height of: a movie frame, sprocket hole, pitch, or another standardized movie tape element; and then it can be expressed as the drift, wow, and flutter characteristic. After the characteristic determination both the soundtrack and movie frames can be corrected. The paper presents the description of the image based drift, wow, and flutter determination method and the experiments confirming the theoretical findings.
Convention Paper 7392

14:00

P12-2 The Norwegian Institute of Recorded Sound: From Collection to Archive to Public Private Partnership—*Mark Drews*,[1] *Jacqueline von Arb*[2]
[1]University of Stavanger, Stavanger, Norway
[2]Norwegian Institute of Recorded Sound, Stavanger, Norway

In 2006, the Norwegian Institute of Recorded Sound (NIRS) entered into a partnership with Memnon Audio Archiving Services to form Mem-Nor, a commercial audio archiving service based in Stavanger, Norway. This paper traces the evolution of the Norwegian Institute of Recorded Sound from a private collection of music recordings to a municipally funded audio archive to a public private partnership and discusses the past, the current, and the future challenges involved. Details of ongoing activities is included.
Convention Paper 7393

14:00

P12-3 Cable-free Audio Delivery for Home Theater Entertainment Systems—*Andreas Floros*,[1] *Nicolas-Alexander Tatlas,*[2] *John Mourjopoulos,*[2] *Dimitris Grimanis*[2]
[1]Ionian University, Corfu, Greece
[2]University of Patras, Patras, Greece

Real time, multichannel audio content delivery over the air is expected to significantly simplify the interconnection complexity required for setting up typical home theater applications. However, despite the technological advantages of wireless networking standards related to high transmission rates and Quality-of-Service support, a number of issues has to be additionally addressed, such as multiple loudspeaker synchronization and packet delay/losses containing compressed quality and multiplexed audio data. In this paper further developments in the area of wireless audio delivery are presented by considering in detail multichannel reproduction for wireless home theater applications. Using both subjective and objective performance evaluation criteria, it is shown that cable-free multichannel audio playback is feasible under specific networking and audio coding conditions.
Convention Paper 7394

14:00

P12-4 Adaptive Playout for VoIP Based on the Enhanced Low Delay AAC Audio Codec—*Jochen Issing, Nikolaus Färber, Manfred Lutzky*, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

The MPEG-4 Enhanced Low Delay AAC (AAC-ELD) codec extends the application area of the Advanced Audio Coding (AAC) family toward high quality conversational services. Through the support of the full audio bandwidth at low delay and low bit rate, it offers excellent support

for enhanced VoIP applications. In this paper we provide a brief overview of the AAC-ELD codec and describe how its codec structure can be exploited for IP transport. The overlapping frames and excellent error concealment make it possible to use frame insertion/deletion in order to adjust the playout time to varying network delay. A playout algorithm is proposed that estimates the jitter on the network and adapts the size of the de-jitter buffer in order to minimizes buffering delay and late loss. Considering typical network conditions and the same average delay, it is shown that the playout algorithm can reduce the loss rate by more than one magnitude compared to fixed playout.
Convention Paper 7395

Master Class 1 **Sunday, May 18**
14:00 – 16:00 **Room B**

LOUDSPEAKER PARAMETERS

Presenters: **Richard Small**, Harman-Motive Inc., USA
Neville Thiele, University of Sydney, Sydney, NSW, Australia

The loudspeaker parameters provide a procedure that is now established universally for specifying loudspeaker drivers and designing their enclosures and associated equalizers. Their derivation and measurement involves simplifications and approximations to what is, in fact, a complex acoustical/mechanical/electrical system. They apply most accurately to the response of the system to small signals but nevertheless allow measurements to be made with surprisingly simple equipment and the performance of loudspeaker systems to be predicted with high precision. Since the initial publications between 1961 and 1973, their effectiveness has been increased through additions and refinements in understanding the mechanisms involved and in measuring and calculating them, made by a number of researchers and reported in a number of places. This master class aims to present them in a unified whole, correct some misconceptions, and answer some frequently-asked questions that, over the years, have arisen and mystified some students and designers.

Live Sound Seminar 5 **Sunday, May 18**
14:00 – 18:00 **Forum**

LIVE MIXING WORKSHOP

Presenter: **Gregor Zielinsky**, Sennheiser Electronic Corporation

This session presents a live sound check with a band including In Ear Check Live on Stage. Similar workshops have been presented in many places worldwide, e.g., at AES conventions in Barcelona and Vienna, a.o. The focus this time will be on playback sound quality by loudspeakers for drums, guitar, and piano. The drum sound, especially, is probably the most ambitious —live or in the studio. A team of experienced presenters will show many professional tricks.

- How to record percussion
- How to get a piano loud
- How to get highest sound quality from a guitar

Historical Event
M3 RECORDING TECHNIQUE
Sunday, May 18, 14:00 – 15:00
Room N

Presenter: **Hans Lauterslager**

Hans Lauterslager, retired balance engineer with Polygram-Philips, gives a description of 3-microphone recording technique as applied by Mercury and Philips in the sixties.

Sunday, May 18 **14:00** **Room K**
Technical Committee Meeting on Electro Magnetic Compatibility

Student Event/Career Development
RECORDING COMPETITION—STEREO
Sunday, May 18, 14:30 – 18:00
Room L

The Student Recording Competition is a highlight at each convention. A distinguished panel of judges participates in critiquing finalists in each of the categories in an interactive presentation during the convention. Student members can submit stereo and surround recordings in the categories classical, jazz, folk/world music, and pop/rock. Meritorious awards will be presented at the closing Student Delegate Assembly Meeting on Tuesday.

Judges:
Classical (14:30–15:20): Ken Blair, Leo de Klerk, Theo Wubbolts
Jazz/Blues (15:20–16:10): Jim Anderson, Ulrike Schwartz, Johannes Wohlleben
Folk/World Music (16:20–17:10): Eva Bauer-Oppelland, Darcy Proper, Mark Derksen
Pop/Rock (17:10–18:00): Carlos Abrect, Alex Case, Werner Penasaert

Recording competition sponsors: Schoeps, PMC, Sennheiser, Neumann

Historical Event
100 YEAR OF THE TRIODE
Sunday, May 18, 15:00 – 16:00
Room N

Presenter: **Guido Tent**

“100 years of the Triode” is a short journey through the history of electronics, which started in 1907 with the invention of the triode. This lecture gives an overview of how electronics developed, shows a documentary of the (re)making of the triode, and ends with an audio demonstration.

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

ROOM ACOUSTICS MODELING

Presenters: **Tapio Lokki**, Helsinki University of Technology, Espoo, Finland
Lauri Savioja, Helsinki University of Technology, Espoo, Finland

This tutorial will cover commonly applied wave-based and ray-based methods in room acoustics modeling. The main emphasis is to explain basic principles of physically-based modeling methods. In addition, pros and cons of each method are discussed. Examples are given with several modeling methods. The modeling results, namely room acoustical parameters and auralizations, are also discussed in the light of each method.

Sunday, May 18 15:00 Room K
Technical Committee Meeting on Microphones and Applications

Sunday, May 18 15:30 Room H
Standards Committee Meeting on SC-03-02 Transfer Technologies

Session P13 Sunday, May 18
16:00 – 17:30 Topaz Lounge

POSTERS: SIGNAL PROCESSING, SOUND QUALITY DESIGN

16:00

P13-1 Time-Alignment of Multiway Loudspeakers with Group Delay Equalization: Part I—*Sunil Bharitkar, Chris Kyriakakis, Tom Holman*, University of Southern California and Audyssey Labs, Los Angeles, CA, USA

In this paper, a first of two-parts, a technique for time-aligning the driver responses (viz., woofer, mid-range, and tweeter responses) in a multiway loudspeaker system is presented. Generally, woofers exhibit a much larger time-of-arrival delay at a listening position, compared to the mid-range and high-frequency drivers due to the presence of crossover networks. Moreover, the time-of-arrival delay for all drivers is frequency dependent exhibiting a large variation over the audible frequency domain. Due to these differences a two-part study was undertaken to understand the effects of these variations, quantitatively and qualitatively, in the direct as well as reverberant field in typical listening rooms with diverse content. Various multiway loudspeakers, measured in one of the anechoic chambers at Audyssey, were selected to provide a diverse corpus of responses. In this first part we present the motivation behind the system used for applying all-pass filters to process audio signals being delivered to the multiway speaker and propose a time-delay difference equalization technique, between the drivers of the multiway loudspeakers showing group-delay equalization while retaining a flat magnitude response. Clearly, applying all-pass filters will result in temporal-smearing of the measured response. Furthermore, an investigation using perceptually motivated variable-octave complex smoothing of responses, and designing all-pass filters based on this phase-smoothed data, will also be undertaken. Quantitative results obtained will be presented in this paper whereas the next part of the two-part paper will present results from listening tests.
Convention Paper 7396

16:00

P13-2 Singing Voice Separation Combining Panning Information and Pitch Tracking—*Maximo Cobos, Jose J. Lopez*, Technical University of Valencia, Valencia, Spain

Source Separation techniques applied to music mixtures are able to extract relevant information that can be very useful for many applications, such as music remixing and reprocessing, lyrics recognition or music information retrieval. Among all the sources present in modern music themes, the singing voice has a especial interest because it is the only one that combines music, lyrics, and expression. In this paper we propose a system designed for extracting the singing voice from stereo recordings in different steps. This system combines panning information and pitch tracking, allowing the refinement of the time-frequency mask applied for extracting a vocal segment, and thus, improving the separation. An application example is discussed.
Convention Paper 7397

16:00

P13-3 The Downsampling Dilemma: Perceptual Issues in Sample Rate Reduction—Brett Leonard, New York University, New York, NY, USA

Many options currently exist for sample rate conversion. With sample rate reduction playing an integral part in the modern production world, downsampling algorithm quality is more important than ever. This paper presents data exploring the differences in sample rate reduction algorithms. While certain tests clearly display differences in the quality of the algorithms, listening test data shows the average listener is unable to repeatedly discern the difference in sample rate reduction methods.
Convention Paper 7398

16:00

P13-4 NU-Tech: The Entry Tool of the hArtes Toolchain for Algorithms Design—*Ariano Lattanzi,*[1] *Ferruccio Bettarelli*,[1] *Stefania Cecchi*[2]
[1]Leaff Engineering, Porto Potenza Picena (MC), Italy
[2]Univerista Politecnica delle Marche, Ancona, Italy

The aim of the hArtes project is to facilitate and automate the rapid design and development of heterogeneous embedded systems, targeting a combination of a general purpose embedded processor, digital signal processing, and reconfigurable hardware. In this paper we present the NU-Tech platform, the main entry tool from the hArtes toolchain, which has the role of assisting the designers in tuning and possibly improving the input algorithm at the highest level of abstraction. A brief description of the project itself will be given and its vocation to audio highlighted through a case study application.
Convention Paper 7399

16:00

P13-5 Recovery of Missing Signals Utilizing (GHA) Generalized Harmonic Analysis—Applied Interpolation—*Teruo Muraoka, Takahiro Miura, Tohru Ifukube*, University of Tokyo, Tokyo, Japan

For archiving damaged historical recordings, recoveries of missing portions are as essentially important as noise reduction. Conventional counter-measures with functional interpolation

are not effective when the missing interval is long. Inharmonic frequency analysis GHA is profitable for this purpose, because the recomposed signal with frequency components obtained by GHA exhibits very long periods. Length of the period is given as an inverse number of the least common multiple of the rendered inharmonic frequency's periods. This feature is very advantageous for signal recovery, and the authors devised an extrapolation simply extending a re-synthesizing waveform obtained through GHA analysis/re-synthesis. The authors got satisfactory results as a whole by applying interpolation combined with forward and backward extrapolations based upon the abovementioned method. Results of recovery highly depend upon characters of signals (such as music), and the authors did not find definite rules for setting GHA's analyzing conditions. Those are given through auditory examination this time.
Convention Paper 7400

16:00

P13-6 Combination of Warped and Linear Filter Structures for Loudspeaker Equalization—*German Ramos,*[1] *Jose J. Lopez,*[1] *Basilio Pueo*[2]
[1]Technical University of Valencia, Valencia, Spain
[2]University of Alicante, Alicante, Spain

The warping filters were introduced years ago for loudspeaker equalization in order to solve the lack of resolution of the linear filters at low frequencies, and also to follow the frequency resolution of psycho-acoustic scales like the Bark scale, with a more logarithmic than linear behavior. However, this improvement in the frequency resolution at low and mid frequencies is done at the expense of loosing resolution at high frequencies and increasing the complexity of the filter and its implementation computational cost. In this paper a smart combination of linear and warped filter structures previously developed by the authors for FIR filters is presented with new contributions and extended to IIR filters. This combination saves computational cost and obtains a proper frequency resolution at the whole frequency band, obtaining better results for the same computational cost than when using linear or warped filters alone. The results have been subjectively tested using the ABX methodology with successfully results. The presented filter structures, methodology, and apparatus to do the filtering are patent pending.
Convention Paper 7401

16:00

P13-7 Multichannel Dereverberation System Using Modified Correlation-Based Blind Deconvolution and Multi-Microphone Spectral Subtraction—*Jae-Woong Jeong*,[1] *Young-Cheol Park*,[2] *Seok-Pil Lee*,[3] *Dae-Hee Youn*[1]
[1]Yonsei University, Seoul, Korea
[2]Yonsei University, Wonju, Korea
[3]Korea Electronics Technology Institute (KETI), Sungnam, Korea

This paper presents a new multichannel dereverberation system combining modified correlation-based blind deconvolution with multi-microphone spectral subtraction. In the proposed system, we make M combinations of observed signals and apply them to the correlation-based blind deconvolution. The deconvolved signals are then used as inputs to the multi-microphone spectral subtraction. These spectral subtractions with the multiple deconvolved signals estimate the reverberant energy by using both a frame delay and a frequency-dependent weight. Due to the accurate estimation of the reverberant energy, the combination of correlation-based blind deconvolution with the multi-microphone spectral subtraction provides improved dereverberation performance. Performance improvement of the proposed system has been confirmed through experiments.
Convention Paper 7402

16:00

P13-8 Harmonic and Intermodulation Analysis of Nonlinear Devices Used in Virtual Bass Systems—*Nay Oo, Woon-Seng Gan*, Nanyang Technological University, Singapore

Nonlinear Devices (NLD) are used in virtual bass system. NLD generates harmonics which in turn create the pitch perception and are used in audio bass enhancement systems using psychoacoustics. This paper presents the mathematical derivations and analysis of five different NLD devices, together with intermodulation analysis of harmonics generated by these NLDs. The five NLDS are half-wave rectifier, full-wave rectifier, square wave, polynomial function, and exponential function. The derivation of harmonic analysis equations are based on Fourier Theorems, Chebyshev Polynomials, and Taylor Series expansions. Besides the harmonics, intermodulation components also resulted from NLDs. Both mathematical analysis and simulation results are presented for the intermodulation effects of harmonics generated by NLDS.
Convention Paper 7403

Tutorial 10 **Sunday, May 18**
16:00 – 18:00 **Room B**

TINNITUS. JUST ANOTHER BUZZ WORD?

Presenters: **Neil Cherian**, Center for Performance Medicine, Cleveland Clinic, Cleveland, OH, USA
Michael Santucci, Sensaphonics Hearing Conservation, Inc., Chicago, IL, USA
Jan Voetmann, Voetmann-Acoustics, Copenhagen, Denmark

Tinnitus is a common, yet poorly understood, disorder where sounds are perceived in the absence of an external source (phantom). Significant sound exposure with or without hearing loss is the single most common risk factor. Tinnitus can be debilitating, affecting quality of life or even one's ability to function. Given the potential harm of sound in the development of tinnitus, more aggressive and proactive attitudes must be taken. In-ear monitoring strategies further necessitate meaningful conversations regarding hearing awareness, hearing protection, safe standards for listening, and appropriate safeguards for products. This tutorial introduces the concept of tinnitus, the pertinent anatomy and physiology, the audiologic parameters of tinnitus, current research, guidelines for ➥

identifying high risk behaviors, and how to determine that you have a problem.

Historical Event
RECORDING HISTORY
Sunday, May 18, 16:00 – 17:00
Room N

Presenter: **Bert van der Wolf**

"The importance of technical specifications of recording equipment, and the sound quality of a recording, in the communication of the Emotion in a recorded Musical Performance . . . and how our brain and ears respond to reproduced acoustical information in several different audio formats." A personal report of 20 years practical experience in Record Producing, and development of recording equipment through research of, and fascination for the human audio/music perception.

Sunday, May 18 16:00 Room K
Technical Committee Meeting on Loudspeakers and Headphones

Workshop 14 Sunday, May 18
16:30 – 18:00 Room N101

CHANGING TO FILE-BASED PRODUCTION

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

Panelists: *Steinar Bjørlykke*, NRK – Norwegian Broadcasting
Dominique Brulhart, Merging Technologies
Jean-Christophe Liechti, TSR – Télévision Suisse Romande
Brendan Mallon, BBC Scotland

While the radio world has already several years of experience with file-based production and transmission, the television side is just at the beginning of the transition toward a fully file-based workflow. This is a significant change—one that poses many challenges on the engineers, the infrastructure, the equipment, and the system that ties it all together. Buzzwords like Content Management System, MXF (Material Exchange Format), Buffer Storage, or Playout Server enter the world of the audio engineer—often causing considerable panic. In this workshop representatives of broadcasters who have already gained some experiences in this field will share their viewpoints, introduce their specific solutions to common problems, and generally raise the awareness and urgency to dive into this world. It will soon be the daily life of us all.

Sunday, May 18 17:00 Room H
Standards Committee Meeting on SC-03-04 Storage and Handling of Media

Special Event
OPEN HOUSE OF THE TECHNICAL COUNCIL AND THE RICHARD C. HEYSER MEMORIAL LECTURE
Sunday, May 18, 18:00 – 19:30
Room B

Lecturer: **A. J. (Guus) Berkhout**

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 124th AES Convention is A. J. (Guus) Berkhout. He has written several hundred scientific papers and a number of books in the fields of acoustics, geophysics, and innovation. In the late eighties he introduced the concept of wave field synthesis (WFS) and wave field analysis (WFA) in audio engineering. In 2003 he received the highest international award in the field of Exploration Geophysics.

Berkhout is a member of the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Netherlands Academy of Engineering (AcTI). He received honorary memberships of the Society of Exploration Geophysicists (SEG) and the European Association of Geophysicists and Engineers (EAGE). He also serves on a number of Governing Boards in the scientific and business communities. The title of his lecture is, "The Big Challenges in Audio: A Glance into the Future."

Sound is an important information carrier. In speech and music the source carries the message, but in acoustical imaging the message is given by the medium. In all these applications, it is important to realize that sound is a wave phenomenon, often with complex wave fronts and complex time signals. For example, in enclosed spaces wavefields represent an intricate interference pattern of multi-source signals and multi-boundary reflections. Here, audio systems have the important task to enhance the social function of these spaces. If we want to make the next big step in improving audio solutions for demanding listening environments, we should challenge traditional believes and rethink current design methods.

On the one hand, there is the dimension of new technological capability. Wavefield knowledge should have a major impact on the way we develop the next generation of audio products. This technological expedition into the future will lead us to the exciting world of transducer arrays and matrix processors, both for analysis (WFA) and synthesis (WFS) purposes. New functionality will include variable acoustics, focused sound delivery, and selective signal enhancement.

On the other hand, there is the dimension of improved user value. For commercial success, technological excellence is necessary but the perceived value by the market is of overriding importance. Knowledge of the changing soft values in society should inspire the new solutions. This human-centered expedition into the future will bring us to the innovative world of integrated audio-optical systems. Examples are the combination of variable acoustics with variable lighting (WFS-plus), the integration of hearing aids with spectacles ("hearing glasses"), and the connection of optical cameras with highly directional microphones ("forensic audio products").

In conclusion, the message of the 2008 Richard C. Heyser Memorial Lecture is: *Do not try to predict the future, but have the ambition to create the future*.

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

Special Event
ORGAN RECITAL BY GRAHAM BLYTH
Sunday, May 18, 20:30 – 21:30
Thomas Aquinen Kerk
Hunzestraat 87, Amsterdam

Graham Blyth's traditional organ concert will be given at Thomas Aquinen Kerk (walking distance fom the convention center). The organ was built by the local firm of Flentrop in 1959, one year after the famous organ they built for the Bush Reisinger Museum at Harvard University, and very similar in stoplist.

The organ is a so-called straight organ with wind chests and mechanical tracker action. There are 30 stops totally, divided over Great Organ (in the middle), Choir Organ (the little organ in front), Brustwerk (above the manuals), and Pedal Organ (on both sides of the Great Organ). The organ case was made in the style common the fifties and sixties.

The featured works *Toccata, Adagio, and Fugue in C* by Bach; *Variations on Mein junges Leben hat ein End*, by Sweelink, *2nd Organ Sonata* by Mendelssohn, plus works by Dubois, Vierne, and Mushel.

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

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

Session P14 **Monday, May 19**
09:00 – 12:00 **Room C-D**

PSYCHOACOUSTICS, PERCEPTION, AND LISTENING TESTS, PART 1

Chair: **Jan de Laat**, LUMC, Leiden, The Netherlands

09:00

P14-1 Speech Quality Measurement for the Hearing Impaired on the Basis of PESQ—*John G. Beerends,*[1] *Jan Krebber,*[2] *Rainer Huber,*[3] *Koen Eneman,*[4] *Heleen Luts*[4]
[1]TNO Information and Communication Technology, Delft, The Netherlands
[2]Technische Universität Dresden, Dresden, Germany, now with Sysopendigia PLC, Helsinki, Finland
[3]HörTech GmbH, Oldenburg, Germany
[4]Katholieke Universiteit Leuven, Leuven, Belgium

One of the research topics within the HearCom project, a European project that studies the impact of hearing loss on communication, is to find methods with which the speech quality as perceived by the hearing impaired can be measured objectively. ITU-T Recommendation P.862 PESQ and its wideband extension P.862.2, are obvious candidates for this despite the fact that they were developed for normal hearing subjects. This paper investigates the extent to which PESQ and possible simple extensions can be used to measure the quality of speech signals as perceived by hearing impaired subjects.
Convention Paper 7404

09:30

P14-2 Subjective Evaluation of Speech Quality in a Conversational Context—*Emilie Geissner,*[1] *Valérie Gautier-Turbin,*[1] *Marie Guéguin,*[2,3] *Laetitia Gros*[1]
[1]France Télécom R&D, Lannion, France
[2]Laboratoire Traitement du Signal et de l'Image, Rennes, France
[3]Université de Rennes, Rennes, France

Within the framework of ITU-T, an objective conversational model is developed to predict the impact of network impairments on the conversational quality experienced by a end-user. To train and validate such a model, subjective scores are required. Assuming that a conversation is made of talking, listening, and inter-action activities, a subjective test protocol is specially designed to take into account these multidimensional aspects of the speech quality in a conversation. Subjects are asked to evaluate speech quality in talking, listening, and conversational contexts separately during three successive tasks. The analyses of several tests show that this method is valid for the assessment of listening, talking, and conversational quality.
Convention Paper 7405

10:00

P14-3 Contribution of Interaural Difference to Obstacle Sense of the Blind While Walking—*Takahiro Miura, Teruo Muraoka, Shuichi Ino, Tohru Ifukube*, University of Tokyo, Tokyo, Japan

Most blind people can recognize some measure of objects existing around them only by hearing. This ability is called "obstacle sense" or "obstacle perception." It is known that this ability is facilitated while the subjects are moving, however, the exact reason of the facilitation has been unknown. It is apparent that some differences of sounds reaching between both ears significantly change while approaching the obstacles. We focused on this phenomenon called interaural difference in order to analyze the facilitation mechanism of the obstacle sense. We investigated how the interaural differences change depending on the head rotation while walking and then measured the DL (Difference Limen) of the interaural difference. Furthermore, we compared the measurement data and the DL with the relationship between the subject-to-obstacle distance and then discussed one of the factors

of the facilitating the obstacle sense.
Convention Paper 7406

10:30

P14-4 The Accuracy of Localizing Virtual Sound Sources: Effects of Pointing Method and Visual Environment—*Piotr Majdak, Bernhard Laback, Matthew Goupell, Michael Mihocic*, Austrian Academy of Sciences, Vienna, Austria

The ability to localize sound sources in a 3-D-space was tested in humans. The subjects listened to noises filtered with subject-specific head-related transfer functions. In the experiment using naïve subjects, the conditions included the type of visual environment (darkness or structured virtual world) presented via head mounted display and pointing method (head and manual pointing). The results show that the errors in the horizontal dimension were smaller when head pointing was used. Manual pointing showed smaller errors in the vertical dimension. Generally, the effect of pointing method was significant but small. The presence of structured virtual visual environment significantly improved the localization accuracy in all conditions. This supports the benefit of using a visual virtual environment in acoustic tasks like sound localization.
Convention Paper 7407

11:00

P14-5 Perceived Spatial Distribution and Width of Horizontal Ensemble of Independent Noise Signals as Function of Waveform and Sample Length—*Toni Hirvonen, Ville Pulkki*, Helsinki University of Technology, Espoo, Finland

This paper investigates the perceived sound distribution and width of a horizontal loudspeaker ensemble as a function of signal length, as all loudspeakers emit simultaneous, white Gaussian noise bursts. In Experiment 1, subjects indicated the perceived distribution of 10 frozen cases where signal length was 2.5 ms. In Experiment 2, two cases from the previous test were investigated with signal lengths of 5-640 ms. The results indicate that (1) ensembles consisting of different short noise bursts vary in perceived distribution between cases and (2) when the length of the signal is increased, the produced sound event is generally perceived more wide. In perceiving such cases, the hearing system possibly utilized some temporal integration and/or adaptive processes.
Convention Paper 7408

11:30

P14-6 Effect of Minimizing Spatial Separation and Melodic Variations in Simultaneously Presented Two-Syllable Words—*Jon Allan, Jan Berg*, Luleå University of Technology, Piteå, Sweden

This paper will examine two important factors for the conception Auditory Streaming defined by Bregman, pitch, and localization. By removing one or two of these factors as possible identifiers to separate sound sources, the importance of each of them and the effect of reducing both of them will be studied. Stimuli with combinations of two-syllable words will be presented simultaneously in speakers to subjects, and the number of correct identifications will be measured. In one category of stimuli speech, melody will be removed and replaced with a monotonous pitch, equal for all words. One category will have all words presented from one speaker only. Conclusions will be related to earlier studies and common theories, the Cocktail party effect among others.
Convention Paper 7409

Session P15 **Monday, May 19**
09:00 – 12:00 **Room E-F**

SIGNAL PROCESSING, SOUND QUALITY DESIGN

Chair: **Jan Abildgaard Pedersen**, Lyngdorf Audio, Skive, Denmark

09:00

P15-1 Characterization of the Multidimensional Perceptive Space for Current Speech and Sound Codecs—*Thierry Etame,*[1,2] *Laetitia Gros,*[1] *Catherine Quinquis,*[1] *Gérard Faucon,*[2,3] *Régine Le Bouquin Jeannes*[2,3]
[1]France Télécom R&D, Lannion Cedex, France
[2]University of Rennes, Rennes, France
[3]INSERM, Rennes, France

The purpose of our work is to produce a reference system that can simulate and calibrate degradations of speech and audio codecs which are currently used on telecommunications networks, for subjective assessment tests of voice quality. At first, 20 wideband codecs are evaluated through subjective tests with the general goal of producing the multidimensional perceptive space underlying the perception of current degradations. Then, from a verbalization task, it appears that the identified attributes are clear/muffle, high-frequency noise, noise on speech, and hiss. Finally, these dimensions are characterized with correlates such as spectral centroid, spectral flatness measure, Mean Opinion Score, and correlation coefficient.
Convention Paper 7410

09:30

P15-2 An Automatic Maximum Gain Normalization Technique with Applications to Audio Mixing—*Enrique Perez Gonzalez, Joshua D. Reiss*, Queen Mary, University of London, London, UK

A method for real-time magnitude gain normalization of a changing linear system has been developed and tested with a parametric filter design. The method is useful in situations where the maximum gain before feedback is needed. The method automatically calculates the appropriate gain that should be applied in order to maintain maximum unitary gain. The method uses an impulse measurement of a mathematical model of the system to be normalized. This is particularly useful for mixing engineers, who have to continually revise their gain structure in order to maximize gain before feedback. The system is also useful in many other situations where solving the analytical solution from the

mathematical model is not possible.
Convention Paper 7411

10:00

P15-3 An Alternative Approach for the Convolution in Time-Domain: The Taches-Algorithm—*Laurent Millot, Gérard Pelé*, ENS Louis-Lumière, Noisy-le-Grand cedex France

We present an alternative temporal approach for convolution, providing a new algorithm, called the taches-algorithm. Based on interferences between the successive delayed and amplified output signals associated respectively with the impulses constituting the input signal, the taches-algorithm can give access immediately to the new output sample and have a low latency response using vector-based optimization of the calculation. With the taches-algorithm it is easy to change (even in real time) the impulse response while running the calculation, simply by updating the impulse response to use it for next samples, a task rather difficult to achieve using FFT convolution. Real time audio demonstrations using Pure Data and simple explanations of the taches-algorithm will be given.
Convention Paper 7412

10:30

P15-4 Performance of Independent Component Analysis when Used to Separate Competing Acoustic Sources in Anechoic and Reverberant Conditions—*Ben Shirley, Paul Kendrick*, University of Salford, Salford, Greater Manchester, UK

A review of existing methods for independent component analysis was carried out and a series of experiments conducted assessing the use of existing independent component analysis (ICA) methods to separate microphone sources in varied acoustic environments. Specifically the research looked at how effectively ICA could perform in a broadcast context using standard microphone techniques such as spaced omni and coincident crossed cardioid pairs. Experiments were carried out in an anechoic chamber and also in a listening room conforming to the ITU-R BS.1116-2 standard. Results clearly indicate the limitations of ICA when performed on audio material recorded in a reverberant environment; however it was still shown possible to achieve separation of signals of up to 12 dB even in these conditions.
Convention Paper 7413

11:00

P15-5 A Cross-Platform Audio Signal Processing Environment for Real-Time Audio Algorithm Development—*Mika Ristimäki*,[1] *Matti Hämäläinen*,[2] *Julia Turku*,[1] *Riitta Väänänen*[2]
[1]Nokia Research Center, Helsinki, Finland
[2]Nokia Research Center, Tampere, Finland

This paper presents a real-time audio algorithm development environment for experimental audio system research. The backbone of the system is Pure Data audio signal processing platform, which enables flexible implementation of real-time audio systems. With the proposed development environment the user can concentrate on real-time audio algorithm development and performance evaluation in the workstation environment. We present the proposed algorithm design method and environment, and its application to an experimental Voice over Internet Protocol (VoIP) system development.
Convention Paper 7414

11:30

P15-6 New Enhancements to the Automatic Noise Removal (ANR) System Utilizing Improved Noise Statistics and Multi-Band Processing—*Shamail Saeed,*[1] *Harinarayanan E. V.*,[1] *Deepen Sinha*,[2] *Anibal Ferreira*[2,3]
[1]ATC Labs, Noida, India
[2]ATC Labs, Chatham, NJ, USA
[3]University of Porto, Porto, Portugal

We recently introduced a novel Automatic Noise Reduction (ANR) algorithm for the removal of wideband stationary/nonstationary noise from audio. Current noise reduction techniques exhibit certain undesirable characteristics. Distortion and/or alteration of the audio characteristics is a common problem. User intervention in identifying the noise profile is sometimes necessary. ANR uses a novel framework employing dominant component subtraction and restoration and performs better than conventional techniques in subjective tests. Here we describe three enhancements to ANR. The first of these increases the level of noise removal for the special case of stationary background noise. The second is a new tool for improving the temporal envelope coherence and yields additional noise removal. The third is a multi-band processing tool for conditioning time-frequency envelope for reduced listener fatigue.
Convention Paper 7415

Workshop 15 **Monday, May 19**
09:00 – 10:30 **Room B**

PRACTICAL USES OF ACOUSTIC MODELING

Chair: **Ben Kok**, Nelisse Ingenieursbureau B.V., Eindhoven, The Netherlands

Panelists: *Peter Mapp*, Peter Mapp Associates, Colchester, Essex, UK
Henrik Møller, Akukon Oy Consulting Engineers, Helsinki, Finland
Dirk Noy, WSDG, Basel, Switzerland

What can be achieved by acoustic simulations for various types of environments? The panel members will explain how they used software simulation in their consulting and design practice. Examples will be given for performance spaces, small rooms (like studios), and harsh environments (like stadiums and tunnels).

Tutorial 11 **Monday, May 19**
09:00 – 11:00 **Room L**

THE GAME AUDIO PROCESS

Presenters: **Mario Lavin**
Lucas Van Tol

➥

Andreas Varga
Anton Woldhek

This tutorial provides an overview of the process of developing audio for a modern computer game, from initial conception through to final release. The session is presented by the audio team from Guerrilla Games who are responsible for a number of PlayStation titles. The talk illustrates the various aspects of game audio by considering the development of a single title from the point of view of the audio director, sound designer, and programmer. The panel will also draw on their separate and collective experience from other titles to contrast different approaches to game audio. The tutorial offers a detailed but technical introduction to game audio and is suitable for newcomers and experienced professionals alike.

Monday, May 19 09:00 Room H
Standards Committee Meeting on SC-03-06 Digital Library and Archive Systems

Monday, May 19 09:00 Room K
Technical Committee Meeting on Acoustics and Sound Reinforcement

Session P16 Monday, May 19
09:30 – 11:00 Topaz Lounge

POSTERS: ANALYSIS AND SYNTHESIS OF SOUND, PART 1

09:30

P16-1 A Channel Vocoder Using Wavelet Packets over a Reconfigurable Device—*César Daniel Salvador Castañeda*, Pontificia Universidad Católica del Perú, Lima, Peru

A channel vocoder using wavelet packets for computer music applications is proposed. The inputs are a modulating signal, which is choice to be voice, and a carrier signal, which can be music or noise. The Wavelet Packets Channel Vocoder transforms windowed frames of both signals to a symmetric multiresolution representation, mixes the envelope of the modulating signal with the carrier, and transforms back the result to the original domain. Simulations run with Simulink. Real time implementations are presented for Pure Data and Xilinx Virtex II Pro FPGA. Appropriate choices of window, overlap, wavelets, decomposition levels, and envelope detector are presented to achieve different sound effects. Finally, new ideas to improve transmission and compression rates in future works are also proposed.
Convention Paper 7416

09:30

P16-2 The Effects of Lossy Audio Encoding on Genre Classification Tasks—*Kurt Jacobson*,[1] *Ben Fields*,[2] *Mark Sandler*,[1] *Michael Casey*[2]
[1]Queen Mary University of London, London, UK
[2]Goldsmith's College, University of London, London, UK

In large audio collections, it is common to store audio content using perceptual encoding. However, encoding parameters may vary from collection to collection or even within a collection—using different bit rates, sample rates, codecs, etc. We evaluate the effect of various lossy audio encodings on the application of audio spectrum projection features to the automatic genre classification tasks. We show that decreases in mean classification accuracy, while small, are statistically significant for bit-rates of 96-kbps or lower. Also, a heterogeneous collection of audio encodings has statistically significant decreases in mean classification accuracy compared to a pure PCM collection.
Convention Paper 7417

09:30

P16-3 Loop Region Detection in Music Signals—*Bee Suan Ong, Sebastian Streich*, Centre for Advanced Sound Technologies, Yamaha Corporation, Japan

Spotting loops within a music recording seems to be an easy task for human listeners. Nevertheless it becomes highly time and effort consuming when loop segments are to be identified from a large music collection. The process can be greatly facilitated with an audio editing tool that highlights regions where loops appear and suggests loop durations respectively. This paper proposes a method for computing both types of information from the music signals. Our approach is based on identifying sequential and regular repetitions of tonal features. In addition, we present a prototype implementation featuring the proposed method to facilitate the audio browsing and searching process. Finally, we discuss other possible applications of this technology in the audio content description context.
Convention Paper 7418

09:30

P16-4 Music-Inspired Harmony Search Algorithm Applied to Feature Selection for Sound Classification in Hearing Aids—*Javier Amor, Enrique Alexandre, Roberto Gil-Pita, Lorena Álvarez, Ester Huerta*, Universidad de Alcalá, Alcalá, Spain

This paper explores the application of the music-inspired harmony search algorithm to the problem of feature selection for sound classification in digital hearing aids. The importance of this problem is given by the strong computational constraints inherent to the DSPs used in modern digital hearing aids. The goal of the feature selection algorithm is to select a subset of features in order to reduce the computational complexity of the system while maintaining a low probability of error. A set of experiments will be performed to test the performance of the proposed system, using a total of 74 different features. The results will be compared with those obtained using other widely-used algorithms, such as a genetic algorithm, a sequential search algorithm or random search.
Convention Paper 7419

09:30

P16-5 Analysis of the Effects of Finite Precision in Sound Classifiers for Digital Hearing Aids—*Ester Huerta, Enrique Alexandre, Roberto Gil-Pita, Lorena Álvarez, Javier Amor*, Universidad de Alcalá, Alcalá de Henares, Madrid, Spain

This paper deals with the analysis of quantization effects in an automatic sound classification system for DSP-based hearing aids. The results obtained in this work will be used to find out the impact of finite accuracy determined by the digital signal processor (DSP) on the users of hearing aids. The DSP has a finite word length that affects the main ability of these systems: the automatic adaptation to the changing acoustic environment. The goal of this work is to model a quantized Neural Network-based classifier in order to compare the probability of error obtained with those nonfinite precision systems.
Convention Paper 7420

09:30

P16-6 A Constructive Algorithm for Multilayer Perceptrons for Speech/Non-Speech Classification in Hearing Aids—*Lorena Álvarez, Enrique Alexandre, Raúl Vicen, Lucas Cuadra, Manuel Rosa*, Universidad de Alcalá, Alcalá de Henares, Spain

Constructive learning algorithms offer an attractive approach for the incremental construction of near-minimal neural-network architectures for pattern classification. This paper explores the feasibility of using a constructive algorithm for multilayer perceptrons (MLPs) applied to the problem of speech/non-speech classification in hearing aids. When properly designed and trained, MLPs are able to generate an arbitrary classification frontier with a relatively low computational complexity. The paper will focus on the design of a constructive algorithm for MLPs that attempts to converge to the minimum complexity network for the given problem. The results obtained will be compared with those cases in which the constructive algorithm is not considered.
Convention Paper 7421

09:30

P16-7 Seeing the Inaudible. Descriptors Used for Generating Objective and Reproducible Data in Real-Time for Musical Instrument Playing Standard Situations—*Tobias Grosshauser, Diemo Schwarz, Norbert Schnell*, IRCAM, Paris, France

This paper describes a method to generate objective and reproducible data to assist instrument teaching and practicing. the method is based on using audio descriptors and their efficient visualization that assist in the perception of musical parameters difficult to hear. To aid comparison, we defined and recorded a comprehensive database of positive and negative sound examples from the violin that encompasses frequent mistakes made by students and a wide variety of playing styles.
Convention Paper 7422

Workshop 16 — Monday, May 19
09:30 – 11:00 — Room O

IS THE CHANNEL DEAD?

NEW WAYS TO BROADCAST YOUR MESSAGE

Chair: **Eirik Solheim**, NRK Oslo, Oslo, Norway

Co-Chair: *Karl Petermichl*, ORF Radio

Eirik Solheim and Karl Petermichl will go through the latest trends on the internet. And they'll show you case studies that include the use of YouTube, Facebook, Twitter, and a slightly controversial distribution method called BitTorrent.

And possibly one more thing . . .

Live Sound Seminar 6 — Monday, May 19
09:30 – 12:00 — Forum

FROM RIDER TO SHOW, PART 1: LOGISTICS

Presenter: **Eberhard Müller**, Neumann & Müller, Hamburg, Germany

• About the rider

• How to choose the correct equipment by using a software program

• How to calculate within a given budget

Monday, May 19 — 10:00 — Room K
Technical Committee Meeting on Audio Forensics

Workshop 17 — Monday, May 19
11:00 – 12:30 — Room B

AUDIO NETWORKING FOR THE PROS

Chair: **Umberto Zanghieri**, ZP Engineering srl, Rome, Italy

Panelists: *Steve Gray*, Cirrus Logic, Inc., Austin, TX, USA
Greg Shay, Axia Audio, Cleveland, OH, USA
Jérémie Weber, Auvitran, Meylan, France
Aidan Williams, Audinate, Broadway, Australia

Several solutions are available on the market today for digital audio transfer over conventional data cabling, but only some of them allow usage of standard networking equipment. This workshop presents some commercially available solutions (Cobranet, Livewire, Ethersound, Dante), with specific focus on noncompressed, low-latency audio transmission for pro-audio and live applications, using standard IEEE 802.3 network technology. The main challenges of digital audio transport will be outlined, including compatibility with common networking equipment, reliability, latency, and deployment. Typical scenarios will be proposed, with panelists explaining their own approaches and solutions.

Exhibitor Seminar 3 — Monday, May 19
11:00 – 12:00 — Room P

SENNHEISER ELECTRONIC CORPORATION

Presenters: **Sebastian Schmitz**
Gregor Zielinsky

MKH 800 Twin with Remote Controllable Directivity

MKH 800 Twin The MKH 800 Twin is a Doublecapsule Microphone with remote-controllable pattern for stereo ➥

and surround recording. Each of the capsules can be recorded separately as they have separate outputs. This enables you to seamlessly adjust the pickup pattern at post-production stage. The presentation shows several examples of how to use this technique in stereo and surround examples.

Monday, May 19 11:00 Room H
Standards Committee Meeting on SC-03-07 Audio Metadata

Monday, May 19 11:00 Room K
Technical Committee Meeting Advisory on Regulations

Session P17 Monday, May 19
11:30 – 13:00 Topaz Lounge

POSTERS: ANALYSIS AND SYNTHESIS OF SOUND, PART 2

11:30

P17-1 Structural Segmentation of Music Using Set Accented Tones—*Cillian Kelly, Mikel Gainza, David Dorran, Eugene Coyle*, Dublin Institute of Technology, Dublin, Ireland

An approach that efficiently segments Irish Traditional Music into its constituent structural segments is presented. The complexity of the segmentation process is greatly increased due to melodic variation existent within this music type. In order to deal with these variations, a novel method using "set accented tones" is introduced. The premise is that these tones are less susceptible to variation than all other tones. Thus, the location of the accented tones is estimated and pitch information is extracted at these specific locations. Following this, a vector containing the pitch values is used to extract similar patterns using heuristics specific to Irish Traditional Music. The robustness of the approach is evaluated using a set of commercial Irish Traditional recordings.
Convention Paper 7423

11:30

P17-2 AnClaS3: A Blackboard-Based Cooperative Framework for Sound Separation—*Antonio Pena,*[1] *Norberto Degara-Quintela,*[1] *Manuel Sobreira-Seoane,*[1] *Soledad Torres-Guijarro*[2]
[1]Universidade de Vigo, Vigo, Spain
[2]Laboratorio Oficial de Metroloxía de Galicia (LOMG), Tecnópole, Ourense, Spain

Blackboard modeling provides a great flexibility in structuring complex problems and a robust adaptation to the conditions of the signal to be analyzed, adding both bottom-up and top-down capabilities to the system. AnClaS3 (Analysis, Classification, and Synthesis for Sound Separation) is a cooperative project where five research groups collaborate integrating algorithms and developing new separation methods. This contribution defines a blackboard-based framework where four blackboard-based systems interact to integrate the expertise of independent research groups in order to solve a sound separation problem.
Convention Paper 7424

11:30

P17-3 Analysis and Synthesis of Audio Vibrato Using Harmonic Sinusoids—*Wen Xue, Mark Sandler*, Queen Mary, University of London, London, UK

This paper introduces the analysis and synthesis of vibrato in music audio. The analyzer separates frequency modulators from their carriers using a demodulation process. It then describes the frequency variations of a vibrato using a period-synchronized parameter set and the accompanying amplitude variations using a source-filter model, both of which can be regarded slow-varying. The synthesizer, on the other hand, reconstructs a vibrato from a given set of parameters. Using this system we are able to retrieve specific characteristics of vibratos, or modify them to implement various audio effects.
Convention Paper 7425

11:30

P17-4 Distortion Analysis and Reduction for the Parametric Array—*Ee-Leng Tan,*[1] *Woon-Seng Gan,*[1] *PeiFeng Ji,*[2] *Jun Yang*[2]
[1]Nanyang Technological University, Singapore
[2]Chinese Academy of Sciences, Beijing, China

In this paper distortion analysis and reduction for the parametric array loudspeaker is being presented. The parametric loudspeaker has been found useful in generating a highly directional sound beam. However, due to the nonlinear interaction of ultrasonic wave in air, several undesired harmonic distortions have been generated. Conventional approaches in reducing the distortion have not created satisfying solutions. A new approach capable of further reducing the distortion has been proposed in this paper. Several simulation results are being carried out in this work to test and compare the effectiveness of this proposed solution with conventional approaches.
Convention Paper 7426

11:30

P17-5 Piano "Forte Pedal" Analysis and Detection—*Antony Schutz,*[1] *Valentin Emiya,*[2] *Dirk T. M. Slock,*[1] *Bertrand David,*[2] *Roland Badeau*[2]
[1]EURECOM Institute, Valbonne Sophia-Antipolis, France
[2]Telecom Paris Tech, Paris, France

In this paper we describe some features of the Forte Pedal piano effect and propose a method for detecting it through signal analysis. The detection method is applied to single tones recorded for this purpose. The Forte Pedal is found to increase the decay time of partials. In fact, this effect dominates the behavior of the partials, in not only the duration, but also the evolution. When the sustain pedal is used, a floor noise appears for all the notes of the piano. Here, after the analysis of some relevant characteristics we provide a method based on harmonic plus

noise decomposition for analyzing the residual and decide if the pedal is pressed or not.
Convention Paper 7427

Tutorial 12 **Monday, May 19**
11:30 – 14:00 **Room L**

BROADCAST CASE STUDIES OF PRODUCTIONS MADE IN CHALLENGING CIRCUMSTANCES

Presenters: **Florian Camerer**, ORF Vienna
Gaute Nistov, NRK Oslo
Teemu Tanskanen, YLE Helsinki

A follow up from the last European AES convention in Vienna in 2007, this workshop assembles 3 sound engineers each presenting an especially demanding television production needing tailor-made solutions for unusual challenges. An in-depth look to the nuts and bolts of these productions will be provided, from initial planning, to diverse microphone setups, signal flow, and redundancy and also including recording concepts for possible postproduction. The programs are: the Nobel Peace Prize Concert in Oslo, the Eurovision Song Contest in Helsinki, and the New Year's Concert 2008 in Vienna.

Tutorial 13 **Monday, May 19**
11:30 – 13:30 **Room N101**

ELECTROACOUSTIC MEASUREMENTS, PART 2

Presenter: **Christopher J. Struck**, CJS Labs, San Francisco, CA, USA

This tutorial focuses on applications of electroacoustic measurement methods, instrumentation, and data interpretation as well as practical information on how to perform appropriate tests. Linear system analysis and alternative measurement methods are examined. The topic of simulated free field measurements is treated in detail. Non-linearity and distortion measurements and causes are described. Last, a number of advanced tests are introduced. This tutorial is intended to enable the participants to perform accurate audio and electroacoustic tests and provide them with the necessary tools to understand and correctly interpret the results.

Tutorial 14 **Monday, May 19**
11:30 – 13:00 **Room O**

SMALL ROOM ACOUSTICS

Presenter: **Ben Kok**, Nelissen Ingenieursbureau B.V., Eindhoven, The Netherlands

Acoustic basics of small rooms will be discussed. Specific issues related to the size of the room (room-modes) will be addressed. Absorption, reflection, diffraction, diffusion and how to use it, and low frequency treatment. Although this will not be a studio design class, specifics and differences of recording rooms and control rooms will be identified, including considerations for loudspeaker and microphone placement.

Student Event/Career Development
MENTORING SESSIONS
Monday, May 19, 12:00 – 14:00
Topaz Lounge

Students are invited to sign-up for an individual meeting with a distinguished mentor from the audio industry. The opportunity to sign up will be given at the end of the opening SDA meeting. Any remaining open spots will be posted in the student area. All students are encouraged to participate in this exciting and rewarding opportunity for individual discussion with industry mentors.

Exhibitor Seminar 4 **Monday, May 19**
12:00 – 13:00 **Room P**

MINNETONKA AUDIO SOFTWARE (GERMANY)

Presenters: **Markus Hintz**
Jayson Tomlin

SurCode for Dolby E Workflow on Pro Tools, VST, and Other Platforms

Dolby E has become a de facto standard in surround audio even outside of broadcast facilities. Minnetonka Audio will present how the SurCode for Dolby E software products improve existing workflows on Pro Tools, VST, and on Minnetonka's AudioTools *Audio Workflow Engine* (AWE), helping studios to be productive and efficient.

Session P18 **Monday, May 19**
12:30 – 17:00 **Room C-D**

ROOM AND ARCHITECTURAL ACOUSTICS AND SOUND REINFORCEMENT

Chair: **Jan Voetmann**, DELTA Acoustics, Hoersholm, Denmark

12:30

P18-1 Diffusing Boundary Implementations in the 2-D Digital Waveguide Mesh—*Simon Shelley*,[1] *Damian Murphy*[2]
[1]Technische Universiteit Eindhoven, Eindhoven, The Netherlands
[2]York University, Heslington, York, UK

The digital waveguide mesh is a wave-based time-domain approach to the simulation of sound wave propagation in an acoustic system. The implementation of diffuse reflection is an important consideration in such an application, as the presence of diffuse reflection has a significant effect on an acoustic environment. The scattering effect of diffuse boundaries on reflected sounds, both in simulation and the real world, can be described using a technique that results in the formulation of frequency dependent diffusion coefficients. In this paper a number of different approaches to modeling diffuse reflection in a 2-D digital waveguide mesh are presented as well as a detailed analysis and comparison of the local scattering effect of the diffuse boundary models using this technique.
Convention Paper 7428

13:00

P18-2 ***RenderAIR*—Room Acoustics Simulation Using a Hybrid Digital Waveguide Mesh**—*Damian Murphy,*[1] *Mark Beeson*,[1] *Simon Shelley,*[2] *Alex Southern,*[1] *Alastair Moore*[1]
[1]University of York, Heslington, York, UK ➥

[2]Eindhoven University of Technology, Eindhoven, The Netherlands

The digital waveguide mesh (DWM) is a numerical simulation technique used to model signal propagation through a regular grid of spatio-temporal sampling points, and has been demonstrated as appropriate for modeling the acoustics of an enclosed space, particularly at low frequencies. The RenderAIR DWM application allows intuitive definition of parameters associated with geometry, boundary surface, and source/receiver parameters required to generate spatially encoded Room Impulse Responses (RIRs). In this paper the expectations and limitations of DWM-based room acoustics modeling are explored through the use of the RenderAIR application in a number of situations. ISO3382 metrics are used as the main benchmark for the results obtained, which compare well with both real-world measurements and more traditional geometric acoustic approaches.
Convention Paper 7429

13:30

P18-3 Volumetric Diffusers—*Richard Hughes, Jamie A. S. Angus, Trevor Cox, Olga Umnova*, University of Salford, Salford, Greater Manchester, UK

Although many types of diffusers have been proposed, they are predominantly surface treatment. This paper places the diffuser in the volume of the room rather than on the surfaces, forming a volume-based diffuser. In particular, we examine suitable sequences for their implementation. We also consider suitable metric's to evaluate their performance. At first single layer volumetric diffusers are examined, and then multi-layer volumetric diffusers are investigated. In particular, the effects of varying the spacing, and number of layers, is more closely examined. The Boundary Element Method (BEM) model is used to gain accurate predictions of the diffuser's performance. Finally, we demonstrate a diffusion structure that has a similar performance to that of a Primitive Root Diffuser (PRD).
Convention Paper 7432

14:00

P18-4 Commercial Low Frequency Absorbers —A Comparative Study—*Gabriel Hauser*,[1] *Dirk Noy*,[1] *John Storyk*[2]
[1]Walters-Storyk Design Group, Basel, Switzerland
[2]Walters-Storyk Design Group, Highland, NY, USA

This paper ties in to a previous Convention Paper by the same authors (AES 115th Convention, 2003, #5944) and presents a current set of commercially available passive and active low frequency absorbing devices. One item in particular is of an experimental nature—a wood box loaded with conventional membrane loudspeakers. These are not connected to an amplifier, but to a variety of different passive electronics networks (parallel, serial). Reproducible acoustical measurements have been taken in a completely untreated rectangular concrete room, sequentially with and without a total of eight different absorbing devices. Results are compared and conclusions are presented.
Convention Paper 7431

14:30

P18-5 Modeling Frequency-Dependent Boundaries as Digital Impedance Filters in FDTD and K-DWM Room Acoustic Simulations—*Konrad Kowalczyk, Maarten van Walstijn*, Queen's University Belfast, Belfast, Northern Ireland, UK

This paper presents a new method for modeling frequency-dependent boundaries in finite difference time domain (FDTD) and Kirchhoff variable digital waveguide mesh (K-DWM) room acoustics simulations. The proposed approach allows direct incorporation of a digital impedance filter (DIF) in the multi-dimensional (i.e., 2-D or 3-D) FDTD boundary model of a locally reacting surface. An explicit boundary update equation is obtained by carefully constructing a suitable recursive formulation. The method is analyzed in terms of pressure wave reflectance for different wall impedance filters and angles of incidence. Results obtained from numerical experiments confirm the high accuracy of the proposed digital impedance filter boundary model, the reflectance of which closely matches locally reacting surface (LRS) theory. Furthermore, a numerical boundary analysis (NBA) formula is provided as a technique for analytic evaluation of the numerical reflectance of the proposed digital impedance filter boundary formulation.
Convention Paper 7430
Winner of the Student Paper Award

15:00

P18-6 Loudspeaker Time Alignment Using Live Sound Measurements—*Wolfgang Ahnert, Stefan Feistel, Thorsten Maier, Alexandru Radu Miron*, Ahnert Feistel Media Group, Berlin, Germany

The authors previously introduced the measurement software EASERA SysTune, which can be used for measurements with live music and speech signals. In this paper we discuss specifically the use of real-time measurements for the time alignment of loudspeaker arrays and distributed systems and for the optimal adjustment of their phase relationships. Being capable of deriving impulse responses of up to 12 seconds length, the measuring process with EASERA SysTune is simpler and more accurate as the real-time function provides a more immediate view on the tuning process. Because measurements can be performed with standard stimulus signals as well as with external speech and music signals, fine-tuning loudspeaker settings becomes possible even during the rehearsal time of the musicians. Required measurement conditions and limitations are given
Convention Paper 7433

15:30

P18-7 INR as an Estimator for the Decay Range of Room Acoustic Impulse Responses—

Constant Hak,[1] *Jan Hak,*[2] *Remy Wenmaekers*[3]
[1]Eindhoven University of Technology, Eindhoven, The Netherlands
[2]Acoustics Engineering, Boxmeer, The Netherlands
[3]Level Acoustics, Eindhoven, The Netherlands

A room acoustic impulse response can be used to derive the reverberation time and other parameters. To this end a certain minimum energy decay range or effective signal to noise ratio is required, which relates to the difference between the integrated signal level and the noise level. An impulse response parameter called INR is presented as an estimator for the decay range and shown to be a useful qualifier in practical measurements.
Convention Paper 7434
[Paper presented by Remy Wenmaekers]

16:00

P18-8 Musical-Inspired Features for Automatic Sound Classification in Digital Hearing Aids—*Pedro Vera-Candeas,*[1] *Francisco J. Cañadas-Quesada,*[1] *Enrique Alexandre,*[2] *Manuel Rosa*[2]
[1]University of Jaén, Linares, Jaén, Spain
[2]University of Alcalá, Alcalá, Spain

This paper proposes the use of some musical-inspired features for the automatic classification of sounds in digital hearing aids. This kind of application is characterized by very strong constraints in terms of computational complexity. The proposed features are based on fundamental frequency detection and exhibit a low computational complexity while providing good results in terms of probability of correct classification. The performance of the system will be tested using a 1-NN classifier, the goal being to distinguish among speech, noise, and music. For the experiments a sound database, obtained using a hearing aid simulator, will be used.
Convention Paper 7435

16:30

P18-9 Assessing the Potential Intelligibility of Assistive Audio Systems for the Hard of Hearing and Other Users—*Peter Mapp*, Peter Mapp Associates, Colchester, Essex, UK

Around 14% of the European population suffer from a noticeable degree of hearing loss and would benefit from some form of hearing assistance or deaf aid. Recent DDA legislation and requirements mean that many more hearing assistive systems are being installed—yet there is evidence to suggest that many of these systems fail to perform adequately and provide the benefit expected. This paper reports on the results of some trial acoustic performance testing of such systems. In particular the effects of system microphone type, distance, and location are shown to have a significant effect on the resultant performance. The potential of using the Sound Transmission Index (STI) and in particular STIPa, for carrying out installation surveys has been investigated, and a number of practical problems are highlighted. The requirements for a suitable acoustic test source to mimic a human talker are discussed as is the need to the need to adequately assess the effects of both reverberation and noise. The findings discussed in the paper are also relevant to the installation and testing of classroom “sound field” systems and also boardroom type reinforcement systems and conferencing / teleconferencing systems.
Convention Paper 7436

Session P19 — Monday, May 19
12:30 – 15:30 — Room E-F

SOFTWARE, INSTRUMENTATION, AND MEASUREMENT

Chair: **John Vanderkooy**, University of Waterloo, Waterloo, Ontario, Canada

12:30

P19-1 Graphical Control of a Parametric Equalizer
—*Jörn Loviscach*, Hochschule Bremen (University of Applied Sciences), Bremen, Germany

Graphic equalizers allow the user to define a filter’s magnitude response virtually free of restrictions. Parametric equalizers are much more limited. However, they offer some vital advantages over graphic equalizers, such as consuming less computational power and operating minimally invasively with naturally soft magnitude and phase responses. This paper aims at combining the best of both worlds. It presents a range of methods to control a digital parametric equalizer graphically through a curve or a collection of anchor points. While the user is editing the graphical input, an optimization process runs in the background and adjusts the equalizer’s parameters to reflect the input. In addition, the number of bands and their type (shelving/peak) can be adjusted automatically to produce a simple solution.
Convention Paper 7437

13:00

P19-2 Audio Software Development—An Audio Quality Perspective—*Jonas Ekeroot, Jan Berg*, Luleå University of Technology, Piteå, Sweden

When developing audio applications, different choices on software implementation aspects influence the total audio software signal path and can be of importance from an audio quality perspective. The field is not well documented in the literature. A study was carried out aiming at identifying relevant questions that must be considered. The general development perspective was on audio software written in C++ to be run on general purpose CPUs. A research review, comprising literature from different fields such as audio engineering, computer science, and software engineering was conducted to summarize and integrate an overview of the field. The result can be viewed as a map of questions for future research activities, consisting of further literature studies and experiments with software prototypes.
Convention Paper 7438

➥

13:30

P19-3 Multi Carrier Modulator for Switch-Mode Audio Power Amplifiers—*Arnold Knott*,[1,2] *Gerhard Pfaffinger*,[1] *Michael A. E. Andersen*[2]
[1]Harman/Becker Automotive Systems GmbH, Straubing, Germany
[2]Technical University of Denmark, Lyngby, Denmark

While switch-mode audio power amplifiers allow small implementations and high output power levels due to their high power efficiency, they are very well known for creating electromagnetic interference (EMI) with other electronic equipment, in particular radio receivers. Lowering the EMI of switch-mode audio power amplifiers while keeping the performance measures to excellent levels is, therefore, of high general interest. A modulator utilizing multiple carrier signals to generate a two level pulse train will be shown in this paper. The performance of the modulator will be compared in simulation to existing modulation topologies. The lower EMI as well as the preserved audio performance will be shown in simulation as well as measurement results of a prototype.
Convention Paper 7439

14:00

P19-4 A Comparison of Theoretical, Simulated, and Experimental Results Concerning the Stability of Sigma Delta Modulators—*Georgi Tsenov,*[1] *Valeri Mladenov*,[1] *Joshua D. Reiss*[2]
[1]Technical University of Sofia, Sofia, Bulgaria
[2]Queen Mary, University of London, London, UK

Sigma delta modulation is a popular form of audio analog-to-digital and digital-to-analog conversion, but suffers from stability problems for many designs and many input signals. A general theory of stability in sigma delta modulators has been developed that predicts the stability of a high order one-bit sigma-delta modulator (SDM) under a variety of designs. In this paper the theoretical approach to stability as it applies to boundedness of states is explained. Several low pass SDM designs are developed that are intended for audio analog to digital conversion, and predicted results for stability of these designs are given. Stability is examined both in terms of the maximum allowable DC input amplitude and the theoretical sufficient conditions for stable behavior. Theoretical results are compared with simulated results, and where possible, with experimental results from a realization of a third order SDM with adjustable parameters. Practical observations are then made concerning the effect of noiseshaping, pole/zero placement, and cut-off frequency on the stability.
Convention Paper 7440

14:30

P19-5 A New Method for Identification of Nonlinear Systems Using MISO Model with Swept-Sine Technique: Application to Loudspeaker Analysis—*Antonín Novák*,[1,2] *Laurent Simon,*[2] *Pierrick Lotton*,[2] *Frantisek Kadlec*[1]
[1]Czech Technical University in Prague, Prague, Czech Republic
[2]Université du Main, Le Mans, France

This paper presents a Multiple Input Single Output (MISO) nonlinear model in combination with sine-sweep signals as a method for nonlinear system identification. The method is used for identification of loudspeaker nonlinearities and can be applied to nonlinearities of any audio components. It extends the method based on nonlinear convolution presented by Farina, providing a nonlinear model that allows simulation of the identified nonlinear system. The MISO model consists of a parallel combination of nonlinear branches containing linear filters and memory-less power-law distortion functions. Once the harmonic distortion components are identified by the method of Farina, the linear filters of the MISO model can be derived. The practical application of the method is demonstrated on a loudspeaker.
Convention Paper 7441

15:00

P19-6 Junction Identification Using Acoustic Reflectometry—*Adam Kestian, Agnieszka Roginska*, New York University, New York, NY, USA

Acoustic reflectometry is a non-invasive, time-domain method of identifying the geometry of an acoustical space. A sound pulse is injected into a space and the resulting impulse response details particular changes of impedance. In the present paper acoustic reflectometry is utilized to identify scattering junctions of geometric spaces. Most notably, the four most common types of scattering junctions are identified: a cross-sectional increase, cross-sectional decrease, L-intersection, and T-intersection.
Convention Paper 7442

Workshop 18 **Monday, May 19**
12:30 – 14:30 **Room B**

AUDIO OVER IP SYSTEMS FOR REPORTERS

Chair: **Lars Jonsson**, Swedish Radio

Panelists: *Lee Chaundy*, BBC, UK
Gerald List, Tranxtel, Germany
Ulrich Loebbert, WDR, Germany

This workshop will present case studies on the use of audio over IP systems for reporters out in the field. Audio over IP units have gradually become an alternative to ISDN and other means of reporting. Sometimes well-managed IP-networks can be used, or the unreliable but widely spread Internet. Wireless alternatives for reporters will also be discussed and evaluated. Examples of using the SIP protocol for connecting between AoIP units and VoIP telephones will be discussed.

Live Sound Seminar 7 **Monday, May 19**
13:00 – 17:00 **Forum**

FROM RIDER TO SHOW, PART 2: LIVE MIXING EVENT

Presenter: **Eberhard Müller**, Neumann & Müller, Hamburg, Germany

• Stage set up, calibrating

• Microphone set up and sound check

• Show

This session includes practical demonstrations with a jazz combo (piano, percussion, guitar, bass, vocal) and is presented in combination with an OB-Van. Attendees can move between auditorium and OB-van to calculate the different sound patterns.

Exhibitor Seminar 5 **Monday, May 19**
13:00 – 14:00 **Room P**

ODEON

Presenters: **Claus Lynge Christensen**
Jens Holger Rindel

ODEON Room Acoustic Simulations

ODEON is the state-of-the-art software for room acoustic simulations and auralizations, used in major concert hall and theater projects. The seminar will show examples of modeling including CAD-import, source directivity, array modelling, tools for reflection analysis, and ODEON's multichannel high fidelity auralization.

Monday, May 19 **13:00** **Room H**
Standards Committee Meeting on SC-02-12 Audio Applications of IEEE 1394

Monday, May 19 **13:00** **Room K**
Technical Committee Meeting on Multichannel and Binaural Audio Technologies

Student Event/Career Development
EDUCATION FORUM PANEL
Monday, May 19, 13:30– 15:30
Room P

Moderator: **Jason Corey**, University of Michigan, MI, USA

Panelists: *Mark Drews*, University of Stavanger, Norway
Theresa Leonard, The Banff Centre, Canada
Konrad Strauss, Indiana University, USA
Cornelis Van der Gragt, The Hague Royal Conservatoire, The Netherlands

Guiding and Evaluating Student Learning in Audio Engineering

As educators we are concerned with aiding and facilitating students' acquisition of knowledge and experience. We also usually need to find some method of evaluating whether students have gained the appropriate skill level and knowledge base for a given course. As audio is somewhat unique in terms of how students increase their knowledge of the field, we as educators are faced with unique challenges as well. During the discussion the panel will consider questions such as: What are the best ways to evaluate how well students are learning audio engineering? How should lectures, labs, and projects be structured to facilitate student learning in this field? How should we best deal with differences in learning among students in a class? How much should curricula change as technology and the industry changes?

Session P20 **Monday, May 19**
14:00 – 15:30 **Topaz Lounge**

POSTERS: PSYCHOACOUSTICS, PERCEPTION, AND LISTENING TESTS

14:00

P20-1 **Loss of Subjective Localization Cues in Virtual Acoustic Opening**—*Elena Blanco-Martín, Francisco Javier Casajus-Quiros, Juan Jose Gomez-Alfageme, Luis I. Ortiz-Berenguer*, Universidad Politécnica de Madrid, Madrid, Spain

The reproduced sound event quality is very important in a WFS configuration that is used for an acoustic opening. One way of checking the subjective quality perceived by a listener is the ITU R. BS.1387 "Method for objective measurements of perceived audio quality," but this method does not provide information about the listener's ability to localize the sound. A Matlab application has been implemented (SEL, Sound Event Localization) simulating an acoustic opening configuration. The number of microphones and loudspeakers in the arrays is selectable, just as the sound source position, the gap between array transducers and the listener position. This simulation has been verified against a real configuration of acoustic opening. Moreover, the loss of localization cues has been analyzed with different multichannel codifications.
Convention Paper 7443

14:00

P20-2 **Effect of Interaural Differences on Loudness of Narrowband Noise Bursts**—*Toni Hirvonen, Ville Pulkki*, Helsinki University of Technology, Espoo, Finland

This paper investigates the effects of interaural time and level differences (ITDs and ILDs, respectively) on loudness. Dichotic samples containing various amounts of interaural differences were compared to a diotic reference. The subjects adjusted the relative threshold gain of the test sample using a two-alternative, forced choice adaptive procedure (2AFC). The test signals were Gaussian noise samples with a bandwidth of one critical band and center frequencies of 150, 600, and 2400 Hz. The results imply that ILD is prominently responsible for changes in directional loudness, which is in agreement with present binaural loudness models that consider only ILD. The experiments revealed significant individual differences between subjects even when matching two identical signals.
Convention Paper 7444

14:00

P20-3 **Perception of Movements of a Focused Sound Generated with a Linear Loudspeaker Array System**—*Ichiki Manon,*[1] *Daiki Sato,*[1] *Tomoaki Tanno*,[1] *Kaoru Ashihara*,[2] *Shogo Kiryu*[1]
[1]Musashi Institute of Technology, Setagaya-ku, Tokyo, Japan
[2]National Institute of Advanced Science and Technology, Tsukuba, Japan

A special loudspeaker array system was ➥

developed for an experiment on perception of movements of a focused sound. The spatial patterns of the sound pressure level for the focused sounds were measured. The patterns were improved compared to the previous preliminary experiment using commercial devices. A psychoacoustic experiment on perception of movements of the focused sound was conducted using the developed system.
Convention Paper 7445

14:00

P20-4 Subjective Evaluation for Music Recording Positions in a Coherent Region of a Reverberant Field—*Yoshifumi Hara,*[1] *Hiroaki Nomura,*[2] *Mikio Tohyama,*[3] *Kazunori Miyoshi*[1]
[1]Kogakuin University, Shinjuku-ku, Tokyo, Japan
[2]Kure College of Technology, Kure-city, Hiroshima, Japan
[3]Waseda University, Shinjuku, Tokyo, Japan

In this paper we describe the most preferable frequency characteristics of the early reflections for music recording positions. We recorded short passages from two pieces of music (Handel, "Water Music Suite" and Brahms, "Symphony No. 4") at various distances from a sound source in a coherent region in a reverberation chamber. Subjects evaluated the preference and the subjective loudness through headphones under the diotic condition by paired comparison tests. As a result, we found that the most preferable distance indicated the distance where the loudness became maximum. The preferable recording condition could be also characterized by narrow-band envelope spectrum analysis.
Convention Paper 7446

14:00

P20-5 Efficient Individualization of HRTF Using Critical-Band-Based Spectral Cues Control—*Yoomi Hur,*[1] *Young-Cheol Park,*[2] *Dae-Hee Youn,*[1] *Seok-Pil Lee*[3]
[1]Yonsei University, Seoul, Korea
[2]Yonsei University, Wonju, Korea
[3]Korea Electronics Technology Institute, Bundang-Gu, Sungnam-Si, Korea

Recently, 3-D audio technologies have been commonly implemented through headphones. A major problem of the headphone-based 3-D audio is in-the-head localization, which occurs due to the inaccurate head-related transfer function (HRTF). Since the individual measurements of HRTFs are impractical, there have been several researches for HRTF customization. In this paper we propose an efficient method of customizing HRTFs. In the proposed method, spectral notches and envelopes are controlled based on a critical-band rate. Thus, the structure of the proposed algorithm is much simpler than that of previous methods, but still effective. The proposed method was evaluated in the problem of externalization, and the results showed that the customized HRTF using the proposed method could greatly improve the externalization performance.
Convention Paper 7447

14:00

P20-6 How to Widen the Sweet Spot in Monitoring 5.1—*Julien Bassères, Patrick Thevenot,* Taylor Made System, Nangis, France

Generally speaking, sound reproduction tends to achieve the widest sweet spot. But it's seldom realized and more than that, the restricted sweet spot has become rather usual and well accepted by the audio community. This paper proposes to find a new approach in order to get a wider sweet spot, up to a certain extent, in multichannel. By optimizing the directivity of each loudspeaker in order to compensate the position of the listener, this method aims at creating a coherent and homogeneous acoustic field. Special care will be given to the directivity pattern (amplitude and phase) of the loudspeaker system.
Convention Paper 7448

14:00

P20-7 Auditory Modeling via Frequency Warped Transforms—*Alexey Petrovsky,*[1] *Marek Parfieniuk,*[2] *Adam Borowicz,*[2] *Alexander Petrovsky*[1,2]
[1]Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus
[2]Bialystok Technical University, Bialystok, Poland

The goal of this paper is to show and compare four different versions of auditory modeling based on frequency warped transforms: bark-scaled wavelet packet decomposition, bark-scaled adapted wavelet packet decomposition, warped discrete Fourier transform, and four-band wavelets paraunitary filter bank, useful for perceptual audio coding, speech enhancement, and parametric audio coding matching pursuit procedure based on the psychoacoustic optimized wavelet packet dictionary. A practical implementation of the audio signal processing based on the given auditory modeling approaches are in details considered and analyzed from positions: depth of a compression, perceptual perception, a structural realizability, an opportunity to build embedded systems.
Convention Paper 7449

14:00

P20-8 The Role of Spectral Features in Sound Localization—*Daniela Toledo, Henrik Møller,* Aalborg University, Aalborg, Denmark

Spectral components of head-related transfer functions (HRTFs) are highly dependent on the anthropometric characteristics of subjects. In the low frequency range, a common structure is often found in HRTFs from different subjects. However, individual differences are seen at high frequencies. In binaural synthesis with non-individual HRTFs, localization errors occur if the spectral characteristics of the directional filters used do not match the individual characteristic of the listener. This investigation is focused on the spectral characteristics of HRTFs that are relevant as localization cues and how to parameterize them. This is done by cross-matching individual and non-individual HRTFs from different subjects according to the results of localization experiments.
Convention Paper 7450

14:00

P20-9 Multichannel Loudness Listening Test—*Ian Dash*,[1] *Luis Miranda*,[2] *Densil Cabrera*[2]
[1]Australian Broadcasting Corporation, Sydney, NSW, Australia
[2]University of Sydney, Sydney, NSW, Australia

As part of ongoing research for ITU Recommendation BS.1770 Algorithms to measure audio programme loudness and true-peak audio level, listening tests were conducted using a standard five-channel geometry in a standard listening room to confirm the channel gains and the spectral weightings for equal loudness contribution. Most ITU-related work to date has used broadcast program as a test signal. In this test, octave band noise was used as a test signal. Twenty-seven listeners participated. Results were analyzed for statistical consistency as well as for average and variance. Agreement between the test results and various broadband loudness models, including ITU-R BS.1770, is examined.
Convention Paper 7451

Workshop 19 Monday, May 19
14:00 – 15:30 Room L

BRINGING DYNAMICS BACK INTO THE MIX

Chair: **Ronald Prent**, Galaxy Studios

Panelists: *Thomas Lund*, TC Electronic A/S
Darcy Proper, Galaxy Studios

Pop music and the fight for level—this is an ongoing battle of even a hint of dynamics. Peak measurement and normalization has caused the fight for "who is the loudest" and with it intersample peaks well beyond 0dBFS. Severe distortion and listener fatigue is the result. The switch of a few broadcasters from peak normalization to loudness normalization is the straw on which proponents of a dynamic, lively, and "breathing" mix cling to in order to bring this topic center-stage again. Key figures from the mixing and mastering side as well as from the manufacturers and like-minded initiatives will discuss ways out of the loudness race, good practices, spreading the word, etc., and will illustrate their points through examples.

Tutorial 15 Monday, May 19
14:00 – 15:30 Room N101

SOUND ARCHIVING: STANDARDS RELATED TO THE LONG-TERM PRESERVATION OF A SOUND RECORDING

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

The stocks of sound recordings held in repositories worldwide are estimated to amount to 100 million hours. Even if we want to keep only a fraction of this legacy, logistical and financial challenges to preserve these documents of artistry, history, and cultural diversity are of significant dimensions. Original carriers, whether analog or digital, are prone to decay, and—more importantly—replay equipment becomes increasingly unavailable for those holdings that have not yet been transferred to professional digital repositories. The time slot left for this transfer is estimated to be 20 years only.

The tutorial describes the global situation and surveys standards, guidelines, and recommended practices related to principles and practical aspects of sound archiving, issued by AES, IASA (International Association of Sound and Audiovisual Archives), and UNESCO.

Historical Event
VINTAGE LOUDSPEAKERS
Monday, May 19, 14:00 – 15:00
Room N

Presenter: **Garry Margolis**

An informal review of recording studio monitoring practices in North America from the 1950s through the 1970s as seen and heard by a recording engineer who joined a loudspeaker company and learned how studio monitors really work.

Monday, May 19 14:00 Room K
Technical Committee Meeting on Automotive Audio

Master Class 2 Monday, May 19
14:30 – 16:00 Room B

CONCEPTS IN SOUND QUALITY

Presenter: **Jens Blauert**, Ruhr-Universitaet Bochum, Bochum, Germany

The lecture presents a fundamental consideration of the nature of sound quality and offers ways for structuring different aspects of it. It aims at making audio engineers more aware of the various components and complex processes involved in the formation of sound-quality judgments. Thinking of sound quality means thinking of percepts, which brings subjectivity to the fore. To deal with subjectivity scientifically, the first part of the lecture will present a line of epistemological arguments that is stringently based on actual perception and not on fiction. In this way, the place and role of subjectivity in acoustics and audio engineering will be determined. To further elucidate the formation process of sound quality, the second part of the talk starts with the introduction of the character of sounds and then moves on to different aspects of sound quality, starting from sound quality (as such) through sound-transmission quality and auditory-scene quality to product-sound quality. A generalized, system-oriented approach toward sound-quality evaluation is thereupon presented. Finally, the third part of the talk deals exemplarily with sound quality in spaces for musical performances. The question of proper references turns out to be crucial for any further analysis. Consequently, efforts are taken to explore and assess these references systematically. To this end, a hierarchical order for references, based on the amount of abstraction, is proposed and discussed in some detail. Issues like typicalness, functional adequacy, aesthetic form, and tradition are touched upon.

Historical Event
COMPACT CASSETTE
Monday, May 19, 15:00 – 16:00
Room N

Presenter: **Willem Andriessen**

The "Compact Cassette," an incredible success story with an incredible sudden end. The technical and market environment in the late fifties and early sixties, which led to its birth, its technical and product development over the years, its impact on magnetic tape development, its "explosive" marketing development, its rivals, such as DC international, Elcassette, Microcassette, DAT, and DCC, and at last its ultimate killing factor: the total integrated digitizing of private and personal audio consumption,thanks to CDR for downloading and MP3 for "ideal" portability and mobility will be discussed by Willen Andriesse, a retired tape specialist at BASF.

Exhibitor Seminar 6 — Monday, May 19
15:00 – 16:00 — Room P

TRINNOV AUDIO (FRANCE)

Presenters: **Arnaud Laborie**
Felipe Avila-Reyes

Digital Room Correction for the Control Room

The main topic of this presentation is the challenge of achieving consistent audio monitoring in the control room and throughout the production chain. Is digital room correction relevant to professional control rooms? How can it help to correct acoustical/ monitoring problems? What are its limitations? Real world examples and a short demonstration are provided.

Monday, May 19 — 15:00 — Room K
Technical Committee Meeting on Transmission and Broadcasting

Student Event/Career Development
RECORDING COMPETITION—SURROUND
Monday, May 19, 15:30 – 18:30
Room L

The Student Recording Competition is a highlight at each convention. A distinguished panel of judges participates in critiquing finalists in each of the categories in an interactive presentation during the convention. Student members can submit stereo and surround recordings in the categories classical, jazz, folk/world music, and pop/rock. Meritorious awards will be presented at the closing Student Delegate Assembly Meeting on Tuesday.

Judges:

Sound for Picture (15:30–16:30): Florian Camerer, Bernhard Maisch, Wilfried van Baelen

Classical (16:30–17:30): David Griesinger, Jared Sacks, Cornelis Van der Gragt

Non-Classical (17:30–18:30): Ronald Prent, Bosse, Ternstrom, Wilfried van Baelen

Recording competition sponsors: Schoeps, PMC, Sennheiser, Neumann

Session P21 — Monday, May 19
16:00 – 17:30 — Topaz Lounge

POSTERS: MULTICHANNEL SOUND

16:00

P21-1 Challenges in Reproduction and Evaluation of Upmixed Audio in an Automotive Environment—*Oliver Hellmuth,*[1] *Steffen Bergweiler,*[2] *Manfred Neumann,*[2] *Stefan Holzhäuser,*[2] *Andreas Walther*[1]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]Lear Corporation GmbH, Kronach, Germany

Audio systems with high quality sound reproduction capabilities are becoming more and more popular in the car. The need to create a pleasant sound field has lead to an increased number of loudspeakers combined with digital signal processing. To benefit from the advantages of surround sound reproduction also for two-channel legacy content an upmixing algorithm is required. In this paper challenges and requirements for high quality surround sound reproduction and upmixing are first introduced separately and then discussed jointly with the specific focus on the automotive environment. Finally a test method for the evaluation of different upmixing algorithms in the car is suggested.
Convention Paper 7452

16:00

P21-2 A General Approach to Methods for Loudspeaker Array Synthesis—*Juan Miguel Navarro Ruiz*, Universidad Católica San Antonio de Murcia, Guadalupe, Murcia, Spain

Loudspeaker arrays are often used as sound reinforcement in large concert halls and outdoor events to provide increased directivity. Contrary to what happens in loudspeaker systems, there is an entrenched theory in antenna array synthesis, which has been used extensively over the past few years. This paper focuses on discussing several consolidated antenna array synthesis methods. Then, a simulation software is implemented to show pros and cons of using on loudspeaker arrays. Finally, an efficient synthesis method is proposed to achieve the required characteristics.
Convention Paper 7453

16:00

P21-3 On Large Multiactuator Panels for Wave Field Synthesis Applications—*Basilio Pueo,*[1] *José Escolano,*[2] *José Javier López,*[3] *Germán Ramos*[3]
[1]University of Alicante, Alicante, Spain
[2]University of Jaén, Linares (Jaén), Spain
[3]Technical University of Valencia, Valencia, Spain

Wave Field Synthesis (WFS) is a spatial sound rendering technique that generates a true sound field using loudspeaker arrays. Multiactuator Panels (MAPs) are an alternative technology to the dynamic piston loudspeakers, based on the distribute mode operation. Because of its low visual profile and negligible vibration of the panel, MAPs are very suitable for WFS reproduction. However, the size of current prototypes does not allow its use for real immersing environments in which the loudspeaker must be integrated as walls or as projection screens. In addition, the extra area of a large panel can be used to accommodate extra exciters with which to generate sound fields at another elevation level. In this paper a very large MAP prototype is presented that has been designed and built to fulfill the requirements of immersive audio appli-

cations. It represents a step forward in the applications of MAPs for immersing scenarios. The panel size enhances its acoustic behavior in the low frequency range. Also, it can be employed for relatively large projection screens for videoconferencing and for virtual reality.
Convention Paper 7454

16:00

P21-4 Temporal Changes of Psychological Impressions Regarding Microphone Arrays for Multichannel Recording—*Toru Kamekawa, Atsushi Marui*, Tokyo University of the Arts, Adachi-ku, Tokyo, Japan

Microphone technique for surround sound recording of an orchestra is discussed. Seven types of surround microphone sets recorded in a concert hall were compared in subjective listening test on attributes such as powerfulness and spaciousness using a method inspired by MUSHRA (MUltiple Stimuli with Hidden Reference and Anchor). To minimize temporal change in music, Phase Randomized Signal (PRS) was proposed. From the average score of the listening test, the impression difference between original source and PRS was found in some microphone arrays consisting of directional microphones at some pieces. It means that the impression of these arrays depend on temporal changes in music. The data from the listening test between the original source and PRS showed that impressions of powerfulness had slightly higher correlation. The relations of the physical factors of each array were also compared, such as SC (Spectral Centroid), LFC (Lateral Fraction Coefficient), and S/O (Side/Omni Ratio) of each array. The correlation of these physical factors and the attribute scores show that the contribution of these physical factors depends on music and its temporal change.
Convention Paper 7455

Tutorial 16 **Monday, May 19**
16:00 – 18:00 **Room N101**

QUANTIZATION EFFECTS IN AUDIO SIGNAL PROCESSING

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

In this tutorial we will first present the principles of quantization and the basic effects of finite precision in digital filter implementations (both fixed and floating point) on the audio signal. We explain the effects of finite precision on frequency response (using an audio equalization task as an example). We shall see how different forms of filter structures offer advantages when finite word length is considered. We shall also look at how different structures affect the audio signal. Finally we will discuss the effect of different types of dither and the various pitfalls that can occur.

Tutorial 17 **Monday, May 19**
16:00 – 18:00 **Room E-F**

MUSIC-INDUCED HEARING DISORDERS

Presenter: **Jan de Laat**, LUMC, Leiden, The Netherlands

Frequent exposure to loud sounds often causes hearing damage. Recent studies in Sweden demonstrated that 74% of all musicians complain of hearing disorders such as: hearing loss, tinnitus (whistling or noise in the ear), hyperacusis (suffering from sounds that are too loud), distortion (attentive) and diplacusis (perception of different pitches on the left and right sides). It appears that the sound volume of concerts has gradually become louder and louder over the last decades, and this corresponds to the extent of hearing problems in musicians, even attended by unfitness for work.

In The Netherlands these troubles are becoming recognized little by little. A few years ago the government (Department of Social Affairs) established an agreement with the employers (administrations of professional symphony orchestras) and employees (union of musicians) aimed to manage these problems. Attention and consideration will be focused on awareness, professional information (even at early ages, starting at music colleges), prevention of and protection from hearing loss (not only by hearing protectors such as personal ear moulds), and rehabilitation of hearing handicaps in musicians.

In a research project with 259 musicians of three professional symphony orchestras in Amsterdam, we combined results of the outcomes of questionnaires, pure tone audiograms, speech audiometry, transient and distortion-product oto-acoustic emissions, tinnitus matching, and frequency selectivity measures in order to achieve more insight in the individual hearing sensitivity related to the time of exposure to loud music. Remarkable results of this investigation will be discussed.

Monday, May 19 **16:00** **Room K**
Technical Committee Meeting on Perception and Subjective Evaluation of Audio

Workshop 20 **Monday, May 19**
16:30 – 18:00 **Room B**

WOW, THIS SOUNDS GREAT!

Chair: **Christian Hugonnet,** Consultant

Panelists: *Morten Lindberg*, 2L Oslo, Oslo, Norway
George Massenburg, GML Inc.
Eirik Solheim, NRK – Norwegian Broadcasting
Helmut Voitl, Director

What parameters do we have and do/will our children have to say this in the age of MP3, Skype, and general data-reduction? Does high-quality audio have a future? Will new high-resolution music-only formats ever be a commercial success? Is generally the sensitivity for intricacies and detail, for dynamics and resolution deteriorating? Do social developments have an impact on this situation? What is the effect of increasing environmental acoustic pollution? What are our criteria for “quality”?

These and other questions will be discussed by a panel of experts with an acute sensitivity for those changes and a fascinating insight into this development.

➥

Special Event
NEW REVENUE STREAMS FOR STUDIOS
Monday, May 19, 16:30 – 17:30
Room O

In this special event, the European Sound Directors' Association, ESDA in collaboration with the APRS have invited a distinguished panel of prominent European studio owners and producers including Robin Millar, Malcolm Atkin, and Dave Harries from the UK; Eric van Tijn and Chris Pilgram from the Netherlands; and others to discuss the potential opportunities to establish new business models and revenue streams for the recording communities.

Studios

Times are hard for studios and producers and, for that matter, professional equipment manufacturers. We all wallow toward the bottom of the music business food chain suffering the rebound from the commercial uncertainties that are attacking the record companies and the music business in general. Is the time right for studios to become participants in the royalty structure? What possibilities are there for new studio-based services and other diversification such as digital distribution, video recording, DVD authoring and extended deliverables.

Producers

Producers have long waged a campaign to share in performance royalties. Progress in the UK, a few EU territories, and the USA is encouraging, however, the treatment of producers is inconsistent and needs to be harmonized. Chairman of ESDA, Peter Filleul, will report on progress across the territories.

Monday, May 19 17:00 Room K
Technical Committee Meeting on Semantc Audio Analysis

Special Event
BANQUET
Monday, May 19, 20:00 – 23:30
Kompaszaal

The traditional AES banquet offers an excellent opportunity to meet and/or invite your colleagues, (potential) customers, officials, and all who are involved in audio. This year's event will take place in the so-called Kompaszaal, an Arrival/Departure hall of the KNSM, a shipping company in the east harbor region of Amsterdam. This hall was built in 1956 under the supervision of architect Johan van Tienhoven and radiates the style of the fifties. The company no longer exists. If the weather cooperates, we have the possibility to have a wide look over the water while drinking our aperitifs. Here we also would like to remember the 60 years existence of the AES!

75 Euros for AES members; 80 Euros for nonmembers
Tickets will be available at the Special Events desk.

Session P22 Tuesday, May 20
09:00 – 11:00 Room E-F

AUDIO ARCHIVING, STORAGE, RESTORATION, AND CONTENT MANAGEMENT

Chair: Tin Jonker, NOB, Hilversum, The Netherlands

09:00

P22-1 Manufacturing Recordings from 100-Year-Old Masters—*Sean Davies*,[1] *Rinus Hooning*[2]
[1]S.W. Davies Ltd., Aylesbury, UK
[2]Record Industry bv, Haarlem, The Netherlands

Most work on the 78 rpm analog recording format concentrates on pressings made near to the time of the recording and the best ways to retrieve the information from these for future storage and reproduction. However, a considerable number of metal master plates have been preserved from the earliest days to the end of the format's active period. This paper describes a project to manufacture new pressings from the original plates, the reasons for doing so, and the technical challenges involved.
Convention Paper 7456

09:30

P22-2 Replay of Digital Original Tapes: Practical Experiences with Video Tape Based PCM Adapters and R-DAT—*Nadja Wallaszkovits*,[1] *Heinrich Pichler*,[2] *Johannes Spitzbart*[1]
[1]Austrian Academy of Sciences, Vienna, Austria
[2]Audio Consultant, Vienna, Austria

As many of the early digital formats are already obsolete and support of these formats cannot be guaranteed much longer by the manufacturers, archives should presently give priority to the replay of original recordings on such material. Based on a short theoretical discussion and outlining the format-specific characteristics, the paper discusses a variety of practical problems of signal retrieval from PCM (Pulse Code Modulation) encoded signals on a VTR (video tape recorder) and R-DAT (Rotary-Head Digital Audio Tape), such as mechanical problems, tracking problems and playback incompatibilities, data integrity checking, extraction and incompatibility of sub-code-information, pre-emphasis, as well as other problems occurring from irregular recording conditions (typically with field recordings produced on portable devices) or format peculiarity
Convention Paper 7457

10:00

P22-3 A New System for File-Based Audio Recording, Preservation, and Access at Indiana University's Jacobs School of Music—*Konrad Strauss, Travis Gregg*, Indiana University, Bloomington, IN, USA

The Indiana University Jacobs School of Music has been making live concert recordings on a variety of formats since the 1940s, and we continue to record approximately 500 concerts each year. Recent industry trends and changes in technology have led us to investigate the possibility of creating high-resolution digital files rather than continuing to use physical media as the archival format for our recordings. Our goal was to develop a system for the creation, access, and long-term preservation of high-resolution audio recordings and associated metadata that conformed to emerging standards for digital audio preservation. We began building such a system in July of 2006 and reached full imple-

mentation in February of 2007. This paper gives an overview of the development process, presents hardware and software solutions, and discusses workflow and data management issues.
Convention Paper 7458

10:30

P22-4 A Fast Feature Extraction System on Compressed Audio Data—*Tobias Friedrich, Matthias Gruhne, Gerald Schuller*, Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany

We describe an efficient system that directly extracts features from compressed audio material. It consists of a time/frequency conversion method and a feature extraction algorithm. The conversion method provides the feature extraction algorithm with a suitable complex spectral representation directly from the compressed domain. It further allows a trade-off between computational complexity and conversion accuracy. Several operating points using different conversion accuracies were tested with an MPEG audio identification system in order to evaluate the identification confidence. Based on these results it is possible to reduce the computational complexity from O(N log N) to O(N) compared to the conventional approach (complete decoding followed by a frequency analysis).
Convention Paper 7459

Workshop 21 **Tuesday, May 20**
09:00 – 10:30 **Room N101**

JOAN OF ARC [CANCELLED]

Session P23 **Tuesday, May 20**
09:30 – 12:00 **Room C-D**

PSYCHOACOUSTICS, PERCEPTION, AND LISTENING TESTS, PART 2

Chair: **Joerg Bitzer**, University of Applied Science Oldenburg, Oldenburg, Germany

09:30

P23-1 A Proposed Audio Visual Product Measure—*Joe Peters*, National University of Singapore, Singapore

The Multimedia Section at the Centre for Instructional Technology at the National University of Singapore has developed an audio visual assessment index (AVAI) to serve as a tool for clients to measure their evaluation of audio and video products. AVAI is based on a listing of indicators and variables that make up the fundamental elements in the capture and processing of AV products (video production): image, color, light, audio, form, aesthetics, and delivery. AVAI is currently being used by professionals for internal evaluation. A series of simulator-based AVAI courses are also underway, the purpose of which is to enable lay persons to understand the indicators and variables through simulated explanations. The thesis is, that in order to keep product value high the information gap between the producers and the lay clients must be narrowed. The sub-set of this thesis is that this narrowing can be achieved through even a singular simulator training session. What is presented in this paper is the conceptual framework and some preliminary tests. The tests are not substantial as studies are slow. AVAI is not a core area of the work of the Multimedia Section handling this study. Nevertheless, it is important to have some response from AES on this preliminary presentation.
Convention Paper 7460
[This paper was not presented but may be purchased]

10:00

P23-2 Nonexistence of Frontal Signal Unmasking from Spatially Wide Masker—*Ville Pulkki, Jukka Ahonen*, Helsinki University of Technology, Espoo, Finland

The masking of a frontal signal by spatially wide noise sources was investigated in a listening experiment. The noise sources consisted of a single or multiple symmetrically positioned loudspeakers in the frontal horizontal plane in anechoic conditions. It is shown that the detection threshold of the signal does not depend on masker width, which suggests that frontal unmasking does not exist in loudspeaker listening. In additional tests with signal source positioned in side it is shown that moderately small binaural unmasking occurs in that case from wide masker, and that increasing the width of masker source decreases binaural unmasking effect.
Convention Paper 7461

10:30

P23-3 Reaction Times and Performances in Recognition Tasks to Assess Speech Quality—*Virginie Durin,*[1] *Laetitia Gros*,[1] *Gilles Hericher*[2]
[1]Orange Labs, Lannion, France
[2]Laboraoire Psychologie et Neurosciences de la cognition, Mont Saint Aignan, France

This paper deals with perceptive test methodologies to assess speech quality of telecommunication systems. Faced with drawbacks of typical methodologies recommended by ITU-T, a new way to assess speech quality is investigated, by collecting reaction times and performances when subjects are achieving tasks involving degraded speech signals. A duel task with a digit recognition memory task and a letter recognition task is proposed. Three different quality levels are applied to audio signals describing digits and letters. The results show significant differences of performances and reaction times between the three quality levels.
Convention Paper 7490

11:00

P23-4 Evaluation of Stereophonic Images with Listening Tests and Model Simulations—*Munhum Park,*[1] *Philip Nelson*,[1] *Kyeong Ok Kang*[2]
[1]University of Southampton, Highfield, Southampton, UK
[2]Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea

➥

A binaural hearing model has recently been suggested for the evaluation of the performance of virtual acoustic imaging systems. The model considers excitation-inhibition (EI) cell activity patterns as the internal representation of sound localization cues and a pattern-matching procedure with a frequency-weighting scheme produces the estimate of source location in the horizontal plane. Given the reasonable prediction of some important features in human sound localization and lateralization, this paper presents a further verification and application of the model in actual listening tests. In this paper participants' responses to stereophonic images have been compared with the predictions of the model, individually established from the subject's own HRTF. Model predictions have been found to be both qualitatively and quantitatively consistent with the test results, and in particular, the agreement between 2 and 3 kHz gave a good indication that, unlike some similar models, the current model can effectively incorporate both ITD and ILD information according to their relative importance.
Convention Paper 7463

11:30

P23-5 The Sound Character Space of Spectrally Distorted Telephone Speech and its Impact on Quality—*Marcel Wältermann, Alexander Raake, Sebastian Möller*, Berlin University of Technology, Berlin, Germany

Spectral distortions of speech transmitted over a telephone channel may stem from linear channel filtering, codecs, electro-acoustic properties of end-user terminals, or the acoustic environment at send side. In this paper a study is presented that aims at revealing the perceptual space of spectrally distorted telephone speech and establishing a link to the overall quality of the speech. Two dimensions were identified as relevant for explaining the perceived quality: indirectness and brightness. Whereas brightness is related to the center frequency of a transfer function, indirectness is correlated with the equivalent rectangular bandwidth and constitutes the dominating factor in the perceptual space in terms of covered variance. The concept of the bandwidth impairment factor that fits into the framework of the so-called E-Model and that is based on these simple parameters for computing the integral quality of spectrally distorted speech could successfully be applied to the given data.
Convention Paper 7464

Tuesday, May 20 09:00 Room H
Standards Committee Meeting: AESSC Plenary

Session P24 Tuesday, May 20
09:30 – 11:00 Topaz Lounge

POSTERS: ROOM AND ARCHITECTURAL ACOUSTICS AND SOUND REINFORCEMENT

09:30

P24-1 Objective Evaluation of a Non-Environmental Control Room for 5.1 Surround Listening—*Soledad Torres-Guijarro,*[1] *Antonio Pena,*[2] *Norberto Degara-Quintela*[2]
[1]Laboratorio Oficial de Meroloxía de Galicia (LOMG), Tecnópole, Ourense, Spain
[2]Universidad de Vigo, Vigo, Spain

The control room of the Universidad de Vigo was built for the purpose of assessing small audio artifacts, such as listening analytically to coded material with different data rates. It follows a non-environment design that minimizes the influence of the room. The use of such a room as a 5.1 surround listening room will be analyzed according to international recommendations. This research includes the study of the electroacoustic behavior of loudspeakers, geometric and acoustic properties of the room, and sound field conditions. A discussion of some divergences and implications for its use when performing surround listening tests follows the measurement results.
Convention Paper 7465

09:30

P24-2 A Case Study of Sound Reproduction and Acoustic Enhancement in Concert Halls Using Wave Field Synthesis—*Clemens Kuhn-Rahloff, Matthias Rosenthal, Max Casdorff, Roger Moser*, sonic emotion ag, Obergltt (Zurich), Switzerland

This paper presents the wave fields synthesis system under construction at the National Conservatory of Music Detmold (Germany). The system is dedicated to sound reproduction for artistic purposes at the Tonmeister department (Erich Thienhaus Institute) and to an enhancement of room acoustics. The system comprises 346 independent loudspeaker channels, including a horizontal loudspeaker array all around the auditory (500 seats) and ceiling loudspeakers. Since the hall is used for a broad repertoire comprising chamber music, romantic orchestra instrumentations, organ concerts, contemporary music, etc., the hall will be equipped with a variable room acoustic system. The paper presents perceptual aspects of system design concerning the direct sound and diffuse field as well as practical implementations for WFS rendering.
Convention Paper 7466

09:30

P24-3 Small Studios with Gypsum Board Sound Insulation: A Review of Their Room Acoustics, Details at the Low Frequencies—*Lorenzo Rizzi, Francesco Nastasi*, Rizzi Acustica, Lecco LC, Italy

At the present time most music preproduction

and production is often carried out in very small, privately owned rooms, which are called "project studios." Gypsum board technology is very common in the construction of these rooms because of its high insulation capabilities compared to low monetary and time costs. The paper discusses sweet spot impulse response measurements that have been carried out in three different but acoustically small rooms built with gypsum board sound insulating structures comparing it to a masonry built one. The room modal behavior is underlined, continuing with the analysis of decaying in time at low frequencies related to insights on perception and analysis. A different methodology of study is proposed.
Convention Paper 7467

09:30

P24-4 On the Measurement of Electro Acoustic Enhanced Sound Fields—*Florian Walter*,[1] *Frank Melchior*[2]
[1]Mueller-BBM GmbH, Munich, Germany
[2]Fraunhofer Institute for Digital Media Technology, Ilmenau, Germany

The installation and optimization of acoustic enhancement systems require a large amount of experience. The verification in terms of measurement is most of the time done using conventional reverberation and acoustic parameter measurements according to ISO 3382. This is a good solution for diffuse sound analysis and general examination of early reflections, but in terms of direction dependant analysis the results are not satisfying. In this paper a room equipped with an acoustic enhancement system was measured using a circular array. The effects of adding specific early reflections and direction-dependent diffuse energy generated by the acoustic enhancement system are investigated. The results are compared to standard measurements according to ISO 3382.
Convention Paper 7468

09:30

P24-5 Applying Cochlear Modeling and Psychoacoustics in Room Acoustics—*Jasper van Dorp Schuitman, Diemer de Vries*, Delft University of Technology, Delft, The Netherlands

The acoustical qualities of a concert hall or any other room are generally expressed using acoustical parameters. These parameters are determined from impulse responses, as measured from single positions in a room or along a line array. However, from array measurements it turned out that parameters can fluctuate severely between small distance steps, something which does not agree with human perception. Applying cochlear modeling and psychoacoustics in this process seems a promising technique to reach results that do not suffer from these fluctuations and thus are much closer to human perception compared with conventional techniques.
Convention Paper 7469

09:30

P24-6 Empirical Evaluation of the Frequency-Dependent Boundary Conditions in a Digital Waveguide Mesh—*José Escolano*,[1] *Basilio Pueo*,[2] *José J. López,*[3] *Máximo Cobos*[3]
[1]University of Jaén, Linares, Jaén, Spain
[2]University of Alicante, Alicante, Spain
[3]Technical University of Valencia, Valencia, Spain

The digital waveguide mesh is a popular method for time domain acoustic system simulation such as room acoustics. One of the main reasons to choose this paradigm relies in the ease to include boundary conditions in the simulation. This paper is focused on the comparison of the simulation with real-world measurements, where a particular scenario is physically built and the corresponding simulation, according to their physical parameters, is carried out. The main scope of this paper is the validation and discussion of a boundary condition model and their correspondence with the measurements through an example.
Convention Paper 7470

09:30

P24-7 Subjective Effects of Dispersion in the Simulation of Room Acoustics Using Digital Waveguide Mesh—*Jose J. Lopez*,[1] *José Escolano*,[2] *Maximo Cobos*,[1] *Basilio Pueo*[3]
[1]Technical University of Valencia, Valencia, Spain
[2]University of Jaén, Linares (Jaén), Spain
[3]University of Alicante, Alicante, Spain

The simulation of room acoustics using the Digital Waveguide Mesh method has gained interest in the last few years. One of the problems of this method is the frequency and angle dependent dispersion. In order to reduce this effect, an oversampling is usually employed but at the cost of highly increasing the resulting computational cost and restricting the simulation to lower frequencies. In this paper a subjective analysis is carried out; where different oversampling factors in voice band simulations have been performed and evaluated by a set of listeners. Some listening tests employing ABX methodology have been used to evaluate the subjective effects, obtaining some preliminary results that, although not being conclusive, they represent a first approach to the problem.
Convention Paper 7471

Workshop 22 — Tuesday, May 20
09:30 – 11:00 — Room B

SOS—SAVE OUR SPECTRUM!

Chair: **Malcolm Johnson**, Institute of Broadcast Sound, UK

Panelists: *Matthias Fehr*, Sennheiser – Initiative Digital Dividend
Jürgen Hanelt, VDT Germany
Jan Outters, IRT Mnich

The sell-out of the UHF radio spectrum poses a serious threat not only to the broadcast industry but on most other audio fields as well. Coordinated international action is required in a situation where time is running out. The legal situation will be examined and possible and necessary steps will be discussed. Current activities will be presented, hoping to raise the importance of immediate action to still counteract this dramatic development.

Workshop 23 **Tuesday, May 20**
09:30 – 10:00 **Room P**

MUSICAL TELEPRESENCE WITH NARROW BAND NETWORKS

Chair: **Alexander Carôt**

Panelists: *Ulrich Kramer*
Gerald Schuller
Christian Werner

In this presentation Alexander Carôt will demonstrate a live music performance with professional musicians in Germany. He will play bass guitar live and in real time over the Internet as if in the same room. In that context he will use his Soundjack software in combination with the ULD codec. While Soundjack offers the appropriate low delay audio streaming conditions, the ULD codec (developped at Fraunhofer IDMT) introduces a total encoding/decoding delay of only 6 ms and provides decent bit rates for the use in narrow band networks.

Historical Committee Meeting

The **Historical Committee meeting** will be held Tuesday, May 20, at 11:00 in Room N.

Tutorial 18 **Tuesday, May 20**
10:30 – 12:30 **Room L**

WHAT IS THE MATRIX?

Presenters: **Geoff Martin**, Bang & Olufsen a/s, Struer, Denmark
Helmut Wittek, Schoeps Mikrofone GmbH, Karlsruhe, Germany

This tutorial will be an introduction to the use of matrixing in microphone and mixing techniques for stereo and multichannel. The basics of microphone polar patterns will be explained, followed by the fundamentals of techniques such as "textbook" M/S, double M/S, and soundfield recording. Included in the tutorial will be a discussion of how to use matrixing to "re-mix" an existing recording, to modify microphone configurations in post-production, and to manipulate spatial characteristics of stereo mixes. In addition, information on the exciting possibilities in the fast-paced world of karaoke soundtrack production will be presented.

Workshop 24 **Tuesday, May 20**
11:00 – 13:00 **Room N101**

FINDING AND EVALUATING MUSIC FOR DOWNLOADING OR STREAMING

Chair: **Steve Harris**, BridgeCo AG

Panelists: *Karlheinz Brandenbur*g, Fraunhofer Institute
James Cridland, BBC Audio and Music Interactive
Jurgen Jaron, mufin GmbH
Liz Rice, Last.fm

The continuing rapid evolution of the online digital music scene has meant that a wide variety of independent music sources now exist, many catering to indie artists or specific genres or niches, and often with individualized business models and distribution methods. In the process, the exploration, discovery, and critiquing of music has become decidedly more complex, if potentially more interesting, for the consumer. Online music sources range from massive data bases, some available only by membership or subscription and tied to hardware or ISPs; to social networking sites; to home websites of labels, bands, orchestras, artists, and radio stations; to blogs; a/v streaming; internet radio, among many others. Methods to adequately search out music among all of these, to judge quality, and to find the bit rates, formats, and software compatible with one's playback system are still evolving. This workshop looks at current approaches to the search and evaluation problem, including algorithmic methods, the exchange of opinion via social networks, online reviews, etc., and also considers whether online reviews, blogs, and chat rooms are adequate as the new determinants of success and quality.

Session P25 **Tuesday, May 20**
11:30 – 16:00 **Room E-F**

MULTICHANNEL SOUND

Chair: **Günther Theile**, Institut für Rundfunktechnik, Munich, Germany

11:30

P25-1 Bitstream Format for Spatio-Temporal Wave Field Coder—*Francisco Pinto, Martin Vetterli*, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

We present a non-parametric method for compressing multichannel audio data for reproduction through Wave Field Synthesis. The method consists of applying a two-dimensional filterbank to the input multichannel signal, in both time and channel dimensions, and coding the two-dimensional spectra using a spatio-temporal frequency masking model. The coded spectral data is organized into a bitstream together with side information containing scale factors and Huffman codebook information. We demonstrate how this coding method can be applied to any smooth distribution of loudspeakers in space, while obtaining a stable bit rate that is 15% lower compared to coding each channel independently
Convention Paper 7472

12:00

P25-2 The Design of Ambisonic Decoders for the ITU 5.1 Layout with Even Performance Characteristics— *David Moore, Jonathan Wakefield*, University of Huddersfield, Huddersfield, West Yorkshire, UK

All previously published Ambisonic decoders for irregular loudspeaker layouts have localization

performance that varies significantly by angle around the listener. This contrasts with decoders designed for evenly spaced arrangements of loudspeakers where performance characteristics are isotropic. Furthermore, even localization performance around the listener is desirable for a number of application areas of 5.1 surround sound. New decoder design criteria are presented that aim to reduce this variation in localization performance. These criteria are added to a multi-objective fitness function, based on auditory localization theory, which guides a heuristic search algorithm to derive decoder parameter sets for the ITU5.1 layout. The derived decoders exhibit a significant improvement in localization performance variation by angle around the 360-degree sound stage.
Convention Paper 7473

12:30

P25-3 Methods for Sharing Stereo and Multichannel Recordings among Planetariums—*Leslie Gaston, Peter Dougall, Erick D. Thompson*, University of Colorado at Denver, Denver, CO, USA

There is a demand for research on the transferability of surround sound audio from one planetarium to another, so that (1) audiences have similar experiences and (2) audio engineers can easily create this experience. This paper will consider: acoustics, production, delivery, equipment, and seating arrangements. Our recent survey of over 100 planetariums worldwide in the fall of 2007 will provide a look at current practices. The University of Colorado Denver and Gates Planetarium have collaborated in order to explore the potential of current audio technology, and to discover what similarities and differences exist between planetariums in order to achieve this goal of transferability.
Convention Paper 7474

13:00

P25-4 Optimal Hierarchical Bandwidth Limitation of Surround Sound—*Yu Jiao, Slawomir Zielinski, Francis Rumsey*, University of Surrey, Guildford, Surrey, UK

In order to save the transmission bandwidth of surround sound, a technique named Hierarchical Bandwidth Limitation (HBL) was proposed by the authors. In HBL, a psychoacoustically hierarchical transform is used as the preprocessing algorithm prior to bandwidth limitation. In our former experiments we found that the Karhunen-Lòeve transform (KLT) is a suitable hierarchical transform for HBL. Besides the hierarchical transform, the choice of an appropriate strategy for bandwidth allocation is also essential from the point of view of the resultant audio quality. In order to find the optimal bandwidth allocation strategy that achieves the best audio quality, the authors attempted to build up the mathematical relationship between audio quality and the bandwidth allocation strategy using a MUSHRA listening test. The experiment design and results of this listening test are reported in this paper.
Convention Paper 7475

13:30

P25-5 Frequency-Dependent Signal-Correlation in Surround- and Stereo-Microphone Systems and the Blumlein-Pfanzagl-Triple (BPT)—*Edwin Pfanzagl-Cardone*,[1] *Robert Höldrich*[2]
[1]Salzburg Festival, Salzburg, Austria
[2]Institute of Electronic Music and Acoustics, Graz, Austria

With the aim to recreate the original concert-hall sound field as faithfully as possible in the control- or living-room, recordings were made simultaneously with an artificial head and several surround microphone techniques (among them the new BPT method). The surround recordings were rerecorded using the same dummy-head as in the concert hall. The results of subjective listening tests (loudspeaker as well as binaural) were assessed using ANOVA and correlation analysis. Acoustical analysis of the dummy-head recordings was performed by measuring the Frequency-Dependent Inter Aural Cross-Correlation Coefficient (FIACC): the low-correlation AB-PC microphone system was capable of reproducing the original sound field better than any of the other systems under test (DECCA, KFM, OCT). A microphone systems Critical Frequency, below which correlation raises toward 1, is defined.
Convention Paper 7476

14:00

P25-6 Holographic Design of Source Array for Achieving a Desired Sound Field—*Wan-Ho Cho,*[1] *Jeong-Guon Ih*,[1] *Marinus M. Boone*[2]
[1]Korea Advance Institute of Science and Technology (KAIST), Daejeon, Korea
[2]Delft University of Technology, Delft, The Netherlands

For realizing a desired complicated sound field, an acoustic source array should be designed appropriately to obtain the acoustic source parameters. To this end, we suggest a method utilizing the acoustical holography technique based on the inverse boundary element method. Acoustical analogy between the problems of source reconstruction and source design was the initial motivation of the study. In the design of the source array, the pressure distribution at specific field points is the constraint of the problem and the signal distribution at the source surface points is the object function of the problem. The whole procedure of the application consists of three stages. First, a condition of the desired sound field should be set as the constraint. Second, the geometry and boundary condition of the source array system and the target field, i.e., points in the sound field of concern, are modeled by the boundary elements. Actual characteristics of source and space can be considered to generate the accurate condition of the target field. Finally, the source parameters are inversely calculated by the backward projection. As an example, a source array to fulfill the plane wave propagating zone and another quiet zone near the propagation zone was designed and tested by simulation and measurement.
Convention Paper 7477

14:30

P25-7 New Dimensions for Ambisonics—*Michael Chapman,* Culoz, France

Both two-dimensional (pantophonic) and three-dimensional (periphonic) representations of soundfields are common place in ambisonics. Reproducing either on rigs essentially designed for the other is common place. What though if one synthesizes a four (or more) dimensional soundfield and reproduces this on a standard rig? As there appears to be no source on hyper-spherical harmonics applicable to ambisonics, the mathematical basis is first set out. The manipulation of hyperambisonic soundfields (rotation, mirroring, dominance) is then discussed. During that discussion various "proofs" are advanced as to the finite range of transformations that can be applied to ambisonic soundfields, of whatever dimension.
Convention Paper 7478

15:00

P25-8 Improving Spherical Microphone Arrays—*Nicolas Epain, Jérome Daniel*, France Télécom R&D, Lannion, France

Spherical microphone arrays are useful for numerous applications, such as spatial audio capture and beamforming. However, these sensor arrays are known to have a limited frequency range, due to poor directivity at low frequencies and spatial aliasing at high frequencies. In this paper we study two methods aiming at enhancing the frequency range of spherical microphone arrays without using more sensors. First, the benefit of locating the sensors at the end of cavities within the sphere is assessed through measurements and simulations. Second, we study the influence of using large membrane microphones. Finally, results show that the frequency range could be increased in both cases studied.
Convention Paper 7479

15:30

P25-9 Migration of 5.0 Multichannel Microphone Array Design to Higher Order MMAD (6.0, 7.0, and 8.0) with or without the Inter-Format Compatibility Criteria—*Michael Williams*, Sounds of Scotland, Paris, France

The severe limitations of the 5.0 Multichannel Reproduction Standard in reproducing good quality audio-visual or stand-alone audio surround sound reproduction has increased the pressure on recording and reproduction system designers to increase the number of channels in an attempt to give an even more satisfactory envelopment experience. This paper extends the MMAD process to show how higher order channel array designs (6.0, 7.0, and 8.0) can be developed from the existing data on 4.0 or 5.0 Multichannel Front Sound Stage Coverage Array Designs with almost perfectly seamless and linear surround sound reproduction. Designing for inter-format compatibility can also be accommodated from the existing multi-format array design data described in a previous paper on Multichannel Arrays Generating Inter-format Compatibility (MAGIC arrays).
Convention Paper 7480

Session P26 **Tuesday, May 20**
11:30 – 13:00 **Topaz Lounge**

POSTERS: LOW BIT-RATE CODING

11:30

P26-1 Autoregressive Modeling of Hilbert Envelopes for Wide-Band Audio Coding—*Sriram Ganapathy,*[1,2] *Petr Motlicek,*[1] *Hynek Hermansky,*[1,2] *Harinath Garudadri*[3]
[1]IDIAP Research Institute, Matigny, Switzerland
[2]Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[3]Qualcomm Inc., San Diego, CA, USA

Frequency Domain Linear Prediction (FDLP) represents the technique for approximating temporal envelopes of a signal using autoregressive models. In this paper we propose a wide-band audio coding system exploiting FDLP. Specifically, FDLP is applied on critically sampled sub-bands to model the Hilbert envelopes. The residual of the linear prediction forms the Hilbert carrier, which is transmitted along with the envelope parameters. This process is reversed at the decoder to reconstruct the signal. In the objective and subjective quality evaluations, the FDLP-based audio codec at 66 kbps provides competitive results compared to the state-of-art codecs at similar bit-rates.
Convention Paper 7481

11:30

P26-2 On Locality of Spectral Oriented Tree for Bit-Plane Based Low-Bit Rate Audio Coding—*Yu-Lin Wang, Alvin W. Y. Su*, National Cheng-Kung University, Tainan, Taiwan

For Spectral Oriented Trees (SOT) based coders such as SPIHT and CEIHT, locality is usually related to the locations of coefficients within a SOT and its effect to coding efficiency. How to construct a SOT to achieve better locality is very important. This paper presents a diagnostic aspect of the localities of different ordering techniques for low bit-rate audio coding. We used several coefficient ordering schemes to construct SOTs with the same set of MDCT coefficients and observed their effects. Both objective and subjective results are presented.
Convention Paper 7482

11:30

P26-3 Perceptual Matching Pursuit for Audio Coding— *Hossein Najaf-Zadeh, Ramin Pichevar, Hassan Lahdili, Louis Thibault*, Communications Research Centre Canada, Ottawa, Ontario, Canada

This paper introduces a Perceptual Matching Pursuit (PMP) algorithm for audio coding. A masking model has been developed and integrated into the matching pursuit algorithm to account for the characteristics of the hearing system. By doing so, only an audible kernel is extracted at each iteration. Moreover, contrary to

the matching pursuit algorithm, PMP will stop decomposing an audio signal once there is no audible part left in the residual. We have used ITU_R PEAQ to compare audio materials decomposed by PMP and by matching pursuit. Objective scores for PMP increase by up to 1 unit. A semi-formal listening test has verified the objective scores and shown the perceptual superiority of PMP over the matching pursuit algorithm.
Convention Paper 7483

11:30

P26-4 A Unifying Approach to Transform and Sinusoidal Coding of Audio—*Maciej Bartkowiak*, Poznan University of Technology, Poznan, Poland

The paper describes a new scenario for low bit rate audio compression that combines two classical techniques: transform coding and sinusoidal coding into a united framework. The main idea is to adaptively decompose the audio signal into subbands whose central frequencies follow continuously the local instantaneous frequencies of certain signal components (formants or individual harmonic partials). The content in each subband is encoded in the baseband after frequency shift toward DC. The technique may be considered either as modified transform coding, i.e., coding along instantaneous frequencies or as extended sinusoidal coding, i.e., modeling with partial envelopes that are represented by transform coefficients. In other words, it is a hybrid scheme offering a continuous operating mode between purely transform and purely sinusoidal compression.
Convention Paper 7484

11:30

P26-5 Low Bit Rate Audio Coding for Digital Wireless Systems—*Stephen Wray*, APT Ltd., Belfast, Northern Ireland, UK

With the transition from analog to digital television, the available spectrum for wireless microphones, in-ear monitors, and other wireless devices could be under threat. Spectrum is a valuable commodity, and it is the responsibility of governments to manage it appropriately. Much has been made recently of the Spectrum Squeeze both sides of the Atlantic with discussions on White Spaces and the Digital Dividend. With bandwidth at such a premium the audio industry has been forced to consider new technologies that make efficient use of spectrum without sacrificing quality or service. Within this context, we need a new revolutionary approach to maximizing bandwidth efficiency. The author will present a new and novel coding solution to overcome the prevailing technical limitations and industry requirements for wireless applications.
Convention Paper 7485

11:30

P26-6 Bit Allocation for Linear Prediction Coefficients with Application to Lossless Audio Compression —*Florin Ghido, Ioan Tabus*, Tampere University of Technology, Tampere, Finland

We propose a novel technique of using bit allocation for linear prediction coefficients in asymmetric lossless audio compression. We show how to determine the optimal bit allocation using a new closed-form formula for the excess error from quantization, and describe a recently introduced algorithm (Optimization-Quantization Least Squares), which computes the optimal quantized prediction coefficients applied for the allocation. The proposed method, implemented as a modified asymmetrical OptimFROG, obtains small (but consistent) signal dependent compression improvements with virtually no decoder complexity increase (on a 847 MB audio corpus, up to 0.27%, on average 0.06%). Compared to MPEG-4 ALS, it obtained 0.38% better compression, while being at the same time approximately 5 times faster at decoding.
Convention Paper 7486

11:30

P26-7 Design of Framing in MPEG Surround Based on Dynamic Programming Algorithm—*Chi-Min Liu, Chung-Han Yang, Han-Wen Hsu, Wen-Chieh Lee*, National Chiao Tung University, Hsinchu, Taiwan

MPEG Surround (MPS) defined by ISO/IEC is the audio coding standard of multichannel signals based on the down-mixed signal and the spatial parameters. In MPEG Surround, the time-frequency tiles decide the units to share the same spatial parameters among the multichannel signals. Hence, the decision of the tiles is the critical module deciding the required quality and bits. However, the large number of combination in the time regions, frequency bands, and multichannel signal statistics has spanned the huge search space for deciding the tiles. Our previous work at AES 119 has proposed the dynamic programming method to efficiently decide the time-frequency units for the parameter stereo coding in HE-AAC. This paper will extend the dynamic programming method to the MPS coding.
Convention Paper 7487

11:30

P26-8 New Enhancements to the Audio Bandwidth Extension Toolkit (ABET)—*Harinarayanan E. V.,*[1] *Raghuram Annadana*,[1] *Deepen Sinha*,[2] *Anibal Ferreira*[2,3]
[1]ATC Labs, Noida, India
[2]ATC Labs, Chatham, NJ, USA
[3]University of Porto, Porto, Portugal

Audio bandwidth extension has emerged as a key low bit rate coding tool. In continuation with our on going research on audio bandwidth extension, this paper presents new enhancements to the Audio Bandwidth Extension Toolkit (ABET). ABET consists of three primary tools Accurate Spectral Replacement (ASR), Fractal Self Similarity Model (FSSM), and Multi-band Temporal Envelope Amplitude Coding (MBTAC). Additionally we have also introduced a blind bandwidth extension mode into ABET. We discuss several new ideas / improvements to ABET. Specifically, enhancements to the blind bandwidth extension architecture that allow it to work with signals with only 3.5–4.0 kHz au- ➥

dio bandwidth are described. We also elaborate on a new tool for efficient coding of time-frequency envelope that cuts the overhead by 0.75–1.0 kbps/channel. We also address a practical issue, i.e., the computational complexity and describe a new low decoder complexity mode of ABET
Convention Paper 7488

Workshop 25 **Tuesday, May 20**
11:30 – 13:30 **Room O**

10 THINGS TO GET RIGHT IN PA AND SOUND REINFORCEMENT

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

Panelists: *Mark Bailey*, Harman/JBL
Luke Jenks, Meyer Sound Laboratories, Berkeley, CA, USA
Evert Start, Duran Audio

The workshop will discuss the 10 most important things to get right when designing / operating sound reinforcement and PA systems. However, as attendees at the workshop will learn there are many more things to consider than just the 10 golden rules, and that the order of importance of these often changes depending upon the venue and type of system. The workshop aims to provide a practical approach to sound systems design and operation and will be illustrated with many practical examples and case histories. Each workshop panelist has many years of practical experience and between them can cover just about any aspect of sound reinforcement and PA systems design, operation, and technology. Come along to a workshop that aims to answer questions you never knew you had—but of course, to find out the ten most important ones, you will have to attend the session!

Master Class 3 **Tuesday, May 20**
12:00 – 14:00 **Room B**

LINEAR AUDIO POWER AMPLIFICATION

Presenter: **Douglas Self**

Audio power amplifiers are one of the few fields where analog still has the upper hand when high quality is required. In this presentation I will review some of the configurations available and examine their advantages and otherwise. The classical amplifier configuration still remains extremely useful and is capable of very good results in terms of linearity when the design details are handled properly. I hope to show that this can be done easily. Some new findings in minimizing distortion will be presented. I shall also be looking at the issue of noise in power amplifier circuits, and how that too can be reduced as much as possible.

Session P27 **Tuesday, May 20**
13:00 – 16:00 **Room C-D**

PSYCHOACOUSTICS, PERCEPTION, AND LISTENING TESTS, PART 3

Chair: **John Beerends**, TNO Information and Communication Technology, Delft, The Netherlands

13:00

P27-1 Perceptual Evaluation of Numerically Simulated Head-Related Transfer Functions
—*Julia Turku, Miikka Vilermo, Eira Seppälä, Monika Pölönen, Ole Kirkeby, Asta Kärkkäinen, Leo Kärkkäinen*, Nokia Research Center, Helsinki, Finland

Head-related transfer functions (HRTFs) produced by numerical simulations were compared to measured HRTFs through two listening tests. The purpose was to determine whether the numerically simulated HRTFs, which do not contain any of the artifacts associated with acoustic measurements, capture the detail necessary for reproducing convincing 3-D sound. The results suggest that when virtual sound sources are presented to listeners binaurally over headphones, the measured and modeled HRTF sets perform equally well in terms of perception of direction. Regarding preference of binauralization methods, the simulated HRTFs performed slightly better.
Convention Paper 7489

13:30

P27-2 Evaluating Perception of Salient Frequencies: Do Mixing Engineers Hear the Same Thing?
—*Joerg Bitzer*,[1] *Jay LeBoeuf*,[2] *Uwe Simmer*[1]
[1]University of Applied Science Oldenburg, Oldenburg, Germany
[2]Imagine Research, Inc., San Francisco, CA, USA

In this paper we analyze the agreement of mixing engineers when finding salient frequencies in recorded audio tracks. Twenty-two mixing engineers were asked to use an equalizer with a high-Q and high-gain setting. Using this tool to sweep through the files' frequencies, they analyzed sixteen audio tracks and reported the most perceptually salient frequencies. The results show that the agreement depends on the analysis bandwidth. Most mixing engineers agree with a wide frequency range. However, only a few engineers agree if the matching bandwidth is below or equal to one-third octave. In this paper we try to explain these results and give a detailed analysis.
Convention Paper 7462

14:00

P27-3 Influence of Visual Appearance on Loudspeaker Sound Quality Evaluation—
Alex Karandreas, Flemming Christensen, Aalborg University, Aalborg, Denmark

Product sound quality evaluation aims to identify relevant attributes and assess their influence on the overall auditory impression. Extending this sound specific rationale, the present paper evaluates overall impression in relation to audition and vision, specifically for loudspeakers. In order to quantify the bias that the loudspeaker appearance has on the sound quality evaluation of a naive listening panel, audio stimuli of varied degradation are coupled with actual loudspeakers of different visual appearance.
Convention Paper 7491

14:30

P27-4 Comparison of Loudspeaker/Room Equalization Preferences for Multichannel, Stereo, and Mono Reproductions: Are Listeners More Discriminating in Mono?—*Sean Olive, Sean Hess, Allan Devantier*, Harman International Industries, Inc., Northridge, CA, USA

Automated digital loudspeaker/room correction products are more popular than ever despite the general lack of perceptual studies on their performance measured over a range of different playback conditions. This paper describes the first of several perceptual experiments designed to explore how different loudspeaker-room correction methods affect the sound quality of reproduction given a range of different listening rooms, loudspeakers, setups, and programs that might influence their perceived performance. A panel of trained listeners gave comparative preference ratings for three different loudspeaker equalizations evaluated in a semi-reflective room using three multichannel music recordings reproduced in surround, stereo, and mono playback modes. The three equalizations were based on either anechoic or in-room measurements with different perceptual weighting given to the direct versus the direct + reflected sounds radiated by the loudspeaker. The different equalizations were identical below 400 Hz to focus on perceptual effects occurring above the room's transition frequency. The results are summarized as follows: all three equalizations were equally preferred over the unequalized system; the difference in preference increased monotonically as the number of playback channels was reduced from 5 (surround) to 1 (mono).
Convention Paper 7492

15:00

P27-5 Caution and Warning Alarm Design and Evaluation for NASA CEV Auditory Displays
—*Durand Begault*,[1] *Martine Godfroy*,[2] *Aniko Sandor*,[3] *Kritina Holden*[4]
[1]NASA Ames Research Center, Moffett Field, CA, USA
[2]San Jose State University Foundation, NASA Ames Research Center, Moffett Field, CA, USA
[3]LZ Technology, NASA Johnson Space Center, Houston, TX, USA
[4]Lockheed Martin Corporation, NASA Johnson Space Center, Houston, TX, USA

The design of caution-warning signals for NASA's Crew Exploration Vehicle (CEV) and other future spacecraft will be based on both best practices based on current research and evaluation of current alarms. A design approach is presented based upon cross-disciplinary examination of psychoacoustic research, human factors experience, aerospace practices, and acoustical engineering requirements. A listening test with thirteen participants was performed involving ranking and grading of current and newly developed caution-warning stimuli under three conditions: (1) alarm levels adjusted for compliance with ISO 7731, "Danger signals for work places—Auditory Danger Signals"; (2) alarm levels adjusted to an overall 15 dBA s/n ratio; and (3) simulated codec low-pass filtering. The resulting analyses include determination of sounds that were judged as inappropriate, independent of condition.
Convention Paper 7493

15:30

P27-6 Loudness Calculation for Individual Acoustical Objects within Complex Temporally Variable Sounds—*Cornelius Bradter, Klaus Hobohm*, Hochschule für Film und Fernsehen, Potsdam, Germany

Models used for loudness calculation normally treat their input signal as an integral whole. For sounds consisting of two or more distinguishable acoustical objects this contradicts the listening experience. Auditory perception analyzes and identifies acoustical objects and may treat them differently. By expanding principles used in excitation synthesis-based loudness models, we developed a procedure to calculate loudness of a time-varying acoustical object while a second object is simultaneously present. When signals of both objects are available individually and in combination, the procedure reflects effects of one object on the other as well as changes of loudness perception due to signal features of one or both objects.
Convention Paper 7494

Workshop 26 — Tuesday, May 20
13:00 – 14:30 — Room L

MIX YOUR OWN NEW YEAR'S CONCERT—PART 1

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

Panelists: *Jean-Marie Geijsen*, Polyhymnia
Morten Lindberg, 2L Oslo - Oslo, Norway
Ronald Prent, Galaxy Studios
Josef Schütz, ORF Radio

A selection of experienced mixers with different backgrounds (classical music, pop, rock) have about 5-7 minutes to set up a 5-channel mix of the current New Year's Concert with the original multitrack recording. The audience will be able to witness the individual approaches to that task. A second session (30 minutes later) will compare those mixes all aligned so that on-the-fly switching is possible. The different approaches and results will be discussed.

Session P28 — Tuesday, May 20
13:30 – 15:00 — Topaz Lounge

POSTERS: SOFTWARE, INSTRUMENTATION, AND MEASUREMENT

13:30

P28-1 An Anatomy of Graph-Based User Interfaces for Media Processing—*Christopher Schultz*,[1] *Jörn Loviscach*,[2] *Shailendra Mathur*,[3] *Jay LeBoeuf*[4]
[1]Universität Bremen, Bremen, Germany, now at mediaclipping, Bremen, Germany
[2]Hochschule Bremen, Bremen, Germany
[3]Softimage Corp., Avid Technology, Inc., Montreal, Quebec, Canada

[4]Digidesign, Daly City, CA, USA, now at Imagine Research, Inc., San Francisco, CA, USA

Graph-based user interfaces are employed in a variety of software such as audio synthesizers, video compositing tools, and database application builders. All of these uses afford the graphical metaphor of a graph: "Nodes" such as sound generators or filters are tied together by "links," which may represent signal flow or conceptual relations. Focusing on media production tools, we have examined a large range of current software products to find out which de-facto standards have evolved in the field of graph-based interfaces and which features can be considered unique. We categorize a multitude of interface concepts employed in actual graph-based interfaces and describe differences in their implementation. The findings provide guidelines for developers of media production software.
Convention Paper 7495

13:30

P28-2 A Framework for Automatic Mixing Using Timbral Similarity Measures and Genetic Optimization—*Bennett Kolasinski*, New York University, New York, NY, USA

A novel method is introduced for automatic mix recreation using timbral classification techniques and an optimization algorithm. This approach uses the Euclidean distance between modified Spectral Histograms to calculate the distance between a mix and a target sound and uses a genetic optimization algorithm to figure out the best coefficients for that mix. The implementation has been shown to successfully recreate multitrack mixes accurately and may pave the way toward the automatic mixing of novel multitrack sessions based on a desired target sound.
Convention Paper 7496

13:30

P28-3 Delta-Sigma DAC Topologies for Improved Jitter Performance—*Ivar Løkken, Anders Vinje, Trond Sæther*, Norwegian University of Science and Technology, Trondheim, Norway

Specifications for audio digital-to-analog-converters (DACs) place requirements on the analog circuit design that contradict physical design conditions in a modern, digital-oriented system on a chip process. Because of low supply voltages, use of current-steering DACs has become the dominant choice for high resolution applications. Fed by a delta-sigma modulator that requantizes the digital signal to a manageable number of bits, the current-steering DAC is a continuous time type converter without any discrete time filtering. This makes it very susceptible to sampling clock jitter. In this paper jitter distortion is addressed at a topology level, investigating design choices for the delta-sigma requantizer and the possible use of semidigital multi-bit current-steering filter DACs to reduce problems with jitter susceptibility.
Convention Paper 7497

13:30

P28-4 New Measurement Methods for Anechoic Chamber Characterization—*Juan Gómez-Alfageme, José Luis Sánchez-Bote, Elena Blanco-Martín*, Universidad Politécnica de Madrid, Madrid, Spain

As a continuation of the work presented at the 122nd AES Convention (Paper 7153), this paper tries to study in depth the anechoic chambers qualification. The purpose of this paper is to find parameters that allow the characterization of this type of enclosure. The proposal tries to obtain data of the anechoic chambers absorption by means of the transfer functions between pairs of microphones or by means of the impulse response between pairs of microphones. The

results of the transfer functions between pairs of microphones can be easily checked by the agreement of the inverse squared law, allowing determination of the chamber cut-off frequency. Making a band filtering confirmed the anechoic chamber's qualifications.
Convention Paper 7498

13:30

P28-5 Acoustic Feedback Reduction Based on LMS and Normalized LMS Algorithms in WOLA Filters Bank Based Digital Hearing Aids—*Raúl Vicen-Bueno*,[1] *Almudena Martínez-Leira*,[2] *Manuel Rosa-Zurera*,[1] *Lucas Cuadra-Rodríguez*[1]
[1]Universidad de Alcalá, Alcalá de Henares, Madrid, Spain
[2]Dimetronic Signals, San Fernando de Henares, Madrid, Spain

Acoustic feedback phenomenon can disturb a digital hearing aid performance at high gains, causing instability in the haring aid and degradation in the speech. In order to restore a stable situation, an acoustic feedback reduction (AFR) subsystem using adaptive algorithms such as the least-mean square (LMS) algorithm is needed. This algorithm has a reduced computational cost, but it is very unstable. In order to avoid this situation, another feedback reduction system based on a modified version of the LMS algorithm is used. Such algorithm is: the Normalized LMS (NLMS). These two algorithms are tested in two digital hearing aid categories: the In-The-Ear and the In-The-Canal. These categories are selected because they have great feedback effects, so robust AFR subsystems are needed. The added stable gain (ASG) over the limit gain when an AFR subsystem is working in the digital hearing aid is obtained for each category. The ASG is determined as a trade-off between two measurements: the segmented signal-to-noise ratio (objective measurement) and the speech quality (subjective measurement). The results show how the digital hearing aids working with a feedback reduction adaptive filter adapted with the NLMS algorithm is able to achieve up to 18 dB of increase over the limit gain.
Convention Paper 7499

13:30

P28-6 Nonlinear Distortions in Capacitors—*Menno van der Veen*,[1] *Hans van Maanen*[2]
[1]ir.bureau Vanderveen bv, Zwolle, The Netherlands
[2]Temporal Coherence, The Netherlands

Many people have claimed that capacitors have a notable influence on the audible quality of systems. We have identified one of the major causes of nonlinear distortions in capacitors. Charging the capacitor will result in an attractive force acting on the conducting plates. As no material is infinitely stiff, this force will reduce the thickness of the dielectricum and thus increase the capacitance. This process occurs in both phases of an AC signal in the same way and is thus nonlinear. In this paper the consequences of this process are discussed. It should be noted that other passive components like resistors and inductors can also show similar nonlinear behavior.
Convention Paper 7500

Workshop 26 **Tuesday, May 20**
13:00 – 14:30 **Room L**

MIX YOUR OWN NEW YEAR'S CONCERT—PART 1

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

Panelists: *Jean-Marie Geijsen*, Polyhymnia
Morten Lindberg, 2L Oslo - Oslo, Norway
Ronald Prent, Galaxy Studios
Josef Schütz, ORF Radio

A selection of experienced mixers with different backgrounds (classical music, pop, rock) have about 5-7 minutes to set up a 5-channel mix of the current New Year's Concert with the original multitrack recording. The audience will be able to witness the individual approaches to that task. A second session (30 minutes later) will compare those mixes all aligned so that on-the-fly switching is possible. The different approaches and results will be discussed.

Tutorial 19 **Tuesday, May 20**
14:00 – 16:30 **Room N101**

LARGE ROOM ACOUSTICS

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

In the tutorial the physical principles of room acoustics will be explained. It will be explained how, by measuring or calculating impulse responses along arrays of receiver positions, the temporal and spatial properties of a sound field can be analyzed and understood, and how these properties are related to perceptual quality cues. Several models to calculate impulse responses will be discussed. Architectural as well as electro-acoustical measures to correct for acoustical shortcomings will be proposed. All these aspects will be illustrated with examples from practice.

Master Class 4 **Tuesday, May 20**
14:30 – 16:30 **Room B**

A UNIVERSAL GRAMMAR OF CLASS D AUDIO AMPLIFICATION

Presenter: **Bruno Putzeys**, Hypex Electronics and Grimm Audio, The Netherlands

At first sight, class D amplifiers exhibit a baffling array of disjoint design philosophies and topologies. Yet, all of these can be seen as permutations of the same basic structure. This master class explores the "universal grammar of class D audio amplification" the building blocks and design choices underlying all class D designs, like power stage topology, loop control strategy, and modulation method. This generalized view makes clear how certain designs turn out to be uncontroversially suboptimal choices while some other apparently meaningless rearrangements like localized error feedback can be demonstrated to have real performance advantages.

Student Event/Career Development
STUDENT DELEGATE ASSEMBLY MEETING —PART 2

Tuesday, May 20, 14:30 – 16:00
Room O

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 European 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.

Workshop 27 **Tuesday, May 20**
15:00 – 16:30 **Room L**

MIX YOUR OWN NEW YEAR'S CONCERT—PART 2

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

Panelists: *Jean-Marie Geijsen*, Polyhymnia
Morten Lindberg, 2L Oslo, Oslo, Norway
Ronald Prent, Galaxy Studios
Josef Schütz, ORF Radio

A selection of experienced mixers with different backgrounds (classical music, pop, rock) have about 5-7 minutes to set up a 5-channel mix of the current New Year's Concert with the original multitrack recording. The audience will be able to witness the individual approaches to that task. A second session (30 minutes later) will compare those mixes all aligned so that on-the-fly switching is possible. The different approaches and results will be discussed.