AES 126th Convention Program

May 7 – 10, 2009

M,O,C, Center, Munich

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 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 126th 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 126th AES Convention Student Paper Award is:
Modeling of External Ear Acoustics for Insert Headphone Usage—*Marko Hiipakka (Presenting Author), Miikka Tikander, Matti Karjalainen*
Convention Paper 7739

To be presented on Friday, May 8 in Session P15 —Posters: Hearing

* * * * *

Student Event/Career Development
STUDENT SCIENCE SPOT
Thursday, May 7 through Sunday, May 10
Hall 2 Foyer

Ongoing presentations of student's non-commercial scientifically challenging works will be given throughout the Convention in an accessible way. They will offer prototypes or demonstrations, including a student poster session. Within the scope of their presentation, the students will have the opportunity to present themselves and their institutes in an appropriate way. The SSS will be hosted by the AES Student Section Graz (AES SSG), supported by the SDA and managed in cooperation with the students who share their projects at the booth.

Session P1 — **Thursday, May 7**
09:00 – 11:30 — **Room K3**

AUDIO FOR TELECOMMUNICATIONS

Chair: **Damian Murphy**, University of York, York, UK

09:00

P1-1 20 Things You Should Know Before Migrating Your Audio Network to IP—*Simon Daniels*, APT, Belfast, Northern Ireland, UK

For many years, synchronous networks have been considered the industry standard for audio transport worldwide. Balanced analog copper circuits, microwave, and synchronous based systems such as V.35/X.21 or T1/E1 have been the traditional choice for studio transmitter and inter-studio links in professional audio broadcast networks. Readily available from all major service providers, the popularity of synchronous links has been largely due to the fact that they offer dedicated, reliable, point-to-point and bidirectional communication at guaranteed data and error rates. However, the reign of synchronous links as the preferred choice for STLs is currently coming under threat from a new challenger, in the form of IP-based network technology.
Convention Paper 7651

09:30

P1-2 Deploying Large Scale Audio IP Networks—*Kevin Campbell*, APT, Belfast, Northern Ireland, UK

This paper will examine the key considerations for those interested in deploying large-scale IP audio networks. It will include an overview of the main challenges and draw on the experience of national public broadcasters who have already migrated to IP. We will provide an overview of the key concerns such as jitter, delay, and link reliability that are valid for an IP network of any size. However, this paper will focus mainly on the issues arising from the greater complexity and scale of large national and country-wide deployments. The paper will use illustrations and network applications from real-world deployments to illustrate the points.
Convention Paper 7652
Paper presented by Hartmut Foerster

10:00

P1-3 A Spatial Filtering Approach for Directional Audio Coding—*Markus Kallinger, Henning Ochsenfeld, Giovanni Del Galdo, Fabian Kuech, Dirk Mahne, Richard Schultz-Amling, Oliver Thiergart*, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

In hands-free telephony, spatial filtering techniques are employed to enhance intelligibility of speech. More precisely, these techniques aim at reducing the reverberation of the desired speech signal and attenuating interferences. Additionally, it is well-known that the spatially separate reproduction of desired and interfering sources enhance intelligibility of speech. For the latter task, Directional Audio Coding (DirAC) has proven to be an efficient method to capture and reproduce spatial sound. In this paper we propose a spatial filtering processing block, which works in the parameter domain of DirAC. Simulation results show that compared to a standard beamformer the novel technique offers significantly higher interference attenuation, while introducing comparably low distortion of the desired signal. Additional subjective tests of speech intelligibility confirm the instrumentally obtained results.
Convention Paper 7653

10:30

P1-4 A New Bandwidth Extension for Audio Signals without Using Side-Information—*Kha Le Dinh, Chon Tam Le Dinh, Roch Lefebvre*, Université de Sherbrooke, Sherbrooke, Quebec, Canada

The use of narrow bandwidth (300 – 3400 Hz) in the current telephone network limits the perceptual quality of telephone conversations. Changing to wideband network is a solution that can help to improve quality, but it will need a long time to upgrade. Thus, bandwidth extension can be seen as an alternative solution during the transition time. A new bandwidth extension method is presented in this paper. Without using any side-information, the proposed method can be applied as a post-processing step at the terminal devices, maintaining the compatibility to the current telephone network, and thus, no modification is needed in the network nodes. Experimental results show that the proposed solution can help to improve significantly the perceptual quality of narrowband telephone signal.
Convention Paper 7654

11:00

P1-5 Feature Selection vs. Feature Space Transformation in Music Genre Classification Framework—*Hanna Lukashevich*, Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

Automatic classification of music genres is an important task in music information retrieval research. Nearly all state-of-the-art music genre recognition systems start from the feature extraction block. The extracted acoustical features often could tend to be correlated or/and redundant, which can cause various difficulties in the classification stage. In this paper we present a comparative analysis on applying supervised Feature Selection (FS) and Feature Space Transformation (FST) algorithms to reduce the feature dimensionality. We discuss pros and cons of the methods and weigh the benefits of each one against the others.
Convention Paper 7655

Session P2 **Thursday, May 7**
09:00 – 11:30 **Room K4**

AUDIO FOR GAMES AND INTERACTIVE MEDIA

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

09:00

P2-1 Viable Distribution of Multichannel Audio-over-IP for Live and Interactive "Voice Talent"-Based Gaming Using High-Quality, Low-Latency Audio Codec Technology—*Gregory Massey*, APT, Belfast, Northern Ireland, UK

The delivery of multichannel audio—from mono to surround sound—in real-time over public IP networks for the purpose of interactive crowd-participant gaming presents a significant design engineering challenge to games developers, console manufacturers, ISPs, and CDNs. Leveraging expertise gained in professional broadcasting and recording studio postproduction, APT has developed a robust and scalable audio codec technology that meshes with popular gaming systems to realize low-latency distribution of high-quality audio for immersive, instantaneous audio experiences in massively multiplayer online games involving interactive audience responses to vocal/singing talent.
Convention Paper 7656
Paper presented by David Trainor

09:30

P2-2 Elevator: Emotional Tracking Using Audio/Visual Interaction—*Basileios Psarras, Andreas Floros, Marianna Strapatsakis*, Ionian University, Corfu, Greece

The research interest on modeling everyday human emotions and controlling them through typical multimedia content (i.e., audio and video data) has recently increased. In this paper an interactive methodology is introduced for detecting, controlling, and tracking emotions. Based on the above methodology, an interactive audiovisual installation termed "Elevator" was realized, aiming to analyze and manipulate simple emotions of the participants (such as anger) using simplified emotion detection audio signal processing techniques and specifically selected combined audio/visual content. As a result, the human emotions are "elevated" to pre-defined levels and appropriately mapped to visual content, which corresponds to the emotional "thumbnail" of the participants.
Convention Paper 7657

10:00

P2-3 Applications of Bending Wave Technology in Human Interface Devices—*Neil Harris*, New Transducers Ltd. (NXT), Cambridge, UK

The application of bending-waves to so-called "flat panel loudspeakers" has often been the topic of papers at AES Conventions. This paper looks at other interesting applications of the technology that have, or are beginning to

have commercial pull. These applications are also part of the interface between human and machine, but focus on the sense of touch rather than of hearing. The idea of a touch screen is not new but is only now becoming ubiquitous with a new generation of devices, typified by the i-Phone. If touch sensors are the analog of the microphone, then haptic feedback generators are the analog of the loudspeaker. Bending waves are beginning to find application here too.
Convention Paper 7658

10:30

P2-4 Designing Auditory Display Menu Interfaces —Cues for Users' Current Location in Extensive Menus—*Erik Sikström, Jan Berg*, Luleå University of Technology, Luleå, Sweden

This paper reviews the current research in auditory display in search for design guidelines for presenting the contents in audio-only menu interfaces. The aim of the review is to find new directions for auditory display menu interface design. Among several techniques for representing individual menu items the preliminary results show that the spearcon seems to be the most suitable method. For the layout of menu items, studies have shown that spatial separation, different timbres, and staggering onset between the items improves recognition rates, particularly for concurrently presented items. A remaining issue to be investigated is how to remind the users of their current location in the menus of extensive menu interfaces.
Convention Paper 7659

11:00

P2-5 Symmetry Model-Based Key Finding—*Markus Mehnert,*[1] *Gabriel Gatzsche,*[2] *Daniel Arndt*[3]
[1]Technische Universität Ilmenau, Ilmenau, Germany
[2]Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany
[3]Fraunhofer IIS, Ilmenau, Germany

In this paper we introduce a new key finding algorithm that is based on the symmetry model introduced by Gatzsche et al. The algorithm consists of two parts. First, the most probable diatonic pitch class set of the musical piece is recognized. Second, using one of the subspaces of the symmetry model the mode of the piece is estimated. The algorithm is evaluated with 100 Beatles songs, 90 newer “Pop and Rock” songs, and 252 classical pieces from the Naxos database. The results will be compared to the algorithms of Lerch, Zhu et al., and an algorithm based on binary major and minor chord profiles. The new algorithm has the highest overall key finding MIREX’05 score of 82.9 percent.
Convention Paper 7660

Workshop 1 **Thursday, May 7**
09:00 – 11:00 **Room D123**

LIES, DAMN LIES, AND STATISTICS

Chair: **Sylvain Choisel**, Philips Consumer Lifestyle, Leuven, Belgium

Panelists: *Thomas Sporer*, Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany
Florian Wickelmaier, University of Tübingen, Tübingen, Germany

Listening tests have become an important part of the development of audio systems (CODECS, loudspeakers, etc.). Unfortunately, even the simplest statistics (mean and standard deviation) are often misused.

This workshop will start with a basic introduction to statistics, but room will be given to discuss the pertinence of some commonly-used tests, and alternative methods will be proposed, thereby making it interesting for more experienced statisticians as well. The following topics will be covered (among others): experimental design, distributions, hypothesis testing, confidence intervals, analysis of paired comparisons and ranking data, and common pitfalls potentially leading to wrong conclusions.

Tutorial 1 **Thursday, May 7**
09:00 – 11:00 **Room K1**

A NEW DOWNMIX ALGORITHM OPTIMIZING COMB FILTER PERFORMANCE

Presenter: **Jörg Deigmöller**, IRT Munich

Downmixing 5.1 surround sound to 2.0 stereo is a necessity in current challenging production environments. Straightforward downmix coefficients, as simple as they are to execute, result in comb filtering effects dependent on the correlation of the signals that are downmixed. A new algorithm is presented that tackles this issue with optimized decorrelation pre-processing, leading to much improved timbre of the downmixed signal.

Tutorial 2 **Thursday, May 7**
09:00 – 11:00 **Room K2**

AN INTRODUCTION TO DIGITAL AUDIO EFFECTS

Chair: **Christoph M. Musialik**, Algorithmix GmbH

Panelists: *Joshua D. Reiss*, Queen Mary, University of London, London, UK
Udo Zölzer, Helmut-Schmidt-Universität, Hamburg, Germany

In this tutorial we discuss the ways by which signal processing techniques are used to produce effects acting on digital audio signals. The audio effects are systematically classified and discussed, with emphasis on how and why they are used. Practical examples of common effects are provided, along with block diagrams, pseudo-code, and sound examples. During the tutorial, a few effects will be created from scratch and the audience will be provided with the basic background knowledge to design their own effects.

Thursday, May 7 **09:00** **Room D120**
Technical Committee Meeting on Acoustics and Sound Reinforcement

Session P3 **Thursday, May 7**
10:00 – 11:30 **K4 Foyer**

POSTERS: RECORDING, REPRODUCTION, AND DELIVERY

10:00

P3-1 Audio Content Annotation, Description, and Management Using Joint Audio Detection,

Segmentation, and Classification Techniques —*Christos Vegiris, Charalambos Dimoulas, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The current paper focuses on audio content management by means of joint audio segmentation and classification. We concentrate on the separation of typical audio classes, such as silence/background noise, speech, crowded speech, music, and their combinations. A compact feature-vector subset is selected by a Correlation feature selection subset evaluation algorithm after the use of EM clustering algorithm on an initial audio data set. Time and spectral parameters are extracted using filter-banks and wavelets in combination with sliding windows and exponential moving averaging techniques. Features are extracted on a point-to-point basis, using the finest possible time resolution, so that each sample can be individually classified to one of the available groups. Clustering algorithms like EM or Simple K-means are tested to evaluate the final point-to-point classification result, therefore the joint audio detection-classification indexes. The extracted audio detection, segmentation, and classification results can be incorporated into appropriate description schemes that would annotate audio events/segments for content description and management purposes.
Convention Paper 7661

10:00

P3-2 Ambience Sound Recording Utilizing Dual MS (Mid-Side) Microphone Systems Based upon Frequency Dependent Spatial Cross Correlation (FSCC) [Part 3: Consideration of Microphones' Locations]—*Teruo Muraoka, Takahiro Miura, Tohru Ifukube*, University of Tokyo, Tokyo, Japan

In order to achieve ambient and exactly sound-localized musical recording with fewer number of microphones, we studied sound acquisition performances of microphone arrangements utilizing their Frequency Dependent Spatial Cross Correlation (FSCC). The result is that an MS microphone is best for this purpose. The setting of the microphone's directional azimuth at 132 degrees is the best for ambient sound acquisition and setting of that at 120 degrees is best for on-stage sound acquisition. We conducted actual concert recordings with a combination of those MS microphones (Dual MS microphone systems) and obtained satisfactory results. Successively, we studied the proper setting positions of those microphones. For ambient sound acquisition, suspending the microphone at the center of a concert hall is favorable, and for on-stage sound acquisition, locating it at almost above the conductor's position will also be satisfactory. Process of the studies will be reported.
Convention Paper 7662

10:00

P3-3 A Comparative Approach to Sound Localization within a 3-D Sound Field—*Martin J. Morrell, Joshua D. Reiss*, Queen Mary, University of London, London, UK

In this paper we compare different methods for sound localization around and within a 3-D sound field. The first objective is to determine which form of panning is consistently preferred for panning sources around the loudspeaker array. The second objective and main focus of the paper is localizing sources within the loudspeaker array. We seek to determine if the sound sources can be located without movement or a secondary reference source. The authors compare various techniques based on ambisonics, vector base amplitude panning and time delay based panning. We report on subjective listening tests that show which method of panning is preferred by listeners and rate the success of panning within a 3-D loudspeaker array.
Convention Paper 7663

10:00

P3-4 The Effect of Listening Room on Audio Quality in Ambisonics Reproduction—*Olli Santala*,[1] *Heikki Vertanen*,[1,2] *Jussi Pekonen*,[1] *Jan Oksanen*,[1] *Ville Pulkki*[1]
[1]Helsinki University of Technology, Espoo, Finland
[2]University of Helsinki, Helsinki, Finland

In multichannel reproduction of spatial audio with first-order Ambisonics the loudspeaker signals are relatively coherent, which produces prominent coloration. The coloration artifacts have been suggested to depend on the acoustics of the listening room. This dependency was researched with subjective listening tests in an anechoic chamber with an octagonal loudspeaker setup. Different virtual listening rooms were created by adding diffuse reverberation with 0.25 seconds RT60 using a 3-D 16-channel loudspeaker setup. In the test, the subjects compared the audio quality in the virtual rooms. The results suggest that optimal audio quality was obtained when the virtual room effect and the direct sound were on equal level at the listening position.
Convention Paper 7664

10:00

P3-5 Ontology-Based Information Management in Music Production—*Gyorgy Fazekas, Mark Sandler*, Queen Mary, University of London, London, UK

In information management, ontologies are used for defining concepts and relationships of a domain in question. The use of a schema permits structuring, interoperability, and automatic interpretation of data, thus allows accessing information by means of complex queries. In this paper we use ontologies to associate metadata, captured during music production, with explicit semantics. The collected data is used for finding audio clips processed in a particular way, for instance, using engineering procedures or acoustic signal features. As opposed to existing metadata standards, our system builds on the Resource Description Framework, the data model of the Semantic Web, which provides flexible and open-ended knowledge representation. Using this model, we demonstrate a framework for managing information, relevant in music production.
Convention Paper 7665

Thursday, May 7 10:00 Room D120
Technical Committee Meeting on Archiving, Restoration, and Digital Libraries

Thursday, May 7 11:00 Room D120
Technical Committee Meeting on Coding of Audio Signals

Student Event/Career Development
OPENING AND STUDENT DELEGATE ASSEMBLY MEETING – PART 1
Thursday, May 7, 11:15 – 12:00 Room K2

Chair: **Misato Yamada**

Vice Chair: **Miroslav Jakovljevic**

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–Part 2 on Saturday, May 9, at 17:30.

Special Event
AWARDS PRESENTATION AND KEYNOTE ADDRESS
Thursday, May 7, 12:00 – 13:3
Room K3

Opening Remarks:
- Executive Director Roger Furness
- President Jim Anderson
- Convention Chair Martin Wöhr

Program:
- AES Awards Presentation
- Introduction of Keynote Speaker by Convention Chair Martin Wöhr
- Keynote Address by Gerhard Thoma

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

Bronze Medal Award:
- **Ivan Stamac**

Fellowship Award:
- **Martin Wöhr**

Board of Governors Award
- **Jan Berg**
- **Kimio Hamasaki**
- **Shinji Koyano**
- **Tapio Lokki**
- **Jiri Ocenasek**
- **John Oh**
- **Jan Abildgaard Pedersen**
- **Joshua Reiss**

Keynote Speaker
This year's Keynote Speaker is **Gerhard Thoma**. Thoma has been leading the Department of Acoustics Projects at BMW for more than 20 years. His speech will highlight many aspects of perception and acoustics from an unusual point of view: What does a driver in a car need to hear, what should not hear, and how can the acoustics and sounds of a car help to significantly enhance driving pleasure and safety? The title of his address is *"Human-Oriented Acoustics—Targets for Future Developments in Vehicle Acoustics."*

Thursday, May 7 13:30 Room D104
Standards Committee Meeting on SC-02-02 Digital Input/Output Interfacing

Workshop 2 Thursday, May, 7
13:30 – 15:30 Room K2

New Technologies for Audio Over IP

Chair: **Jeremy Cooperstock**, McGill University, Montreal, Quebec, Canada

Panelists: *Steve Church*, Telos
Christian Diehl, Mayah
Manfred Lutzky, Fraunhofer IIS
Greg Massey, APT

This workshop is intended to provide an "under the hood" discussion of various low-latency codecs as well as a comparison of their pros and cons for different applications. Codecs including AAC-ELD and ULD will be discussed, along with techniques such as adaptive jitter buffer management.

Session P4 Thursday, May 7
14:00 – 18:30 Room K3

RECORDING, REPRODUCTION, AND DELIVERY

Chairs: **Jörg Wuttke**, Jörg Wuttke Consultancy, Pfinztal, Germany, Schoeps GmbH, Karlsruhe, Germany
Siegfried Linkwitz, Linkwitz Lab, Corte Madera, CA, USA

14:00

P4-1 An Expert in Absentia: A Case-Study for Using Technology to Support Recording Studio Practice —*Andrew King*, University of Hull, Scarborough, North Yorkshire, UK

This paper examines the use of a Learning Technology Interface (LTI) to support the completion of a recording workbook with audio examples over a ten-week period. The VLE provided contingent support to studio users for technical problems encountered in the completion of four recording tasks. Previous research has investigated how students collaborate and problem-solve during a short session in the recording studio using technology as a contingent support tool. In addition, online message boards have been used to record problems encountered when completing a prescribed task (critical-incident recording). A mixed-methods case study approach was used in this study. The students interactions within the LTI were logged (i.e., frequency, time, duration, type of support) and their feedback was elicited via a user questionnaire at the end of the project. Data for this study demonstrates that learning technology can be a successful support tool and also highlights the frequency and themes ➥

concerning the types of recording practice information accessed by the learners.
Convention Paper 7669

14:30

P4-2 Recording and Reproduction over Two Loudspeakers as Heard Live—Part 1: Hearing, Loudspeakers, and Rooms— *Siegfried Linkwitz*,[1] *Don Barringer*[2]
[1]Linkwitz Lab, Corte Madera, CA, USA
[2]Linkwitz Lab, Arlington, CA, USA

Innate hearing processes define the realism that can be obtained from reproduced sound. An unspecified system with two loudspeakers in a room places considerable limitations upon the degree of auditory realism that can be obtained. It has been observed that loudspeakers and room must be hidden from the auditory scene that is evoked in the listener's brain. Requirements upon the polar response and the output volume capability of the loudspeaker will be discussed. Problems and solutions in designing a three-way, open baffle loudspeaker with piston drivers will be presented. Loudspeakers and listener must be symmetrically placed in the room to minimize the effects of reflections upon the auditory illusion.
Convention Paper 7670

15:00

P4-3 Recording and Reproduction over Two Loudspeakers as Heard Live—Part 2: Recording Concepts and Practices— *Don Barringer*,[1] *Siegfried Linkwitz*[2]
[1]Linkwitz Lab, Arlington, CA, USA
[2]Linkwitz Lab, Corte Madera, CA, USA

For a half century, the crucial interaction between recording engineer and monitor loudspeakers during two-channel stereophonic recording has not been resolved, leaving the engineer to cope with uncertainties. However, recent advances in defining and improving this loudspeaker-room-listener interface have finally allowed objectivity to inform and shape the engineer's choices. The full potential of the two-channel format is now accessible to the recording engineer, and in a room that is just as normal as most consumers' rooms. The improved reproduction has also allowed a deeper understanding of the merits and limits of spaced and coincident/near-coincident microphone arrays. As a result of these and earlier observations, a four-microphone array was conceived that exploits natural hearing processes to achieve greater auditory realism from two loudspeakers. A number of insights have emerged from the experiments.
Convention Paper 7671

15:30

P4-4 Vision and Technique behind the New Studios and Listening Rooms of the Fraunhofer IIS Audio Laboratory— *Andreas Silzle,*[1] *Stefan Geyersberger,*[1] *Gerd Brohasga*,[1] *Dieter Weninger*,[1,2] *Michael Leistner*[3]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]Innovationszentrum für Telekommunikation stechnik GmbH IZT, Erlangen, Germany
[3]Fraunhofer Institute for Building Physics IBP, Stuttgart, Germany

The new audio laboratory rooms of the Fraunhofer IIS and their technical design are presented here. The vision behind them is driven by the very high demands of a leading edge audio research organization with more than 100 scientists and engineers. The 300 m^2 sound studio complex was designed with the intention of providing capabilities that are in combination far more extensive than those available in common audio research or production facilities. The reproduction room for listening tests follows the strict recommendations of ITU-R BS 1116. The results of the qualification measurements regarding direct sound, reflected sound, and steady state sound field will be shown and the construction efforts needed to achieve these values are explained. The connection from all the computers in the server room to more than 70 loudspeakers in the reproduction rooms, other audio interfaces, and the projection screens is done by an audio and video routing system. The architecture of the advanced control software of this routing system is presented. It allows easy and flexible access for each class of user to all the possibilities made available by this completely new system.
Convention Paper 7672

16:00

P4-5 Advances in National Broadcaster Networks: Exploring Transparent High Definition IPTV— *Matthew O'Donnell*, British Sky Broadcasting, Upminster, UK

British commercial broadcasters are increasing their ability to determine the quality of distribution of audio-over-IP by acquiring and installing next generation national Gigabit networks. This paper explores how broadcasters can use the advances in broadband technology to transparently integrate supplemental on-demand IPTV services with traditional broadcasting transport, which has led to broadcasters being confident in achieving scalable carrier-class quality of service for delivery of high definition media direct to the customer's set top box.
Convention Paper 7673

16:30

P4-6 Multi-Perspective Surround Sound Audio Recording— *Mark J. Sarisky*, The University of Texas at Austin, Austin, TX, USA

With the advent of Blu-Ray Disc Audio (BD-Audio), high resolution uncompressed audio recordings can be presented as a consumer product in a variety of surround sound formats. This paper proposes a new take on the recording of live and studio music in surround sound that allows the consumer to benefit from the large capacity of the BD-Audio disc and enjoy the recording from multiple listening perspectives.
Convention Paper 7674

17:00

P4-7 Sound Intensity-Based Three-Dimensional Panning— *Akio Ando, Kimio Hamasaki*, NHK Science and Technical Research Laboratories, Setagaya, Tokyo, Japan

Three-dimensional (3-D) panning equipment is essential for the production of 3-D audio content. We have already proposed an algorithm to enable such panning. It generates the input signal to be fed into multichannel loudspeakers so as to realize the same physical properties of sound at the receiving point as those created by a single loudspeaker model of the virtual source. A sound pressure vector is used as the physical property. This paper proposes a new method that uses sound intensity instead of the sound pressure vector and shows that both conventional "vector base amplitude panning" and our previous method come very close to achieving coincidence of sound intensity. A new panning method using four loudspeakers is also proposed.
Convention Paper 7675

17:30

P4-8 A Practical Comparison of Three Tetrahedral Ambisonic Microphones—*Dan Hemingson, Mark Sarisky*, The University of Texas at Austin, Austin, TX, USA

This paper compares two low-cost tetrahedral ambisonic microphones, an experimental microphone, and a Core Sound TetraMic with a Soundfield MKV or SPS422B serving as a standard for comparison. Recordings were made in natural environments of live performances, in a recording studio, and in an anechoic chamber. The results of analytical and direct listening tests of these recordings are discussed in this paper. A description of the experimental microphone and the recording setup is included.
Convention Paper 7676

18:00

P4-9 A New Reference Listening Room for Consumer, Professional, and Automotive Audio Research— *Sean Olive*, Harman International, Northridge, CA, USA

This paper describes the features, scientific rationale, and acoustical performance of a new reference listening room designed for the purpose of conducting controlled listening tests and psychoacoustic research for consumer, professional, and automotive audio products. The main features of the room include quiet and adjustable room acoustics, a high-quality calibrated playback system, an in-wall loudspeaker mover, and complete automated control of listening tests performed in the room.
Convention Paper 7677

Session P5 Thursday, May 7
14:00 – 18:30 Room K4

LOUDSPEAKERS

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

14:00

P5-1 Estimating the Velocity Profile and Acoustical Quantities of a Harmonically Vibrating Loudspeaker Membrane from On-Axis Pressure Data—*Ronald M. Aarts*,[1,2] *Augustus J. Janssen*[1]
[1]Philips Research Europe, Eindhoven, The Netherlands
[2]Technical University of Eindhoven, Eindhoven, The Netherlands;

Formulas are presented for acoustical quantities of a harmonically excited resilient, flat, circular loudspeaker in an infinite baffle. These quantities are the sound pressure on-axis, far-field, directivity and the total radiated power. These quantities are obtained by expanding the velocity distribution in terms of orthogonal polynomials. For rigid and non-rigid radiators, this yields explicit, series expressions for both the on-axis and far-field pressure. In the reverse direction, a method of estimating velocity distributions from (measured) on-axis pressures by matching in terms of expansion coefficients is described. Together with the forward far-field computation scheme, this yields a method for assessment of loudspeakers in the far-field and of the total radiated power from (relatively near-field) on-axis data (generalized Keele scheme).
Convention Paper 7678

14:30

P5-2 Testing and Simulation of a Thermoacoustic Transducer Prototype—*Fotios Kontomichos, Alexandros Koutsioubas, John Mourjopoulos, Nikolaos Spiliopoulos, Alexandros Vradis, Stamatis Vassilantonopoulos*, University of Patras, Patras, Greece

Thermoacoustic transduction is the transformation of thermal energy fluctuations into sound. Devices fabricated by appropriate materials utilize such a mechanism in order to achieve acoustic wave generation by direct application of an electrical audio signal and without the use of any moving components. A thermoacoustic transducer causes local vibration of air molecules resulting in a proportional pressure change. The present paper studies an implementation of this alternative audio transduction technique for a prototype developed on silicon wafer. Measurements of the performance of this hybrid solid state device are presented and compared to the theoretical principles of its operation, which are evaluated via simulations.
Convention Paper 7679

15:00

P5-3 Analysis of Viscoelasticity and Residual Strains in an Electrodynamic Loudspeaker—*Ivan Djurek,*[1] *Antonio Petosic*,[1] *Danijel Djurek*[2]
[1]University of Zagreb, Zagreb, Croatia
[2]Alessandro Volta Applied Ceramics (AVAC), Zagreb, Croatia

An electrodynamic loudspeaker was analyzed in three steps: (a) as a device supplied by the market, (b) removed upper suspension, and (c) dismantled assembly consisting only of vibrating spider and voice coil. In three steps, resonant frequency and stiffness were measured dynamically for driving currents up to 100 mA, whereas stiffness was also measured quasi-statically by the use of calibrated masses. It was found that widely quoted effect of decreasing resonant frequency, as plotted against driving current, comes from the residual strain in the vibrating material, and ➥

significant contribution is associated with the spider. When driving current increases residual strain is gradually compensated, giving rise to the minimum of stiffness, and further increase of resonant frequency is attributed to a common nonlinearity in the forced vibrating system.
Convention Paper 7680

15:30

P5-4 Integrated Conventional and Directional Loudspeaker—*Changuang Cai*,[1] *Koray Ozcan*,[2] *Jani Nurminen*,[3] *Xuesheng Li*,[4] *Limei Xu*,[4] *Salo Antti*,[5] *Xia Wang*,[1] *Leon Xu*[1]
[1]Nokia Research Center (Beijing), Beijing, China
[2]Nokia (UK) Ltd., Hampshire, UK
[3]Nokia Devices R&D, Tampere, Finland
[4]University of Electronic Science and Technology of China, Chengdu Sichuan, China
[5]Nokia Research Center, Helsinki, Finland

This paper presents a novel loudspeaker concept that realizes both conventional and directional loudspeaker functions using a single transducer component. The transducer component works in such a way that its low frequency characteristics satisfy the requirements of the conventional loudspeaker, while one of its natural frequencies is located in the ultrasonic range. When the input exciting signal is a conventional audio signal, the transducer can produce conventional audible waves. When the input signal is an ultrasonic signal modulated by an audible signal, it can produce the ultrasound by working near the specific natural frequency of the transducer component. Then, the nonlinear interaction of ultrasound in air can produce directional audible sound. In this study a demonstrator is built for testing the feasibility of the concept. The experiment results indicate that this concept is feasible even though the performance of this demonstrator should be further improved and optimized.
Convention Paper 7681
Paper not presented but available for purchase.

16:00

P5-5 Forces in Cylindrical Metalized Film Audio Capacitors—*Philip J. Duncan*,[1] *Nigel Williams*,[2] *Paul S. Dodds*[2]
[1]University of Salford, Greater Manchester, UK
[2]ICW Ltd., Wrexham, Wales, UK

This paper is concerned with the analysis of forces acting in metalized polypropylene film capacitors in use in loudspeaker crossover circuits. Capacitors have been subjected to rapid discharge measurements to investigate mechanical resonance of the capacitor body and the electrical forces that drive the resonance. The force due to adjacent flat current sheets has been calculated in order that the magnitude of the electro-dynamic force due to the discharge current can be calculated and compared with the electrostatic force due to the potential difference between the capacitor plates. The electrostatic force is found to be dominant by several orders of magnitude, contrary to assumptions in previous work where the electro-dynamic force is assumed to be dominant. The capacitor is then modeled as a series of concentric cylindrical conductors and the distribution of forces within the body of the capacitor is considered. The primary outcome of this is that the electrostatic forces act predominantly within the inner and outer turn of the capacitor body, while all of the forces acting within the body of the capacitor are balanced almost to zero. Experimental results where resonant acoustic emissions have been measured and analyzed are presented and discussed in the context of the model proposed.
Convention Paper 7682

16:30

P5-6 On the Use of Motion Feedback as Used in 4th Order Systems—*Stefan Willems*,[1] *Guido D'Hoogh*[2]
[1]Denon & Marantz Holding, Premium Sound Solutions, Leuven Belgium
[2]Retired

Class D amplification allows the design of compact very high power amplifiers with a high efficiency. Those amplifiers are an excellent candidate for being used in compact high-powered subwoofers. The drawback of compact subwoofers is the nonlinear compression of the air inside the (acoustically) small box. Fourth order systems are beneficial over 2nd order systems due to their increased efficiency. To combine the best of both worlds, 4th order design and acoustically small enclosures, a feedback mechanism has been developed to reduce the nonlinear distortion found in compact high-powered subwoofers. Acceleration feedback on woofer systems is traditionally used in 2nd order systems. This paper discusses the use of an acceleration and velocity feedback system applied to a 4th order system.
Convention Paper 7683

17:00

P5-7 Mapping of the Loudspeaker Emission by the Use of Anemometric Method—*Danijel Djurek*,[1] *Ivan Djurek*,[2] *Antonio Petosic*[2]
[1]Alessandro Volta Applied Ceramics (AVAC), Zagreb, Croatia
[2]University of Zagreb, Zagreb, Croatia

Lateral wire anemometry (LWA) has been developed for recording of air vibration. Standard anemometry is founded upon the hot wire method, and wire temperature changes in the oscillating air velocity in the range 800–1000 °C, which is less suitable because of the proper heat emission from the wire. LWA deals only with the initial slope of the changing wire resistance, and subsequent Fourier analysis enables measurements of periodic air velocity. The probe has been developed for precise mapping of the air velocity field in the front of the membrane, and local power emission of the membrane may be evaluated in the region fitted to 0.15 cm^2.
Convention Paper 7684

17:30

P5-8 Flat Panel Loudspeaker Consisting of an Array of Miniature Transducers—*Daniel Beer*,[1] *Stephan Mauer*,[1] *Sandra Brix*,[1] *Jürgen Peissig*[2]
[1]Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

[2]Sennheiser Electronic GmbH & Co. KG, Wedemark, Germany

Multichannel audio reproduction systems like the Wave Field Synthesis (WFS) use a large number of small and closely spaced loudspeakers. The successful use of WFS requires, among other things, the ability of an "invisible" integration of loudspeakers in a room. Flat panel loudspeakers compared with conventional cone loudspeakers provide advantages in the space saved room integration because of their low manufactured depth. In this way flat panel loudspeakers can be found in furniture, media devices, or like pictures hung on the wall. Besides the integration, flat loudspeakers should provide at least the same good acoustical performance as conventional loudspeakers. This is indeed a problem, because the low depth negatively influences the acoustical quality of reproduction in the lower and middle frequency range. This paper demonstrates a new flat panel loudspeaker consisting of an array of miniature transducers.
Convention Paper 7685

18:00

P5-9 Subwoofer Loudspeaker System with Dynamic Push-Pull Drive—*Drazenko Sukalo*, DSLab–Device Solution Laboratory, Munich, Germany

This paper examines the influence of mutual coupling between two driver-diaphragms driven by two electrical signals, each with a 90° phase shift on the voice-coil impedance curve. A new model of the system is described, and the effects are observed using the electrical circuit simulator PSpice. Finally, predicted and measured values are presented.
Convention Paper 7686

Session P6 Thursday, May 7
14:00 – 15:30 K4 Foyer

POSTERS: MULTICHANNEL CODING

14:00

P6-1 Adaptive Predictive Modeling of Stereo LPC with Application to Lossless Audio Compression—*Florin Ghido, Ioan Tabus*, Tampere University of Technology, Tampere, Finland

We propose a novel method for exploiting the redundancy of stereo linear prediction coefficients by using adaptive linear prediction for the coefficients themselves. We show that an important proportion of the stereo linear prediction coefficients, on both the intrachannel and the interchannel parts, still contains important redundancy inherited from the signal. We can therefore significantly reduce the amplitude range of those LP coefficients by using adaptive linear prediction with orders up to 4, separately on the intrachannel and intrachannel parts. When integrated into asymmetrical OptimFROG, the new method obtains on average 0.29 percent improvement in compression with negligible increase in decoder complexity.
Convention Paper 7666

14:00

P6-2 A Study of MPEG Surround Configurations and Its Performance Evaluation—*Evelyn Kurniawati, Samsudin Ng, Sapna George*, ST Microelectronics Asia Pacific Pte. Ltd., Singapore

The standardization of MPEG Surround in 2007 opens a new range of possibility for low bit rate multichannel audio encoding. While ensuring backward compatibility with legacy decoder, MPEG Surround offers various configurations to upmix to the desired number of channels. The downmix stream, which can be in mono or stereo format, can be passed to transform, hybrid, or any other types of encoder. These options give us more than one possible combination to encode a multichannel stream at a specific bit rate. This paper presents a comparative study between those options in terms of their quality performance that will help us choose the most suitable configuration of MPEG Surround in a range of operating bit rate.
Convention Paper 7667

14:00

P6-3 Lossless Compression of Spherical Microphone Array Recordings—*Erik Hellerud, U. Peter Svensson*, Norwegian University of Science and Technology, Trondheim, Norway

The amount of spatial redundancy for recordings from a spherical microphone array is evaluated using a low delay lossless compression scheme. The original microphone signals, as well as signals transformed to the spherical harmonics domain, are investigated. It is found that the correlation between channels is, as expected, very high for the microphone signals, in several different acoustical environments. For the signals in the spherical harmonics domain, the compression gain from using inter-channel prediction is reduced, since this conversion results in many channels with low energy. Several alternatives for reducing the coding complexity are also investigated.
Convention Paper 7668

Workshop 3 Thursday, May, 7
14:00 – 16:00 Room D123

INTELLIGENT DIGITAL AUDIO EFFECTS

Chair: **Christian Uhle**, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Panelists: *Alexander Lerch*, Z-Plane, Berlin, Germany
Josh Reiss, Queen Mary, University of London, London, UK
Udo Zölzer, Helmut-Schmidt-Universität, Hamburg, Germany

Intelligent Digital Audio Effects (I-DAFx) process audio signals in a signal-adaptive way by using some kind of high-level analysis of the input. Beat tracking, for example, enables the automated adaption of time delays or of the LFO rate in tremolos, auto-wahs, and vibrato effects. A harmonizer can adapt the additional intervals to the melody that is played. Automatic mixing is approached by analyzing the signal content in all channels to control the panning. These are examples of techniques that are ➥

in the scope of this workshop. It presents an overview of I-DAFx and of methods of semantic audio analysis used in these devices. Practical examples are described and sound examples are demonstrated.

Tutorial 3 **Thursday, May 7**
14:00 – 16:30 **Room K1**

DESIGN OF HIGH-PERFORMANCE BALANCED AUDIO INTERFACES

Presenter: **Bill Whitlock**, Jensen Transformers, Inc., Chatsworth, CA, USA

High signal-to-noise ratio is an important goal for most audio systems. However, AC power connections unavoidably create ground voltage differences, magnetic fields, and electric fields. Balanced interfaces, in theory, are totally immune to such interference. For 50 years, virtually all audio equipment used transformers at its balanced inputs and outputs. Their high noise rejection was taken for granted and the reason for it all but forgotten. The transformer's extremely high common-mode impedance—about a thousand times that of its solid-state equivalents —is the reason. Traditional input stages will be discussed and compared. A novel IC that compares favorably to the best transformers will be described. Widespread misunderstanding of the meaning of balance as well as the underlying theory has resulted in all-too-common design mistakes in modern equipment and seriously flawed testing methods. Therefore, noise rejection in today's real-world systems is often inadequate or marginal. Other topics will include tradeoffs in output stage design, effects of non-ideal cables, and the pin 1 problem.

Thursday, May 7 **14:00** **Room D120**
Technical Committee Meeting on Fiber Optics (Formative Meeting)

Thursday, May 7 **15:00** **Room D120**
Technical Committee Meeting on Audio for Games

Thursday, May 7 **16:00** **Room D120**
Technical Committee Meeting on Hearing and Hearing Loss Prevention

Session P7 **Thursday, May 7**
16:30 – 18:00 **K4 Foyer**

POSTERS: SPATIAL AUDIO PROCESSING

16:30

P7-1 **Low Complexity Binaural Rendering for Multichannel Sound—***Kangeun Lee, Changyong Son, Dohyung Kim*, Samsung Advanced Institute of Technology, Suwon, Korea

The current paper is concerned with an effective method to emulate the multichannel sound in a portable environment where low power is required. The goal of this paper is to show the complexity of binaural rendering of the multichannel to stereo sound systems in cases of portable devices. To achieve this, we proposed the modified discrete cosine transform (MDCT) based binaural rendering, combined with the Dolby Digital decoder (AC-3) that is a multichannel audio decoder. A reverberation algorithm is added to the proposed algorithm for closing to real sound. This combined structure is implemented on a DSP processor. The complexity and quality are compared with a conventional head-related transfer function (HRTF) filtering method and Dolby headphone that are the most current in commercial binaural rending technology, demonstrating significant complexity reduction and comparable sound quality to the Dolby headphone.
Convention Paper 7687

16:30

P7-2 **Optimal Filtering for Focused Sound Field Reproductions Using a Loudspeaker Array—** *Youngtae Kim, Sangchul Ko, Jung-Woo Choi, Jungho Kim*, SAIT, Samsung Electronics Co., Ltd., Gyeonggi-do, Korea

This paper describes audio signal processing techniques in designing multichannel filters for reproducing an arbitrary spatial directivity pattern with a typical loudspeaker array. In designing the multichannel filters, some design criteria based on, for example, least-squares methods and the maximum energy array are introduced as non-iterative optimization techniques with a lower computational complexity. The abilities of the criteria are first evaluated with a given loudspeaker configuration for reproducing a desired acoustic property in a spatial area of interest. Also, additional constraints are considered to impose for minimizing the error between the amplitudes of actual and the desired spatial directivity pattern. Their limitations in practical applications are revealed by experimental demonstrations, and finally some guidelines are proposed in designing optimal filters.
Convention Paper 7688

16:30

P7-3 **Single-Channel Sound Source Distance Estimation Based on Statistical and Source-Specific Features —***Eleftheria Georganti*,[1,2] *Tobias May*,[3] *Steven van de Par,*[1] *Aki Härmä*,[1] *John Mourjopoulos*[2]
[1]Philips Research Europe, Eindhoven, The Netherlands
[2]University of Patras, Patras, Greece
[3]Technische Universiteit Eindhoven, Eindhoven, The Netherlands;

In this paper we study the problem of estimating the distance of a sound source from a single microphone recording in a room environment. The room effect cannot be separated from the problem without making assumptions about the properties of the source signal. Therefore, it is necessary to develop methods of distance estimation separately for different types of source signals. In this paper we focus on speech signals. The proposed solution is to compute a number of statistical and source-specific features from the speech signal and to use pattern recognition techniques to develop a robust distance estimator for speech signals. Experiments with a database of real speech recordings showed that the proposed model is capable of estimating source distance with acceptable performance for applications such as ambient telephony.
Convention Paper 7689

16:30

P7-4 **Implementation of DSP-Based Adaptive Inverse Filtering System for ECTF**

Equalization—*Masataka Yoshida, Haruhide Hokari, Shoji Shimada*, Nagaoka University of Technology, Nagaoka, Niigata, Japan

The Head Related Transfer Function (HRTF) and the inverse Ear Canal Transfer Function (ECTF) must be accurately determined if stereo earphones are realized out-of-head sound localization (OHL) with high presence. However, the characteristics of ECTF depend on the type of earphone used and the number of earphone mounting and demounting operations. Therefore, we present a DSP-based adaptive inverse filtering system for ECTF equalization in this paper. The buffer composition and size of DSP were studied so as to implement operation processing. As a result, we succeeded in constructing a system that was able to work in the audio-band of 15 kHz with the sampling frequency of 44.1 kHz. Listening tests clarified that the effective estimation error of the adaptive inverse-ECTF for OHL was less than –11 dB with convergence time of about 0.3 seconds.
Convention Paper 7690

16:30

P7-5 Improved Localization of Sound Sources Using Multi-Band Processing of Ambisonic Components —*Charalampos Dimoulas, George Kalliris, Konstantinos Avdelidis, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The current paper focuses on the use of multi-band ambisonic-processing for improved sound source localization. Energy-based localization can be easily delivered using soundfield microphone pairs, as long as free field conditions and the single omni-directional-point-source model apply. Multi-band SNR-based selective processing improves the noise tolerance and the localization accuracy, eliminating the influence of reverberation and background noise. Band-related sound-localization statistics are further exploited to verify the single or multiple sound-sources scenario, while continuous spectral fingerprinting indicates the potential arrival of a new source. Different sound-excitation scenarios are examined (single /multiple sources, narrow-band / wideband signals, time-overlapping, noise, reverberation). Various time-frequency analysis schemes are considered, including filter-banks, windowed-FFT and wavelets with different time resolutions. Evaluation results are presented.
Convention Paper 7691

16:30

P7-6 Spatial Audio Content Management within the MPEG-7 Standard of Ambisonic Localization and Visualization Descriptions—*Charalampos Dimoulas, George Kalliris, Kostantinos Avdelidis, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The current paper focuses on spatial audio video/imaging and sound field visualization using ambisonic-processing, combined with MPEG-7 description schemes for multi-modal content description and management. Sound localization can be easily delivered using multi-band ambisonic processing under free-field and single point-source excitation conditions, offering an estimate on the achieved accuracy. Sound source forward propagation models can be applied in case that confident localization accuracy has achieved, to visualize the corresponding sound field. Otherwise, 3-D audio/surround sound reproduction simulation can be used instead. In any case, sound level distribution colormap-videos and highlighting images can be extracted. MPEG-7 adapted description schemes are proposed for spatial-audio audiovisual content description and management, facilitating a variety of user-interactive postprocessing applications.
Convention Paper 7692

Workshop 4 **Thursday, May 7**
16:00 – 18:30 **Room K2**

MICROPHONES—WHAT TO LISTEN FOR—WHAT SPECS TO LOOK FOR

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

Panelists: *Jean-Marie Greijsen*, Polyhymnia International
David Josephson, Josephson Engineering
Douglas McKinnie, Middle Tennessee State University
Mikkel Nymand, DPA Microphones
Ossian Ryner, DR, Danish Broadcasting

When selecting microphones for a specific music recording, it is worth knowing what to expect and what to listen for. Accordingly it is good to know what specifications that would be optimum for that microphone. This workshop explains the process of selecting a microphone both from the aesthetical as well as the technical point of view. Also explained and demonstrated: what to expect when placing the microphone. This is not a “I feel like . . .” presentation. All presenters on the panel are serious and experienced engineers and tonmeisters. The purpose of this workshop is to encourage and teach young engineers and students to take advantage by taking a closer look at the specifications the next time they are going to pick a microphone for a job.

Workshop 5 **Thursday, May 7**
16:30 – 18:30 **Room D123**

PROFESSIONAL AUDIO NETWORKING IN SOUND REINFORCEMENT AND BROADCAST APPLICATIONS

[CANCELLED]

Tutorial 4 **Thursday, May 7**
17:00 – 18:30 **Room K1**

THE GROWING IMPORTANCE OF MASTERING IN THE HOME STUDIO ERA

Presenter: **Andrés Mayo**, Andrés Mayo Mastering & DVD, Argentina

Artists and producers are widely using their home studios for music production, with a better cost/benefit ratio. But they usually lack technical resources and the acoustic response of their rooms is unknown. Therefore, there is greater need for a professional mastering service in order to achieve the so-called "standard commercial quality." This tutorial presents a list of common mistakes that can be found in homemade mixes with real-life audio examples taken directly from recent mastering sessions. What can and what can’t be fixed at the mastering stage. ➥

Thursday, May 7 16:00 Room D120
Technical Committee Meeting on Semantic Audio Analysis

Special Event
OPEN HOUSE OF THE TECHNICAL COUNCIL AND THE RICHARD C. HEYSER MEMORIAL LECTURE
Thursday, May 7, 18:30 – 19:30
Room K3

Lecturer: **Gunnar Rasmussen**

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 126th AES Convention is Gunnar Rasmussen, a pioneer in the construction of acoustic instrumentation, particularly of microphones, transducers, vibration and related devices. He was employed at Brüel & Kjær Denmark as an electronics engineer immediately after his graduation in 1950. After holding various positions in development, testing, and quality control, he spent one year in the United States working for Brüel & Kjær in sales and service.

After his return to Denmark in the mid-1950s he began the development of a new measurement microphone. This resulted in a superior mechanical stability, increased temperature, and long term stability. The resulting one-inch pressure microphone soon became the de facto standard microphone for acoustical measurements to replace the famous W.E. 640AA standardized microphone.

The optimized mechanical design of the new generation of measurement microphones opened up the possibility for reducing the size of the microphones, first to a ½” microphone and then to ¼” and 1/8” microphones with essentially the same superior mechanical, temperature and long term stability. Notably the ½” microphone is still the most widely used measurement tool today. Since the beginning of the 1960s, this microphone design has been preferred for all types of acoustic measurements and has formed the basis for the IEC 1094 series of international standards for measurement microphones.

Gunnar Rasmussen received the Danish Design Award in 1969 for his novel design of the microphones that were exhibited at the New York Museum of Modern Art. He also developed the first acoustically optimized sound level meter, where the shape of the body was designed to minimize the effect of reflections from the casing to the microphone. This type 2203 Sound Level meter was for many years seen as the archetype of sound level meters and its characteristic shape became the symbol of a sound level meter.

Other major inventions and designs include the Delta Shear accelerometer, the dual piston pistonphone calibrator for precision calibration, the face-to-face sound intensity probe and hydrophones, occluded ears, artificial mouth, etc. Rasmussen is also the author of numerous papers on acoustics and vibration and has served as chairman and vice-chairman of various international organizations and standard committees. In 1990 he received the CETIM medal for his contribution to the field of intensity techniques. He is also a Fellow of the Acoustical Society of America.

In 1994 Rasmussen started his own company, G.R.A.S. Sound and Vibration. Originally a company specializing in precision Outdoor Microphones for permanent noise monitoring around airports, it is now one of the world’s leading companies in acoustic front-ends and transducers forming a wide range of general purpose and specialized microphones, electro-acoustic measurement devices such as ear couplers, precision calibration tools and multi-dimensional sound intensity probes. The title of his lecture is, “The Reproduction of Sound Starts at the Microphone.”

The microphones may be developed for many specific purposes: for communication, recording or precision measurements. Quality may have different meaning for different applications. Price may be a dominating factor. Carbon microphones were dominating up to the 1950s. Electret microphones have taken the place of carbon microphones with great improvement in quality and performance at low prices. The MEMS microphones are on the way.

The challenge in the high quality microphone development is to match or exceed the human ear in perception of sound for measurement purposes. Without measurements we cannot qualify our progress. We are still trying to match the frequency band, the dynamic range, the phase linearity of the human ear and to obtain very good reproducibility in all situations where humans are involved. We need microphones for development, for standardized measurements and for legal related measurements. Where are we today?

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

Session P8 Friday, May 8
09:00 – 11:30 Room K3

ROOM ACOUSTICS

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

09:00

P8-1 Phase Velocity and Group Velocity in Cylindrical and Spherical Waves—*Ian M. Dash*,[1] *Fergus R. Fricke*[2]
[1]Australian Broadcasting Corporation, Sydney, NSW, Australia
[2]University of Sydney, Sydney, NSW, Australia

Closed-form expressions are derived for phase velocity and group velocity in cylindrical and spherical sound waves. These are plotted and compared for orders 0, 1, and 2, but the expressions are general and may be applied to waves of any order. Dispersion characteristics of these waves are examined and discussed. The implications for thermodynamic applicability of the wave equation and for application of Huygens’ principle are discussed.
Convention Paper 7693

9:30

P8-2 Selection of Loudspeaker Positions for Reverberation Time and Sound Field Measurements—*Elena Prokofieva*, Napier University, Edinburgh, UK

According to the various building standards, the source loudspeakers and receiving microphones during the internal noise level measurements can be placed “in any convenient position,” with just some distance restrictions from the nearest reflecting surfaces. In rooms of different shapes and volumes the location of the source and receiver microphone may significantly affect the measured results. If the difference between the reverberation times or noise levels measured for two positions in the same room exceeds 10 percent, they cannot be averaged. The simulation

program is created to recommend the most suitable locations for microphone and loudspeakers in tested room for reverberation time measurements. The results of series of tests are analyzed to confirm the results of the simulation.
Convention Paper 7694

10:00

P8-3 A Rehearsal Hall with Virtual Acoustics for Symphony Orchestras—*Tapio Lokki,*[1] *Jukka Pätynen,*[1] *Timo Peltonen,*[2] *Olli Salmensaari*[2]
[1]Helsinki University of Technology, Espoo, Finland
[2]Akukon Consulting Engineers Ltd., Helsinki, Finland

A solution for constructing a small rehearsal hall, the acoustics of which resembles the stage of a large concert hall is presented. The implemented system was evaluated both objectively with measurements and subjectively by collecting feedback from musicians. The subjective opinions were very positive and encouraging and the main objective was achieved. The electroacoustically enhanced rehearsal space sounded like a much bigger hall, although the sound pressure level increased less than one decibel. The presented solution is applicable in all spaces, which are not very reverberant by nature and where the height of the room is at least twice the standard room height.
Convention Paper 7695

10:30

P8-4 Sound Field Characterization and Absorption Measurement of Wideband Absorbers— *Soledad Torres-Guijarro,*[1] *Antonio Pena,*[2] *Alfonso Rodríguez-Molares,*[2] *Norberto Degara-Quintela*[2]
[1]Laboratorio Oficial de Metroloxía de Galicia (LOMG), Ourense, Spain
[2]Universidad de Vigo, Vigo, Spain

Wideband absorbers are a fundamental part of non-environment control rooms. They consist of huge angled hanging panels in conjunction with a multilayer wall or ceiling. Their absorption capacity is very noticeable, mostly in the low frequency range. In this paper the mechanisms of absorption of the wideband absorbers of the rear wall of the control room at the Universidad de Vigo will be studied. Conclusions will be drawn from the analysis of pressure, velocity volume, and intensity measurements performed in the vicinity of the panels, and from the computation of the normal specific acoustic impedance and the normal absorption coefficient.
Convention Paper 7696

11:00

P8-5 Temporal Matching of 2-D and 3-D Wave-Based Acoustic Modeling for Efficient and Realistic Simulation of Rooms—*Jeremy J. Wells, Damian T. Murphy, Mark Beeson,*
University of York, York, UK

Methods for adapting the output of a two-dimensional Kirchoff-variable digital waveguide mesh to better match that of a 3-D mesh, both of which are intended to model the same acoustic space, are presented. Details of the methods, including quality of output and computational demands, are given along with the details of how they are incorporated into the hybrid system within which they are employed.
Convention Paper 7697

Session P9 — Friday, May 8
09:00 – 12:30 — Room K4

SIGNAL ANALYSIS, MEASUREMENTS, RESTORATION

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

09:00

P9-1 Some Improvements of the Playback Path of Wire Recorders—*Nadja Wallaszkovits,*[1] *Heinrich Pichler*[2]
[1]Phonogrammarchiv Austrian Academy of Sciences, Vienna, Austria
[2]Audio Consultant, Vienna, Austria

The archival transfer of wire recordings to the digital domain is a highly specialized process that incorporates a wide range of specific challenges. One of the basic problems is the format incompatibility between different manufacturers and models. The paper discusses the special design philosophy, using the tone control network in the record path as well as in the playback path. This tone control circuit causes additional phase and group delay distortions. The influence and characteristics of the tone control (which was not a priori present with every model) is discussed and analog phase correction networks are described. The correction of phase errors is outlined. As this format has been obsolete for many decades, a high quality archival transfer can only be reached by modifying dedicated equipment. The authors propose some possible main modifications and improvements of the playback path of wire recorders, such as signal pickup directly after the playback head, introducing a high quality preamplifier, followed by analog phase correction and correction of the amplitude characteristics. Alternatively signal pickup directly after the playback head, introducing a high quality preamplifier, followed by digital signal processing to optimize the output signal is discussed.
Convention Paper 7698

09:30

P9-2 Acoustics of the Crime Scene as Transmitted by Mobile Phones—*Eddy B. Brixen,*
EBB-consult, Smorum, Denmark

One task for the audio forensics engineer is to extract background information from audio recordings. A major problem is the assessment of analyzed telephone calls in general and mobile phones (LPC-algorithms) in particular. In this paper the kind of acoustic information to be extracted from a recorded phone call is initially explained. The parameters used for the characterization of the various acoustic spaces and events in question are described. It is discussed how the acoustical cues should be assessed. The validity of acoustic analyses carried out in the attempt to provide crime scene information like reverberation time is presented.
Convention Paper 7699

10:00

P9-3 Silence Sweep: A Novel Method for Measuring Electroacoustical Devices—*Angelo Farina*, University of Parma, Parma, Italy

This paper presents a new method for measuring some properties of an electroacoustical system, for example a loudspeaker or a complete sound system. Coupled with the already established method based on Exponential Sine Sweep, this new Silence Sweep method provides a quick and complete characterization of not linear distortions and noise of the device under test. The method is based on the analysis of the distortion products, such as harmonic distortion products or intermodulation effects, occurring when the system is fed with a wideband signal. Removing from the test signal a small portion of the whole spectrum, it becomes possible to collect and analyze the not-linear response and the noise of the system in that "suppressed" band. Changing continuously the suppressed band over time, we get the Silence Sweep test signal, which allows for quick measurement of noise and distortion over the whole spectrum. The paper explains the method with a number of examples. The results obtained for some typical devices are presented, compared with those obtained with a standard, state-of-the-art measurement system.
Convention Paper 7700

10:30

P9-4 Pitch and Played String Estimation in Classic and Acoustic Guitars—*Isabel Barbancho, Lorenzo Tardón, Ana M. Barbancho, Simone Sammartino*, Universidad de Málaga, Málaga, Spain

In classic and acoustic guitars that use standard tuning, the same pitch can be produced at different strings. The aim of this paper is to present a method based on the time and frequency-domain characteristics of the recorded sound to determine, not only the pitch but also the string of the guitar that has been played to produce that pitch. This system will provide information not only of the pitch of the notes played, but also about how those notes were played. This specific information can be valuable to identify the style of the player and can be used in teaching to play the guitar.
Convention Paper 7701

11:00

P9-5 Statistical Properties of Music Signals—*Miomir Mijic, Drasko Masovic, Dragana Sumarac-Pavlovic*, Faculty of Electrical Engineering, Belgrade, Serbia

This paper is concerned with the results of a complex approach to statistical properties of various music signals based on 412 musical pieces classified in 12 different genres. Analyzed signals contain more than 24 hours of music. For each piece time variation of the signal level was found, performed with a 10 ms period of integration in rms calculation and with 90 percent overlap, making a new signal representing the level as a function of time. For each piece the statistical analysis of signal level has been performed by its statistical distribution, cumulative distribution, effective value within complete duration of piece, mean level value, and level value corresponding to maximum of the statistical distribution. The parameter L1, L10, L50, and L99 were extracted from cumulative distributions as numerical indicators of dynamic properties. The paper contains detailed statistical data and averaged data for all observed genres, as well as quantitative data about dynamic range and crest factor of various music signals.
Convention Paper 7702

11:30

P9-6 Multi-Band Generalized Harmonic Analysis (MGHA) and its Fundamental Characteristics in Audio Signal Processing—*Takahiro Miura, Teruo Muraoka, Tohru Ifukube*, University of Tokyo, Tokyo, Japan

One of the main problems in sound restoration of valuable historical recordings includes the noise reduction. We have been proposing and continuing to improve the noise reduction method utilized by inharmonic analysis such as GHA (Generalized Harmonic Analysis). Algorithm of GHA frequency extraction enables us to extract arbitrary frequency components. In this paper we aimed at more accurate frequency identification from noisy signals to divide analyzed frequency section into multi-bands before analysis: this algorithm is named as Multi-Band GHA (MGHA). The simulation of frequency analysis in a noise-free condition indicated that MGHA is more effective than GHA for the extraction of low frequency components in the condition of both lower window length and amount of frequency components. However, excluding the case of both lower window length and amount of frequency components, GHA identifies frequency components more precisely. Furthermore the result of frequency analysis in condition with steady noise shows that MGHA can be more effectively applied to the case of short window length, many frequency components, and low S/N.
Convention Paper 7703

12:00

P9-7 Automatic Detection of Salient Frequencies—*Joerg Bitzer*,[1] *Jay LeBoeuf*[2]
[1]University of Applied Science Oldenburg, Oldenburg, Germany
[2]Imagine Research, Inc., San Francisco, CA, USA

In this paper we present several techniques to find the most significant frequencies in recorded audio tracks. These estimated frequencies could be used as a starting point for mixing engineers in the EQing process. In order to evaluate the results, we compare the detected frequencies with a list of reported salient frequencies from audio engineers. The results show that automatic detection is possible. Thus, one of the more boring tasks of a mixing engineer can be automated, which gives the mixing engineer more time to do the artistic part of the mixing process.
Convention Paper 7704

Workshop 6 **Friday, May 8**
09:00 – 11:00 **Room K1**

THE OPERA BEHIND THE OPERA—CHALLENGES AND REAL WORLD SOLUTIONS IN TRANSMISSIONS OF BIG OPERA PRODUCTIONS

Presenters: **Billy Henningsen**, Senior Sound Engineer, Norwegian TV, Oslo, Norway
Joseph Schütz, Senior Engineer, Austrian Radio, Vienna, Austria
Peter Urban, Senior Engineer, Bavarian Radio, Munich, Germany

In this workshop the aesthetics and concepts of an ideal opera sound will be introduced, theoretically. Be swept away by the realities of awkward acoustics, directors and producers with challenging stage directions and last minute ideas. Both Bayreuth and Vienna, the new opera of Oslo, and all the other big stages have had their transmissions and you will hear how the challenges have been met. In stereo and surround.

Workshop 7 **Friday, May 8**
09:00 – 11:00 **Room K2**

AUDIO CODING FOR BROADCASTING—STATUS REPORT

Chair: **Gerhard Stoll**, IRT Munich

Panelists: *Bjørn Aarseth*, NRK, Norway
Udo Appel, Bayerischer Rundfunk, Germany
Dave Marston, BBC, UK
Roger Miles, EBU Switzerland
Thomas Saner, SRG SSR idée suisse, Switzerland
Karl M. Slavik, Broadcast Consultant Dolby Laboratories, Inc./Artecast, Austria

Both Digital Radio and Digital TV are relying on new, efficient audio codecs in order to supply the listener with lots of audio channels. Lots of audio channels may have a different meaning for TV and radio. Whereas TV, in particular HDTV, is going toward mainly multichannel audio, digital radio, here in particular terrestrial radio, is aiming for more and more radio programs with stereo as the main audio format. In this context several questions need to be answered, e.g.:

- Do we have the appropriate audio coding systems for stereo and multichannel audio?
- What kind of audio quality do we need together with HDTV?
- Is 5.1 multichannel audio still the suitable format for HDTV?
- What happens with cascading multichannel audio codecs, which occur in real-life situations?
- Is broadcast quality, as defined by EBU in 1990, when DAB had been developed, still an issue for radio broadcasting?
- How many radio programs can be provided by new radio broadcast systems, e.g., DAB+?
- What is the role of 5.1 multichannel audio in radio broadcasting?

Several broadcast experts will talk about their experience with low bit-rate audio in a modern broadcast infrastructure and will try to provide answers to the questions above.

Tutorial 5 **Friday, May 8**
09:00 – 10:30 **Room D123**

LIPSYNC IN THE LATENCY AGE

Presenter: **Friedrich Gierlinger**, IRT Munich

Synchronization of audio and video through the whole broadcast chain became ever more challenging with the switch to fully digital production methods, wireless picture acquisition practices, omnipresent different processing latency, and an unavoidable delay in consumer displays that may or may not be able to be compensated. This tutorial systematically examines the various reasons why lipsync problems occur, quantifies the possible error at any given stage, and recommends best practices to improve the situation.

Friday, May 8 **09:00** **Room D120**
Technical Committee Meeting on Spatial Audio

Friday, May 8 **09:00** **Room D104**
Standards Committee Meeting on SC-02-08 Audio-File Transfer and Exchange

Student Event/Career Development
MENTORING SESSIONS
Friday, May 8, 10:00 – 12:00
K4 Foyer

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.

Friday, May 8 **10:00** **Room D120**
Technical Committee Meeting on Signal Processing

Session P10 **Friday, May 8**
10:30 – 12:00 **K4 Foyer**

POSTERS: AUDIO FOR TELECOMMUNICATIONS

10:30

P10-1 Harmonic Representation and Auditory Model-Based Parametric Matching and its Application in Speech/Audio Analysis— *Alexey Petrovsky,*[1] *Elias Azarov*,[1] *Alexander Petrovsky*[2]
[1]Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus
[2]Bialystok Technical University, Bialystok, Poland

The paper presents new methods for the selection of sinusoids and transients components in hybrid sinusoidal modeling of speech/audio. The instantaneous harmonic parameters (magnitude, frequency, and phase) are calculated as the result of the narrow band filtering of speech/audio. The frequency-modulated filters synthesis with the closed form impulse response has been proposed. The filter frequency bounds can be determined during the components frequency tracking and can be adjusted according to the fundamental frequency modulations. It can be implemented speech/audio harmonic/noise decomposition. The transient components modeling are presented by matching pursuit with frame-based psychoacoustic optimized wavelet packet dictionary. The choice of most relevant coefficients is based on maximizing the matching between the auditory excitation scalograms of original and modeled signals.
Convention Paper 7705

10:30

P10-2 Perceptual Compression Methods for Metadata in Directional Audio Coding Applied to Audiovisual Teleconference—
Toni Hirvonen,[1] *Jukka Ahonen,*[2] *Ville Pulkki*[2]
[1]Institute of Computer Science (ICS) of the Foundation for Research and Technology, Hellas, Greece
[2]TKK, Espoo, Finland

In teleconferencing application of Directional Audio Coding, the transmitted data consists of monophonic audio signal and directional metadata measured in frequency bands depending on time. In reproduction, each frequency channel of the signal is reproduced to corresponding direction with corresponding diffuseness. This paper examines methods for reducing the data rate of the metadata. The compression methods are based on psychoacoustic studies about the accuracy of directional hearing, and further developed and validated. Informal tests with one-way reproduction, as well as usability testing where an actual teleconference was arranged, were utilized for this purpose. The results indicate that the data rate can be as low as approximately 3 kbit/s without a significant loss in the reproduced spatial quality.
Convention Paper 7706

10:30

P10-3 Speaker Detection and Separation with Small Microphone Arrays—*Maximo Cobos, Jose J. Lopez, David Martinez*, Universidad Politécnica de Valencia, Valencia, Spain

Small microphone arrays are desirable for many practical speech processing applications. In this paper we describe a system for detecting several sound sources in a room and enhancing a predominant target source using a pair of close microphones. The system consists of three main steps: time-frequency processing of the input signals, source localization via model fitting, and time-frequency masking for interference reduction. Experiments and results using recorded signals in real scenarios are discussed.
Convention Paper 7707

10:30

P10-4 Directional Audio Coding with Stereo Microphone Input—*Jukka Ahonen,*[1] *Ville Pulkki,*[1] *Fabian Kuech,*[2] *Giovanni Del Galdo,*[2] *Markus Kallinger,*[2] *Richard Schultz-Amling*[2]
[1]TKK,Espoo, Finland
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

The use of stereo microphone configuration as input to teleconference application of Directional Audio Coding (DirAC) is presented. DirAC is a method for spatial sound processing, in which the direction of the arrival of sound and diffuseness are analyzed and used for different purposes in reproduction. So far, omnidirectional microphones arranged in an array have been used to generate input signals for one- and two-dimensional sound field analysis in DirAC processing. In this study the possibility to use domestic stereo microphones with DirAC analysis is investigated. Different methods to derive omnidirectional and dipole signals from stereo microphones for directional analysis are presented and their applicability is discussed.
Convention Paper 7708

10:30

P10-5 Robust Noise Reduction Based on Stochastic Spatial Features—*Mitsunori Mizumachi*, Kyushu Institute of Technology, Fukuoka, Japan

This paper proposes a robust noise reduction method relying on stochastic spatial features. Almost all of noise reduction methods have both strong and weak sides in the real world. In this paper time evolution of direction of arrival (DOA) and its stochastic reliability are the clues for selecting a suitable approach of noise reduction under time-variant noisy environments, where a DOA is an important spatial feature in beamforming for noise reduction. On the other hand, single channel approaches for noise reduction may be reasonable when DOA estimates are not reliable. Then, either spectral subtraction or beamforming is selected out for achieving robust noise reduction depending on a DOA estimate and its reliability. The proposed method had an advantage in noise reduction compared with a conventional approach.
Convention Paper 7709

Tutorial 6 **Friday, May 8**
11:00 – 13:30 **Room D123**

LOUDNESS—LIGHT AT THE END OF THE TUNNEL

Chair: **Florian Camerer**, ORF, EBU Group P/LOUD

Presenters: *Eelco Grimm*, Dutch Loudness Committee
Mike Kahsnitz, RTW
Ralph Kessler, Pinguin Engineering
Thomas Lund, tc electronic
Andrew Mason, BBC R&D

Audio levels in broadcasting have become increasingly diverse and different over the last decades. Despite clear guidelines and recommended practices the general use of peak measurement in audio metering and the development of more and more sophisticated level processors have led to over-compression of audio signals with the questionable aim of being louder than the competition. This attitude has especially impacted the audio quality of advertisements and promos with very little dynamic range. Already considered a hopeless situation, the introduction of loudness level metadata and especially the introduction of an international standard of loudness measurement (ITU-R BS.1770) is a light at the end of the tunnel. A few broadcasters and even whole countries have addressed the loudness issue thoroughly, and their experience shows that it is possible to solve that problem to the advantage of the consumer. It is long overdue to establish a new paradigm in audio leveling: the switch from peak normalization to loudness normalization. With widespread adoption of this approach consistent loudness not only within a channel, but also between different channels will be within reach—thus finding the “Holy Grail” of audio broadcasting.

In this session the current situation from the perspective of the EBU Group “P/LOUD” will be examined. Vendors will present their approaches to loudness metering.

Friday, May 8 11:00 Room D120
Technical Committee Meeting on Microphones and Applications

Friday, May 8 11:00 Room D105
Standards Committee Meeting on SC-03-02 Transfer Technologies

Exhibitor Seminar Friday, May 8
11:00 – 12:00 Exhibit Floor Booth 2303
ES1 OPTOCORE – A

TECHNICAL INTRODUCTION TO OPTOCORE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

The OPTOCORE Synchronous Fibre Network is introduced. Content will include an Introduction/Overview, design of the network, synchronicity, latency, topology, redundancy, protocol, and devices.

Workshop 8 Friday, May 8
11:30 – 13:30 Room K1

WE HAVE YOU SURROUNDED—MASTERING FOR MULTICHANNEL

Chair: **Darcy Proper**, Senior Mastering Engineer, Galaxy Studios, Mol, Belgium

Panelists: *Simon Heyworth*, Owner/Chief Engineer, Super Audio Mastering, Devon, UK
Thor Levgold, Owner/Chief Engineer, Sonovo Mastering, Stavanger, Norway

With the increasing adoption of home cinema systems, DVD releases of performing artists as well as the EBU specifying 5.1 surround audio as standard for HDTV, the need for professional mastering in multichannel formats is growing. For many, working in surround represents a change of paradigm and brings with it unique challenges and requirements. The panelists in this workshop will share some of the types of challenges they face in the course of their work and some practical tips for handling them when they arise in stereo, surround, and multimedia productions.

Tutorial 7 Friday, May 8
11:30 – 13:30 Room K2

BINAURAL AUDIO TECHNOLOGY—HISTORY, CURRENT PRACTICE, AND EMERGING TRENDS

Presenter: **Robert Schulein**, RBS Consultants

During the winter and spring of 1931-32, Bell Telephone Laboratories, in cooperation with Leopold Stokowski and the Philadelphia Symphony Orchestra, undertook a series of tests of musical reproduction using the most advanced apparatus obtainable at that time. The objectives were to determine how closely an acoustic facsimile of an orchestra could be approached using both stereo loudspeakers and binaural reproduction. Detailed documents discovered within the Bell Telephone archives will serve as a basis for describing the results and problems revealed while creating the binaural demonstrations. Since these historic events, interest in binaural recording and reproduction has grown in areas such as sound field recording, acoustic research, sound field simulation, audio for electronic games, music listening, and artificial reality. Each of these technologies has its own technical concerns involving transducers, environmental simulation, human perception, position sensing, and signal processing. This tutorial will cover the underlying principles germane to binaural perception, simulation, recording, and reproduction. It will include live demonstrations as well as recorded audio/visual examples.

Friday, May 8 11:30 Room D104
Standards Committee Meeting on SC-02-12 Audio Applications of IEEE 1394

Exhibitor Seminar Friday, May 8
11:30 – 12:30 Room D124
ES2 KRK (USA)

ROOM OPTIMIZATION WITH ERGO AND STUDIO MONITOR DESIGN

Presenter: *Nils Karsten*

Every room has his own characteristics. Especially these days it is important, to optimize the acoustics of the listening environment. An acoustically well treated environment is needed to ensure a well translating mix. ERGO makes any monitor system more accurate by reducing "bad" room sound. Besides the room acoustic also the speaker design has a fundamental impact on the radiation of the sound. The KRK design is known to built bullet proof tools for the recording and mastering engineer to ensure to receive a well translating mix.

Friday, May 8 12:00 Room D120
Technical Committee Meeting on Loudspeakers and Headphones

Exhibitor Seminar Friday, May 8
12:00 – 13:00 Exhibit Floor Booth 2303
ES3 OPTOCORE – B

BASICS OF FIBRE OPTICS

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

Technical, economical, safety aspects, and advantages of fibre optics is the topic of this seminar. Content will include fibre advantages, reliability, types of fibre, parameters of fibres (attenuation, bandwidth length product, latency), connector types, cleaning, and testing.

Friday, May 8 12:30 Room D105
Standards Committee Meeting on SC-03-04 Storage and Handling of Media

Session P11 Friday, May 8
13:00 – 16:30 Room K4

AUDIO CODING

Chair: **Nick Zacharov**, DELTA, Hørsholm, Denmark ➥

13:00

P11-1 A Novel Scheme for Low Bit Rate Unified Speech and Audio Coding—MPEG RM0—
Max Neuendorf,[1] *Philippe Gournay,*[2] *Markus Multrus,*[1] *Jérémie Lecomte,*[1] *Bruno Bessette,*[2] *Ralf Geiger,*[1] *Stefan Bayer,*[1] *Guillaume Fuchs,*[1] *Johannes Hilpert,*[1] *Nikolaus Rettelbach,*[1] *Frederik Nagel,*[1] *Julien Robilliard,*[1] *Redwan Salami,*[3] *Gerald Schuller,*[4] *Roch Lefebvre,*[2] *Bernhard Grill*[1]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany;
[2]Université de Sherbrooke, Sherbrooke, Quebec, Canada
[3]VoiceAge Corporation, Montreal, Quebec, Canada
[4]Fraunhofer IDMT, Ilmenau, Germany

Coding of speech signals at low bit rates, such as 16 kbps, has to rely on an efficient speech reproduction model to achieve reasonable speech quality. However, for audio signals not fitting to the model this approach generally fails. On the other hand, generic audio codecs, designed to handle any kind of audio signal, tend to show unsatisfactory results for speech signals, especially at low bit rates. To overcome this, a process was initiated by ISO/MPEG, aiming to standardize a new codec with consistent high quality for speech, music, and mixed content over a broad range of bit rates. After a formal listening test evaluating several proposals MPEG has selected the best performing codec as the reference model for the standardization process. This paper describes this codec in detail and shows that the new reference model reaches the goal of consistent high quality for all signal types.
Convention Paper 7713

13:30

P11-2 A Time-Warped MDCT Approach to Speech Transform Coding—*Bernd Edler,*[1] *Sascha Disch,*[1] *Stefan Bayer,*[2] *Guillaume Fuchs,*[2] *Ralf Geiger*[2]
[1]Leibniz Universität Hannover, Hannover, Germany
[2]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

The modified discrete cosine transform (MDCT) is often used for audio coding due to its critical sampling property and good energy compaction, especially for harmonic tones with constant fundamental frequencies (pitch). However, in voiced human speech the pitch is time-varying and thus the energy is spread over several transform coefficients, leading to a reduction of coding efficiency. The approach presented herein compensates for pitch variation in each MDCT block by application of time-variant re-sampling. A dedicated signal adaptive transform window computation ensures the preservation of the time domain aliasing cancellation (TDAC) property. Re-sampling can be designed such that the duration of the processed blocks is not altered, facilitating the replacement of the conventional MDCT in existing audio coders.
Convention Paper 7710

14:00

P11-3 A Phase Vocoder Driven Bandwidth Extension Method with Novel Transient Handling for Audio Codecs—*Frederik Nagel,*[1] *Sascha Disch,*[2] *Nikolaus Rettelbach*[1]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]Leibniz Universitaet Hanover, Hanover, Germany

Storage or transmission of audio signals is often subject to strict bit-rate constraints. This is accommodated by audio encoders that encode the lower frequency part in a waveform preserving way and approximate the high frequency signal from the lower frequency data by using a set of reconstruction parameters. This so called bandwidth extension can lead to roughness and other unpleasant auditory sensations. In this paper the origin of these artifacts is identified, and an improved bandwidth extension method called Harmonic Bandwidth Extension (HBE) is outlined avoiding auditory roughness in the reconstructed audio signal. Since HBE is based on phase vocoders, and thus intrinsically not well suited for transient signals, an enhancement of the method by a novel transient handling approach is presented. A listening test demonstrates the advantage of the proposed method over a simple phase vocoder approach.
Convention Paper 7711

14:30

P11-4 Efficient Cross-Fade Windows for Transitions between LPC-Based and Non-LPC-Based Audio Coding—*Jérémie Lecomte,*[1] *Philippe Gournay,*[2] *Ralf Geiger,*[1] *Bruno Bessette,*[2] *Max Neuendorf*[1]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]Université de Sherbrooke, Sherbrooke, Quebec, Canada;

The reference model selected by MPEG for the forthcoming unified speech and audio codec (USAC) switches between a non-LPC-based coding mode (based on AAC) operating in the transform-domain and an LPC-based coding (derived from AMR-WB+) operating either in the time domain (ACELP) or in the frequency domain (wLPT). Seamlessly switching between these different coding modes required the design of a new set of cross-fade windows optimized to minimize the amount of overhead information sent during transitions between LPC-based and non-LPC-based coding. This paper presents the new set of windows that was designed in order to provide an adequate trade-off between overlap duration and time/frequency resolution, and to maintain the benefits of critical sampling through all coding modes.
Convention Paper 7712

15:00

P11-5 Low Bit-Rate Audio Coding in Multichannel Digital Wireless Microphone Systems—
Stephen Wray, APT Licensing Ltd., Belfast, Northern Ireland, UK

Despite advances in voice and data communications in other domains, sound production for live events (concerts, theater, conferences, sports, worship, etc.) still largely depends on spectrum-inefficient forms of analog wireless microphone technology. In these live scenarios, low-latency transmission of high-quality audio is mission critical. However, while demand increases for wireless audio channels (for microphones, in-ear monitoring, and talkback systems), some of the radio bands available for "Program Making and Special Events" are to be re-assigned for new wireless mobile telephony and Internet connectivity services: the FCC recently decided to permit so-called White Space Devices to operate in sections of UHF spectrum previously reserved for shared use by analog TV and wireless microphones. This paper examines the key performance aspects of low bit-rate audio codecs for the next generation of bandwidth-efficient digital wireless microphone systems that meet the future needs of live events.
Convention Paper 7714

15:30

P11-6 Krasner's Audio Coder Revisited—*Jamie Angus, Chris Ball, Thomas Peeters, Rowan Williams*, University of Salford, Salford, Greater Manchester, UK

An audio compression encoder and decoder system based on Krasner's work was implemented. An improved Quadrature Mirror Filter tree, which more closely approximates modern critical band measurements, splits the input signal into sub bands that are encoded using both adaptive quantization and entropy coding. The uniform adaptive quantization scheme developed by Jayant was implemented and enhanced through the addition of non-uniform quantization steps and look ahead. The complete codecs are evaluated using the perceptual audio evaluation algorithm PEAQ and their performance compared to equivalent MPEG-1 Layer III files. Initial, limited, tests reveal that the proposed codecs score Objective Difference Grades close to or even better than MPEG-1 Layer III files encoded at a similar bit rate.
Convention Paper 7715

16:00

P11-7 Inter-Channel Prediction to Prevent Unmasking of Quantization Noise in Beamforming—*Mauri Väänänen*, Nokia Research Center, Tampere, Finland

This paper studies the use of inter-channel prediction for the purpose of preventing or reducing the risk of noise unmasking when beamforming type of processing is applied to quantized microphone array signals. The envisaged application is the re-use and postprocessing of user-created content. Simulations with an AAC coder and real-world recordings using two microphones are performed to study the suitability of two existing coding tools for this purpose: M/S stereo coding and the AAC Long Term Predictor (LTP) tool adapted for inter-channel prediction. The results indicate that LTP adapted for inter-channel prediction often gives more coding gain than mere M/S stereo coding, both in terms of signal-to-noise ratio and perceptual entropy.
Convention Paper 7716

Friday, May 8 13:00 Room D120
Technical Committee Meeting on Audio for Telecommunications

Session P12 Friday, May 8
13:30 – 15:00 K4 Foyer

POSTERS: LOUDSPEAKERS

13:30

P12-1 Reduction of Distortion in Conical Horn Loudspeakers at High Levels—*Sverre Holm*,[1] *Rune Skramstad*[2]
[1]University of Oslo, Oslo, Norway
[2]Paragon Arrays, Drammen, Norway

Many horns have audible distortion at high levels. We measured a horn consisting of 6 conical sections with a 10-inch element at 99 dB SPL. A closed back gave maximum 2.4 percent second harmonic and 3.4 percent third harmonic distortion in the 100–1000 Hz range, while an open construction had 1.25 percent and 0.6 percent. A new semi-permeable back chamber reduced this to 0.7 percent and 0.35 percent. We hypothesize that the distortion is partly due to the nonlinear compliance of air in the back chamber, and partly is due to the element's interaction with the front and back loading of the horn, and that the new construction loads the element in a more optimal way.
Convention Paper 7717

13:30

P12-2 Comparison of Different Methods for the Subjective Sound Quality Evaluation of Compression Drivers—*José Martínez*,[1] *Joan Croañes*,[2] *Jorge Francés Monllor*,[3] *Jaime Ramis*[3]
[1]Acustica Beyma S.L., Valencia, Spain
[2]Escola Politecnica Superior de Gandia, Valencia, Spain
[3]Universidad de Alicante, Alicante, Spain

In this paper an approach to the problem of sound quality evaluation of radiating systems is considered, applying a perceptual model. One of the objectives is to use the parameter proposed by Moore to test if it provides satisfactory results when it is applied to the quality evaluation of indirect radiation loudspeakers. Three compression drives have been used for these proposals. Recordings with different test signals at different input voltages have been done. Using this experimental base, an approach to the problem from different points of view is done: Taking in consideration classic sound quality parameters such as roughness, sharpness, and tonality. Applying the parameter suggested by Moore obtained from the application of a perceptual model. Moreover, a psy-

choacoustic experiment has been made on a population of 25 people. The results, although preliminary and strongly dependant on the reference signal used to obtain Rnonlin, show a good correlation with the Rnonlin values.
Convention Paper 7718

13:30

P12-3 Membrane Modes in Transducers with the Direct D/A Conversion—*Libor Husník*, Czech Technical University in Prague, Prague, Czech Republic

Operating principle of systems with the direct acoustic D/A conversion, which are sometimes called digital loudspeakers, brings new features to the field of transducer design. There are many design possibilities to these systems, using different transduction principles and spatial arrangement of constituting parts. This paper deals with the single-acting condenser transducer, suitable for micromachining applications, in which the membrane is driven by a partitioned back electrode. While in conventional transducers the electric force between the back electrode and the membrane is evenly distributed, in digital transducers it is no longer the case. Consequences to membrane vibrations for some cases of excitation by various distributions of forces representing given binary combinations from the dynamic level are presented.
Convention Paper 7719

13:30

P12-4 Increasing Active Radiating Factor of High-Frequency Horns by Using Staggered Arrangement in Loudspeaker Line Array— *Kang An, Yong Shen, Aiping Zhang*, Nanjing University, Nanjing, China

Active Radiating Factor (ARF) is an important parameter to analyze the loudspeaker line array when considering the gap between each two transducers, especially for high-frequency horns. As ARF is desired to be as high as possible, the staggered arrangement of horns is introduced in this paper. The responses in vertical direction and horizontal direction are analyzed. Compared with the conventional arrangement, the negative effects of gaps are reduced and responses are improved in simulation.
Convention Paper 7720

Exhibit Seminar — Friday, May 8
13:00 – 14:00 — Exhibit Floor Booth 2303
ES4 OPTOCORE – C

OPTOCORE IN PRACTICE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

This seminar will cover practical exercises with OPTOCORE system/devices. Content will include practical setup of a network, remote pre-amp control through third party devices, transport of control data and Ethernet, software control, and firmware upgrades.

Student Event/Career Development
RECORDING COMPETITION—STEREO
Friday, May 8, 13:30 – 17:00
Room K2

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 Saturday.

Judges include: Jim Anderson, Akira Fukada, Andres Mayo, Ronald Prent, and Darcy Proper.

Session P13 — Friday, May 8
14:00 – 18:30 — Room K3

SPATIAL AUDIO AND SPATIAL PERCEPTION

Chair: **Tapio Lokki**, Helsinki University of Technology, Espoo, Finland

14:00

P13-1 Evaluation of Equalization Methods for Binaural Signals—*Zora Schärer, Alexander Lindau*, TU Berlin, Berlin, Germany

The most demanding test criterion for the quality of binaural simulations of acoustical environments is whether they can be perceptually distinguished from a real sound field or not. If the simulation provides a natural interaction and sufficient spatial resolution, differences are predominantly perceived in terms of spectral distortions due to a non-perfect equalization of the transfer functions of the recording and reproduction systems (dummy head microphones, headphones). In order to evaluate different compensation methods, several headphone transfer functions were measured on a dummy head. Based upon these measurements, the performance of different inverse filtering techniques re-implemented from literature was evaluated using auditory measures for spectral differences. Additionally, an ABC/HR listening test was conducted, using two different headphones and two different audio stimuli (pink noise, acoustical guitar). In the listening test, a real loudspeaker was directly compared to a binaural simulation with high spatial resolution, which was compensated using seven different equalization methods.
Convention Paper 7721

14:30

P13-2 Crosstalk Cancellation between Phantom Sources —*Florian Völk, Thomas Musialik, Hugo Fastl*, Technical University of München, München, Germany

This paper presents an approach using phantom sources (resulting from the so-called summing localization of two loudspeakers) as sources for crosstalk cancellation (CTC). The phantom sources can be rotated synchronously with the

listener's head, thus demanding significantly less processing power than traditional approaches using fixed CTC loudspeakers, as an online re-computation of the CTC filters is (under certain circumstances) not necessary. First results of localization experiments show the general applicability of this procedure.
Convention Paper 7722

15:00

P13-3 Preliminary Evaluation of Sweet Spot Size in Virtual Sound Reproduction Using Dipoles— *Yesenia Lacouture Parodi, Per Rubak*, Aalborg University, Aalborg, Denmark

In a previous study, three crosstalk cancellation techniques were evaluated and compared under different conditions. Least square approximations in frequency and time domain were evaluated along with a method based on minimum-phase approximation and a frequency independent delay. In general, the least square methods outperformed the method based on minimum-phase approximation. However, the evaluation was only done for the best-case scenario, where the transfer functions used to design the filters correspond to the listener's transfer functions and his/her location and orientation relative to the loudspeakers. In this paper we present a follow-up evaluation of the performance of the three inversion techniques when these conditions are violated. A setup to measure the sweet spot of different loudspeaker arrangements is described. Preliminary measurement results are presented for loudspeakers placed at the horizontal plane and an elevated position, where a typical 60-degree stereo setup is compared with two closely spaced loudspeakers. Additionally, two- and four-channel arrangements are evaluated.
Convention Paper 7723

15:30

P13-4 The Importance of the Direct to Reverberant Ratio in the Perception of Distance, Localization, Clarity, and Envelopment— *David Griesinger*, Consultant, Cambridge, MA, USA

The Direct to Reverberant ratio (D/R)—the ratio of the energy in the first wave front to the reflected sound energy—is absent from most discussions of room acoustics. Yet only the direct sound (DS) provides information about the localization and distance of a sound source. This paper discusses how the perception of DS in a reverberant field depends on the D/R and the time delay between the DS and the reverberant energy. Threshold data for DS perception will be presented, and the implications for listening rooms, hall design, and electronic enhancement will be discussed. We find that both clarity and envelopment depend on DS detection. In listening rooms the direct sound must be at least equal to the total reflected energy for accurate imaging. As the room becomes larger (and the time delay increases) the threshold goes down. Some conclusions: typical listening rooms benefit from directional loudspeakers, small concert halls should not have a shoe-box shape, early reflections need not be lateral, and electroacoustic enhancement of late reverberation may be vital in small halls.
Convention Paper 7724

16:00

P13-5 Frequency-Domain Interpolation of Empirical HRTF Data— *Brian Carty, Victor Lazzarini*, National University of Ireland, Maynooth, Ireland

This paper discusses Head Related Transfer Function (HRTF)-based artificial spatialization of audio. Two alternatives to the minimum phase method of HRTF interpolation are suggested, offering novel approaches to the challenge of phase interpolation. A phase truncation, magnitude interpolation technique aims to avoid complex preparation, manipulation or transformation of empirical HRTF data, and any inaccuracies that may be introduced by these operations. A second technique adds low frequency nonlinear frequency scaling to a functionally based phase model. This approach aims to provide a low frequency spectrum more closely aligned to the empirical HRTF data. Test results indicate favorable performance of the new techniques.
Convention Paper 7725

16:30

P13-6 Analysis and Implementation of a Stereophonic Play Back System for Adjusting the "Sweet Spot" to the Listener's Position— *Sebastian Merchel, Stephan Groth*, Dresden University of Technology, Dresden, Germany

This paper focuses on a stereophonic play back system designed to adjust the "sweet spot" to the listener's position. The system includes an optical face tracker that provides information about the listener's x-y position. Accordingly, the loudspeaker signals are manipulated in real-time in order to move the "sweet spot." The stereophonic perception with an adjusted "sweet spot" is theoretically investigated on the basis of several models of binaural hearing. The results indicate that an adjustment of signals corresponding to the center of the listener's head does improve the localization over the whole listening area. Although some localization error remains due to asymmetric signal paths for off-center listening positions, which can be estimated and compensated for.
Convention Paper 7726

17:00

P13-7 Issues on Dummy-Head HRTFs Measurements— *Daniela Toledo, Henrik Møller*, Aalborg University, Aalborg, Denmark

The dimensions of a person are small compared to the wavelength at low frequencies. Therefore, at these frequencies head-related transfer functions (HRTFs) should decrease asymptotically until they reach 0 dB—i.e., unity gain—at DC. This is not the case in measured HRTFs: the limitations of the equipment used result in a wrong—and random—value at DC and the effect ➥

can be seen well within the audio frequencies. We have measured HRTFs on a commercially available dummy-head Neumann KU-100 and analyzed issues associated to calibration, DC correction, and low-frequency response. Informal listening tests suggest that the ripples seen in HRTFs with a wrong DC value affect the sound quality in binaural synthesis.
Convention Paper 7727

17:30

P13-8 Binaural Processing Algorithms: Importance of Clustering Analysis for Preference Tests—*Andreas Silzle, Bernhard Neugebauer, Sunish George, Jan Plogsties*, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

The acceptability of a newly proposed technology for commercial application is often assumed if the sound quality reached in a listening test surpasses a certain target threshold. As an example, it is a well-established procedure for decisions on the deployment of audio codecs to run a listening test comparing the coded/ decoded signal with the uncoded reference signal. For other technologies, e.g., upmix or binaural processing, however, the unprocessed signal only can act as a "comparison signal." Here, the goal is to achieve a significant preference of the processed over the comparison signal. For such preference listening tests, we underline the importance of clustering the test results to obtain additional valuable information, as opposed to using the standard statistic metrics like mean and confidence interval. This approach allows determining the size of the user group that significantly prefers to use the proposed algorithm when it would be available in a consumer device. As an example, listening test data for binaural processing algorithms are analyzed in this investigation.
Convention Paper 7728

18:00

P13-9 Perception of Head-Position-Dependent Variations in Interaural Cross-Correlation Coefficient—*Russell Mason, Chungeun Kim, Tim Brookes*, University of Surrey, Guildford, Surrey, UK

Experiments were undertaken to elicit the perceived effects of head-position-dependent variations in the interaural cross-correlation coefficient of a range of signals. A graphical elicitation experiment showed that the variations in the IACC strongly affected the perceived width and depth of the reverberant environment, as well as the perceived width and distance of the source. A verbal experiment gave similar results and also indicated that the head-position-dependent IACC variations caused changes in the perceived spaciousness and envelopment of the stimuli.
Convention Paper 7729

Tutorial 8 Friday, May 8
14:00 – 16:00 Room D123

MICROPHONE HISTORY

Chair: Jörg Wuttke, Schoeps, Technical Director Emeritus

Presenters: *Ulrich Apel*, Microtech Gefell GmbH
Sean Davies, S.W. Davies
Stephan Peus, Neumann GmbH

This tutorial will be presented in 3 parts.

Stephan Peus' presentation, "35 Years of Microphone Development at Neumann—What Touched Us, What Moved Us," gives an insight to specific development topics and to some very special test procedures including: microphone's transient response: insights beyond frequency response or polar pattern; RF susceptibility: already a topic before the era of mobile phones; capsule distortion measurement: difficult procedure giving a lot of interesting results; dynamic range and self noise level of studio microphones: a remarkable development within the 35 years in question.

Ulrich Apel will report on "The Importance of Vacuum for Condenser Microphones." He will speak on such topics as: the electron-tube was and is still an important step in the development of condenser microphones; the construction of special-made tubes for use in mics such as RE084k, Hiller MSC2, Telefunken AC701k, EF804, Valvo EF86, 6072, etc.; and special measuring capabilities to select tubes regarding noise, stability. and sound"

Sean Davies's presentation is "Microphone History: The Why, The How, and The Who." The developments in microphone technology are reviewed from the earliest telephone based type through the decades as far as the 1970s. The "Why" section looks at the reasons behind the different designs, e.g., directional characteristics, output signal levels, diffraction effects, frequency range. The "How" examines the solutions proposed for the "Why" section, and the "Who" identifies the landmark designs and the designers behind them.

Friday, May 8 14:00 Room D120
Technical Committee Meeting on Network Audio Systems

Friday, May 8 14:00 Room D104
Standards Committee Meeting on SC-04-03 Loudspeaker Modeling and Measurement

Workshop 9 Friday, May 8
15:00 – 16:30 Room K1

MIXING SPORTS IN 5.1, PART 1

Chair: **Gerhard Stoll**, IRT Munich

Panelists: *Dennis Baxter*, Sound for the Olympics, USA
Beat Joss, tpc AG, Switzerland
Ales Koman, Slovenia TV, Slovenia

Sports has proven to be a major driving force for the introduction of HDTV into the market. Events like the Football World Championships 2006, the Football European Championships in 2008, and the 2008 Bejing Olympics provide an excellent opportunity to showcase the strengths of high resolution pictures. Teaming up with HD picture is High Definition Surround Sound and a generally more elaborate and creative sound design. The challenges are manifold:

- from capturing a roaring stadium crowd of 80,000 to hearing the details of every ball kick, the so-called "close ball";
- from producing a surround mix to still serving your

stereo-viewers with an appropriate downmix and an intelligible commentator;

• from getting your signal to and through your broadcast center unharmed to arriving at the consumer in sync with the picture;

• from making meaningful use of LFE to using the center channel not only for commentary but for the "sound of sports" as well.

A panel of experienced protagonists will discuss these issues and other challenges. Examples will give the audience the chance to judge the effectiveness themselves.

In **Part 2** of this workshop a multitrack recording of one of the finals of the Swiss Ice-Hockey Championship 2009 will be mixed live for the audience to witness the approach to creating a compelling surround sound field, which includes the atmosphere of thousands of bawling, booing, and applauding fans and the "sounds of the match," which you see on your screen.

Friday, May 8 15:00 Room D120
Technical Committee Meeting on Human Factors in Audio Systems

Exhibitor Seminar Friday, May 8
15:00 – 16:00 Exhibit Floor Booth 2303
ES5 OPTOCORE – A

TECHNICAL INTRODUCTION TO OPTOCORE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

The OPTOCORE Synchronous Fibre Network is introduced. Content will include an Introduction/Overview, design of the network, synchronicity, latency, topology, redundancy, protocol, and devices.

Friday, May 8 16:00 Room D120
Technical Committee Meeting on Automotive Audio

Exhibitor Seminar Friday, May 8
16:00 – 17:00 Exhibit Floor Booth 2303
ES6 OPTOCORE – B

BASICS OF FIBRE OPTICS

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

Technical, economical, safety aspects, and advantages of fibre optics is the topic of this seminar. Content will include fibre advantages, reliability, types of fibre, parameters of fibres (attenuation, bandwidth length product, latency), connector types, cleaning, and testing.

Session P14 Friday, May 8
16:30 – 18:30 Room K4

MULTICHANNEL CODING

Chair: **Ville Pulkki**, Helsinki University of Technology, Espoo, Finland

16:30

P14-1 Further EBU Tests of Multichannel Audio Codecs—*David Marston,*[1] *Franc Kozamernik,*[2] *Gerhard Stoll,*[3] *Gerhard Spikofski*[3]
[1]BBC R&D, Tadworth, Surrey, UK
[2]EBU, Geneva, Switzerland
[3]IRT, Munich, Germany

The European Broadcasting Union technical group D/MAE has been assessing the quality of multichannel audio codecs in a series of subjective tests. The two most recent tests and results are described in this paper. The first set of tests covered 5.1 multichannel audio emission codecs at a range of bit-rates from 128 kbit/s to 448 kbit/s. The second set of tests covered cascaded contribution codecs, followed by the most prominent emission codecs. Codecs under test include offerings from Dolby, DTS, MPEG, APT, and Linear Acoustics. The conclusions observe that while high quality is achievable at lower bitrates, there are still precautions to be aware of. The results from cascading of codecs have shown that the emission codec is usually the bottleneck of quality.
Convention Paper 7730

17:00

P14-2 Spatial Parameter Decision by Least Squared Error in Parametric Stereo Coding and MPEG Surround —*Chi-Min Liu, Han-Wen Hsu, Yung-Hsuan Kao, Wen-Chieh Lee*, National Chiao Tung University, Hsinchu, Taiwan

Parametric stereo coding (PS) and MPEG Surround (MPS) are used to reconstruct stereo or multichannel signals from down-mixed signals with a few spatial parameters. For extracting spatial parameters, the first issue is to decide a time-frequency (T-F) tile that controls the resolution of reconstructed spatial scenes and highly determines the amount of consumed bits. On the other hand, according to the standard syntax, the up-mixing matrices for time slots not on time borders are reconstructed by interpolation in the decoder. Therefore, the second issue is to decide the transmitted parameter values on the time borders for confirming the minimum reconstruction error of matrices. For both PS and MPS, based on the criterion of least squared error, this paper proposes a generic dynamic programming method for deciding the two issues under the tradeoff of audio quality and limited bits.
Convention Paper 7731

17:30

P14-3 The Potential of High Performance Computing in Audio Engineering—*David Moore, Jonathan Wakefield*, University of Huddersfield, West Yorkshire, UK

High Performance Computing (HPC) resources are fast becoming more readily available. HPC hardware now exists for use in conjunction with standard desktop computers. This paper looks at what impact this could have on the audio engineering industry. Several potential applications of HPC within audio engineering research are discussed. A case study is also presented that highlights the benefits of using the Single ➥

Instruction, Multiple Data (SIMD) architecture when employing a search algorithm to produce surround sound decoders for the standard 5-speaker surround sound layout.
Convention Paper 7732

18:00

P14-4 Efficient Methods for High Quality Merging of Spatial Audio Streams in Directional Audio Coding—*Giovanni Del Galdo*,[1] *Ville Pulkki*,[2] *Fabian Kuech*,[1] *Mikko-Ville Laitinen*,[2] *Richard Schultz-Amling*,[1] *Markus Kallinger*[1]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]Helsinki University of Technology, Espoo, Finland

Directional Audio Coding (DirAC) is an efficient technique to capture and reproduce spatial sound. The analysis step outputs a mono DirAC stream, comprising an omnidirectional microphone pressure signal and side information, i.e., direction of arrival and diffuseness of the sound field expressed in time-frequency domain. This paper proposes efficient methods to merge multiple mono DirAC streams to allow a joint playback at the reproduction side. The problem of merging two or more streams arises in applications such as immersive spatial audio teleconferencing, virtual reality, and online gaming. Compared to a trivial direct merging of the decoder outputs, the proposed methods are more efficient as they do not require the synthesis step. From this it follows the benefit that the loudspeaker setup at the reproduction side does not have to be known in advance. Simulations and listening tests confirm the absence of any artifacts and that the proposed methods are practically indistinguishable from the ideal merging.
Convention Paper 7733

Session P15 — Friday, May 8
16:30 – 18:00 — K4 Foyer

POSTERS: HEARING

16:30

P15-1 Psychoacoustics and Noise Perception Survey in Workers of the Construction Sector—*Marcos D. Fernández, Bálder Vitón, José Antonio Ballesteros, Samuel Quintana, Isabel González*, Escuela Universitaria Politécnica de Cuenca, Universidad de Castilla-La Mancha, Cuenca, Spain

Noise levels are not enough to assess completely the influence of the noise. Therefore, psychoacoustics and perception surveys should be taken into account. The noise that the construction workers produce in their tasks is recorded with a HATS. Later, those recordings are processed to derive different parameters: spectrum, weighted equivalent levels, and the main psychoacoustics parameters. After that, a specific survey has been developed to assess the perception of such activity noises during the working time to correlate adjectives of perception with those parameters mentioned. The survey has been designed to be answered by the workers that are exposed to the noise, so that conclusions could be derived about the feelings and annoyance that the noise can cause.
Convention Paper 7734

16:30

P15-2 On the Design of Automatic Sound Classification Systems for Digital Hearing Aids—*Enrique Alexandre, Lorena Álvarez-Perez, Roberto Gil-Pita, Raúl Vicen-Bueno, Lucas Cuadra*, University of Alcalá, Alcalá de Henares, Spain

The design of digital hearing aids able to carry out advanced functionalities (such as, for instance, classify the acoustic environment and automatically select the best amplification program for the user's comfort) exhibits a great difficulty. Since hearing aids have to work at very low clock frequency in order to minimize power consumption and maximize life battery, the number of available instructions per second is actually very small. This enforces the design of efficient algorithms with a reduced number of instructions. In particular, this paper will focus on three extremely related topics: (1) the design of low-complexity features; (2) the use of automatic feature selection algorithms to optimize the performance of the classifier; and (3) the critical analysis of a variety of different classification algorithms, basically based on their complexity and performance and determining whether or not they are feasible to be implemented.
Convention Paper 7735

16:30

P15-3 Pruning Algorithms for Multilayer Perceptrons Tailored for Speech/Non-Speech Classification in Digital Hearing Aids—*Lorena Álvarez, Enrique Alexandre, Manuel Rosa-Zurera*, University of Alcalá, Alcalá de Henares, Spain

This paper explores the feasibility of using different pruning algorithms for multilayer perceptrons (MLPs) applied to the problem of speech/non-speech classification in digital hearing aids. A classifier based on MLPs is considered the best option in spite of its *presumably* high computational cost. Nevertheless, its implementation has been proven to be feasible: it requires some trade-offs involving a balance between reducing the computational demands (that is, the number of neurons) and the quality perceived by the user. In this respect, this paper will focus on the design of three *novel* pruning algorithms for MLPs, which attempt to converge to the minimum complexity network (that is, the lowest number of neurons in the hidden layer) without degrading the performance of it. The results obtained with the proposed algorithms will be compared with those obtained when using another pruning algorithm proposed in the literature.
Convention Paper 7736

16:30

P15-4 Evolutionary Optimization for Hearing Aids of Computational Auditory Scene Analysis—

Anton Schlesinger, Marinus M. Boone, Technical University of Delft, Delft, The Netherlands

Computational auditory scene analysis (CASA) provides an excellent means to improve speech intelligibility in adverse acoustical situations. In order to utilize algorithms of CASA in hearing aids, sets of algorithmic parameters need to be adjusted to the individual auditory performance of the listener and the acoustic scene in which they are employed. Performed manually, the optimization is an expensive procedure. We therefore developed a framework in which algorithms of CASA are automatically optimized by the principles of evolution, i.e., by a genetic algorithm. By using the speech transmission index (STI) as an objective function, the presented framework presents a holistic routine that is solely based on psychoacoustical and physiological models to improve and to assess speech intelligibility. The initial listening test revealed a discrepancy between the objective and subjective assessment of speech intelligibility, which suggests a review of the objective function. Once the objective function is in accordance with the individual perception of speech intelligibility, the presented framework could be applied in the optimization of all complex speech processors and therewith accelerate their assessment and application.
Convention Paper 7737

16:30

P15-5 Enhanced Control of On-Screen Faders with a Computer Mouse—*Michael Hlatky,*[1] *Kristian Gohlke,*[1] *David Black*,[1] *Jörn Loviscach*[2]
[1]Hochschule Bremen (University of Applied Sciences), Bremen, Germany
[2]Fachhochschule Bielefeld (University of Applied Sciences), Bielefeld, Germany

Input devices of the audio studio that formerly were physical have mostly been converted into virtual controls on the computer screen. Whereas this transition saves space and cost, it has reduced the performance of these controls, as virtual controls adjusted using the computer mouse do not exhibit the accuracy and accessibility of their physical counterparts. Previous studies show that interaction with scrollable timelines can be enhanced by an intelligent interpretation of the mouse movement. We apply similar techniques to virtual faders as used for audio control, leveraging such approaches as controllable zoom levels and pseudo-haptic interaction. Tests conducted on five such methods provide insight into how to decouple the fader from the mouse movement to improve accuracy without impairing the speed of the interaction.
Convention Paper 7738

16:30

P15-6 Modeling of External Ear Acoustics for Insert Headphone Usage—*Marko Hiipakka, Miikka Tikander, Matti Karjalainen*, Helsinki University of Technology, Espoo, Finland

Although acoustics of the external ear has been studied extensively for auralization and hearing aids, the acoustic behavior with insert headphones is not as well known. Our research focused on the effects of outer ear physical dimensions, particularly on sound pressure at the eardrum. The main parameter was the length of the canal, but eardrum's damping of resonances was also studied. Ear canal simulators and a dummy head were constructed. Measurements were also performed from human ear canals. The study was carried out both with unblocked ear canals and when the canal entrance was blocked with an insert earphone. Special insert earphones with in-ear microphones were constructed for this purpose. Physics-based computational models were finally used to validate the approach.
Convention Paper 7739

Workshop 10 — Friday, May 8
16:30 – 18:30 — Room D123

AUDIO NETWORK CONTROL PROTOCOLS

Chair: **Richard Foss**, Rhodes University, Grahamstown, South Africa

Panelists: *John Grant*, Nine Tiles
Robby Gurdan, UMAN
Rick Kreifeldt, Harman Professional
Philip Nye, Engineering Arts

With the advent of digital audio networking, there has been a need to manage the connection of devices on networks and also to control and access various parameters of the devices. A number of standard and proprietary protocols have been developed. In this workshop a panel of experts who have helped develop some of these protocols will discuss their approaches and attempt to define a way forward, whereby devices with differing protocol implementations can communicate.

Friday, May 8 — 16:30 — Room D104
Standards Committee Meeting on SC-04-04 Microphone Measurement and Characterization

Workshop 11 — Friday, May 8
17:00 – 18:30 — Room K1

MIXING SPORTS IN 5.1, PART 2

Chair: **Gerhard Stoll**, IRT Munich

Panelists: *Dennis Baxter*, Sound for the Olympics, USA
Beat Joss, tpc AG, Switzerland
Ales Koman, Slovenia TV, Slovenia

Sports has proven to be a major driving force for the introduction of HDTV into the market. Events like the Football World Championships 2006, the Football European Championships in 2008, and the 2008 Bejing Olympics provide an excellent opportunity to showcase the strengths of high resolution pictures. Teaming up with HD picture is High Definition Surround Sound and a generally more elaborate and creative sound design. The challenges are manifold:

- From capturing a roaring stadium crowd of 80,000 to hearing the details of every ball kick, the so-called "close ball";

➥

• From producing a surround mix to still serving your stereo-viewers with an appropriate downmix and an intelligible commentator;

• From getting your signal to and through your broadcast center unharmed to arriving at the consumer in sync with the picture;

• From making meaningful use of LFE to using the center channel not only for commentary but for the "sound of sports" as well.

A panel of experienced protagonists will discuss these issues and other challenges. Examples will give the audience the chance to judge the effectiveness themselves.

In **Part 2** of this workshop a multitrack recording of one of the finals of the Swiss Ice-Hockey Championship 2009 will be mixed live for the audience to witness the approach to creating a compelling surround sound field, which includes the atmosphere of thousands of bawling, booing, and applauding fans and the "sounds of the match," which you see on your screen.

Exhibitor Seminar **Friday, May 8**
17:00 – 18:00 **Exhibit Floor Booth 2303**
ES7 OPTOCORE – C

OPTOCORE IN PRACTICE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

This seminar will cover practical exercises with OPTOCORE system/devices. Content will include practical set-up of a network, remote pre-amp control through third party devices, transport of control data and Ethernet, software control, and firmware upgrades.

Tutorial 9 **Friday, May 8**
17:30 – 18:30 **Room K1**

TECHNIQUES OF AUDIO LOCALIZATION FOR GAMES

Presenters: **Fabio Minazzi**, Binari Sonori Srl
Francesco Zambon, Binari Sonori Srl

Videogaming is a new form of communication, which heavily relies on audio. The production of game soundtracks shares many technologies and techniques with the production of audio for other media. At the same time, videogames are software products, with nonlinear and dynamic behaviors, and these features of the media affect the way audio is produced and localized. This presentation reviews a set of specific audio production techniques that are applied in games localization, like the pre- and postproduction methods for localizing voices, the A/V asset management, the team that attend the local recording sessions, as well as the international quality assurance processes. Some theoretical aspects are also described, like the typical constraints that need to be accounted for during the speech voice recording phase, in order to allow the software to seamlessly load, mix, and combine audio assets coming from various countries, recorded in different times.

Special Event
BANQUET
Friday, May 8, 20:00 – 23:30
Isar Brau, Munchen Pullach

This year the Banquet will take place in a small old railway station, above the valley of the River Isar. The railway opened in 1891 and steam trains took people from the city to many beautiful places in the south of Munich. Today the steam trains have been replaced and the line is now part of the S-Bahn, so the old station is not needed anymore and has been turned into a traditional Bavarian style restaurant with its own micro-brewery. What could be more natural than making this location a pleasant place for a "get together" in a lovely atmosphere?

The welcome beer from the micro brewery and other drinks will be followed by a fine buffet with Bavarian delicacies. At the end of a long day at the Convention, these *Schmankerl* will be a good way to relax and enjoy the evening with old and new friends and colleagues. Come and savour Munich's lifestyle. The ticket price includes all food and drinks and the bus to the restaurant and back.

55 Euros for AES members; 65 Euros for nonmembers
Tickets will be available at the Special Events desk.

Session P16 **Saturday, May 9**
09:00 – 11:30 **Room K4**

SPATIAL RENDERING, PART 1

Chair: **Andreas Silzle**, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

09:00

P16-1 An Alternative Ambisonics Formulation: Modal Source Strength Matching and the Effect of Spatial Aliasing—*Franz Zotter,*[1] *Hannes Pomberger*,[1] *Matthias Frank*[2]
[1]University of Music and Dramatic Arts, Graz, Austria
[2]Graz University of Technology, Graz, Austria

Ambisonics synthesizes sound fields as a sum over angular (spherical/cylindrical harmonic) modes, resulting in the definition of an isotropically smooth angular resolution. This means, virtual sources are synthesized with outstanding smoothness across all angles of incidence, using discrete loudspeakers that uniformly cover a spherical or circular surface around the listening area. The classical Ambisonics approach models the fields of these discrete loudspeakers in terms of a sampled continuum of plane-waves. More accurately, the contemporary concept of Ambisonics uses a continuous angular distribution of point-sources at finite distance instead, which is considerably easier to imagine. This also improves the accuracy of holophonic sound field synthesis and the analytic description of the sweet spot. The sweet spot is a limited area of faultless synthesis emerging from angular harmonics truncation. Additionally, playback with loudspeakers causes spatial aliasing. In this sense, it allows for a successive consideration of the major shortcomings of Ambisonics: the limited sweet spot size and spatial aliasing. To elaborate on this concept this paper starts with the solution of the nonhomogeneous wave equation for a spherical point-source distribution, and ends with a novel study on spatial aliasing in Ambisonics.
Convention Paper 7740

09:30

P16-2 Sound Field Reproduction Employing Non-Omnidirectional Loudspeakers—*Jens Ahrens, Sascha Spors*, Deutsche Telekom Laboratories, Techniche Universität Berlin, Berlin, Germany

In this paper we treat sound field reproduction via circular distributions of loudspeakers. The general formulation of the approach has been recently published by the authors. In this paper we concentrate on the employment of secondary sources (i.e., loudspeakers) whose spatio-temporal transfer function is not omnidirectional. The presented approach allows us to treat each spatial mode of the secondary source's spatio-temporal transfer function individually. We finally outline the general process of incorporating spatio-temporal transfer functions obtained from microphone array measurements.
Convention Paper 7741

10:00

P16-3 Alterations of the Temporal Spectrum in High-Resolution Sound Field Reproduction of Different Spatial Bandwidths—*Jens Ahrens, Sascha Spors*, Deutsche Telekom Laboratories, Techniche Universität Berlin, Berlin, Germany

We present simulations of the wave field reproduced by a discrete circular distribution of loudspeakers. The loudspeaker distribution is driven either with signals of infinite spatial bandwidth (as it happens in wave field synthesis), or the loudspeaker distribution is driven with signals of finite spatial bandwidth (as it is the case in near-field compensated higher order Ambisonics). The different spatial bandwidths lead to different accuracies of the desired component of the reproduced wave field and to spatial aliasing artifacts with essentially different properties. Our investigation focuses on the potential consequences of the artifacts on human perception.
Convention Paper 7742

10:30

P16-4 Cooperative Spatial Audio Authoring: Systems Approach and Analysis of Use Cases—*Jens-Oliver Fischer,*[1] *Francis Gropengiesser,*[2] *Sandra Brix*[1]
[1]Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany
[2]TU Ilmenau, Ilmenau, Germany

Today's audio production process is highly parallel and segregated. This is especially the case in the field of audio postproduction for motion pictures. The introduction of spatial audio systems like 5.1, 22.2 or Wave Field Synthesis results in even more production steps, namely the spatial authoring, to accomplish a rich experience for the audience. This paper proposes a system that enables the audio engineers to work together on the same project. The proposed system is planned to be implemented for an existing spatial authoring software but can be utilized by any other application that organizes its data in a tree structured way. Three major use cases, i.e., Single User, Work Space, and Work Group, are introduced and analyzed.
Convention Paper 7743

11:00

P16-5 Spatial Sampling Artifacts of Wave Field Synthesis for the Reproduction of Virtual Point Sources— *Sascha Spors, Jens Ahrens*, Deutsche Telekom Laboratories, Techniche Universität Berlin, Berlin, Germany

Spatial sound reproduction systems with a large number of loudspeakers are increasingly being used. Wave field synthesis is a reproduction technique using a large number of densely placed loudspeakers (loudspeaker array). The underlying theory, however, assumes a continuous distribution of loudspeakers. Individual loudspeakers placed at discrete positions constitute a spatial sampling process that may lead to sampling artifacts. These may degrade the perceived reproduction quality and will limit the application of active control techniques like active room compensation. The sampling artifacts for the reproduction of plane waves have already been discussed in previous papers. This paper derives the spatial sampling artifacts and anti-aliasing conditions for the reproduction of virtual point sources on linear loudspeaker arrays using wave field synthesis techniques.
Convention Paper 7744

Session P17 | **Saturday, May 9**
09:00 – 11:00 | **Room K3**

ROOM ACOUSTICS AND LOUDSPEAKER INTERACTION

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

09:00

P17-1 Effects of Loudspeaker Directivity on Perceived Sound Quality—A Review of Existing Studies—*William Evans,*[1] *Jakob Dyreby,*[2] *Søren Bech,*[2] *Slawomir Zielinski,*[1] *Francis Rumsey*[1]
[1]University of Surrey, Guildford, Surrey, UK
[2]Bang & Olufsen A/S, Struer, Denmark

The directivity of a loudspeaker system is often regarded as a prominent factor in the overall subjective quality of the reproduced sound experience. Much literature is available on the topic, and currently a broad field of opinion exists among designers. This paper provides an overview of the available literature, as well as an extended investigation into listener-based research. Results indicate that for such a widely debated topic, conclusive measurement data with regard to human listeners is limited, and, therefore, a proposal for more informative listening tests is presented.
Convention Paper 7745

09:30

P17-2 Subjective Validity of Figures of Merit for Room Aspect Ratio Design—*Matthew*

Wankling, Bruno Fazenda, University of Huddersfield, Huddersfield, West Yorkshire, UK

Attempts have long been made to classify a room's low frequency audio reproduction capability with regard to its aspect ratio. Common metrics used have relied on the homogeneous distribution of modal frequencies and from these a number of "optimal" aspect ratios have emerged. However, most of these metrics ignore the source and receiver coupling to the mode shapes—only a few account for this in the derivation of a figure of merit. The subjective validity of these attempts is tested and discussed. Examples are given of supposedly good room ratios with bad performance and vice versa. Subjective assessment of various room scenarios is undertaken and a ranking order has been obtained to correlate with a proposed figure of merit.
Convention Paper 7746

10:00

P17-3 A Study of Low-Frequency Near- and Far-Field Loudspeaker Behavior— *John Vanderkooy*,[1,2] *Martial Rousseau*[2]
[1]University of Waterloo, Waterloo, Ontario, Canada
[2]B&W Group Ltd., Steyning, West Sussex, UK

Low-frequency loudspeaker measurements are difficult. Room reflections, mediocre anechoic chambers, and random noise play havoc with the quest. Diffraction is different in nearfield and farfield. This paper covers a range of topics that bear on these problems, such as boundary element diffraction simulations, an approximate theory for low frequencies, methods to shorten the impulse response, and nearfield characteristics. A few points are illustrated with measurements. An earlier simplified diffraction theory of Kessel is checked for axisymmetric cylindrical and rectangular boxes by boundary-element simulations, in an attempt to pin down the diffractive 4pi to 2pi transition. It turns out to have a strong connection to the acoustic center of a loudspeaker. Some measurements are made under various conditions. Shortening methods are used to minimize the deleterious effect of truncating room reflections from the impulse response.
Convention Paper 7747

10:30

P17-4 Subwoofers in Symmetrical and Asymmetrical Rooms—*Juha Backman*, Nokia Corporation, Espoo, Finland

A theoretical study of behavior of single and multiple subwoofers, taking also geometrical and acoustical asymmetry of practical listening environments into account, is presented. The results indicate that configurations aimed at precise cancellation of individual modes have a high sensitivity to deviations from the ideal. However, with multiple subwoofers it is possible to find robust placements that both reduce the spatial variation of the sound field and the frequency variation of the response. This, however, requires loudspeaker placements where also the height of the source from the floor is varied.
Convention Paper 7748

Workshop 12 Saturday, May 9
09:00 – 11:00 Room K1

REALITY IS NOT A RECORDING / A RECORDING IS NOT REALITY

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

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

Student Event/Career Development
EDUCATION FORUM PANEL
Saturday, May 9, 09:00 – 11:00
Room K2

Moderator: **John Krivit**, New England Institute of Art, Brookline, MA, USA

Teaching Quality

John Krivit invites all educators to join him for this open discussion on how we teach students about sound quality and high production standards. In an age when many graduates will not have the chance to work with the finest equipment in the finest rooms, what are the responsibilities of the faculty and what are the expectations for the facilities at institutions of higher learning? Come share your approaches and learn from others as we all seek to teach quality. An open forum for educators, students are also welcome to attend and participate.

Saturday, May 9 09:00 Room D120
Technical Committee Meeting on Audio Recording and Mastering Systems

Saturday, May 9 09:00 Room D104
Standards Committee Meeting on SC-05-05 Grounding and EMC Practices

Saturday, May 9 10:00 Room D120
Technical Committee Meeting on High Resolution Audio

Session P18 Saturday, May 9
10:30 – 12:00 K4 Foyer

POSTERS: ASSESSMENT, EVALUATION

10:30

P18-1 WhisPER—A New Tool for Performing Listening Tests—*Simon Ciba, André Wlodarski, Hans-Joachim Maempel*, Technical University of Berlin, Berlin, Germany

A software tool is presented for performing experiments in the field of perceptual audio evalua-

tion and psychoacoustic measurement, controlling the interaction with both the subject and the playback environment. For this purpose a repertoire of test procedures has been implemented, including popular qualitative and quantitative approaches. By using OpenSound Control commands, not only traditional multichannel reproduction is supported, but also advanced spatial audio reproduction such as dynamic binaural synthesis or wavefield synthesis. WhisPER has been written in MATLAB to facilitate its further development within the scientific community. As opposed to existing libraries it provides a coherent graphical user interface system allowing easier access and configuration also for users without advanced programming experience.
Convention Paper 7749
[Paper presented by Hans-Joachim Maempel]

10:30

P18-2 Psychoacoustic Assessment of the Noise Emitted by the Machines. The Case of the Grinders—*Marcos D. Fernández, José Antonio Ballesteros, Iván Suárez, Samuel Quintana, Isabel González*, Escuela Universitaria Politécnica de Cuenca, Universidad de Castilla-La Mancha, Cuenca, Spain

Sound quality is used as the suitability of the sound emitted by a machine, depending on the characteristics of that sound and the perceptual sensation received that reflects the degree of acceptance of the machine by the user. In order to evaluate the sound quality of the grinders under study, binaural recordings, are required of the emitted sounds to determine the objective psychoacoustic parameters, and then, to make subjective tests to a representative number of people about the impression made by that particular sound.
Convention Paper 7750

10:30

P18-3 Investigations of the Effects of Nonlinear Distortions on Psychoacoustical Measures—*Stephan Herzog*, Technical University Kaiserslautern, Kaiserslautern, Germany

The perception of nonlinear distortions of audio devices, in particular the perception of nonlinear distortions of digital audio, is only insufficiently described by typical measures like THD. To provide a better insight into perceptual effects of nonlinear distortions, their audibility, and the impact on the psychoacoustical measures like loudness and sharpness is examined. For this purpose a test method has been developed. The first step in the test is the measurement of the frequency response of the device under test with an efficient method to enable the separation of linear and nonlinear processing. The second step of the test consists of the computation of the psychoacoustical measures and the thresholds for the audibility of nonlinear distortions. Both computations are based on the same psychoacoustical model to obtain consistent results. Results for several types of distortion obtained with simulations and measurements on analog circuits are presented.
Convention Paper 7751

Live Sound Seminar 1 Saturday, May 9
10:30 – 13:00 Atrium 2

NEUMANN & MULLER AND D & B

Presenters: **Stefan Goertz**
Michael Kennedy

Sound System Design and Commissioning in Critical Acoustic Environments

The interaction of sound systems with special attention to the excitation of diffuse sound will be examined in theory and practical demonstrations using speech intelligibility measurements as an indicator. Software-aided line and subwoofer array designs will be discussed followed by a live demonstration of the tuning and alignment process.

Exhibitor Seminar Saturday, May 9
10:30 – 11:30 Room D124
ES8 MINNETONKA AUDIO SOFTWARE (GERMANY)

SURCODE ENCODER FOR DTS-HD WORKFLOW ON APPLE FINAL CUT PRO – COMPRESSOR

Presenters: *Markus Hintz*, Minnetonka Audio Software
Achim Scherner, DTS
Jayson Tomlin, Minnetonka Audio Software

DTS HD has become a preferred audio format for Blu-ray production. DTS and Minnetonka will explain how authoring houses can increase workflow efficiency and make best use of available disc capacity for both Blu-ray and DVD-Video production.

Saturday, May 9 11:00 Room D120
Technical Committee Meeting on Electronic Distribution of Audio

Saturday, May 9 11:00 Room D104
Standards Committee Meeting on SC-05-02 Audio Connectors

Exhibitor Seminar Saturday, May 9
11:00 – 12:00 Exhibit Floor Booth 2303
ES9 OPTOCORE – A

TECHNICAL INTRODUCTION TO OPTOCORE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

The OPTOCORE Synchronous Fibre Network is introduced. Content will include an Introduction/Overview, design of the network, synchronicity, latency, topology, redundancy, protocol, and devices.

Workshop 13 Saturday, May 9
11:30 – 13:00 Room K1

AES 42 DIGITAL MICROPHONES

Chair: **Gregor Zielinsky**, Sennheiser

Panelists: *Malgorzata Albinska-Frank*
Stephan Flock, Direct Out
Stephan Peus, Neumann
Helmut Wittek, Schoeps

The panel will discuss daily work with digital microphones and their peripheral devices.

While the first digital microphones were launched a few years ago, their use is still within an exclusive community of users.

During the last two years, prices of digital microphones have dropped, while the choice of mics—as well as choices of interfaces—have risen very strongly. More and more companies are incorporating the AES 42 standard.

In this panel manufacturers and users discuss possibilities and workflows of digital microphones. Experienced users will give their view on the AES 42 standard.

Student Event/Career Development
RECORDING COMPETITION—SURROUND
Saturday, May 9, 11:30 – 13:30
Room K2

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 later on Saturday.

Judges include: Jim Anderson, Akira Fukada, Andres Mayo, Ronald Prent, and Darcy Proper.

Exhibitor Seminar Saturday, May 9
11:30 – 12:30 Room D124
ES10 KRK (USA)

ROOM OPTIMIZATION WITH ERGO AND STUDIO MONITOR DESIGN

Presenter: *Nils Karsten*

Every room has his own characteristics. Especially these days it is important, to optimize the acoustics of the listening environment. An acoustically well treated environment is needed to ensure a well translating mix. ERGO makes any monitor system more accurate by reducing "bad" room sound. Besides the room acoustic also the speaker design has a fundamental impact on the radiation of the sound. The KRK design is known to built bullet proof tools for the recording and mastering engineer to ensure to receive a well translating mix.

Saturday, May 9 12:00 Room D120
Technical Committee Meeting on Studio Practices and Production

Exhibitor Seminar Saturday, May 9
12:00 – 13:00 Exhibit Floor Booth 2303
ES11 OPTOCORE – B

BASICS OF FIBRE OPTICS

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

Technical, economical, safety aspects, and advantages of fibre optics is the topic of this seminar. Content will include fibre advantages, reliability, types of fibre, parameters of fibres (attenuation, bandwidth length product, latency), connector types, cleaning, and testing.

Saturday, May 9 13:00 Room D120
Technical Committee Meeting on Audio Forensics

Saturday, May 9 13:00 Room D104
Standards Committee Meeting on SC-02-01 Digital Audio Measurement Techniques

Exhibitor Seminar Saturday, May 9
13:00 – 14:00 Exhibit Floor Booth 2303
ES12 OPTOCORE – C

OPTOCORE IN PRACTICE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

This seminar will cover practical exercises with OPTOCORE system/devices. Content will include practical setup of a network, remote pre-amp control through third party devices, transport of control data and Ethernet, software control, and firmware upgrades.

Session P19 Saturday, May 9
13:30 – 15:00 Room K4

EVENT, STAGE, AND SOUND REINFORCEMENT

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

13:30

P19-1 Comparative Evaluation of Howling Detection Criteria in Notch-Filter-Based Howling Suppression—*Toon van Waterschoot, Marc Moonen*, Katholieke Universiteit Leuven, Leuven, Belgium

Notch-filter-based howling suppression (NHS) is one of the most popular methods for acoustic feedback control in public address and hands-free communication systems. The NHS method consists of two stages: howling detection and notch filter design. While the design of notch filters is based on well-established filter design techniques, there is little agreement in the NHS literature on how the howling detection subproblem should be tackled. Moreover, since the NHS literature mainly consists of patents, only few experimental results have been reported. The aim of this paper is to describe a unifying framework for howling detection and to provide a comparative evaluation of existing and novel howling detection criteria.
Convention Paper 7752

14:00

P19-2 Professional Wireless Microphone Systems: Current Situation and Upcoming Changes in Regulatory Issues in Europe and USA—
Frank Ernst, Beyerdynamic GmbH & Co. KG, Heilbronn, Germany

Professional wireless microphones have been in use for almost 50 years now. The operation is based on a frequency sharing with TV broadcast transmitters. With the transition to digital TV, this situation changes. Digital TV is more spectrum

efficient. After the transition is completed, areas of the spectrum will be cleared from TV broadcast and will be available for new services. These cleared spectrum areas are referred to as "digital dividend" or white spaces. By the allocation of these bands to other services, valuable resources for the operation of professional wireless microphones will be lost. This paper will give an overview on the current situation for professional wireless microphones and the upcoming changes with the transition to digital TV.
Convention Paper 7753

14:30

P19-3 Sound Field Reconstruction: An Improved Approach for Wave Field Synthesis—*Mihailo Kolundzija,*[1] *Christof Faller*,[1] *Martin Vetterli*[1,2]
[1]Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland;
[2]University of California at Berkeley, Berkeley, CA, USA

Wave field synthesis (WFS) is a prevalent approach to multiple-loudspeaker sound reproduction for an extended listening area. Although powerful as a theoretical concept, its deployment is hampered by practical limitations due to diffraction, aliasing, and the effects of the listening room. Reconstructing the desired sound field in the listening area, accounting for the medium propagation characteristic, is another approach termed as sound field reproduction (SFR). It is based on the essential band-limitedness of the sound field, which enables a continuous matching of the reconstructed and the desired sound field by their matching on a discrete set of points spaced below the Nyquist distance. We compare the two approaches in a common single-source free-field setup, and show that SFR provides improved sound field reproduction compared to WFS in a wide listening area around a defined reference line.
Convention Paper 7754

Session P20 — Saturday, May 9
13:30 – 15:00 — K4 Foyer

POSTERS: SIGNAL PROCESSING

13:30

P20-1 Subjective Audio Quality with Multiple Description Coding in a WLAN 802.11-Based Multicast Distribution Network—*Marcus Purat, Tom Ritter*, TFH Berlin, Berlin, Germany

This paper presents a number of results of a study of different methods to mitigate the impact of packet loss on the subjective quality in a wireless distribution network of compressed high fidelity audio. The system was simulated in Matlab based on parameters of an 802.11a WLAN in multicast-mode and the Vorbis codec. The performance of multiple description coding in terms of expected subjective audio quality and its impact on the data rate is quantified and compared to other receiver-based packet loss concealment methods. The optimum set of parameters for a multiple description coder that achieves the best subjective audio quality for a given packet loss rate in the simulated system is presented.
Convention Paper 7755

13:30

P20-2 A Joint Approach to Extract Multiple Fundamental Frequency in Polyphonic Signals Minimizing Gaussian Spectral Distance—*Francisco J. Cañadas-Quesada, Pedro Vera-Candeas, Nicolás Ruiz-Reyes, Julio Jose Carabias-Orti, D. Martínez-Muñoz*, University of Jaén, Jaén, Spain

This paper presents a joint estimation approach to extract multiple fundamental frequency (F0) in monaural polyphonic music signals. In a frame-based analysis, we generate a spectral envelope for each combination of F0 candidates, from non-overlapped partials, under assumption that a harmonic sound is characterized by a Gaussian mixture model (GMM). The optimal F0 candidates combination minimizes a spectral Euclidean distance measure between the original spectrum and Gaussian spectral models. Evaluation was carried out using several piano recordings. Evaluation shows promising results.
Convention Paper 7756

13:30

P20-3 A Mixture-of-Experts Approach for Note Onset Detection—*Norberto Degara,*[1] *Antonio Pena,*[1] *Manuel Sobreira-Seoane*,[1] *Soledad Torres-Gijarro*[2]
[1]Universidade de Vigo, Vigo, Spain
[2]Laboratorio Oficial de Metroloxía de Galicia (LOMG), Tecnópole, Ourense, Spain

Finding the starting time of events (onsets) is useful in a number of applications for audio signals. The goal of this paper is to present a combination of techniques for automatic detection of events in audio signals. The proposed system uses a supervised classification algorithm to combine a set of features extracted from the audio signal and reduce the original signal to a robust detection function. Onsets are obtained by using a simple peak-picking algorithm. This paper describes the analysis system used to extract the features and the details of the neural network algorithm used to combine them. We conclude by comparing the performance of the proposed algorithm with the system that obtained the first place in the 2005 Music Information Retrieval Evaluation eXchange.
Convention Paper 7757
Paper presented by Soledad Torres-Gijarro

13:30

P20-4 Automatic Adjustment of Off-the-Shelf Reverberation Effects—*Sebastian Heise,*[1] *Michael Hlatky*,[1] *Jörn Loviscach*[2]
[1]Hochschule Bremen (University of Applied Sciences), Bremen, Germany
[2]Fachhochschule Bielefeld (University of Applied Sciences), Bielefeld, Germany

Virtually all effect units that process digital audio—software plug-ins as well as dedicated hardware—can be controlled digitally. This allows subjecting their settings to optimization processes. We demonstrate the automatic adap- ➥

tation of reverberation plug-ins to given room impulse responses. This facilitates replacing computationally expensive convolution reverberation units with standard ones, which also are amenable to easier parameter tweaking after their overall setting has been adjusted through our method. We propose optimization strategies for this multi-dimensional nonlinear problem that need no adaptation to the particularities of each effect unit, are sped up using multicore processors and networked computers. The optimization process evaluates the difference between the actual response and the targeted response on the basis of psychoacoustic features. An acoustic comparison with effect parameter settings crafted by professional human operators indicates that the computationally optimized settings yield comparable or better results.
Convention Paper 7758

13:30

P20-5 Improvements on Automatic Parametric Equalization and Cross-Over Alignment of Audio Systems—*German Ramos, Pedro Tomas*, Technical University of Valencia, Valencia, Spain

The idea and algorithm implementation of an automatic parametric equalizer and cross-over alignment of audio systems was proposed previously by one of the authors with proven success. This method designed Infinite Impulse Response (IIR) equalization and cross-over filters directly in a series of second-order-sections (SOS) employing peak filters, and pre-initialized high-pass and low-pass filters, defined by its parameters (frequency, gains, and Q). The method supported the inclusion of constraints (maximum and minimum parameter values) and designed the SOS in order of importance in the equalization, providing thus a scalable filter implementation. In order to lower the order of the needed filter, and looking also for an automatic decision on the selection of filter types and initialization, several improvements are presented. It is now possible for the algorithm to select, configure, and use shelving filters in the SOS chain for equalization. Also, the decision and initialization of the needed high-pass and low-pass SOS filters could be automatic, helping in the cross-over design stage for active audio systems.
Convention Paper 7759

13:30

P20-6 Low Noise Transformer Input Preamp Design—A Solution that Eliminates CMID
—*Milan Kovinic*,[1] *Dragan Drincic*,[2] *Sasha Jankovic*[3]
[1]MMK Instruments, Belgrade, Serbia
[2]Advanced School for Electrical & Computer Engineering, Belgrade, Serbia
[3]OXYGEN-Digital, Parkgate Studio, Sussex, UK

This paper examines more closely the advantages of input transformer-op amplifier configurations, especially those implemented in low-noise designs. Usual transformer input stage topology works in non-inverting architecture, since it allows the transformer to work with optimum loading, to maximize the signal-to-noise ratio. However, this configuration is subject to Common-Mode voltage Induced Distortion—CMID. The susceptibility is further increased if the amplifier source impedance is not perfectly matched. This is illustrated by tests on popular audio op amps. Advanced transformer input stage topology proposed in this paper completely prevents this kind of distortion. Noise performance remains unaffected, yet listening tests in practical application confirm the sound to be more pleasant.
Convention Paper 7760

Workshop 14 — Saturday, May 9
13:30 – 15:30 — Room K1

5.1 HIGH PROFILE MIXING

Cochairs: **Akira Fukada**, Senior Engineer, NHK, Tokyo, Japan
Ulrike Schwarz, Engineer, Bavarian Radio, Munich, Germany

Panelists: *Jean-Marie Geijsen*, Director & Balance Engineer, Polyhymnia, Baarn, The Netherlands
Sascha Paeth, Owner/Engineer, Gate Studios, Wolfsburg, Germany
Ronald Prent, Residential Surround Engineer, Galaxy Studios, Mol, Belgium

It is very interesting how the perception of music can be altered by a mixing engineer. A conductor or a musician changes the figure of written music, the composer's work, by his or her interpretation and expression. For the recording of music it is rather important what the music conveys to the engineer. In the process of recording and mixing the engineer will approach and embrace the music like an artist or musician.

In this workshop engineers who have various cultural and musical backgrounds present their different mixing results. What did each engineer consider and what did they aim at? We believe that considering the result is a very important element for understanding music and art.

Student Event/Career Development
EDUCATION FAIR
Saturday, May 9, 13:30 – 15:00
K4 Foyer

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

Workshop 15 — Saturday, May 9
14:00 – 15:00 — Room K2

WAR ON MUSIC

Chair: **Thomas Lund**, tc electronic

Music is currently under fire from two sides: Hyper-level and data reduction.

Using theory as well as listening examples, the session will demonstrate how this is a lethal combination. 16

bit audio recorded 20 years ago using less perfect conversion and processing equipment than today easily end up sounding better than modern recordings.

Based on years of research, some of the significant "quality bottlenecks" of today's production and delivery system are identified, and ways of improving on the situation are suggested.

Saturday, May 9 14:00 Room D120
Technical Committee Meeting on Perception and Subjective Evaluation of Audio

Exhibitor Seminar Saturday, May 9
14:30 – 15:30 Room D124
ES13 MINNETONKA AUDIO SOFTWARE (GERMANY)

FILE- AND TAPE-BASED WORKFLOWS USING SURCODE FOR DOLBY E

Presenters: *Markus Hintz*, Minnetonka Audio Software
Jayson Tomlin, Minnetonka Audio Software
TBA, Dolby Laboratories

SurCode for Dolby E is quickly becoming the standard for Dolby E software-based encoding and decoding. Minnetonka Audio Software will present how SurCode for Dolby E is the best choice for both file and tape based workflows on a variety of workstation platforms.

Session P21 Saturday, May 9
15:00 – 18:30 Room K4

SPATIAL RENDERING, PART 2

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

15:00

P21-1 Score File Generators for Boids-Based Granular Synthesis in Csound—*Enda Bates, Dermot Furlong*, Trinity College, Dublin, Ireland

In this paper we present a set of score file generators and granular synthesis instruments for the Csound language. The applications use spatial data generated by the Boids flocking algorithm along with various user-defined values to generate score files for grainlet additive synthesis, granulation, and glisson synthesis instruments. Spatialization is accomplished using Higher Order Ambisonics and distance effects are modeled using the Doppler Effect, early reflections, and global reverberation. The sonic quality of each synthesis method is assessed and an original composition by the author is presented.
Convention Paper 7761

15:30

P21-2 Acoustical Rendering of an Interior Space Using the Holographically Designed Sound Array—*Wan-Ho Cho, Jeong-Guon Ih*, KAIST, Daejeon, Korea

It was reported that the filter for the acoustic array can be inversely designed in a holographic way, which was demonstrated in a free-field. In this study the same method using the boundary element method (BEM) was employed to render the interior sound field in an acoustically desired fashion. Because the inverse BEM technique can deal with arbitrary shaped source or bounding surfaces, one can simultaneously consider the effect of irregular radiation surface and reflection boundaries having impedances such as walls, floor, and ceiling. To examine the applicability, a field rendering example was tested to control the relative spatial distribution of sound pressure in the enclosed field.
Convention Paper 7762

16:00

P21-3 Validation of a Loudspeaker-Based Room Auralization System Using Speech Intelligibility Measures—*Sylvain Favrot, Jörg M. Buchholz*, Technical University of Denmark, Lyngby, Denmark

A novel loudspeaker-based room auralization (LoRA) system has been proposed to generate versatile and realistic virtual auditory environments (VAEs) for investigating human auditory perception. This system efficiently combines modern room acoustic models with loudspeaker auralization using either single loudspeaker or high-order Ambisonics (HOA) auralization. The LoRA signal processing of the direct sound and the early reflections was investigated by measuring the speech intelligibility enhancement by early reflections in diffuse background noise. Danish sentences were simulated in a classroom and the direct sound and each early reflection were either auralized with a single loudspeaker, HOA or first-order Ambisonics. Results indicated that (i) absolute intelligibility scores are significantly dependent on the reproduced technique and that (ii) early reflections reproduced with HOA provide a similar benefit on intelligibility as when reproduced with a single loudspeaker. It is concluded that speech intelligibility experiments can be carried out with the LoRA system either with the single loudspeaker or HOA technique.
Convention Paper 7763

16:30

P21-4 Low Complexity Directional Sound Sources for Finite Difference Time Domain Room Acoustic Models—*Alexander Southern, Damian Murphy*, University of York, York, UK

The demand for more natural and realistic auralization has resulted in a number of approaches to the time domain implementation of directional sound sources in wave-based acoustic modeling schemes such as the Finite Difference Time Domain (FDTD) method and the Digital Waveguide Mesh (DWM). This paper discusses an approach for implementing simple regular directive sound sources using multiple monopole excitations with distributed spatial positioning. These arrangements are tested along with a discussion of the characteristic limitations for each setup scenario.
Convention Paper 7764

17:00

P21-5 Binaural Reverberation Using a Modified Jot Reverberator with Frequency-Dependent and Interaural Coherence Matching—*Fritz Menzer,* ➥

Christof Faller, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

An extension of the Jot reverberator is presented, producing binaural late reverberation where the interaural coherence can be controlled as a function of frequency such that it matches the frequency-dependent interaural coherence of a reference binaural room impulse response (BRIR). The control of the interaural coherence is implemented using linear filters outside the reverberator's recursive loop. In the absence of a reference BRIR, these filters can be calculated from an HRTF set.
Convention Paper 7765

17:30

P21-6 Design and Limitations of Non-Coincidence Correction Filters for Soundfield Microphones—*Christof Faller*,[1] *Mihailo Kolundzija*[2]
[1]Illusonic LLC, Lausanne, Switzerland
[2]Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

The tetrahedral microphone capsule arrangement in a soundfield microphone captures a so-called A-format signal, which is then converted to a corresponding B-format signal. The phase differences between the A-format signal channels due to non-coincidence of the microphone capsules cause coloration and errors in the corresponding B-format signals and linear combinations thereof. Various strategies for designing B-format non-coincidence correction filters are compared and limitations are discussed.
Convention Paper 7766

18:00

P21-7 Generalized Multiple Sweep Measurement— *Stefan Weinzieri, Andre Giese, Alexander Lindau*, TU Berlin, Berlin, Germany

A system identification by impulse response measurements with multiple sound source configurations can benefit greatly from time-efficient measurement procedures. An optimized method by interleaving and overlapping of multiple exponential sweeps (MESM) used as excitation signals was presented by Majdak et al. (2007). For single system identifications, however, much higher signal-to-noise ratios (SNR) can be reached with sweeps whose magnitude spectra are adapted to the background noise spectrum of the acoustical environment, as proposed by Müller & Massarani (2001). We investigated on which conditions and to what extent the efficiency of multiple sweep measurements can be increased by using arbitrary, spectrally adapted sweeps. An extension of the MESM approach toward generalized sweep spectra is presented, along with a recommended measurement procedure and a prediction of the efficiency of multiple sweep measurements depending on typical measurement conditions.
Convention Paper 7767

Session P22 Saturday, May 9
15:00 – 18:30 Room K3

MICROPHONES AND HEADPHONES

Chair: **William Evans**, University of Surrey, Guildford, Surrey, UK

15:00

P22-1 Frequency Response Adaptation in Binaural Hearing—*David Griesinger*, Consultant, Cambridge, MA, USA

The pinna and ear canals act as listening trumpets to concentrate sound pressure on the eardrum. This concentration is strongly frequency dependent, typically showing a rise in pressure of 20 dB at 3000 Hz. In addition, diffraction and reflections from the pinna substantially alter the frequency response of the eardrum pressure as a function of the direction of a sound source. In spite of these large departures from flat response, listeners usually report that a uniform pink power spectrum sounds frequency balanced, and loudspeakers are manufactured to this standard. But on close listening frontal pink noise does not sound uniform. The ear clearly uses adaptive correction of timbre to achieve these results. This paper discusses and demonstrates the properties and limits of this adaptation. The results are important for our experience of live music in halls and in reproduction of music through loudspeakers and headphones.
Convention Paper 7768

15:30

P22-2 Concha Headphones and Their Coupling to the Ear—*Lola Blanchard*,[1] *Finn T. Agerkvist*[2]
[1]Bang & Olufsen ICEpower s/a, Lyngby, Denmark
[2]Technical University of Denmark, Lyngby, Denmark

The purpose of the study is to obtain a better understanding of concha headphones. Concha headphones are the small types of earpiece that are placed in the concha. They are not sealed to the ear and therefore, there is a leak between the earpiece and the ear. This leak is the reason why there is a significant lack of bass when using such headphones. This paper investigates the coupling between the headphone and the ear, by means of measurement in artificial ears and models. The influence of the back volume is taken into account.
Convention Paper 7769
Paper presented by Finn Agerkvist

16:00

P22-3 Subjective Evaluation of Headphone Target Frequency Responses—*Gaëtan Lorho*, Nokia Corporation, Finland

The effect of headphone frequency response equalization on listeners' preference was studied for music and speech reproduction. The high-quality circum-aural headphones selected for this listening experiment were first equalized to produce a flat frequency response. Then, a set of filters was created based on two parameters

defining the amplitude and center frequency of the main peak found around 3 kHz in the free-field and diffuse-field equalization curves. Two different listening tests were carried out to evaluate these equalization candidates using a different methodology and a total of 80 listeners. The results of this study indicate that a target frequency response with a 3 kHz peak of lower amplitude than in the diffuse-field response is preferred by listeners for both music and speech.
Convention Paper 7770

16:30

P22-4 Study and Consideration on Symmetrical KEMAR HATS Conforming to IEC60959—
Kiyofumi Inanaga,[1] *Homare Kon*,[1] *Gunnar Rasmussen,*[2] *Per Rasmussen*,[2] *Yasuhiro Riko*[3]
[1]Sony Corporation, Tokyo, Japan
[2]G.R.A.S. Sound & Vibration A/S, Holte, Denmark
[3]Riko Associates, Yokohama, Japan

KEMAR is widely recognized as a leading model of head and torso simulators (HATS) for different types of acoustic measurements meeting requirements of a global industrial standard, ANSI S3.36/ASA58-1985 and IEC 60959:1990. One of the KEMAR HATS pinna models has a reputation for good reproducibility of measured results in examining headphones and earphones. However, it requires free field compensation in order to conduct the measurements; thus, the head-related transfer function (HRTF) of HATS fitted with the pinna model must be corrected. Because headphones and earphones are usually designed symmetrically, we developed a prototype of Symmetrical KEMAR HATS based on the original KEMAR mounted with the pinna model with good reproducibility. We measured and evaluated a set of HRTFs from the sound source to both ears. Our study concluded that the HATS we developed carries symmetrical characteristics and is also suitable to be utilized as a tool to measure the qualities of variety of acoustic devices along with the conventional KEMAR and it can serve as a new common platform for different types of electroacoustic measurements.
Convention Paper 7771

17:00

P22-5 Spatio-Temporal Gradient Analysis of Differential Microphone Arrays—*Mihailo Kolundzija,*[1] *Christof Faller*,[1] *Martin Vetterli*[1,2]
[1]Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland
[2]University of California at Berkeley, Berkeley, CA, USA

The literature on gradient and differential microphone arrays makes a distinction between the two, and nevertheless shows how both types can be used to obtain the same response. A more theoretically sound rationale for using delays in differential microphone arrays has not yet been given. This paper presents a gradient analysis of the sound field viewed as a spatio-temporal phenomenon, and gives a theoretical interpretation of the working principles of gradient and differential microphone arrays. It shows that both types of microphone arrays can be viewed as devices for approximately measuring the spatio-temporal derivatives of the sound field. Furthermore, it also motivates the design of high-order differential microphone arrays using the aforementioned spatio-temporal gradient analysis.
Convention Paper 7772

17:30

P22-6 The Analog Microphone Interface and its History—*Jörg Wuttke*, Jörg Wuttke Consultancy, Pfinztal, Germany, Schoeps GmbH, Karlsruhe, Germany

The interface between microphones and microphone inputs has special characteristics and requires special attention. The low output levels of microphones and the possible need for long cables have made it necessary to think about noise and interference of all kinds. A microphone input is also the electrical load for a microphone and can have an adverse influence on the its performance. Condenser microphones contain active circuitry that required some form of powering. With the introduction of transistorized circuitry in the 1960s, it became practical for this powering to be incorporated into microphone inputs. Various methods appeared in the beginning; 48-Volt phantom powering is now the dominant standard, but this standard method is still not always implemented correctly.
Convention Paper 7773

18:00

P22-7 Handling Noise Analysis in Large Cavity Microphone Windshields—Improved Solution —*Philippe Chenevez*, CINELA, Paris, France

Pressure gradient microphones are well known to be highly sensitive to vibrations. Respectable suspensions are made to create the best isolation possible, but when the microphone is placed inside a large cavity windshield, the external skin behaves as a drum excited by the vibrations of the support (boom or stand). As a consequence structure-borne noise is also transmitted acoustically to the microphone, due to its hard proximity effect. Some theoretical aspects and practical measurements are presented, in conjunction with a proposed improved solution.
Convention Paper 7774

Live Sound Seminar 2 — Saturday, May 9
15:00 – 18:30 — Atrium 2

SENNHEISER & YAMAHA—LIVE SOUND WORKSHOP

Presenters: **Svenja Dunkel**
Wayne "Heights" Gittens
Oliver Voges
Jürgen Wilhelm
Gregor Zielinsky

The workshop will cover two important aspects of PA:
(1) Digital mixing consoles for PA and IT basic requirements, e.g., switches, cables, fiber optics, connectors.
(2) A complete soundcheck of a live rockband on stage. The visitors will experience how high-end professionals do their soundcheck, including FOH, monitors, and wireless techniques.

The special task will be to get the tricky sound of the location under control. The hall Atrium 2 is very reverberant and is not an optimized condition. This is what pros have to work with every day. The band "Rauschenberger" is a new upcoming group from Hannover around singer and leader Rauschenberger, who has a splendid and very characteristic voice. After the soundcheck, there will be a half-hour concert.

Saturday, May 9 15:00 Room D120
Technical Committee Meeting on Electro Magnetic Compatibility

Saturday, May 9 15:00 Room D104
Standards Committee Meeting on SC-03-06 Digital Library and Archive Systems

Saturday, May 9 15:00 Room D105
Standards Committee Meeting on SC-04-01 Acoustics and Sound Source Modeling

Exhibitor Seminar Saturday, May 9
15:00 – 16:00 Exhibit Floor Booth 2303
ES14 OPTOCORE – A

TECHNICAL INTRODUCTION TO OPTOCORE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

The OPTOCORE Synchronous Fibre Network is introduced. Content will include an Introduction/Overview, design of the network, synchronicity, latency, topology, redundancy, protocol, and devices.

Workshop 16 Saturday, May 9
15:30 – 18:30 Room K2

LOUDSPEAKER FE/BE MODELING

Chair: **Steve Hutt**, Equity Sound Investments

Panelists: *Wolfgang Klippel*, Klippel GmbH
Hermann Landes, SIMetris GmbH
Rich Little, Tymphany Inc.
Peter Larsen, Loudsoft Ltd.
Patrick Macey, PACSYS Limited
Joerg Panzer, R and D Team GmbH
Pierre Thieffry, ANSYS Inc.

This workshop will explore FE/BE modeling of loudspeaker drivers. A case study of an existing loudspeaker driver will be modeled by each panelist to benchmark the capabilities of their modeling software package. A loudspeaker driver's dimensions and material properties will be provided to panelists in advance of the convention so that they may develop a thorough model for presentation at the workshop. Results will be discussed along with measurement analysis of the real loudspeaker driver.

Student Event/Career Development
CAREER/JOB FAIR
Saturday, May 9, 15:30 – 17:00
K4 Foyer

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!

Workshop 17 Saturday, May 9
16:00 – 18:30 Room K1

SOUND DESIGN FOR *DIE FALSCHER*

Presenters: **Tobi Fleig**, Dubbing Mixer
Rainer Heesch, Sound Designer
Tatjana Jakob, Sound Designer
Olaf Mierau, Postproduction Sound Supervisor

Die Fälscher (*The Counterfeiters*) was awarded the Oscar in 2008 for the best non-English-speaking film. Key members of the sound team will discuss the concepts of sound designing this movie, communication with the director, freedom and constraints, sound as a strong aesthetic means for the narrative, and other issues. Several examples from the film will be shown, sometimes in different versions and stages of the mix.

Saturday, May 9 16:00 Room D120
Technical Committee Meeting on Transmission and Broadcasting

Exhibitor Seminar Saturday, May 9
16:00 – 17:00 Exhibit Floor Booth 2303
ES15 OPTOCORE – B

BASICS OF FIBRE OPTICS

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

Technical, economical, safety aspects, and advantages of fibre optics is the topic of this seminar. Content will include fibre advantages, reliability, types of fibre, parameters of fibres (attenuation, bandwidth length product, latency), connector types, cleaning, and testing.

Session P23 Saturday, May 9
16:30 – 18:00 K4 Foyer

POSTERS: PSYCHOACOUSTICS AND PERCEPTION

16:30

P23-1 Influence of the Listening Room in the Perception of a Musical Work—*Nelia Valverde, Marcos D. Fernández, José Antonio Ballesteros, Leticia Martínez, Samuel Quintana, Isabel González*, Escuela Universitaria Politénica de Cuenca, Cuenca, Spain

The listening of the same musical composition generates a unique perception for every listener but, simultaneously, the specific acoustic conditions of the chosen room have a decisive influence on the perception. In order to evaluate such differences depending on the listening room, a musical work for choir has been composed and recorded with a HATS in an anechoic room, in a reverberant room, and in a normal room. With those records, surveys to professional musicians and non-expert listeners

have been carried out, once they have previously heard the recording with headphones, and finally, the answers obtained have been evaluated in order to determine the influence of the listening room in the perception of the musical work.
Convention Paper 7775

16:30

P23-2 Comparison of Methods for Measuring Sound Quality through HATS and Binaural Microphones—*José Antonio Ballesteros, Marcos D. Fernández, Samuel Quintana, Isabel González, Laura Rodríguez*, Escuela Universitaria Politécnica de Cuenca, Universidad de Castilla-La Mancha, Cuenca, Spain

Sound quality techniques are currently becoming more important as they take into account the human perception of sound. By now, there is no well established international standards for measuring sound quality and no well recognizable reference index for its assessment. Then, a HATS or a pair of binaural microphones can be used for measuring the typical sound quality parameters. A set of measurements, under the same condition, has been carried out using both devices for assessing the differences and the possible variation in the results. As a consequence of all of this, guidance is given for choosing the device that best fits depending on each measurement context.
Convention Paper 7776

16:30

P23-3 Improving Perceived Tempo Estimation by Statistical Modeling of Higher-Level Musical Descriptors—*Ching-Wei Chen, Markus Cremer, Kyogu Lee, Peter DiMaria, Ho-Hsiang Wu*, Gracenote, Inc., Emeryville, CA, USA

Conventional tempo estimation algorithms generally work by detecting significant audio events and finding periodicities of repetitive patterns in an audio signal. However, human perception of tempo is subjective and relies on a far richer set of information, causing many tempo estimation algorithms to suffer from octave errors, or “double/half-time” confusion. In this paper we propose a system that uses higher-level musical descriptors such as mood to train a statistical model of perceived tempo classes, which can then be used to correct the estimate from a conventional tempo estimation algorithm. Our experimental results show reliable classification of perceived tempo class, as well as a significant reduction of octave errors when applied to an array of available tempo estimation algorithms.
Convention Paper 7777

16:30

P23-4 Perceptually-Motivated Audio Morphing: Softness—*Duncan Williams, Tim Brookes*, University of Surrey, Guildford, Surrey, UK

A system for morphing the softness and brightness of two sounds independently from their other perceptual or acoustic attributes was coded. The system is an extension of a previous one that morphed brightness only, that was based on the Spectral Modeling Synthesis additive/residual model. A Multidimensional Scaling analysis, of listener responses to paired comparisons of stimuli generated by the morpher, showed movement in three perceptually-orthogonal directions. These directions were labeled in a subsequent verbal elicitation experiment that found that the effects of the brightness and softness controls were perceived as intended. A Timbre Morpher, adjusting additional timbral attributes with perceptually-meaningful controls, can now be considered for further work.
Convention Paper 7778

16:30

P23-5 Resolution of Spatial Distribution Perception with Distributed Sound Source in Anechoic Conditions —*Olli Santala, Ville Pulkki*, Helsinki University of Technology, Espoo, Finland

The resolution of directional perception of spatially distributed sound sources was investigated with a listening test in an anechoic chamber using various sound source distributions. Fifteen loudspeakers were used to produce test cases that included sound sources with varying widths and wide sound sources with gaps in the distribution. The subjects were asked to distinguish which loudspeakers emitted sound according to their own perception. Results show that small gaps in the sound source were not perceived accurately and wide sound sources were perceived narrower than they actually were. The results also indicate that the resolution for fine spatial details was worse than 15 degrees when the sound source was wide.
Convention Paper 7779

16:30

P23-6 Perceived Roughness—A Recent Psychoacoustic Measurement—*Robert Mores, Thorsten Smit, Jana-Marie Wiese*, University of Applied Science, Hamburg, Germany

This paper relates to an investigation on perceived roughness from Aures in 1984 where findings are based on psychoacoustic tests with synthetic sounds and a small group of people. The related results have repeatedly been used for modeling roughness perception since then, for instance in the context of noise perception. Roughness is again an issue when investigating the perceived quality or timbre of musical sounds. In this context roughness is one among some ten mid-level features to be extracted. Here, perceived roughness is measured again, but on a wider basis than in the earlier investigation. This paper outlines the psychoacoustic investigation, basically following the method of Aures, but modifying some of the issues under question. The results are reasonable and differ from the earlier findings in various aspects.
Convention Paper 7780

16:30

P23-7 A Physiological Auditory Model—*Václav Vencovsky*, Czech Technical University in Prague, Prague, Czech Republic

A physiological auditory model is described. The model simulates a processing of a sound by an ➥

outer, middle, and inner ear. A nonlinear inner ear model comprises the cochlear frequency selectivity model and the inner hair cells model proposed according to mammalian physiological data. A capability of the auditory model to simulate human psychophysical masking data is verified.
Convention Paper 7781

Saturday, May 9 17:00 Room D104
Standards Committee Meeting on SC-03-07 Audio Metadata

Saturday, May 9 17:00 Room D105
Standards Committee Meeting on SC-03-12 Forensic Audio

Exhibitor Seminar Saturday, May 9
17:00 – 18:00 Exhibit Floor Booth 2303
ES16 OPTOCORE – C

OPTOCORE IN PRACTICE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

This seminar will cover practical exercises with OPTOCORE system/devices. Content will include practical set-up of a network, remote pre-amp control through third party devices, transport of control data and Ethernet, software control, and firmware upgrades.

Student Event/Career Development
STUDENT DELEGATE ASSEMBLY MEETING—PART 2
Saturday, May 9, 17:30 – 18:30
Room D123

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.

Special Event
LIVE CONCERT
Saturday, May 9, 18:30 – 19:00
Atrium 2

The band featured in the Live Sound Workshop LS2, Rauschenberger, will continue to play after the workshop finishes, in a concert open to all attendees.

Special Event
MIXER PARTY
Saturday, May 9, 18:30 – 19:30
The Bistro (between Atriums 3 and 4)

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

Student Event/Career Development
STUDENT PARTY
Saturday, May 7, 20:00 –22:00
saf at Zerwirk

Both dinner and party will take place at the *saf at Zerwirk*, which is restaurant in downtown Munich. The dinner is scheduled for 8 pm, the party will start at 10 pm. The restaurant is not open to the public on that evening, so it is a private event. The restaurant is located in the Ledererstrasse only a two minute walk from Marienplatz, Munich's central square and an important transportation hub.

The food: Saf is a vegan restaurant. There will either be a buffet or a menu for approximately EUR 20 apiece. There will be fifty tickets for the dinner that will be sold after SDA-1 and at the Student Science Spot.

The party: After the dinner there will be the chance to get together and have a good time at the Student Party. The Graz student section will take care of the music. There won't be an entrance fee for the party, but of course everybody will have to pay for their own drinks.

Special Event
ORGAN RECITAL BY GRAHAM BLYTH
Saturday, May 9, 20:30 – 21:30
Cathedral Church of Our Lady (also called the Liebfrauendom)
Frauenplatz 1, Munich

Graham Blyth's traditional organ recital will be given on the organ of the Münchner Dom, the "Frauenkirche." The 99 and 100 meter high respectively Twin Towers of the Frauenkirche with the "Welch" caps are high up in the sky of Munich. This is the oldest landmark of the city. Built between 1468 and 1488 by Jörg von Halback (nicknamed "Ganghofer"), this cathedral is the biggest late-gothic hall church of Europe. Here you will find the splendid tomb of the German Kaiser (emperor) Ludwig the Bavarian.

The featured works are by Bach, plus Mendelssohn's 1st Sonata, and Vierne's 1st Symphony.

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

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

Session P24 Sunday, May 10
09:00 – 12:30 Room K3

ASSESSMENT AND EVALUATION

Chair: **Gaëtan Lorho**, Nokia Corporation, Finland

09:00

P24-1 Influence of Level Setting on Loudspeaker Preference Ratings—*Vincent Koehl, Mathieu Paquier*, Université de Brest, Plouzané, France

The perceived audio quality of a sound-reproduction device such as a loudspeaker is hard to evaluate. Industrial and academic researchers are still focusing on the design of reliable assessment procedures to measure this subjective character. One of the main issues of listening tests is about their validity in regard to real comparison situations (Hi-Fi magazine evaluations, audiophile, sound engineer, customer, etc.). Are the conclusions of laboratory tests consistent with these almost informal comparisons? As an example, one of the main differences between listening tests and real-life comparisons is about the loudness matching. This paper is aimed at comparing paired-comparison tests that are commonly accomplished under laboratory conditions with a procedure assumed to be closer to real-life conditions. It shows that differences in the test procedures led to differences in the subjective assessments.
Convention Paper 7782

09:30

P24-2 Comparing Three Methods for Sound Quality Evaluation with Respect to Speed and Accuracy—*Florian Wickelmaier,[1] Nora Umbach,[1] Konstantin Sering,[1] Sylvain Choisel[2]*
[1]University of Tübingen, Tübingen, Germany
[2]Bang & Olufsen A/S, Struer, Denmark, now at Philips Consumer Lifestyle, Leuven, Belgium

The goal of the present study was to compare three response-collection methods that may be used in sound quality evaluation. To this end, 52 listeners took part in an experiment where they assessed the audio quality of musical excerpts and six processed versions thereof. For different types of program material, participants performed (a) a direct ranking of the seven sound samples, (b) pairwise comparisons, and (c) a novel procedure, called ranking by elimination. The latter requires subjects on each trial to eliminate the least preferred sound; the elimination continues until only the sample with the highest audio quality is left. The methods are compared with respect to the resulting ranking/scaling and the time required to obtain the results.
Convention Paper 7783

10:00

P24-3 Reference Units for the Comparison of Speech Quality Test Results—*Nicolas Côté,[1,2] Vincent Koehl,[3] Valérie Gautier-Turbin,[4] Alexander Raake,[2] Sebastian Möller[2]*
[1]Ecole Nationale d'Ingénieurs de Brest, Plouzané, France
[2]Deutsche Telekom Laboratories, Berlin, Germany
[3]Université de Brest, Plouzané, France
[4]France Telecom R&D, Lannion, France

Subjective tests are carried out to assess the quality of an entity as perceived by a user. However, several characteristics inherent to the subject or to the test methodology might influence the users' judgments. As a result, reference conditions are usually included in subjective tests. In the field of quality of transmitted speech, reference conditions correspond to a speech sample impaired by a known amount of degradation. In this paper several kinds of reference conditions and the process used for their production are presented. Examples of the corresponding normalization procedure of each kind of reference are given.
Convention Paper 7784

10:30

P24-4 The Influence of Sound Processing on Listeners' Program Choice in Radio Broadcasting—*Hans-Joachim Maempel, Fabian Gawlik*, Technische Universität Berlin, Berlin, Germany

Many opinions on broadcast sound processing are founded on tacit assumptions about certain effects on listeners. These, however, have lacked support by internally and ecologically valid empirical data so far. Thus, under largely realistic conditions it has been experimentally investigated to what extent broadcast sound processing influences listeners' program choice. Technical features of stimuli, socio-demographic data of the test persons, and data of listening conditions have been additionally collected. In the main experiment, subjects were asked to choose one out of six audio stimuli varied in content and sound processing. The varied sound processing caused marginal and statistically not significant differences in frequencies of program choice. By contrast, a subsequent experiment enabling a direct comparison of different sound processings of the same audio content yielded distinct preferences for certain sound processings.
Convention Paper 7785

11:00

P24-5 Free Choice Profiling and Natural Grouping as Methods for the Assessment of Emotions in Musical Audio Signals—*Sebastian Schneider,[1] Florian Raschke,[1] Gabriel Gatzsche,[2] Dominik Strohmeier[1]*
[1]Ilmenau University of Technology, Ilmenau, Germany
[2]Fraunhofer Institute for Digital Media Technology, IDMT, Ilmenau, Germany;

To measure the perceived emotions caused by musical audio signals we propose to use "Free Choice Profiling" (FCP) combined with "Natural Grouping" (NG). FCP/NG—originally derived from food research and new to the research of music perception—allow participants to evaluate stimuli using their own vocabulary. To evaluate the proposed methods, we conducted an experiment where 16 participants had to assess major-major and minor-minor chord pairs. Unlike one could expect, allowing participants to express themselves freely does not lead to a degeneration of the quality of the data. Instead, clearly interpretable results consistent with music theory and emotional psychology were obtained. These results encourage further investigations, which could lead to a general method for assessing emotions in music.
Convention Paper 7786

11:30

P24-6 Subjective Quality Evaluation of Audio Streaming Applications on Absolute and Paired Rating Scales—*Bernhard Feiten, Alexander Raake, Marie-Neige Garcia, Ulf Wüstenhagen, Jens Kroll*, Deutsche Telekom Laboratories, Berlin, Germany

In the context of the development of a parametric model for the quality assessment of audiovisual IP-based multimedia applications, audio tests have been carried out. The test method used for the subjective audio tests was aligned to the method used for video tests. Hence, the Absolute Category Ranking (ACR) method was applied. To prove the usability of ACR tests for this purpose MUSHRA and ACR were applied in parallel listening tests. The MPEG audio codecs AAC, HE-AAC, MP2, and MP3 at different bitrates and different packet loss conditions were evaluated. The test results show that the ACR method also reveals the quality differences for higher qualities, even though MUSHRA has superior resolution.
Convention Paper 7787

12:00

P24-7 Assessor Selection Process for Multisensory Applications—*Søren Vase Legarth, Nick Zacharov*, DELTA SenseLab, Hørsholm, Denmark

Assessor panels are used to perform perceptual evaluation tasks in the form of listening and viewing tests. In order to ensure the quality of collected data it is vital that the selected assessors have the desired qualities in terms of discrimination aptitude as well as consistent rating ability. This work extends existing procedures in this field to provide a statistically robust and efficient manner for assessing and evaluating the performance of assessors for listening and viewing tasks.
Convention Paper 7788

Session P25 **Sunday, May 10**
09:00 – 12:30 **Room K4**

SOUND DESIGN AND PROCESSING

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

09:00

P25-1 Hierarchical Perceptual Mixing—*Alexandros Tsilfidis, Charalambos Papadakos, John Mourjopoulos*, University of Patras, Patras, Greece

A novel technique of perceptually-motivated signal-dependent audio mixing is presented. The proposed Hierarchical Perceptual Mixing (HPM) method is implemented in the spectro-temporal domain; its principle is to combine only the perceptually relevant components of the audio signals, derived after the calculation of the minimum masking threshold, which is introduced in the mixing stage. Objective measures are presented indicating that the resulting signals have enhanced dynamic range and lower crest factor with no unwanted artifacts, compared to the traditionally mixed signals. The overall headroom is improved, while clarity and tonal balance are preserved.
Convention Paper 7789

09:30

P25-2 Source-Filter Modeling in Sinusoid Domain—*Wen Xue, Mark Sandler*, Queen Mary, University of London, London, UK

This paper presents the theory and implementation of source-filter modeling in sinusoid domain and its applications on timbre processing. The technique decomposes the instantaneous amplitude in a sinusoid model into a source part and a filter part, each capturing a different aspect of the timbral property. We show that the sinusoid domain source-filter modeling is approximately equivalent to its time or frequency domain counterparts. Two methods are proposed for the evaluation of the source and filter, including a least-square method based on the assumption of slow variation of source and filter in time, and a filter bank method that models the global spectral envelope in the filter. Tests show the effectiveness of the algorithms for isolation frequency-driven amplitude variations. Example applications are given to demonstrate the use of the technique for timbre processing.
Convention Paper 7790

10:00

P25-3 Analysis of a Modified Boss DS-1 Distortion Pedal —*Matthew Schneiderman, Mark Sarisky*, University of Texas at Austin, Austin, TX, USA

Guitar players are increasingly modifying (or paying someone else to modify) inexpensive mass-produced guitar pedals into boutique units. The Keeley modification of the Boss DS-1 is a prime example. In this paper we compare the measured and perceived performance of a Boss DS-1 before and after applying the Keeley All-Seeing-Eye and Ultra mods. This paper sheds light on psychoacoustics, signal processing, and guitar recording techniques in relation to low fidelity guitar distortion pedals.
Convention Paper 7791

10:30

P25-4 Phase and Amplitude Distortion Methods for Digital Synthesis of Classic Analog Waveforms—*Joseph Timoney,[1] Victor Lazzarini,[1] Brian Carty,[1] Jussi Pekonen[2]*
[1]NUI Maynooth, Maynooth, Ireland
[2]Helsinki University of Technology, Espoo, Finland

An essential component of digital emulations of subtractive synthesizer systems are the algorithms used to generate the classic oscillator waveforms of sawtooth, square, and triangle waves. Not only should these be perceived to be authentic sonically, but they should also exhibit minimal aliasing distortions and be computationally efficient to implement. This paper examines a set of novel techniques for the production of the classic oscillator waveforms of analog subtractive synthesis that are derived from using amplitude or

phase distortion of a mono-component input waveform. Expressions for the outputs of these distortion methods are given that allow parameter control to ensure proper bandlimited behavior. Additionally, their implementation is demonstrably efficient. Last, the results presented illustrate their equivalence to their original analog counterparts.
Convention Paper 7792

11:00

P25-5 Soundscape Attribute Identification—*Martin Ljungdahl Eriksson, Jan Berg*, Luleå University of Technology, Luleå, Sweden

In soundscape research, the field's methods can be employed in combination with approaches involving sound quality attributes in order to create a deeper understanding of sound images and soundscapes and how these may be described and designed. The integration of four methods are outlined, two from the soundscape domain and two from the sound engineering domain.
Convention Paper 7793

11:30

P25-6 SonoSketch: Querying Sound Effect Databases through Painting—*Michael Battermann,*[1] *Sebastian Heise,*[1] *Jörn Loviscach*[2]
[1]Hochschule Bremen (University of Applied Sciences), Bremen, Germany
[2]Fachhochschule Bielefeld (University of Applied Sciences), Bielefeld, Germany

Numerous techniques support finding sounds that are acoustically similar to a given one. It is hard, however, to find a sound to start the similarity search with. Inspired by systems for image search that allow drawing the shape to be found, we address quick input for audio retrieval. In our system, the user literally sketches a sound effect, placing curved strokes on a canvas. Each of these represents one sound from a collection of basic sounds. The audio feedback is interactive, as is the continuous update of the list of retrieval results. The retrieval is based on symbol sequences formed from MFCC data compared with the help of a neural net using an editing distance to allow small temporal changes.
Convention Paper 7794

12:00

P25-7 Generic Sound Effects to Aid in Audio Retrieval—*David Black,*[1] *Sebastian Heise,*[1] *Jörn Loviscach*[2]
[1]Hochschule Bremen (University of Applied Sciences), Bremen, Germany
[2]Fachhochschule Bielefeld (University of Applied Sciences), Bielefeld, Germany

Sound design applications are often hampered because the sound engineer must either produce new sounds using physical objects, or search through a database of sounds to find a suitable sample. We created a set of basic sounds to mimic these physical sound-producing objects, leveraging the mind's onomatopoetic clustering capabilities. These sounds, grouped into onomatopoetic categories, aid the sound designer in music information retrieval (MIR) and sound categorization applications. Initial testing regarding the grouping of individual sounds into groups based on similarity has shown that participants tended to group certain sounds together, often reflecting the groupings our team constructed.
Convention Paper 7795

Workshop 18 — Sunday, May 10
09:00 – 11:30 — Room K1

BLU-RAY—THE NEXT (AND ONLY?) CHANCE FOR HIGH RESOLUTION MUSIC MEDIA

Presenters: **Johannes Müller**, msm mastering Munich
Ronald Prent, Freelance Engineer, Galaxy Studios

The concept of utilizing Blu-ray as a pure audio format will be explained, and Blu-ray will be positioned as successor to both SACD and DVD-A. The operational functionality and a double concept of making it usable both with and without screen will be demonstrated by showing a few products that are already on the market.

Workshop 19 — Sunday, May 10
09:00 – 11:30 — Room K2

AUDIO OVER IP

Chair **Heinz-Peter Reykers**, WDR, Germany

Panelists: *Joost Bloemen*, Technica Del Arte BV
Jeremy R. Cooperstock, McGill University, Montreal, Canada
Gerald List, TransTel Communications GmbH
Peter Stevens, BBC R&D

In radio broadcasting most transmissions and communication between remote broadcast venues and the radio house are based on ISDN lines. Within TV the communication lines are also based on ISDN circuits. But the technology of ISDN networks is gradually being exchanged by IP networks and packet- based transmission. What are the consequences of this migration for the broadcasters? In this workshop the focus will be on practical demonstrations. The topics range from SIP-Servers and infrastructure issues to audio examples regarding latency and other related parameters. Quality of Service and current problems as well as possible solutions will be examined.

Sunday, May 10 — 09:00 — Room D104
Standards Committee Meeting: AESSC Plenary

Session P26 — Sunday, May 10
10:30 – 12:00 — K4 Foyer

POSTERS: ROOM ACOUSTICS AND LOUDSPEAKER INTERACTION

10:30

P26-1 Acoustic Design of Classrooms—*Suthikshn Kumar*, PESIT, Bangalore, India

Acoustic principles when used effectively in classroom design can improve the audibility of ➥

the professor in dramatic way. The cost-effective way to enhance the acoustics serves several purposes: Less speaking effort on the part of the lecturer; students can easily hear the lecturer more clearly; improved communication and, hence, improved learning experience. Several improvements can be done to the classroom architecture to enhance the signal-to-noise ratio, reduce reverberation, and background noise. We propose an innovative way of providing parabolic reflectors near the platform for amplifying the lecturer's voice. This paper focuses on the cost-effective, energy efficient acoustic design of classrooms.
Convention Paper 7796

10:30

P26-2 Epidaurus: Comments on the Acoustics of the Legendary Ancient Greek Theater—*Christos Goussios, Christos Sevastiadis, Kalliopi Chourmouziadou, George Kalliris*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The ancient Greek theaters and especially the well preserved theater of Epidaurus are of great interest because of their legendary acoustic characteristics. In the present paper the history and the construction characteristics of the specific theater are presented. The differences between the ancient and modern use of it are explained. Important acoustic parameters calculated using in situ measurements are presented. The conclusions show the relation between its excellent acoustic performance and the obtained results.
Convention Paper 7797

10:30

P26-3 A Matlab Toolbox for the Analysis of Ando's Factors—*Dario D'Orazio, Paolo Guidorzi, Massimo Garai*, University of Bologna, Bologna, Italy

The autocorrelation and crosscorrelation functions analysis, as well-known in literature, obtains remarkable results in different scientific fields. The autocorrelation function (ACF) and the interaural crosscorrelation function (IACF) analysis in architectural acoustics is known thanks to Y. Ando's work. The Toolbox presented in this work has been developed in order to compute Ando's significant and spatial factors (as the factors obtained from ACF and IACF are called), to subjective preference functions and to investigate further applications.
Convention Paper 7798

Live Sound Seminar 3 — Sunday, May 10
10:30 – 12:30 — Atrium 2

YAMAHA AND D & B

Presenters: **Stefan Goertz**
Arthur Koll

d & b

Line Source to Point Source Transformation —Technical Concept of the d & b T-Series

Increasing demands on flexibility, scalability, and efficiency in sound reinforcement applications encouraged one loudspeaker development configurable for point and line source applications by an easy mechanical modification. Several implemented design technologies will be discussed before the listening demonstration of the performance of the system under critical acoustic conditions comparing Q-Series line arrays.

Yamaha

Modern IT-Compatible Audio Networks

IT-compatible audio networks and standards: Cobranet and Ethersound. Both formats will be discussed regarding their advantages and limitations. Advanced network strategies like VLAN programming offers high channel counts and a wide range of additional services via Gigabit audio networks. Now video, intercom, remote control, and DMX services may be included in a modern network infrastructure. Of course, such a network has to be as safe and stable as possible, so the redundancy concepts developed by the IT industry like link aggregation / trunking / spanning tree will be discussed.

Exhibitor Seminar — Sunday, May 10
11:00 – 12:00 — Exhibit Floor Booth 2303
ES17 OPTOCORE – A

TECHNICAL INTRODUCTION TO OPTOCORE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

The OPTOCORE Synchronous Fibre Network is introduced. Content will include an Introduction/Overview, design of the network, synchronicity, latency, topology, redundancy, protocol, and devices.

Workshop 20 — Sunday, May 10
12:00 – 14:00 — Room K2

STANDARDS-BASED AUDIO NETWORKS USING IEEE 802.1 AVB

Presenters: **Ed Clarke**, XMOS Semiconductor, Bristol, UK
Rick Kreifeldt, Harman International, UT, USA

Recent work by IEEE 802 working groups will allow vendors to build a standards-based network with the appropriate quality of service for high quality audio performance and production. This new set of standards, developed by the IEEE 802.1 Audio Video Bridging Task Group, provides three major enhancements for 802-based networks: (1) Precise timing to support low-jitter media clocks and accurate synchronization of multiple streams; (2) A simple reservation protocol that allows an endpoint device to notify the various network elements in a path so that they can reserve the resources necessary to support a particular stream; and (3) Queuing and forwarding rules that ensure that such a stream will pass through the network within the delay specified by the reservation. These enhancements require no changes to the Ethernet lower layers and are compatible with all the other functions of a standard Ethernet switch (a device that follows the IEEE 802.1Q bridge specification). As a

result, all of the rest of the Ethernet ecosystem is available to developers, in particular, the various high speed physical layers (up to 10 gigabit/sec in current standards, even higher speeds are in development), security features (encryption and authorization), and advanced management (remote testing and configuration) features can be used. This workshop will outline the basic protocols and capabilities of AVB networks, describe how such a network can be used, and provide some simple demonstrations of network operation (including a live comparison with a legacy Ethernet network).

Workshop 21 **Sunday, May 10**
12:00 – 14:00 **Room K1**

MULTIMODAL INTEGRATION: AUDIO-VISUAL-HAPTIC-TACTILE

Cochairs: **Durand Begault**, NASA Ames Research Center
Ellen Haas, US Army Research Lab, MD, USA

Panelists: *Ercan Altinsoy*, Technical University of Dresden, Dresden, Germany
Jeremy Cooperstock, McGill University, Montreal, Quebec, Canada
Jürgen Hellbrück, KU Eichstätt, Eichstätt, Germany
Gerhard Mauter, Audi AG, Germany
Alexander S. Treiber, ACC Akustik, Germany

To design the best human interfaces for complex systems such as automobiles, mixing consoles, shared reality spaces, games, military, and aerospace, an increasing number of designers, psychologists, and human factors specialists are reconsidering multimodal displays of information that includes haptic and tactile feedback in addition to auditory-visual communication. This workshop considers in particular the way in which various modalities of information can be best integrated for a particular application, and identifies future requirements for such research.

Exhibitor Seminar **Sunday, May 10**
12:00 – 13:00 **Exhibit Floor Booth 2303**
ES18 OPTOCORE – B

BASICS OF FIBRE OPTICS

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

Technical, economical, safety aspects, and advantages of fibre optics is the topic of this seminar. Content will include fibre advantages, reliability, types of fibre, parameters of fibres (attenuation, bandwidth length product, latency), connector types, cleaning, and testing.

Session P27 **Sunday, May 10**
13:00 – 17:00 **Room K3**

SIGNAL PROCESSING

Chair: **Günther Thiele**

13:00

P27-1 On the Myth of Pulse Width Modulated Spectrum in Theory and Practice—*Arnold Knott*,[1,2] *Gerhard Pfaffinger*,[2] *Michael A. E. Andersen*[1]
[1]Technical University of Denmark, Lyngby, Denmark
[2]Harman/Becker Automotive Systems GmbH, Straubing, Germany

Switch-mode audio power amplifiers are commonly used in sound reproduction. Their well-known drawback is the radiation of high frequent energy, which can disturb radio and TV receivers. The designer of switch-mode audio equipment therefore need to make arrangements to prevent this coupling, which would otherwise result in bad audio performance. A deep understanding of the pulse width modulated (PWM) signal is therefore essential, which resulted in different mythic models as pulse, trapezoidal, or Double Fourier Series (DFS) representations in the past. This paper will clarify these theoretical approaches by comparing them with reality from both the time and the frequency domain perspective. For validation a switch-mode audio power amplifier was built, delivering the contents material with less than 0.06 percent distortion across the audio band at 50 W. The switch-mode signals have been evaluated very precisely in time and spectral domain to enlighten the assumptions about the PWM spectra and decrypt this myth.
Convention Paper 7799

13:30

P27-2 Design Approaches for Psychoacoustical In-Band Noise Shaping Filters—*Jochen Hahn*, University of Kaiserslautern, Kaiserslautern, Germany

Noise shaping is a state-of-the-art technique to preserve the perceived quality of audio signals when requantization happens. Noise shaping filters are special filters because of the nonlinear characteristics in hearing. They have to be taken into account when designing these special digital audio filters. The design approaches presented in this contribution meet these requirements. They minimize or limit the filter magnitude, the unweighted noise amplification, and the group delay characteristics of the filter.
Convention Paper 7800

14:00

P27-3 A New Analog Input Topology for Extreme Dynamic Range Analog to Digital Conversion—*Jamie Angus*, University of Salford, Salford, Greater Manchester, UK

The purpose of this paper is to introduce a novel form of input topology for the analog inputs of oversampled analog to digital converters. This new topology, when used with existing components, can achieve a dynamic range of 28 linear bits but has the potential to achieve even more if suitable technology can be developed. The paper analyzes the current limitations of existing topologies, presents the new topology, and shows how it can achieve much higher dynamic ranges. The optimal application of the topology and means of extending ➥

it for higher dynamic ranges is also discussed.
Convention Paper 7801

14:30

P27-4 Automatic Equalization of Flat TV Loudspeakers Using Parametric IIR Filters—*Herwig Behrends,*[1] *Adrian von dem Knesebeck,*[2] *Werner Bradinal,*[1] *Peter Neumann,*[1] Udo Zölzer[2]
[1]NXP Semiconductors, Hamburg, Germany
[2]Helmut Schimidt University, University of the Federal Armed Forces, Hamburg, Germany

Small loudspeakers used in today's flat television set cabinets and the requirement for "invisible sound" lead to a frequency response that is influenced in a very disadvantageous way by the physical design constraints. Loudspeakers are deeply embedded within the cabinet—the sound is thus forced through narrow vents, funnels or other waveguides. Down- or backfiring placements of the loudspeakers are also common practice, in order to minimize their visibility as much as possible. Generally, this leads to a non-flat frequency response with a strong coloration of the sound. We present an approach to compensate these effects by means of simple second order equalizer sections (biquads), where center frequency, gain, and bandwidth of the equalizer sections are automatically calculated from a measured frequency response. The tool is usable in a laboratory environment, with relatively inexpensive standard PC sound cards and microphones.
Convention Paper 7802

15:00

P27-5 Audio n-Genie: Domain Specific Language for Audio Processing—*Tiziano Leidi,*[1] *Thierry Heeb,*[2] *Marco Colla,*[1] *Jean-Philippe Thiran*[3]
[1]Institute for Applied Computer Science and Industrial Technologies of Southern Switzerland (ICIMSI), Manno, Switzerland
[2]ANAGRAM Technologies SA, Préverenges, Switzerland
[3]Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland

Specialized development environments represent today an important added value for domain specific system providers suffering the lack of a dedicated, ergonomic, efficient, and affordable tool able to boost their core business. This paper describes Audio n-genie, a domain-specific language and its associated development environment supporting the automatic production, by means of component-based model-driven generative programming, of digital audio processing applications.
Convention Paper 7803

15:30

P27-6 Acoustic Echo Cancellation Using MIMO Blind Deconvolution—*Ephraim Gower, Malcolm Hawksford*, University of Essex, Colchester, UK

A new multiple-input multiple-output frame-processing algorithm is introduced that exploits blind deconvolution for acoustic echo cancellation. The channel deconvolution filters, which can be blindly estimated as either finite impulse or infinite impulse responses, are optimized by maximizing the information flow through several nonlinear neurons. The algorithm requires that for every system audio output there be a corresponding microphone for effective feedback signal separation/ cancellation. The desired talker signal from the algorithm outputs is recognized and transmitted while the identified feedback signals are discarded.
Convention Paper 7804

16:00

P27-7 Implementing Audio Algorithms and Integrating Processor-Specific Code Using Model Based Design—*Arvind Ananthan,*[1] *Mark Corless,*[2] *Marco Roggero*[3]
[1]The MathWorks, Natick, MA, USA
[2]The MathWorks, Novi, MI, USA
[3]The MathWorks, Ismaning, Germany

This paper explores the final stages in the model-based design and implementation of an audio algorithm on a fixed-point embedded processor. Once the algorithm, a 3-band parametric equalizer in this example, is designed and simulated using a combination of scripting and graphical modeling tools, real-time C-code is automatically generated from this model. This paper illustrates how algorithmic C-code generated from such a model in Simulink can be integrated into a standalone embedded project as a library and implemented on an Analog Devices Blackfin® 537 processor. It also elaborates how processor specific C-callable assembly code can then be integrated into the model for both simulation and code generation to improve its execution performance on this processor.
Convention Paper 7805

16:30

P27-8 Subjective and Objective Evaluation of the Audio Vacuum-Tube Amplifiers—*Andrzej Dobrucki,*[1] *Stanislaw Maleczek,*[2] *Maurycy Kin*[1]
[1]Wroclaw University of Technology, Wroclaw, Poland
[2]Military Institute of Engineering Technology, Wroclaw, Poland

The subjective and objective evaluation of 5 high-quality vacuum-tube audio amplifiers is presented in this paper. As the reference the professional transistor amplifier has been used. The subjective evaluation has been done by the team of judges as well as with the computer-based psychoacoustic model according with PAQM protocol. The amplifiers' sound quality assessed by the listeners is consistent with the one evaluated with the use of the psychoacoustic model. It was found that the best sound quality is obtained by vacuum-tube amplifiers, the worst—by the reference amplifier. The results of subjective evaluation are inconsistent with quality assessed by measurement of objective parameters: all amplifiers have comparable quality, but the best is the transistor amplifier because of lowest level of THD+N level.
Convention Paper 7806

Session P28 — Sunday, May 10
13:00 – 16:30 — Room K4

PSYCOACOUSTICS AND PERCEPTION

Chair: **Florian Wickelmaier**, University of Tübingen, Tübingen, Germany

13:00

P28-1 Localization of Consecutive Sound Events in Reverberant Environment—*Marko Takanen, Antti Jylhä, Tapani Pihlajamäki, Juha Holm, Ilkka Huhtakallio, Ville Pulkki*, Helsinki University of Technology, Espoo, Finland

A listening test was conducted to assess the localization of consecutive sound events in simulated reverberant conditions. The stimuli consisted of two sound events, which were reverberant wideband harmonic sounds reproduced in a multichannel anechoic chamber. Localization threshold for the latter sound event was measured as the direct-to-reverberant sound level ratio with an adaptive transformed up-down method. The studied factors affecting the localization threshold were the time interval and pitch difference between the two sound events and the time gap between the direct sound and reverberation. The results indicate that all factors have a significant effect on localization.
Convention Paper 7807

13:30

P28-2 The Contrasting and Conflicting Definitions of Envelopment—*Jan Berg*, Luleå University of Technology, Luleå, Sweden

In spatial audio, the term envelopment is not unambiguously defined and the different de facto definitions both overlap and contradict one another. This unclarity may pose a problem where the sensation of being surrounded by sound is subject for investigation and analysis. This paper reviews the different concepts of envelopment in order to point to where possible problems may occur. A tentative suggestion for a terminology that can serve the different contexts of enveloping sounds is also given.
Convention Paper 7808

14:00

P28-3 Apparent Source Width in ITU Surround—*Jorge Medina Victoria, Thomas Görne*, Hamburg University of Applied Sciences, Hamburg, Germany

Apparent Source Widths (ASW) of phantom images in a ITU-R BS.775-1 standard surround loudspeaker configuration have been investigated for different signals by means of a randomized blind test. Test signals were generated from anechoic recordings by amplitude panning between adjacent channels. The listening test showed that an increase of Apparent Source Width coincides with the increase of localization uncertainty at the side and back areas of the ITU setup. Largest ASW values were found between RS and LS channels.
Convention Paper 7809

14:30

P28-4 A New Methodological Approach to the Noise Threat Evaluation Based on the Selected Physiological Properties of the Human Hearing System—*Jozef Kotus*,[1] *Bozena Kostek*,[1,2] *Andrzej Czyzewski*[1]
[1]Gdansk University of Technology, Gdansk, Poland
[2]Excellence Center PROKSIM Institute of Physiology and Pathology of Hearing, Warsaw, Poland

A new way of assessment of noise-induced harmful effects on the human hearing system is presented in this paper. The method takes into consideration properties of the selected physiological human hearing system. On the basis of the hearing examinations and noise measurements results and psychoacoustical noise dosimeter performance the new indicators of the noise harmfulness were proposed. The evaluation of the proposed indicators were conducted on the basis of hearing examinations in the real noise exposure situations and also on the basis of the simulation results using standard test signals (such as white, pink, and brown noise). The performed analysis and obtained results confirmed the practical usefulness and correctness of the proposed indicators.
Convention Paper 7813
Paper presented by Bozena Kostek

15:00

P28-5 Time-Expanded Speech to Improve the Intelligibility of Consonants for the Hard-of-Hearing—*N H Shobha*,[1] *T G Thomas*,[2] *K Subbarao*[1]
[1]Osmania University, Hyderabad, India
[2]BITS-Pilani, Dubai, UAE

Our investigations addressed the efficacy of temporal processing to the clear speech advantage. A case for synthetic clear speech in the context of hearing impairment was constituted. Stop–vowels of English language were used as test stimuli, which were subjected to non-uniform time-expansion. Burst duration, voice onset time, and formant transition duration segments were selectively time-expanded by 50 to 100 percent of their original duration. Hearing impairment was simulated at three SNR levels. The perceptual analysis was accomplished in terms of information transmission analysis measure and statistical measures. Burst duration expansion by 50 percent was reported to be beneficial at a lower SNR level; while VOT and FTD lengthening did not contribute to improved intelligibility.
Convention Paper 7810
Paper not presented but available for purchase

15:30

P28-6 Octave-Band Analysis on ITU-R Listening Test Data—*Ian M. Dash*, Australian Broadcasting Corporation, Sydney, NSW, Australia

Listening test data collected in 2003 on 49 audio program samples were used to formulate the ITU-R BS.1770 program loudness prediction algorithm. The validity of this data at low frequencies was unproven. Octave-band analysis has therefore been performed on the test samples to test for audibility in each band. Results suggest that further listening tests may be needed to obtain reliable low-frequency data. A multiple regression analysis was also performed on the octave-band data to obtain a least-squares weighting curve for comparison with the ➥

BS.1770/RLB2 weighting curve. Results suggest that while the BS.1770 curve performs well, there is still room for improvement.
Convention Paper 7811

16:00

P28-7 Windowed Sine Bursts: In Search of Optimal Test Signals for Detecting the Threshold of Audibility of Temporal Decays—*Andrew Goldberg*, Helsinki University of Technology, Espoo, Finland

A slow decay in an audio signal is perceived as ringing and is commonly caused by room modes. This affects the perception of intelligibility, clarity, definition, and spatial rendering. A method has previously been devised to find the threshold of audibility of the decay in low-frequency narrow-band signals. One of the test signals in the large-scale listening test will be a low-frequency sine burst, but spectral spreading at the start and end of the test signal acts as an additional non-modal cue. This effect is removed by windowing, for example a half Hann. The aim of this paper is to determine the window length required (threshold) to render the end of the test signal free from audible spectral spreading. The Parameter Estimation by Sequential Testing (PEST) method and calibrated headphones (to remove factors associated with the listening environment) are used in subjective listening tests. The window length threshold is found to be constant above 200 Hz but rises exponentially toward low frequencies, and is replay level dependent. Threshold may be related to the absolute threshold of hearing, masking curves and/or auditory filter bandwidth.
Convention Paper 7812

Live Sound Seminar 4 — **Sunday, May 10**
13:00 – 15:00 — **Atrium 2**

SENNHEISER

Chair: **Matthias Fehr**, DKE (German Comission for Electronics and IT-Techn.)

Panelists: *Hubert Eckart*, DTHG (German Society of Theatre, Events, etc.)
Bruno Marx, APWTP (Association of Professional Wireless Production Technologies)
Walter Möller, BLM, TKLM (Technical Conference of State Media Authority)
Florian von Hofen, VPLT (Society of Private Light and Soundtechique)

Loss of Spectrum for Wireless Production Tools

The WRC 07, World Radio Conference 2007, has identified the Spectrum 790 – 862MHz for IMT, International Mobile Telecom. In Germany more than 630,000 users operate in that spectrum: they need a new home for their activities.

The discussion will focus on the alternative spectrum that has to be opened and the transition time that can be expected. For how long can the quality of wireless production tools be guaranteed?

Experts of the Federal Network Agency, the IRT, and the APWPT user organization and from the Manufacturers will give Spectrum users an outlook on expected changes.

Session P29 — **Sunday, May 10**
13:30 – 15:00 — **K4 Foyer**

POSTERS: SIGNAL ANALYSIS, MEASUREMENTS, RESTORATION

13:30

P29-1 Evaluation and Comparison of Audio Chroma Feature Extraction Methods—*Michael Stein*,[1] *Benjamin M. Schubert*,[1] *Matthias Gruhne*,[2] *Gabriel Gatzsche*,[2] *Markus Mehnert*[1]
[1]Ilmenau University of Technology, Ilmenau, Germany
[2] Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

This paper analyzes and compares different methods for digital audio chroma feature extraction. The chroma feature is a descriptor, which represents the tonal content of a musical audio signal in a condensed form. Therefore chroma features can be considered as an important prerequisite for high-level semantic analysis, like chord recognition or harmonic similarity estimation. A better quality of the extracted chroma feature enables much better results in these high-level tasks. In order to discover the quality of chroma features, seven different state-of-the-art chroma feature extraction methods have been implemented. Based on an audio database, containing 55 variations of triads, the output of these algorithms is critically evaluated. The best results were obtained with the Enhanced Pitch Class Profile.
Convention Paper 7814

13:30

P29-2 Measuring Transient Structure-Borne Sound in Musical Instruments—Proposal and First Results from a Laser Intensity Measurement Setup—*Robert Mores*,[1] *Marcel thor Straten*,[2] *Andreas Selk*[3]
[1]Hamburg University of Applied Sciences, Hamburg, Germany
[2]Consultant, Seevetal, Germany
[3]Consultant, Hamburg, Germany

The proposal for this new measurement setup is motivated by curiosity in transients propagating across arched tops of violins. Understanding the impact of edge construction on transient wave reflection back to the top of a violin or on conduction into the rib requires single-shot recordings possibly without statistical processing. Signal-to-noise ratio should be high although mechanical amplitudes at distinct locations on the structure surface are in the range of a few micrometers only. In the proposed setup, the intensity of a laser beam is directly measured after passing a screen attached to the device under test. The signal-to-noise ratio achieved for one micrometer transients in single-shot recordings is significantly more than 60 dB.
Convention Paper 7815

13:30

P29-3 Evaluating Ground Truth for ADRess as a Preprocess for Automatic Musical Instrument Identification —*Joseph McKay, Mikel Gainza,*

Dan Barry, Dublin Institute of Technology, Dublin, Ireland

Most research in musical instrument identification has focused on labeling isolated samples or solo phrases. A robust instrument identification system capable of dealing with polytimbral recordings of instruments remains a necessity in music information retrieval. Experiments are described that evaluate the ground truth of ADRess as a sound source separation technique used as a preprocess to automatic musical instrument identification. The ground truth experiments are based on a number of basic acoustic features, while using a Gaussian Mixture Model as the classification algorithm. Using all 44 acoustic feature dimensions, successful identification rates are achieved.
Convention Paper 7816

13:30

P29-4 Improving Rhythmic Pattern Features Based on Logarithmic Preprocessing—*Matthias Gruhne, Christian Dittmar*, Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

In the area of Music Information Retrieval, the rhythmic analysis of music plays an important role. In order to derive rhythmic information from music signals, several feature extraction algorithms have been described in the literature. Most of them extract the rhythmic information by auto-correlating the temporal envelope derived from different frequency bands of the music signal. Using the auto-correlated envelopes directly as an audio-feature is afflicted with the disadvantage of tempo dependency. To circumvent this problem, further postprocessing via higher-order statistics has been proposed. However, the resulting statistical features are still tempo dependent to a certain extent. This paper describes a novel method, which logarithmizes the lag-axis of the auto-correlated envelope and discards the tempo-dependent part. This approach leads to tempo-invariant rhythmic features. A quantitative comparison of the original methods versus the proposed procedure is described and discussed in this paper.
Convention Paper 7817

13:30

P29-5 Further Developments of Parameterization Methods of Audio Stream Analysis for Security Purposes—*Pawel Zwan, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

The paper presents an automatic sound recognition algorithm intended for application in an audiovisual security monitoring system. A distributed character of security systems does not allow for simultaneous observation of multiple multimedia streams, thus an automatic recognition algorithm must be introduced. In the paper a module for the parameterization and automatic detection of audio events is described. The spectral analysis of sounds of a broken window, gunshot, and scream are performed and parameterization methods are proposed and discussed. Moreover, a sound classification system based on the Support Vector Machines (SVM) algorithm is presented and its accuracy is discussed. The practical application of the system with the use of a monitoring station is shown. The plan of further experiments is presented and the conclusions are derived.
Convention Paper 7818

13:30

P29-6 Estimating Instrument Spectral Envelopes for Polyphonic Music Transcription in a Music Scene-Adaptive Approach—*Julio J. Carabias-Orti, Pedro Vera-Candeas, Nicolas Ruiz-Reyes, Francisco J. Cañadas-Quesada, Pablo Cabañas-Molero*, University of Jaén, Linares, Spain

We propose a method for estimating the spectral envelope pattern of musical instruments in a musical scene-adaptive scheme, without having any prior knowledge about the real transcription. A musical note is defined as stable when variations between its harmonic amplitudes are held constant during a certain period of time. A density-based clustering algorithm is used with the stable notes in order to separate different envelope models for each note. Music scene-adaptive envelope patterns are finally obtained from similarity and continuity of the different note models. Our approach has been tested in a polyphonic music transcription scheme with synthesized and real music recordings obtaining very promising results.
Convention Paper 7819

Exhibitor Seminar **Sunday, May 10**
13:00 – 14:00 **Exhibit Floor Booth 2303**
ES19 OPTOCORE – C

OPTOCORE IN PRACTICE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

This seminar will cover practical exercises with OPTOCORE system/devices. Content will include practical set-up of a network, remote pre-amp control through third party devices, transport of control data and Ethernet, software control, and firmware upgrades.

Workshop 22 **Sunday, May 10**
14:30 – 16:30 **Room K2**

MPEG SAOC: INTERACTIVE AUDIO AND BROADCASTING, MUSIC 2.0, NEXT GENERATION TELECOMMUNICATION

Chair: **Oliver Hellmuth**, Fraunhofer IIS, Erlangen, Germany

Panelists: *Jonas Engdegård*, Dolby, Stockholm, Sweden
Christof Faller, Illusonic LLC, Lausanne, Switzerland
Jürgen Herre, Fraunhofer IIS, Erlangen, Germany
Werner Oomen, Philips Applied Technologies, Eindhoven, The Netherlands

Recently the ISO/MPEG standardization group launched an activity for bit rate-efficient and backward compatible coding of multiple sound objects that heavily exploits the ➥

human perception of spatial sound. On the receiving side, such a "Spatial Audio Object Coding" (SAOC) system renders the transmitted objects interactively into a sound scene on any desired reproduction. Based on the SAOC technology elegant solutions for Interactive Audio and Broadcasting, Music 2.0, Next Generation Telecommunication become feasible. The workshop reviews the ideas and principles behind Spatial Audio Object Coding, especially highlighting its possibilities and benefits for those new types of applications. Additionally, the potential of SAOC is illustrated by means of real-time demonstrations.

Tutorial 10 **Sunday, May 10**
14:30 – 17:00 **Room K1**

AUDIO SYSTEM GROUNDING & INTERFACING—AN OVERVIEW

Presenter: **Bill Whitlock**, Jensen Transformers, Inc., Chatsworth, CA, USA

Although the subject has a black art reputation, this tutorial replaces myth and hype with insight and knowledge, revealing the true causes of system noise and ground loops. Although safety must be the top priority, some widely used cures are both illegal and deadly. Both balanced and unbalanced interfaces are vulnerable to noise coupling, but the unbalanced interface is exquisitely so due to an intrinsic problem. Because balanced interfaces are widely misunderstood, their theoretically perfect noise rejection is severely degraded in most real-world systems. Some equipment, because of an innocent design error, has a built-in noise problem. A simple, no-test-equipment, troubleshooting method can pinpoint the location and cause of system noise. Ground isolators in the signal path solve the fundamental noise coupling problems. Also discussed are unbalanced to balanced connections, RF interference, and power line treatments such as technical power, balanced power, isolation transformers, and surge suppressors.

Exhibitor Seminar **Sunday, May 10**
15:00 – 16:00 **Exhibit Floor Booth 2303**
ES20 OPTOCORE – A

TECHNICAL INTRODUCTION TO OPTOCORE

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

The OPTOCORE Synchronous Fibre Network is introduced. Content will include an Introduction/Overview, design of the network, synchronicity, latency, topology, redundancy, protocol, and devices.

Exhibitor Seminar **Sunday, May 10**
16:00 – 17:00 **Exhibit Floor Booth 2303**
ES18 OPTOCORE – B

BASICS OF FIBRE OPTICS

Presenters: *Martin Barbour*, OPTOCORE Support Engineer
Andreas Kaspar, OPTOCORE Support Engineer
Thorsten Schulze, OPTOCORE Support and Product Manager

Technical, economical, safety aspects, and advantages of fibre optics is the topic of this seminar. Content will include fibre advantages, reliability, types of fibre, parameters of fibres (attenuation, bandwidth length product, latency), connector types, cleaning, and testing.