AES 122nd Convention Program

May 5 – 8, 2007

Austria Center Vienna

Friday, May 4 13:30 Room 142/143
Standards Committee Meeting on SC-02-02 Digital Input/Output Interfacing

Session P1 Saturday, May 5 09:00 – 12:00
Room I

SPATIAL AUDIO PERCEPTION AND PROCESSING, PART 1

Chair: **Thomas Sporer,** Frauenhofer IDMT, Ilmenau, Germany

09:00

P1-1 Using Transient Suppression in Blind Multichannel Up-Mix Algorithms—*Andreas Walther, Christian Uhle, Sacha Disch*, Fraunhofer Institute IIS, Erlangen, Germany

In the field of blind up-mixing, many algorithms exist for generating multichannel sound from mono or stereo sources. One of the important blind up-mix scenarios is the ambience-based up-mix. This approach aims at extracting the ambient parts of a given signal and their reproduction by taking the best possible advantage of a multichannel loudspeaker setup. Depending on the number of channels and the signal characteristics of the input signal, the quality of the extracted ambience can vary. In this paper transient suppression is suggested as a method for improving an extracted ambience signal. Two methods for suppressing transient components are proposed and contrasted to existing techniques. The ability of those methods to improve the perceived quality of the ambience signal and overall up-mix is evaluated in a subjective listening test.
Convention Paper 6990

09:30

P1-2 A Novel Approach to Up-Mix Stereo to Surround Based on MPEG Surround Technology—*Heiko Purnhagen*,[1] *Andreas Ehret*,[2] *Jonas Rödén*,[1] *Alexander Gröschel*[2]
[1]Coding Technologies, Stockholm, Sweden
[2]Coding Technologies, Nuremberg, Germany

With the increasing number of installed surround sound systems in consumer homes and cars, the demand for surround sound content is rising. However, the vast majority of content is still only available in stereo. Furthermore, in many cases it is difficult to create a surround mix for content previously released in mono or stereo due to the high effort necessary or simply because the original multitrack recordings are unavailable. Consequently the need for tools that allow an automated or semi-automated stereo to surround up-mix is growing. In this paper a novel approach is described that is based on technology that is part of the MPEG Surround standard. The basic algorithm and some proposed extensions are outlined and potential use-cases described. Finally, the subjective quality of the presented approach is compared to existing solutions.
Convention Paper 6991

10:00

P1-3 Coding of "2+2+2"Surround Sound Content Using the MPEG Surround Standard—*Andreas Ehret*,[1] *Alexander Gröschel*,[1] *Heiko Purnhagen*,[2] *Jonas Roedén*[2]
[1]Coding Technologies, Nuremberg, Germany
[2]Coding Technologies, Stockholm, Sweden

An increasing number of recordings is available in the so-called 2+2+2 surround sound format, where in addition to two front and two rear loudspeakers, a third pair of loudspeakers is placed at an elevated position above the front speakers. In 2006, the MPEG Surround standard was finalized as an efficient stereo backward-compatible surround sound coding format. The present paper studies the applicability of MPEG Surround for efficient coding of "2+2+2" content. Several alternative approaches are outlined and evaluated by means of subjective listening tests.
Convention Paper 6992

10:30

P1-4 Quality Taxonomies for Auditory Virtual Environments—*Andreas Silzle*, Ruhr-Universität Bochum, Bochum, Germany

The aim of the here developed new taxonomies is to describe the components involved in the quality process of Auditory Virtual Environments (AVE) and to quantify the relations between them for different applications. The taxonomy should allow an overview and identify the relations that are most important in the software development process and the design of listening experiments. For the first time the multivariate relations between the quality elements in the physical domain and the quality features in the perceptual domain of such a quality taxonomy are evaluated for three different AVE applications. This evaluation and quantifica-

tion is done by means of an expert survey (DELPHI method) to objectify the results. Principal component analysis reveals that five dimensions are necessary to describe about 95 percent of the variance in the data. This indicates that the selected seven quality features are clearly distinguishable for the experts, but not orthogonal to each other. Most of the quality features are introduced in meaningful terms in the audio engineering field and therefore usable without training for the participating experts. The results of the expert survey are compared to listening test results, which use the same quality features. The bottom line is that the expert survey is not only a much faster method to get a good overview about a specific application as compared to the listening test, but it also reveals more information about it.
Convention Paper 6993

11:00

P1-5 Individual Localization Behavior for Perception of Virtual Sound Sources—*Cornelius Bradter*,[1] *Klaus Hobohm*[2]
[1]University of Applied Science, Berlin, Germany
[2]Film and Television Academy, Potsdam, Germany

Test results normally indicate very large variability in perception of lateral virtual sound sources with a 5-channel loudspeaker setup. Three recent studies indicate that up to a third of stimuli from a side loudspeaker pair were perceived surprisingly accurately as virtual sound sources. In other cases sound sources were perceived as coming from only one loudspeaker or from its vicinity. Therefore, we specified prototypical localization behaviors. We examined effects on localization by reproduction rooms, exact position of test persons in relation to loudspeakers, test persons' head movements, and trading between time delay and level differences.
Convention Paper 6994

11:30

P1-6 Comparison of Different Sound Capture and Reproduction Techniques in a Virtual Acoustic Window—*Timo Haapsaari, Werner De Bruijn, Aki Härmä*, Philips Research Laboratories, Eindhoven, The Netherlands

In this paper we describe a two-way audio communication system using arrays of microphones and loudspeakers to create a virtual acoustic window. We compare three different methods for capturing the sound: wave field sampling using a line array of microphones, an adaptive beamformer, and close-talk microphones. For sound reproduction, we employ wave field synthesis. In the paper we review the acoustic and perceptual requirements for a real-time virtual acoustic window system and report results of a set of listening experiments performed with the system.
Convention Paper 6995

Workshop 1 Saturday, May 5
09:00 – 11:30 Room P

AUDIO QUALITY EVALUATION—DESCRIPTIVE ANALYSIS 2: INDIVIDUAL VOCABULARY DEVELOPMENT

Chair: **Jan Berg**, Luleå University of Technology, Luleå, Sweden

Panelists: *Sylvain Choisel*, Bank & Olufsen a/s, Struer, Denmark
Garmt Dijksterhuis, Unilever Food & Health Research Institute, Vlaardingen, The Netherlands
Natanya Ford, Bang and Olufsen a/s, Struer, Denmark
William Martens, McGill University, Montreal, Quebec, Canada

As knowledge about the listener experience of audio applications is fundamental in research and product design within the audio field, methods that can be used for evaluation of perceived sound quality are essential to explore. In order to capture and quantify listener experience, different methodological approaches are utilized. One of the approaches involves development of individual vocabulary of test subjects. This considers a collection of techniques that can be used for evaluating the detailed perceptual characteristics of products or systems through listening tests. This workshop aims to provide guidance to the researcher and experimenter regarding the nature of descriptive analysis by means of individual vocabulary development techniques and their application in audio.

Workshop 2 Saturday, May 5
09:00 – 12:00 Room H

THE PRACTICE OF AUDIO FORENSICS

Cochairs: **Syliva Moosmüller**, Acousics Research Institute, Vienna, Austria
Richard Sanders, UCDHSC, Denver Colorado, CO, USA

Panelists: *Durand Begault*, NASA Ames Research Center, Mountainview, CA, USA
Eddy B. Brixen, EBB-consult, Smørum, Denmark
Catalin Grigoras, Forensic Examiner, Bucharest, Romania
Gordon Reid, CEDAR Audio, Cambridge, UK

This workshop will examine the common challenges of an audio forensic professional. Topics will include voice identification, analog and digital media authenticity, audio enhancement, and adaptive filtering and running a forensic business. This three hour workshop will include Q & A during and after each 30-minute presentation. Audio and visual examples will be shown where appropriate. Software will be demonstrated in real-time and input along with questions from the attendees will be encouraged. All of the presenters are considered to be experts in audio forensics and some related fields.

Saturday, May 5 09:00 Room 142/143
Standards Committee Meeting on SC-05-02 Audio Connectors

Session P2 Saturday, May 5 09:30 – 10:30
Room K

AUDIO RECORDING AND REPRODUCTION

Chair: **Nadja Wallaszkovits**, Austrian Academy of Sciences, Vienna, Austria

09:30

P2-1 Recording of Acoustical Concerts Using a

Soundfield Microphone—*Markus Schellstede,*[1] *Christof Faller*[2]
[1]Sunhoe Pro Audio, Bern, Swtizerland
[2]Illusonic LLC, Chavannes, Switzerland

Based on extensive practical experience of one of the authors, recording of acoustical concerts, using a soundfield microphone, without spot or support microphones is discussed. The focus is stereo recording of classical music. Strategies for positioning of the microphone, B-Format decoding, and mastering are presented. The so-obtained final mix is largely based on the natural mix of sound reaching the microphone. This is in contrast to more conventional recording techniques that usually use a large number of spot and support microphones. Last but not least, limitations and cost considerations are discussed.
Convention Paper 6996

10:00

P2-2 Ambience Sound Recording Utilizing Dual MS (Mid-Side) Microphone Systems Based upon Frequency Dependent Spatial Cross Correlation (FSCC)—*Teruo Muraoka, Takahiro Miura, Tohru Ifukube*, University of Tokyo, Tokyo, Japan

In musical sound recording for CD production or broadcasting, a forest of microphones is commonly observed. They are for good sound localization and favorable ambience, however it is desirable to make the forest sparse for less laborious setting up and mixing. Previously, the authors examined ambience representation of stereophonic microphone arrangements utilizing frequency dependent spatial cross correlation (FSCC). FSCC is defined as a cross correlation of outputs by two microphones that of MS microphone system is most favorable. Based upon the result, we devised a combination of two MS microphone systems, one for picking up stage sounds and the other for ambience representation. In addition to minor stage microphones, the authors achieved satisfactory musical recording.
Convention Paper 6997

Workshop 3 Saturday, May 5
09:30 – 11:30 Room G

SURROUND FOR BROADCASTING: AN OVERVIEW

Cochairs: **Kimio Hamasaki**, NHK Science and Technical Research Laboratories, Tokyo, Japan
Lars Jonsson, SR/Swedish Radio

Panelists: *Steve Church*, Telos
Eric Lundbeck, SVT - Swedish Television
Tony Spath, Dolby Laboratories, UK
Gerhard Stoll, IRT, Munich, Germany

This workshop will review proprietary systems and MPEG Surround/spatial standards. Which systems are in use? Which systems will come?

TV and radio broadcasters are using distribution systems for transmitting 5.1 surround sound with different audio coding methods. The workshop will review the current state of the art of used systems and report from broadcasters' experience of starting up new services and give an overview of the receiver penetration in various countries.

The panel will also discuss future directions

TECHNICAL TOUR 1
ORF Radio Broadcasting Studios
Saturday, May 5, 09:30 – 12:30

ORF Radio is the National Public Broadcaster for Austria, with three nationwide, nine regional, and two MW/SW stations. The tour to the Radio Broadcasting House gives access to the music recording studios, continuity studios for three different radio stations, and insight into the practical 5.1-Surround sound operation of the satellite radio channel OE1DD.

Saturday, May 5 09:30 Room 633
Technical Committee Meeting on Electro-Magnetic Compatibility

Session P3 Saturday, May 5 10:00 – 11:30
Foyer IK

POSTERS: LOW BIT-RATE AUDIO CODING

10:00

P3-1 Enhanced MPEG-4 Low Delay AAC—Low Bit-Rate High-Quality Communication—*Markus Schnell,*[1] *Ralf Geiger,*[1] *Markus Schmidt,*[1] *Manuel Jander,*[1] *Markus Multrus,*[1] *Gerald Schuller,*[2] *Jürgen Herre*[1]
[1]Fraunhofer IIS, Erlangen, Germany
[2]Fraunhofer IDMT, Ilmenau, Germany

The MPEG-4 Low Delay Advanced Audio Coding (AAC-LD) scheme has recently evolved into a popular algorithm for audio communication. It produces excellent audio quality at bit rates between 64 kbit/s and 48 kbit/s per channel. This paper introduces an enhancement to AAC-LD that reduces the bit rate demand by 25-to-33 percent. This is achieved by adding both a delay-optimized version of the Spectral Band Replication (SBR) tool and by utilizing a dedicated low delay filterbank. The introduced techniques maintain the high audio quality and offer an algorithmic delay low enough for use in two way communication systems. This paper describes the coder enhancements including a detailed discussion of algorithmic delay issues, a performance assessment, and possible applications.
Convention Paper 6998

10:00

P3-2 On the Design of Low Power MPEG-4 HE-AAC Encoder—*Wen-Chieh Lee, Chung-Han Yang, Cheng-Lun Hu*, National Chiao Tung University, Hsinchu, Taiwan

Spectral Band Replication (SBR) has been combined with MPEG AAC as a bandwidth extension tool. The resulting scheme is referred to as the MPEG-4 High Efficient (HE) AAC or aacPlus. With the SBR module taking care of the high frequency contents, the conventional AAC ➥

encoder can compress the low frequency part using most of the available bits. The SBR parameters are all calculated by SBR encoder in complex domain in the architecture of conventional QMF. If the components in the SBR encoder can be implemented in the real domain, the computational complexity of HE-AAC will be reduced by half. The paper proposes a low power MPEG-4 HE-AAC encoder to reduce the computational complexity. Both subjective and objective experiments are conducted to demonstrate the quality of the low power HE-AAC encoder on critical music tracks.
Convention Paper 6999

10:00

P3-3 High Quality, Low Power QMF Bank Design for SBR, Parametric Coding, and MPEG Surround Decoders—*Hsin-Yao Tseng, Han-Wen Hsu, Chi-Min Liu,* National Chiao Tung University, Hsin-Chu, Taiwan

Due to the alias-free properties, the complex quadrature mirror filter (QMF) bank has been used in MPEG-4 audio standard on SBR, parametric, and surround coding. The high complexity overhead from the complex QMF bank and the complex data processing in the decoder leads to the development of a low power decoder, which adopts the real QMF bank as the basic building module to reduce the complexity. However, the artifacts from the aliasing in the real QMF bank are the major concern. This paper studies the artifacts from the real QMF bank and proposes a novel QMF bank design to achieve both low complexity and high quality. Also, this paper applies the novel QMF bank to develop the high-quality and low-power SBR, parametric, and MPEG surround decoders and shows the merits in complexity and quality.
Convention Paper 7000

10:00

P3-4 Low Power Stereo Perceptual Audio Coding Based on Adaptive Masking Threshold Reuse—*Evelyn Kurniawati, Sapna George*, ST Microelectronics Asia Pacific Pte. Ltd., Singapore

The term perceptual audio coder refers to audio compression schemes that exploit the properties of human auditory perception. The idea is to allocate the quantization noise elegantly below the masking threshold to make it imperceptible to the ear. The process requires considerable computational effort, especially due to the psychoacoustics analysis and bit allocation-quantization process. This paper proposes a new method to simplify the psychoacoustics modeling process by adaptively reusing the computed masking threshold depending on the signal characteristics. The method also devices a scheme to patch the potential spectral hole problems that might occur when the quantization parameters are reused. This proposal can be applied to generic stereo perceptual audio encoders where low computational complexity is required.
Convention Paper 7001

10:00

P3-5 A Hybrid Warped Linear Prediction (WLP) AAC Audio Coding Algorithm—*Jaeseong Lee, Young-Cheol Park, Dae-Hee Youn, Hong-Goo Kang*, Yonsei University, Seoul, Korea

We propose a hybrid warped linear prediction (WLP) AAC audio coding algorithm. The proposed algorithm employs a warped linear prediction (WLP) processor to construct a perceptual pre- and post-filter for the MPEG-4 AAC. The WLP residue is applied to the MDCT filter-bank, and the signal-to-mask ratio (SMR) of the corresponding block is modified to set a masking threshold for the WLP residues. In the decoder, the reconstructed residual signal is passed to a modified WLP synthesis filter to restore the audio signal. Subjective tests show that the proposed audio codec operating at 50 kbps has comparable perceptual quality to the conventional MPEG-4 AAC operating at the 58 kbps.
Convention Paper 7002

10:00

P3-6 Comparison of Stereo Redundancy Reduction Schemes for an Ultra Low Delay Audio Coder—*Tobias Albert,*[1] *Gerald Schuller,*[2] *Stefan Wabnik,*[2] *Ulrich Krämer,*[2] *Jens Hirschfeld*[2]
[1]Fraunhofer IIS, Erlangen, Germany
[2]Fraunhofer IDMT, Ilmenau, Germany

In the Fraunhofer Ultra Low Delay Audio Coder (ULD) a pre-filter that is controlled by a psychoacoustic model is followed by a quantizer and a predictive coder to code signals in the time-domain. The output of the predictor is entropy coded and transmitted. Predictor and entropy coder form the lossless redundancy reduction part of the coder. Our goal is to improve the lossless redundancy-reduction part for stereo signals. We present and evaluate six different alternatives for the stereo redundancy reduction, and we combine those alternatives to obtain a higher compression ratio.
Convention Paper 7003

10:00

P3-7 Speech Codec Enhancements Utilizing Time Compression and Perceptual Coding—*Maciej Kulesza, Andrzej Czyzewski*, Gdansk University of Technology, Gdansk, Poland

A method for encoding wideband speech signal employing standardized narrowband speech codecs is presented as well as experimental results concerning detection of tonal spectral components. The speech signal sampled with a higher sampling rate than it is suitable for narrowband coding algorithm is compressed in order to decrease the amount of samples. Next, the time-compressed representation of a signal is encoded using a narrowband speech codec. The time expansion procedure is applied to the speech signal after transmission and decoding in order to restore original time relations. Finally, the wideband speech signal is presented to the user. The method for spectral envelope estimation involving perceptual criteria is described. The algorithms for tonal components detection

were evaluated and compared during experiments carried-out.
Convention Paper 7004

10:00

P3-8 Design and Implementation of a Web-Based Software Framework for Real Time Intelligent Audio Coding Based on Speech/Music Discrimination—*Jose Enrique Muñoz Exposito, Nicolas Ruiz Reyes, Sebastian Garcia-Galan, Pedro Vera Candeas*, University of Jaén, Jaén, Spain

In this paper a software framework based on client-server architecture is implemented for real time intelligent audio coding. A speech/music discrimination scheme analyzes the input audio signal and takes a decision about the nature of the audio signal (speech or music) on a frame by frame basis. According to the decision of the speech/music discriminator, a suitable coder is selected at each frame. The designed software framework makes use of the speech and audio coders incorporated into the MPEG4 audio standard (HVXC or CELP for speech frames and TwinVQ or AAC for music frames) to evaluate the performance of an intelligent multimode audio coder. The framework supports several types of audio features (timbral texture features and rhythmic content features) and classifiers (classical Statistical Pattern Recognition (SPR) classifiers, Multilayer Perceptron Neural Networks (MLPNN), Support Vector Machines (SVM), Fuzzy Expert Systems (FES), Hidden Markov Models (HMM)). Comparison between a speech/music discrimination based-intelligent audio coder and MPEG4-AAC has been performed using audio signals representative of the two corresponding classes (speech and music). Subjective and objective tests have been accomplished aiming at assessing the behavior of the intelligent audio coding scheme.
Convention Paper 7005

10:00

P3-9 Quantization of Laguerre-Based Stereo Linear Predictors—*Albertus C. den Brinker,*[1] *Arijit Biswas*[2]
[1]Philips Research Laboratories, Eindhoven, The Netherlands
[2]Technische Universiteit Eindhoven, Eindhoven, The Netherlands

Recently a quantization scheme for stereo linear prediction systems was proposed and was tested using random data as input. This research is extended in the current paper by incorporating Laguerre filters in the stereo linear prediction scheme. First, it is shown that the associated normalized reflection matrices (NRM) can be obtained efficiently. Second, the system was tested using stereo audio data in order to gain an insight into the required bit rates for practical applications.
Convention Paper 7006

SPECIAL EVENT
Recording and Mixing in Wave Field Synthesis
Saturday, May 5, 10:30 – 13:30 and 14:30 – 16:45
Sunday, May 6, 10:30 – 13:30 and 14:30 – 16:45
University for Music and Performing Arts, Vienna

CELEBRATE GOOD VIBES!

Hosted by the University for Music and Performing Arts, Vienna and supported by the partners Fraunhofer Institute for Digital Media Technology IDMT, Sennheiser, Studer, and Merging Technologies, this Special Event will present live concerts and their reproduction in Wave Field Synthesis alongside.

Attendees will be able to directly compare the sound inside the hall with its recorded and live mixed counterpart. Therefore, renowned musicians—including Alvaro Pierri (classical guitar) and the tango orchestra Band-O-Neon—will play pieces from their repertoire in the Joseph Haydn Concert Hall. Next door, in the Batiken Hall, the recorded and mixed session will be played back using a WFS system with more than 90 loudspeakers.

Different microphone techniques can be compared and evaluated. Neumann/Sennheiser microphones, a Vista 5 console by Studer, the Merging Technologies Pyramix Digital Audio Workstation, and the Spatial Audio Workstation by Fraunhofer IDMT will provide a professional recording and mixing environment.

This Special Event will be held in four sessions on May 5 and May 6 at the University of Music Vienna. A bus shuttle service will be offered between the Austria Center and the University.

Tickets will be available from the Technical Tour counter. Meeting point is the main entrance of the Austria Center. Please note that the times listed above include travel time.

Additional Demo Sessions in Wave Field Synthesis will also be available at the following times:
- Saturday, May 5, 17:15 – 18:15
- Sunday, May 6, 17:15 – 18:15

Saturday, May 5 10:30 Room 633
Technical Committee Meeting on Archiving, Restoration, and Digital Libraries

Saturday, May 5 11:00 Room 142/143
Standards Committee Meeting on SC-05-05 Grounding and EMC Practices

SPECIAL EVENT
Awards Presentation and Keynote Address
Saturday, May 5, 12:00 – 13:30, Room K

Opening Remarks:
- Executive Director Roger Furness
- President Wieslaw Woszczyk
- Convention Chair Werner Deutsch

Program:
- AES Awards Presentation
- Introduction of Keynote Speaker by Convention Vice Chair Florian Camerer
- Keynote Address by Helmut Voitl, Documentary and Feature Film Director, "Sound and Emotion: The Power of Audio in Storytelling

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

Keynote Speaker

This year's Keynote Speaker is Helmut Voitl. Voitl was born in Dresden and grew up in Vienna, studying photography and film at the Graphische Lehr- und Versuchsanstalt. He scripted and directed his first short film in 1966 and directed his first documentary series in 1968. Voitl teamed up with Elisabeth Guggenberger in 1973, and they began making films for numerous TV productions commissioned by ORF, SRG, ZDF, BR, and NDR. With the productions *Russkiy Chleb* and *Schto delat, Towarisch?* in 1989 and a documentary of the Ukraine in 1990 began an involvement with the former Soviet Union, finding its climax in the re-enactment series *Arctic North-East* (ORF, 1992–1996), a dramatized and staged reconstruction of the Austrian polar expedition of 1872–1874 that led to the discovery of the high arctic archipelago Franz Joseph Land. Shooting took place at original locations in the Russian High Arctic. Another production in Russia followed in 1998, *Russia's Holy War,* and a year later *A-Watch,* which dealt with the environmental state of affairs in Austria. Currently Helmut Voitl is working on a feature film project called *White Clouds Island*, also situated in the High Arctic.

Voitl's keynote address is entitled: "Sound and Emotion: The Power of Audio in Storytelling." With the moment of our birth, our sense of hearing recedes into the second row, into the unconsciousness. Nevertheless it cannot be switched off and is a vital means for our understanding of the world around us, as well as a powerful transport medium for emotional content. Filmmakers have long used these powers for translating their storytelling to the audience in sometimes sublime, sometimes drastic ways. In documentary filmmaking, the situation is no different. In this talk, the various ways how sound can be used to augment, embelish or even contrast the visual impression will be discussed, always with the prerequisite of being an equal partner and a welcome tool for the final result—the experience of watching a film.

Session P4 Saturday, May 5 13:30 – 18:30
Room I

AUDIO ARCHIVING, STORAGE, RESTORATION, AND CONTENT MANAGEMENT

Chair: **Dietrich Schüller**, Austrian Academy of Sciences, Vienna, Austria

13:30

P4-1 Sound Archiving—A Challenge for the Audio Engineering Society [Invited Paper]—*Dietrich Schuller,* Austrian Academy of Sciences, Vienna, Austria

In the course of the past 20 years, issues related to sound archiving and restoration have increasingly conquered a stable position at AES conventions. These issues are also permanently dealt with and further developed by the respective groups within the AES Technical and Standards Committees. In 2001, the 20th AES Conference held in Budapest was devoted to Archiving, Restoration, and New Methods of Recording.

The world's first sound archive, the Phonogrammarchiv, was founded in 1899 by the then Imperial Academy of Sciences, and that, inter alia, may be one of the reasons why the Vienna AES Convention has made "Archiving" one of its main topics.

This paper will introduce to a full afternoon of specialized presentations, reminding us that, at the cradle of sound recording, nobody would have imagined a recording and entertaining industry which today serves one of the greatest markets of the world. Sound recording was the result of scientific interest predominantly aimed at studying the nature of spoken language. And it was scholars—linguists, anthropologists, and musicologists—who systematically employed sound recording technology from its very beginning. Consequently, the academic world played a key role in founding sound archives from around 1900 onward, and the longevity of sound recordings was given special emphasis, specifically in the archives of Vienna and Berlin.

The emerging phonographic industry, however, shaped the development of sound recording technology since that time; yet the permanence of the record that once had attracted the scholarly world was not among the driving forces, particularly not in the development of magnetic tape recording. Only in the late 1950s, when libraries had already been collecting sound recordings as significant cultural sources on a greater scale, did preservation start to become an issue. Today, the world-wide holdings of audio recordings are estimated to amount to some 100 million hours, many of them still on analog or digital single carriers, which sooner or later are prone to decay. Current thinking suggests, however, that obsolescence of replay equipment is an even greater threat to the long-term survival of the audio heritage.

This constitutes substantial challenges to AES, of which the greatest may be: while little can be done to counteract the present terrifying speed of withdrawal from the manufacture of replay equipment and spare parts, how can we maintain the knowledge and the skills needed for the maintenance of equipment and for the optimal retrieval of signals from our audio documents?

This is a presentation only; no Convention Paper will be available for purchase.

14:00

P4-2 150 Years of Time-Base in Acoustic Measurement and 100 Years of Audio's Best Publicity Stunt— 2007 as a Commemorative Year—*George Brock-Nannestad*, Patent Tactics, Gentofte, Denmark

Léon Scott's invention of the phonoautograph in 1857 made a long time-base available for recording of vibrations, and it was also the first time an air-borne sound was recorded. Although his invention formed the basis, both for sound recording and reproduction and for acoustical science as we know it, it has been largely forgotten. Neither Scott nor the instrument maker Koenig are mentioned in the series "Benchmark Papers of Acoustics" Today we take sound archives for granted, but the whole sound archive movement would not have received any attention in the general public, if one particular event had not occurred: the sealed deposit in 1907 of important shellac records and a gramophone in the vaults below the Paris Opera house. They were intended to remain untouched for 100 years, and they have survived to this day. The paper will provide the documentation

for these historical events that form the basis of so many of our professional activities.
Convention Paper 7007

14:30

P4-3 Knowledge: The Missing Element in Archiving and Restoration?—*Sean W. Davies*, SW Davies, Ltd., Aylesbury, UK

This admittedly provocative title nevertheless calls attention to a situation that already exists and may become critical in the future. No sound recording can be considered in isolation from the technical system that produced it. A proper working knowledge of such a system is an essential requirement for any person working on the transfer of such a recording. This paper examines the range of such required knowledge and the means by which it may be taught to personnel likely to be involved with archival material.
Convention Paper 7008

15:00

P4-4 Noncontact Phonographic Disk Digitization Using Structured Color Illumination—*Louis Laborelli, Jean-Hugues Chenot, Alain Perrier*, INA (Institut National de l'Audiovisuel), Bry sur Marne, France

We propose an innovative contact-less optical playing device for 78 rpm and 33 rpm lateral modulation phonographic disks using structured color illumination. An area of the disk is illuminated by a beam of rays, with color depending on the direction of incidence, that are reflected by the groove wall toward a camera image. In contrast with standard methods that measure the velocity of the groove at a single location, direct access to the audio signal value is obtained here directly from pictures through color decoding, and the whole height of the groove wall is exploited. This color coding allows for the detection of occluding dust and automated interpolation of the missing audio signal. Results on distortion, S/N ratio, and bandwidth are presented.
Convention Paper 7009

15:30

P4-5 Improvement of Cylindrical Record Reproduction Utilizing Inharmonic Frequency Analysis GHA—*Teruo Muraoka, Shota Nakagomi, Tohru Ifukube*, University of Tokyo, Tokyo, Japan

Cylindrical records were important sound media around the beginning of 20th century, and a lot of historical recordings were made by using them. We have engaged in the research of cylindrical record reproduction and noise-reduction of damaged SP records utilizing inharmonic frequency analysis of GHA (Generalized Harmonic Analysis). Surface noise of cylindrical records are more serious than SP records, so we challenged its noise reduction by modifying GHA noise-reduction.
Convention Paper 7010

16:00

P4-6 Method Comparison of Pick-Up and Preprocessing of Bias Signal for Wow and Flutter Correction—*Nadja Wallaszkovits*,[1] *Franz Pavuza*,[1] *Heinrich Pichler*[2]
[1]Phonogrammarchiv Austrian Academy of Sciences, Vienna, Austria
[2]Audio Consultant, Vienna, Austria

This paper discusses the practical implementation of high frequency bias signal retrieval from analog magnetic tapes at original replay speeds by using slightly modified standard playback facilities. Based on an implementation within an archival workflow and prior studies of bias signal retrieval from analog magnetic tape, the authors focus on a comparison of reproduction and preprocessing methods of the bias signal. Previous approaches are compared to the authors' proposed method. The signal preprocessing in the analog as well as digital domain is outlined and, based on analysis of bias signals from professional and semi-professional recordings, the various practical problems are discussed: level instability and unknown frequency of the recorded bias signal, frequency variations mainly with semiprofessional devices of older types of recording equipment due to the instability of the bias oscillator, as well as effects of signal distortions, interferences, signal aliasing problems, and ultrasonic artifacts. The practical applicability within a standard archival transfer is discussed.
Convention Paper 7011

16:30

P4-7 Improved Magneto-Optical 1/4-Inch Audio Tape Player for Preservation—*Marcel Guwang*, Hi-Stor Technologies, Colomiers, France

This improved 1/4-inch audio tape player features a multitrack magneto-optical reader to reduce preservation cost through speed, adjustment automation, and compatibility. The benefits of this 32-channel head, connected to a digital signal processor, are high speed capability, compatibility, and automatic detection of any number of audio tracks, real time automatic adjustment of the best digital playback azimuth, and filtering of crosstalk and partial track erasures.
Convention Paper 7012
This paper will be presented by Jean-Hughes Chenot.

17:00

P4-8 Analysis and Restoration of Faulty Audio CDs— *Hélène Galiègue*,[1] *Jean-Marc Fontaine*,[2] *Laurent Daudet*[1]
[1]Université Pierre et Marie Curie, Paris, France
[2]Laboratoire d'Acoustique Musicale, Paris, France

Many audio CDs (mostly CD-Rs but also CD-ROMs) have defects that appear due to bad manufacturing, careless use, or simple aging of its physical constituents. Here, we study such audio CDs that are still readable with a standard player (computer CD/DVD drive or standalone audio player with digital output), but whose defects are not fully handled by error correction codes, resulting in a highly distorted signal. This paper is two-fold: first, we characterize these ➥

errors on a few example discs; and second, we study different means to restore the audio content, by fusion of multiple reads and interpolation schemes.
Convention Paper 7013

17:30

P4-9 Techniques for the Authentication of Digital Audio Recordings—*Eddy B. Brixen*, EBB-consult, Smørum, Denmark

In forensic audio one important task is the authentication of audio recordings. Standards and procedures already exists regarding analog recordings. In the field of digital recording and digital media the conditions are different. A rock solid methodology is needed here, but does not exist yet. This paper reviews existing techniques and presents some results regarding an additional number of tools, the ENF criterion, which should be considered to become a standard within the AES as well as in the forensic community as a whole.
Convention Paper 7014

18:00

P4-10 Using Multiple Feature Extraction with Statistical Models to Categorize Music by Genre—*Benjamin Fields*, Goldsmiths College, University of London, London, UK

In recent years, large capacity portable personal music players have become widespread in their use and popularity. Coupled with the exponentially increasing processing power of personal computers and embedded devices, the way people consume and listen to music is ever changing. To facilitate the categorization of these personal music libraries, a system is employed using MPEG-7 feature vectors as well as Mel-Frequency Cepstral Coefficients classified through multiple trained Hidden Markov Models and other statistical methods. The output of these models is then compared and a genre choice is made based on which model gives the best fit. Results from these tests are analyzed and ways to improve the performance of a genre sorting system are discussed.
Convention Paper 7015

Workshop 4 **Saturday, May 5**
13:30 – 16:30 **Room G**

NEAR SIDE BIAS

Chair: **Tom Nousaine**, TN Communications, MI, USA

Panelists: *David Carlstrom*, DCX
Roger Dressler, Dolby Laboratories
Dan Field, Harman Specialty Group
Jon Lane, Nissan
Geoff Martin, Bang and Olufsen a/s
Richard Stroud, Stroud Audio
Mark Ziemba, Panasonic

In testing nearly 700 OEM autosound systems, over 200 competitive aftermarket autosound systems and several hundred aftermarket loudspeaker systems Tom Nousaine has found that well over 95 percent of passenger seat listeners receive a soundstage with images that are asymmetrically clustered toward the near side of a car with individual images sometimes being placed at or on the near door loudspeakers. Surround system systems with center channel transducers seldom produce an optimal solution, and the majority of driver side listeners receive a smaller degree of the same problem. In the opinion of Nousaine near side spatial bias is a nearly universal problem in autosound design and implementation. In considering this issue, individual panelists will present their individual interpretation of the situation along with methods and techniques for solving the condition.

Exhibitor Seminar **Saturday, May 5**
13:30 **Room 357**

D.A.V.I.D.

Presenter: **Axel Holzinger**

"Visual Radio"— Radio Delivery of Multimedia Content to Multiple Platforms

The upcoming delivery platforms like DMB, DVB-H, DVB-S, and wireless Internet open new dimensions for radio broadcasters to bring content to their listeners. New D.A.V.I.D. broadcast server functions make it easy to extend the classical radio workflow and provide "visible" content to the new mobile devices—at any time.

TECHNICAL TOUR 2
ORF TV Broadcasting Studios
Saturday, May 5, 13:30 – 16:30

ORF TV is the National Public Broadcaster for Austria with two main TV channels (ORF 1+ 2), two dedicated theme channels (TW1 – tourism and weather, ORF Sport Plus), regional programs, and diverse off-air activities, as well as being by far the most popular Web-presence in the country. The focus of the tour to ORF TV will be the three newly refurbished postproduction studios with outstanding acoustic design and full surround sound capability, unique for European broadcast studios. Example clips will be demonstrated, and the tour will also cover the general infrastructure shortly after a major relaunch of a substantial part of ORF's programs.

Saturday, May 5 **13:30** **Room 633**
Technical Committee Meeting on Coding of Audio Signals

Workshop 5 **Saturday, May 5**
14:00 – 16:00 **Room P**

INTERACTIVE AUDIO AND HUMAN PERCEPTION: CAN WE CONNECT KNOWLEDGE WITH PRACTICE?

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

Panelists: *Durand Begault,* NASA Ames Research Center, Mountainview, CA, USA
Douglas S. Brungart, WPAFB
Bill L. Chapin, Ausim-3D, Palo Alto, CA, USA
Jyri Huopaniemi, Nokia, Helsinki, Finland
John-Marc Jot, Creative Advanced Technology Center, Ottawa, Ontario, Canada
William Martens, McGill University, Montreal, Quebec, Canada

Olivier Warusfel, IRCAM, Paris, France

The term "interactive audio" is used here in its most general sense to mean real-time modification of an audio signal in which attributes of a reproduced sound source (such as the source's apparent position, timbral character, etc.) can be controlled by a user's actions (using interaction interfaces such as trackers, game-controllers, joysticks, etc.). Interactive audio is a growing field in today's audio environments. The complexity of interactive environments in computer games and simulations continues to grow. Understanding the perceptual effects of this increase in complexity is becoming a greater challenge. This workshop investigates these perceptual effects by exploring the design of interactive spaces and by highlighting what is already known from established techniques for generating virtual environments.

Tutorial 1 — Saturday, May 5
14:00 – 16:00 — Room 560/561

REALIZATION OF 7.1 MIXES

Presenter: **Jeff Levison**, DTS, Inc., Agoura Hills, CA, USA

The increasing popularity of multichannel playback systems has required the development of a variety of recording and mixing techniques. Artists and engineers are more aware of the impact of speaker placement on the balance between envelopment and imaging— especially for the surround channels d for the reproduction (or illusion) of reverberation and other acoustic environmental aspects. Increasing the number of channels reproduced can ease the dilemma of this envelopment/imaging compromise, and now new high definition playback methods, such as Blu-ray and HD DVD, offer eight channels of discrete reproduction with simultaneous high quality video. Besides using these extra channels for extra surround channels, the possibility exists for the inclusion of height channels for greater spaciousness and vertical pan positioning. This tutorial will present a group of realized examples of 7.1 with four surround channels and alternate mixes with 5.1 plus height.

Comparisons will be made between stereo, 5.1, and these new 7.1 mixes. Proposals of other higher order systems and the possible improvements to three-dimensional audio presentation will be discussed.

Live Sound Seminar 1 — Saturday, May 5
14:00 – 18:30 — Room E

SENNHEISER: The Soundcheck Is the Show!

Chair: **Gregor Zielinsky**

Presenters: *Svenja Dunkel*
Wayne "Heights" Glittens
Oliver Voges

The Sennheiser Live Mixing Event presents a complete soundcheck with a live rock band on stage, with full sound and light equipment. Participants will learn about how professional soundchecks are done in real life. This will include the FOH/PA soundcheck as well as the monitor soundcheck. All participants will receive an In-Ear wireless receiver, to be able to follow the in-ear check. The presentation is held by Wayne "Heights" Gittens (monitor engineer for Herbert Groenemyer, Xavier Naidoo, Söhne Mannheims, etc.), Oliver Voges (FOH engineer for Echo Awards, The Dome, Naturally 7, Scooter, Mousse T), and Svenja Dunkel (radio engineer and production manager for Echo Award, Comet, and productions from Sennheiser). The presenters will show their way of working with both the band and the soundcheck. The rhythm section of "Hot Pants Road Club" plus their lead singer is a top band from Austria, well known for their funky, groovy sound and the always slightly "Vieneese" style of their frontman Harry Ahamer.

Presentation schedule:

14.00 – 15.00 — O. Voges, lecture about FOH sound

15.00 – 16.00 — W. Gittens, lecture about monitor sound and In-Ear

16.15 – 18.30 — Live and practical work

The Live Sound Seminar, led by Gregor Zielinsky, is produced in cooperation with Neumann & Mueller, who support all sound, light, stage, and rigging equipment. This presentation has been held in many places, e.g., Bejing, Moskau, Mumbay, Istanbul, Barcelona.

TECHNICAL TOUR 3
Technical Museum
Saturday, May 5, 14:00 – 17:00

Beside mechanical musical instruments developed in the early twenties of the last century and operated by sophisticated light- or air-controlled devices, this visit will comprise the original first Austrian TV-studio (1950) and a demonstration of a laser vibrometer with a discussion of the results obtained with it (research on fortepiano soundboards and violins; in cooperation with the Vienna University of Music some students used the vibrometer for investigations of cettle drums, violins and cellos), an introduction to a restoration project (a Neo-Bechstein grand) where electroacoustical research was necessary to find the appropriate material for the strings, and a guided tour through the gallery of musical instruments.

Session P5 — Saturday, May 5 — 14:30 – 17:30
Room K

SPATIAL AUDIO PERCEPTION AND PROCESSING, PART 2

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

14:30

P5-1 On the Application of Sound Source Separation to Wave-Field Synthesis—*Máximo Cobos, Jose López*, Technical University of Valencia, Valencia, Spain

Wave-Field Synthesis (WFS) is a spatial sound system that can synthesize an acoustic field in an extended area by means of loudspeaker arrays. Spatial positioning of virtual sources is possible but requires separated signals for each source to be feasible. Despite the fact that most of the music is recorded in separated tracks for each instrument, in the stereo mix-down process this information is lost. Unfortunately, most of the existing recorded material is in stereo format. In this paper we propose to use sound source separation techniques to overcome this problem. Existing algorithms are yet far from perfect resulting in audible artifacts that clearly reduce the quality of the resynthesized sources in practice. Despite these artifacts, when separated ➥

sources are mixed again by a WFS system they are masked by other sounds. The utility of different separation algorithms and the subjective results are discussed.
Convention Paper 7016

15:00

P5-2 Reproduction of Arbitrarily Shaped Sound Sources with Wave Field Synthesis—Physical and Perceptual Effects—*Marije Baalman*, Technische Universität Berlin, Berlin, Germany

Current Wave Field Synthesis (WFS) implementations only allow for point sources and plane waves. In order to reproduce arbitrarily shaped sound sources with WFS several aspects need to be considered, such as the WFS-operator for source points outside of the horizontal plane, discretization of the object surface and diffraction of the sound around the sounding object itself, which can be modeled by introducing secondary sources at the edges of the object. This paper discusses those issues, describes the implementation in software and shows results of both objective and subjective evaluation.
Convention Paper 7017

15:30

P5-3 The Effect of Head Diffraction on Stereo Localization in the Mid-Frequency Range—*Eric Benjamin, Phil Brown*, Dolby Laboratories, San Francisco, CA, USA

In a previous paper, the present author described anomalous localization in intensity stereo at frequencies above the frequency at which the head is approximately one wavelength in diameter. Conventional analysis of stereo localization has usually depended on an asymptotic shadowless model of the head's diffraction. Measurements of the ear signals heard by the subjects in localization experiments showed that there were large differences between what was predicted by the simple model, and what was found in actual circumstances. We present a simple model for the head's diffraction in the range of 1200 Hz to 5000 Hz and show that it produces results which correspond more closely to real-world localization.
Convention Paper 7018

16:00

P5-4 Multiple Exponential Sweep Method for Fast Measurement of Head Related Transfer Functions—*Piotr Majdak, Peter Balazs, Bernhard Laback*, Austrian Academy of Sciences, Vienna, Austria

Presenting sounds in virtual environments requires filtering of free field signals with head related transfer functions (HRTF). The HRTFs describe the filtering effects of pinna, head, and torso measured in the ear canal of a subject. The measurement of HRTFs for many positions in space is a time-consuming procedure. To speed up the HRTF measurement the multiple exponential sweep method (MESM) was developed. MESM speeds up the measurement by interleaving and overlapping sweeps in an optimized way and retrieves the impulse responses of the measured systems. In this paper the MESM and its parameter optimization is described. As an example of an application of MESM, the measurement duration of an HRTF set with 1550 positions is compared to the unoptimized method. Using MESM, the measurement duration could be reduced by a factor of four without a reduction of the signal-to-noise ratio.
Convention Paper 7019

16:30

P5-5 A Fast Multipole Boundary Element Method for Calculating HRTFs—*Wolfgang Kreuzer, Zhensheng Chen*, Austrian Academy of Sciences, Vienna, Austria

Methods for measuring head related transfer functions (HRTFs) for an individual person are rather long and complicated. To avoid this problem a numerical model using the Boundary Element Method (BEM) is introduced. In general, such methods have the drawback that the computations for high frequencies are very time- and resource-consuming. To reduce these costs the BEM-model is combined with a fast multipole method (FMM) and a reciprocal approach.
Convention Paper 7020

17:00

P5-6 A Hybrid Artificial Reverberation Algorithm—*Rebecca Stewart*,[1] *Damian Murphy*[2]
[1]Queen Mary, University of London, London, UK
[2]University of York, York, UK

Convolution based reverberation allows for an accurate reproduction of a space, but yields no flexibility in defining that space, while filterbank-based reverberation allows computational efficiency and flexibility but lacks accuracy. A hybrid artificial reverberation algorithm that uses elements of both convolution and filterbank reverberation is investigated. An impulse response is truncated to contain only the early reflections and is convolved with input audio; the output audio then is combined with audio processed through a filterbank to simulate the late reflections. The parameters defining the filterbank are derived from the impulse response being analyzed. It is shown that this hybrid reverberator can produce a high-quality reverberation comparable to convolution reverberators.
Convention Paper 7021

Exhibitor Seminar **Saturday, May 5**
14:30 **Room 357**

CUBE-TEC INTERNATIONAL

Presenter: **Stefan Zachau**

"World Class Noise Reduction for Pro Tools"

Cube-Tec audio restoration and mastering tools (VPI's) are used by some of the world's most prestigious facilities. This presentation covers a selection of restoration VPI's that have recently been converted for use on the

Digidesign Pro Tools platform. The presentation will include tools for hum & buzz removal, noise reduction, de-crackling, disturbance tone removal, and much more.

Saturday, May 5 14:30 Room 633
Technical Committee Meeting on Audio for Telecommunications

Saturday, May 5 14:30 Room 142/143
Standards Committee Meeting on SC-04-01 Acoustics and Sound Source Modeling

Session P6 Saturday, May 5 15:00 – 16:30
Foyer IK

POSTERS: SIGNAL PROCESSING, SOUND QUALITY DESIGN

15:00

P6-1 On Development of New Audio Codecs—*Imre Varga*, Siemens Networks, Munich, Germany

This paper presents the works recently completed or on-going in 3GPP and ITU-T on the development of new audio codecs. The main applications are wideband speech telephony, audio conferencing, and mobile multimedia applications including Packet-Switched Streaming (PSS), Multimedia Messaging (MMS), and Multimedia Broadcast/Multicast Service (MBMS). In the standardization process, terms-of-reference describing design constraints and performance requirements, test plans, selection rules are finalized first. Next, extensive subjective listening testing is conducted. The codec selection is based on the selection test results and the selection rules. Characterization phase of testing allows obtaining the full amount of information on the codec behavior.
Convention Paper 7022

15:00

P6-2 Fixed-Point Processing Optimization of the MPEG Audio Encoder Using a Statistical Model—*Keun-Sup Lee*,[1] *Young-Cheol Park*,[2] *Dae Hee Youn*[3]
[1]Samsung Electronics Co. Ltd., Suwon, Korea
[2]Yonsei University, Wonju, Korea
[3]Yonsei University, Seoul, Korea

Audio applications for portable devices have two critical restrictions: small size and low power consumption. Therefore, fixed-point implementations are essential for those applications. Even with a fixed-point processor, however, the data width can still be an issue because it can affect both the hardware cost and power consumption. In this paper we propose a statistical model for the MPEG AAC audio encoder that can provide an optimal precision for the implementation. The hardware with the optimal precision, being compared with the floating-point system, is supposed to have perceptually insignificant errors at its output. To have an optimal precision for the AAC encoder, we estimate the maximum allowable amount of fixed-point arithmetic errors in the bit-allocation process using the statistical model. Finally, we present an architecture for the system appropriate for encoding the audio signals with minimum errors by the fixed-point processing. Tests showed that the fixed-point system optimized using the proposed model had sound quality comparable to the floating-point encoding system.
Convention Paper 7023

15:00

P6-3 Enhanced Bass Reinforcement Algorithm for Small-Sized Transducer—*Han-gil Moon, Manish Arora, Chiho Chung, Seong-cheol Jang*, Samsung Electronics Co. Ltd., Suwon, Gyeonggi-Do, Korea

Nowadays, mobile devices such as cell phones or MP3 players using small-sized loudspeaker systems to supply sound events to users is very popular. The main reasons why small-sized transducers are being used are due to the design and the size of the devices. Unfortunately, their design and size restrain the transducers from high quality of low frequency performance. To breakthrough this physical barrier of poor low frequency generation, the well-known psychoacoustical background “missing fundamental illusion” is exploited. In this paper the method of enhancing bass perception using virtual pitch is presented. In our demonstration, listeners can feel the deep bass with fewer artifacts.
Convention Paper 7024

15:00

P6-4 Subordinate Audio Channels—*Tim Jackson*,[1] *Keith Yates*,[1] *Francis Li*[2]
[1]Manchester Metropolitan University, Manchester, UK
[2]University of Salford, Salford, Greater Manchester, UK

In this paper we propose a model for a backward-compatible subordinate audio channel within a host digital audio signal. Embedding and extraction methods are presented and objective-perceptual assessment results reported. The method is designed so as to minimize perceptual degradation to the host signal and maintain compatibility with existing systems. The implementations utilize the discrete cosine transform and the masking properties of the human auditory system. Performance evaluation is assessed using the objective perceptual measure, objective difference grade. Test results support that both the host and subordinate audio channels can maintain good audio fidelity without significant perceptual degradation.
Convention Paper 7025

15:00

P6-5 Room Equalization Based on Acoustic and Human Perceptual Features—*Lae-Hoon Kim*,[1] *Mark Hasegawa-Johnson*,[2] *Jun-Seok Lim*,[3] *Koeng-Mo Sung*[1]
[1]Seoul National University, Seoul, Korea
[2]University of Illinois at Urbana-Champaign, Urbana, IL, USA
[3]Sejong University, Seoul, Korea

Room equalization has the potential to create ➥

improved audio display for homes, cars, and professional applications. In this paper, the signal is inverse filtered using an inverse filter computed by using newly introduced regularized optimal multipoint frequency-warped linear prediction coefficients. We present experimental results that show that the proposed room equalization algorithm improves equalization on the equalizable parts, thus enlarging the region of perceptually effective equalization.
Convention Paper 7026

15:00

P6-6 Parametric Loudspeaker Equalization—Results and Comparison with Other Methods
—*German Ramos, Jose J. Lopez*, Technical University of Valencia, Valencia, Spain

The results obtained by a loudspeaker equalization method are presented and compared with other equalization methods. The main characteristic of the proposed method resides on the fact that the equalizer structure is planned from the beginning as a chain of SOS (Second Order Sections), where each SOS is a low-pass, high-pass, or parametric filter defined by its parameters (frequency, gain, and Q), and designed by a direct search method. This filter structure, combined with the subjectively motivated error function employed, allows obtaining better results from a subjective point of view and requiring lower computational cost. The results have been compared with different FIR (finite impulse response) and IIR (infinite impulse response) filter design methods, with and without warped structures. In all cases, for the same computational cost, the presented method obtains a lower error function value.
Convention Paper 7027

15:00

P6-7 A Zero-Pole Vocal Track Model Estimation Method Accurately Reproducing Spectral Zeros—*Damián Marelli*,[1] *Peter Balazs*[2]
[1]University of Vienna, Vienna, Austria
[2]Austrian Academy of Sciences, Vienna, Austria

Model-based speech coding consists in modeling the vocal tract as a linear time-variant system. The all-pole model produced by the computationally efficiency linear predictive coding method provides a good representation for the majority of speech sounds. However, nasal and fricative sounds, as well as stop consonants, contain spectral zeros, which requires the use of a zero-pole model. Roughly speaking, a zero-pole model estimation method typically does a nonparametric estimation of the vocal tract impulse response and tunes the zero-pole model to fit this estimation in a square sense. In this paper we propose an alternative strategy. We tune the zero-pole model to directly fit the power spectrum of the speech signal in a logarithmic scale, to be consistent with the way the human ear perceives sounds. In this way, we avoid the error introduced by the vocal tract impulse response estimation and obtain a model that is more accurate at reproducing spectral zeros in a logarithmic scale. A drawback of the proposed method, however, is its computational complexity.
Convention Paper 7028

15:00

P6-8 Artificial Speech Synthesis Using LPC—
Manjunath D. Kadaba, Uvinix Computing Solutions Bangalore/Karnataka, Bangalore, India

Speech analysis and synthesis with Linear Predictive Coding (LPC) exploit the predictable nature of speech signals. Cross-correlation, autocorrelation, and autocovariance provide the mathematical tools to determine this predictability. If we know the autocorrelation of the speech sequence, we can use the Levinson-Durbin algorithm to find an efficient solution to the least mean-square modeling problem and use the solution to compress or resynthesize the speech.
Convention Paper 7029

Tutorial 2 — Saturday, May 5
15:00 – 18:00 — Room H

STEREOPHONIC TECHNIQUES

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

What is stereo? Why and how do we hear with spatial acuity? How can we realistically capture and reproduce the stereo sound field with just two microphones and two loudspeakers? These are but a few of the questions discussed in this in-depth tutorial.

The session begins with a discussion and demonstration of how the human ear-brain hearing system works. This is followed by a historical overview of the development stereophonic recording. The main body of the session presents a comprehensive analysis of the various common stereophonic microphone configurations and concludes with numerous recorded examples for evaluation and comparison of the techniques discussed.

Exhibitor Seminar — Saturday, May 5
15:30 — Room 357

LAWO – NETWORKING AUDIO SYSTEMS

Presenter: **Felix Krueckels**

“Integration of Workstation Audio Tools in Modern Broadcast Live Situations”

Lawo will be presenting a new approach in integrating audio editing tools in live broadcast and production desks previously provided only by workstations. You will be enthusiastic about a full integration of audio editing tools in the broadcast mixing desks mc266 and mc290. All common applications of the mc2-series are also available for this integration. See and hear how the merge of two completely independent worlds is realized in the Lawo products. The second part of the lecture is devoted to the networking of several Lawo systems with total studio control.

Saturday, May 5 — 15:30 — Room 633
Technical Committee Meeting on High Resolution Audio

Workshop 6 **Saturday, May 5**
16:30 – 18:30 **Room G**

LOUDNESS METERING SYSTEMS AND INTERNATIONAL ACTIVITIES

Chair: **Lars Jonsson**, SR/Swedish Radio

Panelists: *Kimio Hamasaki*, NHK, Japan
Erik Lundbeck, SVT Swedish TV
Thomas Lundh, TC Electronics, Denmark
Ralph Kessler, Pinguin, Germany
Andrew Mason, BBC, UK
Tony Spath, Dolby, UK
Gerhard Spikofski, IRT, Germany
Matti Zemack, SR, Swedish Radio

This workshop will cover loudness metering and automatic level adjustments and present an overview of systems and international activities. The session will cover the latest progress in standards made by the ITU and the EBU to recommend solutions for broadcasting. The general problem to solve is common to all new digital distribution systems. Old analog audio metering methods have become obsolete and can no longer create a smooth perceived level for listeners. Manufacturers making proposals along with the new standards are presenting their products. A general discussion among the panel members and the audience will point out the direction for the future.

Tutorial 3 **Saturday, May 5**
16:30 – 18:30 **Room 560/561**

REMAPPING AND DOWNMIXING

Presenter: **Jeff Levison**, DTS, Inc., Agoura Hills, CA, USA

For the new Blu-ray and HD-DVD systems remapping of loudspeakers is possible. What this means is that a given mix loudspeaker position can be described and then on playback the room loudspeaker positions are used to balance the power distribution if there is a difference between the source and the playback speaker positions (or the number of speakers). Several examples will be played to demonstrate various remapping and downmixing scenarios in a real-world manner.

Exhibitor Seminar **Saturday, May 5**
16:30 **Room 357**

SCHOEPS

Presenter: **Helmut Wittek**

"Double M/S Recording in Stereo and Surround"

"Double M/S" is a technique for two-channel or multichannel stereo recording. It uses a compact microphone arrangement that requires only three recorder channels, while allowing full postproduction processing. During the seminar, recorded examples from a range of scenarios will be presented. The new Schoeps VST Plug-In will also be introduced and its capabilities demonstrated.

Saturday, May 5 **16:30** **Room 633**
Technical Committee Meeting on SC-04-03 Loudspeaker Modeling Measurement

Saturday, May 5 **16:30** **Room 142/143**
Standards Committee Meeting on SC-04-01 Acoustics and Sound Source Modeling

SPECIAL EVENT
Wave Field Synthesis Demonstration
Saturday, May 5, 16:45 – 18:45
Sunday, May 6, 16:45 – 18:45
University for Music and Performing Arts, Vienna

CELEBRATE GOOD VIBES!

Hosted by the University of Music and Performing Arts Vienna and supported by their partners Fraunhofer Institute for Digital Media Technology IDMT, Sennheiser, Studer, and Merging Technologies, this Demo Session will present the reproduction of preliminarily recorded concerts in Wave Field Synthesis.

Attendees will be able to listen to the recordings that were done in the afternoon.

Different microphone techniques and mixing concepts can be compared and evaluated. Neumann/Sennheiser microphones, a Vista 5 console by Studer, the Merging Technologies Pyramix Digital Audio Workstation, and the Spatial Audio Workstation by Fraunhofer IDMT will provide a professional mixing environment. Additionally Wave Field Synthesis demos including film, radio drama, and music will be presented.

This Special Event will be held in two sessions on May 5 and May 6 at the University of Music Vienna. A bus shuttle service will be offered between the Austria Center and the University.

Tickets will be available from the Technical Tour counter. Meeting point is the main entrance of the Austria Center. Please note that the times listed above include travel time.

Additional Demo Sessions in Wave Field Synthesis will also be available at the following times:

- Saturday, May 5, 17:15 – 18:15
- Sunday, May 6, 17:15 – 18:15

STUDENT EVENT
Opening and Student Delegate Assembly Meeting – Part 1
Saturday, May 5, 17:00 – 18:30 Room P

Chair: **Ainslie Harris**

Vice Chair: **Suzana Jakovic**

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

All students and educators are invited to participate in this meeting. Election results and Recording Competition and Poster Awards will be given at the Student Delegate Assembly Meeting–2 on Tuesday, May 8, at 09:00.

Exhibitor Seminar **Saturday, May 5**
17:30 **Room 357**

MICROTECH GEFELL

Presenter: **Ulrich Apel**

"Surround Microphones Not Only for Concert Halls"

Microtech-Gefell presents microphones with special rain-proof windscreens that are applicable not only in closed premises but for recordings of atmos on locations; even for film and television. Herewith Microtech-Gefell translated their outdoor-and environmental experiences into their studio-microphones. An application of this wind-and rainshield and an operation of new switchable microphone UM930 will be shown.

SPECIAL EVENT
Mixer Party
Saturday, May 5, 18:30 – 19:30

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

Session P7 **Sunday, May 6** **09:00 – 12:00**
Room I

PSYCHOACOUSTICS, PERCEPTION, AND LISTENING TESTS, PART 1

Chair: **Stefan Weinzierl**, Technical University of Berlin, Berlin, Germany

09:00

P7-1 Some Effects of the Torso on Head-Related Transfer Functions—*Ole Kirkeby,*[1] *Eira Seppälä,*[1] *Asta Kärkkäinen,*[1] *Leo Kärkkäinen,*[1] *Tomi Huttunen*[2]
[1]Nokia Research Center, Helsinki, Finland
[2]University of Kuopio, Kuopio, Finland

A numerical method based on the ultra-weak variational formulation (UWVF) is used to calculate three sets of Head-Related Transfer Functions (HRTFs). The three sets are made by combining a hard head with a hard torso, a moderately absorbing torso, and no torso. Each set is sampled for every 50 Hz from DC to 24 kHz at 21,872 points almost evenly distributed in the far-field, thus providing a spatial resolution of approximately one degree everywhere. Since the results of the numerical simulations are not contaminated by the response of an electroacoustic chain it is possible to compare the HRTFs of a head and torso model to the HRTFs of the head only without the risk of interpreting a measurement artifact as a physical phenomenon.
Convention Paper 7030

09:30

P7-2 An Investigation into Head Movements Made When Evaluating Various Attributes of Sound—*Chungeun Kim, Russell Mason, Tim Brookes*, University of Surrey, Guildford, Surrey, UK

This paper extends the study of head movements during listening by including various listening tasks where the listeners evaluate spatial impression and timbre, in addition to the more common task of judging source location. Subjective tests were conducted in which the listeners were allowed to move their heads freely while listening to various types of sound and asked to evaluate source location, apparent source width, envelopment, and timbre. The head movements were recorded with a head tracker attached to the listener's head. From the recorded data, the maximum range of movement, mean position and speed, and maximum speed were calculated along each axis of translational and rotational movement. The effects of various independent variables, such as the attribute being evaluated, the stimulus type, the number of repetition, and the simulated source location were examined through statistical analysis. The results showed that while there were differences between the head movements of individual subjects, across all listeners the range of movement was greatest when evaluating source width and envelopment, less when localizing sources, and least when judging timbre. In addition, the range and speed of head movement was reduced for transient signals compared to longer musical or speech phrases. Finally, in most cases for the judgment of spatial attributes, head movement was in the direction of source direction.
Convention Paper 7031

10:00

P7-3 Binaural Resynthesis for Comparative Studies of Acoustical Environments—*Alexander Lindau, Torben Hohn, Stefan Weinzierl*, Technical University of Berlin, Berlin, Germany

A framework for comparative studies of binaurally resynthesized acoustical environments is presented. It consists of a software-controlled, automated head and torso simulator with multiple degrees of freedom, an integrated measurement device for the acquisition of binaural impulse responses in high spatial resolution, a head-tracked real-time convolution software capable to render multiple acoustic scenes at a time, and a user interface to conduct listening tests according to different test designs. Methods to optimize the measurement process are discussed, as well as different approaches to data reduction. Results of a perceptive evaluation of the system are shown, where acoustical reality and binaural resynthesis of an acoustic scene were confronted in direct A/B comparison. The framework permits, for the first time, to study the perception of a listener instantaneously relocated to different binaurally rendered acoustical scenes.
Convention Paper 7032

10:30

P7-4 Acoustic Factors of Auditory Distance Perception by the Blind While Walking—*Takahiro Miura, Shuichi Ino; Teruo Muraoka, Tohru Ifukube*, University of Tokyo, Tokyo, Japan

The ability by which the blind that can recognize surrounding objects solely by hearing is called

"obstacle sense." By analyzing and modeling its mechanism, this model will be utilized for realizing an acoustic VR environment as well as training systems for the vision-impaired through acoustic analysis. In this paper the authors particularly focused on various sorts of acoustic factors that may contribute to perceive the distance from the subject to the obstacle especially while walking. We also investigated the factors based on the psychophysical experiments and acoustical analysis methods. In addition, the authors discussed the contribution of these factors to the blind persons' auditory distance perception.
Convention Paper 7033

11:00

P7-5 Listener Loudspeaker Preference Ratings Obtained in situ Match those Obtained via a Binaural Room Scanning Measurement and Playback System—*Sean Olive*,[1] *Todd Welti*,[1] *William Martens*[2]
[1]Harman International Industries, Inc., Northridge, CA, USA
[2]McGill University, Montreal, Quebec, Canada

Binaural room scanning (BRS) is a method of capturing, storing, and reproducing via a head-tracking headphone display system the binaural room impulse response of one or more loudspeakers in a listening room. This paper reports the results of the first test in a series of validation tests of a custom BRS system that was developed for research and evaluation of different loudspeakers and different listening spaces. The test examined whether listeners' loudspeaker preference ratings made in a listening room with reflective walls (in situ) were comparable to ratings made in response to BRS reproductions of those loudspeakers located in the same room. Virtually the same results were obtained in these two cases.
Convention Paper 7034

11:30

P7-6 Perceptually-Motivated Audio Morphing: Brightness—*Duncan Williams, Tim Brookes*, University of Surrey, Guildford, Surrey, UK

A system for morphing the brightness of two sounds independently from their other perceptual or acoustic attributes was coded, based on the spectral modeling synthesis additive/residual model. A multidimensional scaling analysis of listener responses showed that the brightness control was perceptually independent from the other controls used to adjust the morphed sound. A timbre morpher, providing perceptually meaningful controls for additional timbral attributes, can now be considered for further work.
Convention Paper 7035

Workshop 7 **Sunday, May 6**
09:00 – 11:00 **Room H**

USER-CENTERED DESIGN OF CONTROLLERS FOR PRO AUDIO

Chair: **William Martens**, McGill University, Montreal, Quebec, Canada

Panelists: *Durand Begault*, NASA Ames Research Center, Mountainview, CA, USA
Florian Camerer, ORF - Austrian TV, Vienna, Austria
George Massenburg, GML, Inc., TN, USA

The design of controllers and/or work surfaces for music recording and mixing will be examined from the experienced user's perspective. This examination is motivated by the belief that there may be better ways to organize controllers, ways that prioritize access to controls in a manner that is based upon what experienced users know about how they do their work. The list of workshop panelists, which includes representatives from live sound, postproduction, and mastering will have a brainstorming session about new control layouts using tools designed to stimulate "thinking outside of the box." Also planned are breakout sessions, within which small groups will be lead to focus on particular applications and/or problems, with results that hopefully contribute some new and useful ideas.

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

SOUND ARCHIVING

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

The world's audio heritage is estimated to amount to 100 million hours of recorded documents. A considerable part is at severe risk of not surviving in the long-term, as it is still kept on analog or digital single carriers, which sooner or later are prone to deterioration. An even greater threat is the fast withdrawal from the manufacture of specific replay equipment and spare parts. This will lead to a situation where still well-preserved recordings cannot be retrieved anymore because of the lack of replay equipment.

This tutorial concentrates on two basic documents for long-term audio preservation, released by the Technical Committee of IASA, the International Association of Sound and Audiovisual Archives:

- IASA-TC 03, The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy
- IASA-TC 04, Guidelines on the Production and Preservation of Digital Audio Objects

The tutorial will also survey the respective AES Standards that concentrate on the storage and handling of various types of audio carriers.

STUDENT EVENT
Recording Competition—Stereo
Sunday, May 6, 09:00 – 12:30 Room G

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

09:00 – Jazz Stereo
09:45 – Pop/ Rock Stereo
10:30 – Folk/ World Music
11:15 – Classical Stereo

Judges:
Classical: Kimio Hamasaki, Ulrich Vette, TBA
Pop: Andres Mayo, Darcy Proper, Thomas Rabitsch
Jazz: Christian Heck, Jan Erik Kongshaug, TBA
World/Folk: Eva Bauer-Oppelland, Akira Fukada, TBA

We would like to acknowledge the generosity of our sponsors for the student recording competition: AKG, Audio-Technica, DTS, Magix, PMC, RME, Schoeps, Sennheiser.

Sunday, May 6 09:00 Room 633
Technical Committee Meeting on Automotive Audio

Sunday, May 6 09:00 Room 142/143
Standards Committee Meeting on SC-02-01 Digital Audio Measurement Techniques

Session P8 Sunday, May 6 09:30 – 11:00
Room K

MULTICHANNEL SOUND, PART 1

Chair: **Karl Petermichl**, ORF, Vienna, Austria

09:30

P8-1 5.1 Radio—Too Much Too Soon?—*Jon McClintock*, APT, Belfast, Ireland, UK

5.1 multichannel audio is arguably the next natural progression for radio. Although digital radio offered an incremental improvement over stereo for listeners, there has not been a fundamental change in radio since the migration from AM to FM. For radio to survive in a highly competitive environment, with an audience that is increasingly judgmental on delivery mediums and content, then as an industry radio needs to embrace 5.1. This paper explores the principles that are vital to the success of multichannel audio for radio, the enabling technology, and outlines various projects that have been undertaken.
Convention Paper 7036

10:00

P8-2 Wide Listening Area with Exceptional Spatial Sound Quality of a 22.2 Multichannel Sound System—*Kimio Hamasaki, Toshiyuki Nishiguchi, Reiko Okumura, Yasushige Nakayama*, NHK Science and Technical Research Laboratories, Tokyo, Japan

While issues regarding the sweet spot in 5.1 surround sound have been discussed, a 22.2 multichannel sound system has been developed for ultrahigh-definition TV. One of its features is expansion of the listening area with exceptional sound quality. Although the wideness of listening area was reported in previous papers, its evaluation was performed using only sound clips of a symphony orchestra without a picture. Therefore, subjective evaluations were performed for comparing the impression of various spatial attributes at different listening positions using contents with pictures in both large and small rooms. These evaluations demonstrate that viewers have better impressions of various spatial attributes in a wider listening area with the 22.2 multichannel sound system than with other sound systems.
Convention Paper 7037

10:30

P8-3 Up-Mixing and Localization—Localization Performance of Up-Mixed Consumer Multichannel Formats—*Ben Shirley*,[1] *Richard Chaffey*[2]
[1]University of Salford, Salford, Greater Manchester, UK
[2]AVE Systems, Hersham, Surrey, UK

A number of listening tests were carried out to assess localization of sound in derived surround sound fields. Two up-mixed consumer multichannel formats that use matrix decoding of 3/2 multichannel surround channels to increase the surround channel array (Dolby Pro Logic IIx and DTS Neo:6) were compared to original 3/2 multichannel material to determine the degree of spatial performance improvement. Noise bursts were panned to 11 different locations, 13 subjects participated in the tests, and results were analyzed to assess any improvement in localization in each of the assessed surround systems.
Convention Paper 7038

Exhibitor Seminar Sunday, May 6
09:30 Room 357

LEAR CORPORATION/DIRAC RESEARCH

Presenters: **Nilo Casimiro Ericsson, Mathias Johansson, Armin Prommersberger**

"Dirac Live™—A New Approach to Room Correction Algorithms"

A method for improving car audio systems through digital correction of the impulse/frequency responses of individual loudspeakers is presented. The ability to simultaneously optimize different time and frequency domain criteria and flexible spatial averaging are the main benefits of the new approach. The solution is featured in the new BMW M5 Individual High End Sound System, recently reviewed by Auto, Motor und Sport (number 8/2007 pp. 146–148) stating: "the most impressive car stereo sound in the world." An A/B comparison to conventional filtering is available in LEAR's demo car in our common exhibition booth.

TECHNICAL TOUR 4
ORF TV Broadcasting Studios
Sunday, May 6, 09:30 – 12:30

ORF TV is the National Public Broadcaster for Austria with two main TV channels (ORF 1+ 2), two dedicated theme channels (TW1 – tourism and weather, ORF Sport Plus), regional programs, and diverse off-air activities, as well as being by far the most popular Web-presence in the country. The focus of the tour to ORF TV will be the three newly refurbished postproduction studios with outstanding acoustic design and full surround sound capability, unique for European broadcast studios. Example clips will be demonstrated, and the tour will also cover the general infrastructure shortly after a major relaunch of a substantial part of ORF"s programs.

Sunday, May 6 10:00 Room 633
Technical Committee Meeting on Multichannel and Binaural Audio Technologies

Session P9 Sunday, May 6 10:30 – 12:00
Foyer IK

POSTERS: AUDIO RECORDING AND REPRODUCTION

10:30

P9-1 Intelligent Editing of Studio Recordings with the Help of Automatic Music Structure Extraction— *György Fazekas, Mark Sandler*, Queen Mary, University of London, London, UK

In a complex sound editing project, automatic exploration and labeling of the semantic music structure can be highly beneficial as a creative assistance. This paper describes the development of new tools that allow the engineer to navigate around the recorded project using a hierarchical music segmentation algorithm. Segmentation of musical audio into intelligible sections like chorus and verses will be discussed briefly followed by a short overview of the novel segmentation approach by timbre-based music representation. Popular sound-editing platforms were investigated to find an optimal way of implementing the necessary features. The integration of music segmentation and the development of a new navigation toolbar in Audacity, an open-source multitrack editor, will be described in more detail.
Convention Paper 7039

10:30

P9-2 Constant Complexity Reverberation for Any Reverberation Time—*Tobias May,*[1,2] *Daniel Schobben*[1]
[1]Philips Research Laboratories, Eindhoven, The Netherlands
[2]Carl-von-Ossietzky University Oldenburg, Oldenburg, Germany

A new artificial reverberation system is proposed, which is based on perceptually relevant components in reverberated audio and, as such, allows for a very efficient implementation. The system first separates the signal into transient and steady-state components. The transient signal is reverberated by using an efficient time-varying recursive filter while the steady-state signal is processed separately with an all-pass filter. In contrast to common reverberation systems, the complexity of the recursive filter is determined solely by the duration of the transients and is therefore independent of the reverberation time.
Convention Paper 7040

10:30

P9-3 Outdoor and Indoor Recording for Motion Picture. A Comparative Approach on Microphone Techniques—*Christos Goussios, Christos Sevastiadis, George Kalliris*, Aristotle University of Thessaloniki, Thessaloniki, Greece

Several recording techniques and equipment are used in outdoor and indoor recordings for motion pictures. The choices are usually characterized from subjectivity and technical limitations irrelevant to the desired final sound quality. Our goal is to present results of comparative recordings in order to give answers to every-day practice problems that arise. Overhead and underneath booming and the use of wireless microphones are compared through third octave frequency analysis.
Convention Paper 7041

10:30

P9-4 Semi-Automatic Mono to Stereo Up-Mixing Using Sound Source Formation—*Mathieu Lagrange,*[1] *Luis Gustavo Martins,*[2] *George Tzanetakis*[1]
[1]University of Victoria, Victoria, British Columbia, Canada
[2]INESC Porto, Porto, Portugal

In this paper we propose an original method to include spatial panning information when converting monophonic recordings to stereophonic ones. Sound sources are first identified using perceptively motivated clustering of spectral components. Correlations between these individual sources are then identified to build a middle level representation of the analyzed sound. This allows the user to define panning information for major sound sources thus enhancing the stereophonic immersion quality of the resulting sound.
Convention Paper 7042

Live Sound Seminar 2 Sunday, May 6
10:30 – 12:30 Room E

NEUMAN & MUELLER: 3 Orchestras & Stars — World Cup Opening in Munich

Presenters: **Michael Kennedy**
Rudolf Pirc

As part of the opening celebrations for the Soccer World Cup in Germany, a concert was given in Munich on June 6, 2006 in the Munich Olympic Stadium. Together on stage were the three principle Munich Orchestras: the Munich Philharmonic, the Bavarian State Opera Orchestra, and the Bavarian Broadcasting Symphony Orchestra. The concert program included Placido Domingo, soprano Diana Damrau, the pianist Lang Lang, and the German band "Die Söhne Mannheims" with Xavier Naidoo.

The presentation describes the realization of the sound reinforcement system, which featured J and Q series D&B loudspeakers, three Yamaha PM1 D systems at Front of House, a PM5D at the monitor position, and two separate DME systems. All interconnections between the mixing consoles and the DME systems were realized in the digital domain.

Exhibitor Seminar Sunday, May 6
10:30 Room 357

D.A.V.I.D.

Presenter: **Ingo Hahn**

"Bi-Medial Production"

DigaSystem offers a wide variety of bi-medial production tools embedded in a scalable content management system. This includes support of standard broadcast audio and video formats, as well as formats for new delivery channels like mobile devices and new media. This lecture gives an overview about bi-medial production workflows within DigaSystem.

Tutorial 5 Sunday, May 6 11:00 – 12:30
Room H

MXF—THE MATERIAL EXCHANGE FORMAT

Presenter: **Tim Harris**, Snell & Willcox

This tutorial takes an audio-oriented look at the MXF file format. It will start off with a general introduction—to the basics of MXF—what MXF is, why it was invented, by whom it was developed, and how it was standardized. The tutorial will then focus on the MXF synchronization model, and the capabilities of the format to combine audio, video, data and metadata in a versatile way.

Carriage of essence within MXF will then be explained with a particular focus on audio. Attention will be drawn to the involvement of the AES in providing the underlying international reference for the Broadcast Wave format. Metadata annotation of essence, particularly recent work on the MXF Master Format Guidelines for multi-lingual annotation, will be explained.

The talk will be interspersed with demonstrations of how MXF software could be used in real-world workflows.

Exhibitor Seminar Sunday, May 6
11:00 Room 441

SENNHEISER

Presenter: **Gregor Zielinsky**

"High Quality Surround Recording with the New MKH 8000 Microphone"

In this seminar Sennheiser presents the brand-new MKH 8000 Studio Microphone Series for the first time ever. On the examples of classical symphony and big band multitrack recordings, this "hands-on" seminar will show several surround microphone techniques using the MKH 8000. The different recordings will be explained and can be compared and judged by the listeners.

Sunday, May 6 11:00 Room 633
Technical Committee Meeting on Acoustics and Sound Reinforcement

Sunday, May 6 11:00 Room 142/143
Standards Committee Meeting on SC-04-04 Microphone Measurement and Characterization

Exhibitor Seminar Sunday, May 6
11:30 Room 357

SCHOEPS

Presenter: **Helmut Wittek**

"Double M/S Recording in Stereo and Surround"

"Double M/S" is a technique for two-channel or multichannel stereo recording. It uses a compact microphone arrangement that requires only three recorder channels, while allowing full postproduction processing. During the seminar, recorded examples from a range of scenarios will be presented. The new Schoeps VST Plug-In will also be introduced and its capabilities demonstrated.

STUDENT EVENT
Education Fair
Sunday, May 6, 12:00 – 14:00 Foyer D

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.

Session P10 Sunday, May 6 12:30 – 17:00
Room I

PSYCHOACOUSTICS, PERCEPTION, AND LISTENING TESTS, PART 2

Chair: **Sean Olive**, Harman International Industries, Inc., Northridge, CA, USA

12:30

P10-1 Verbal Elicitation and Scale Construction for Evaluating Perceptual Differences between Four Multichannel Microphone Techniques—*William L. Martens, Sungyoung Kim*, McGill University, Montreal, Quebec, Canada

A verbal elicitation task using triadic comparisons was completed by eight listeners to explore the adjectives that describe the audible differences between solo piano performances captured using four different multichannel microphone techniques. Although the elicited terms differed somewhat between listeners, a set of five bipolar adjective pairs were found to represent the most salient differences between the auditory imagery associated with multichannel-loudspeaker reproduction of the piano performances. These adjectives were used as the anchors for five attribute rating scales on which the same eight listeners rated each of the 32 stimuli that had been presented for triadic comparison. Stepwise multiple regression showed that ratings on three of the five attributes successfully predicted those listeners' preference ratings for the same stimuli.
Convention Paper 7043

13:00

P10-2 Broadcast Loudness: Mixing, Monitoring, and Control—*Alessandro Travaglini,* FOX International Channels (Italy), Rome, Italy

In the satellite broadcast era, to manage transmission from a digital platform that broadcasts dozens of channels and to guarantee loudness consistency to viewers has became a primary, yet difficult task. In this paper I describe my work in balancing the broadcasted loudness of the

Italian satellite TV platform SKY Italia. In fact, for a few years, its transmissions suffered very audible loudness inconsistency, due to several factors, such as numbers of channels, various content offerings, different mastering levels in between programs, and interstitials, outsourcing productions from many external facilities, etc. The project lasted for over one year, and the results are a much improved audio quality and a more balanced loudness consistency throughout all the channels involved.
Convention Paper 7044

13:30

P10-3 Sound Levels of TV Advertisements Relative to the Adjacent Programs and Cross-National Comparison of the Way of Their Insertion into Programs—*Eiichi Miyasaka, Akiko Kimura*, Musashi Institute of Technology, Yokohama, Kanagawa, Japan

A couple of perceptual experiments were conducted in order to investigate the relationship between the physical sound levels of advertisements (CMs) and the corresponding auditory perception relative to the reference speech. Experiment 1 for three types of CMs aired in Japanese broadcasting with different sound levels show that all CMs sound louder than a reference speech irrespective of large sound level differences. Experiment 2 for three types of CMs with similar sound levels and the standard deviations show that some of them sound louder than the reference. Next, ways of insertion of CMs into programs were investigated for news programs broadcast in Japan, the UK, and the US. The results show that silent periods with different durations were commonly inserted between main programs and CMs in UK and in US, while no silent period was introduced in Japanese commercial broadcasting.
Convention Paper 7045

14:00

P10-4 Influence of Interaction on Perceived Quality in Audio Visual Applications: Subjective Assessment with n-Back Working Memory Task, II—*Ulrich Reiter, Mandy Weitzel*, Technische Universität Ilmenau, Ilmenau, Germany

The mechanisms of human audio visual perception are not fully understood yet. For interactive audio visual applications running on devices with limited computational power it is desirable to know which of the stimuli to be rendered in an audio visual room simulation have the greatest impact upon the perceived quality of the system. We have conducted experiments to determine the effect of interaction upon the precision with which test subjects are able to discriminate between different parameter values of auditory attributes. This paper details one of these experiments and compares different approaches for the analysis of the obtained data. The results show a noticeable bias toward faulty ratings during the involvement in a task, although the analysis using significance tests did not completely confirm this effect.
Convention Paper 7046

14:30

P10-5 On the Audibility of Comb Filter Distortions
—*Stefan Brunner, Hans-Joachim Maempel, Stefan Weinzierl*, Technical University of Berlin, Berlin, Germany

Superpositions of delayed and undelayed versions of the same signal can occur at different stages of the audio transmission chain. Sometimes it is a deliberate measure to provide audio material with certain spatial or timbral qualities. Often it is a result of multiple microphone signals, sound reflections on walls or latencies in digital signal processing leading to comb-filter-shaped, linear distortions. The measurement of a hearing threshold for this type of distortion with its dependence on reflection delay, relative level, and the type of audio content can be the basis for boundaries in everyday recording practice below which undesired timbral distortions can be neglected. Therefore, a listening test was conducted to determine the just noticeable difference for three stimulus categories (speech, a snare drum roll, and a piano phrase) and different time delays between direct and delayed signal from 0.1 ms to 15 ms, equivalent to 0.03 to 5.15 m of sound path difference. The results show that comb-filter distortions can still be audible if the level of the first reflection is more than 20 dB lower than the level of the direct sound.
Convention Paper 7047

15:00

P10-6 VirtualPhone—A Rapid Virtual Audio Prototyping Environment—*Nick Zacharov*, Nokia Corporation, Tampere, Finland

As the complexity of mobile phones increases with the evolution of digital convergence, there is increased demand to ensure high audio quality for all applications. VirtualPhone is a graphical user interface based software environment allowing for the rapid prototyping of mobile phone audio and its subsequent calibrated auralization. This paper describes the framework of the VirtualPhone application, illustrates its usage and performance compared to other conventional prototyping schemes.
Convention Paper 7048

15:30

P10-7 Evaluation of HE-AAC, AC-3, and E-AC-3 Codecs— *Leslie Gaston, Richard Sanders*, University of Colorado at Denver and Health Sciences Center, Denver, CO, USA

The Recording Arts Program at the University of Colorado at Denver and Health Sciences Center (UCDHSC) performed an independent evaluation of three audio codecs: Dolby Digital (AC-3 at 384 kbps), Advanced Audio Coding Plus (HE-AAC at 160 kbps), and Dolby Digital Plus (E-AC-3 at 224 and 200 kbps). UCDHSC performed double-blind listening tests during the summer of 2006, which adhered to the standards of ITU-R BS.1116 (that provides guidelines for multichannel critical listening tests). The results of this test illustrate a clear delineation between the AC-3 codec and the others tested. We will present ➥

the test procedures and findings in this paper.
Convention Paper 7049

16:00

P10-8 Perceptual Evaluation of Mobile Multimedia Loudspeakers—*Gaetan Lorho*, Nokia Corporation, Helsinki, Finland

An experiment was conducted to compare the perceptual characteristics of stereo loudspeaker systems found in mobile multimedia devices. An individual vocabulary development approach was employed for this descriptive analysis. Ten systems and five musical programs were selected for the experiment. Sixteen listeners developed their own set of attributes in three hours and performed a comparative evaluation of the ten systems for several program items using the attribute scales they developed. A total of 111 attributes was generated in this experiment, which could be divided in several perceptual groups relating to spatial, timbral, loudness, sound disturbance and sound articulation aspects. The principle of this sensory profiling method is described and some results of the subjective experiment are presented.
Convention Paper 7050

16:30

P10-9 A Rapid Listening Test Environment—Helping Managers Make Better Decisions—*Nick Zacharov*, Nokia Corporation, Tampere, Finland

As the complexity of mobile phones increases with the evolution of digital convergence, there is increased demand to ensure high audio quality for all applications. This paper presents a set of software applications that allow for the rapid definition, administration, analysis, and reporting of listening tests without the need for extensive technical knowledge of the field. A through description of the concepts behind the client/server architecture of the software is presented followed by some example applications. Last, a performance comparison of listening tests performed using more traditional methods versus the presented method is made.
Convention Paper 7051

Session P11 Sunday, May 6 12:30 – 15:30
Room K

MULTICHANNEL SOUND, PART 2

Chair: **Ulrike Schwarz**, BR, Germany

12:30

P11-1 EBU Tests of Multichannel Audio Codecs—*Andrew Mason*,[1] *David Marston*,[1] *Franc Kozamernik*,[2] *Gerhard Stoll*[3]
[1]British Broadcasting Corporation, Tasworth, Surrey, UK
[2]EBU, Geneva, Switzerland
[3]IRT, Munich, Germany

The latest project of one of the European Broadcasting Union technical groups has been the assessment of the sound quality of multichannel audio bit rate reduction codecs for broadcast applications. Codecs under test include offerings from Dolby, DTS, implementations of MPEG AAC, and of the new MPEG Surround codec. The bit rates ranged from 64 kbit/s to 1.5 Mbit/s. The subjective tests, including choice of method, selection of test material, and statistical analysis of results are described. The conclusions are that the hope for consistently high quality at a relatively low bit rate of, say, 256 kbit/s has not yet been fulfilled, and that some audio material still demands at least 448 kbit/s. It has also been observed that later developments of some codecs perform less well than earlier versions.
Convention Paper 7052

13:00

P11-2 The Design and Analysis of First Order Ambisonic Decoders for the ITU Layout—*David Moore, Jonathan Wakefield*, University of Huddersfield, Huddersfield, West Yorkshire, UK

Ambisonic decoders for irregular layouts can be designed using heuristic search algorithms. These methods provide an alternative to solving complex mathematical equations. New fitness function objectives for search algorithms are presented that ensure derived decoders meet the requirements of the Ambisonic system more closely than previous work. The resulting new decoder coefficients are compared to other published coefficients, and a detailed performance analysis of first order decoders for the ITU layout is given. This analysis highlights common poor performance characteristics that these decoders hold. Proposed future work will attempt to address these issues by looking at techniques for producing decoders with a more even error distribution around the listener and investigating methods for removing the bias toward meeting certain objectives.
Convention Paper 7053

13:30

P11-3 A New Digital Module for Variable Acoustics and Wave Field Synthesis: Design and Applications—*Diemer de Vries*,[1] *Jasper van Dorp Schuitman*,[1,2] *At van den Heuvel*[2]
[1]Delft University of Technology, Delft, The Netherlands
[2]Acoustic Control Systems, Garderen, The Netherlands

A new digital module has been developed that creates variable acoustics for multipurpose halls according to the Acoustic Control Systems (ACS) concept. Additionally, it is capable of generating wave fields according to the Wave Field Synthesis (WFS) concept. The design concepts and criteria, the technical specifications, and some first applications of the module will be explained and discussed.
Convention Paper 7054

14:00

P11-4 Artificial Reverberator with Location Control in Multichannel Recording—*Hwan Shim, Jeong-Hun Seo, Koeng-Mo Sung*, Seoul National University, Seoul, Korea

In this paper a novel artificial reverberator is proposed. Compared with conventional algorithms focused to append appropriate timber and reverberance, the proposed algorithm is designed to produce realistic reverberation by controlling each location of sound sources. The new algorithm proposed in this paper controls perceived direction by panning the direct sound and controls perceived distance by adjusting the energy decay curve of reverberation, which is obtained by a location-clustering method and gain of the direct sound. In addition, the algorithm enhances Listener Envelopment (LEV) to make late reverberation incoherent among channels.
Convention Paper 7055

14:30

P11-5 Spatial Audio Rendering Using Sparse and Distributed Arrays—*Aki Härmä, Steven van de Par, Werner de Bruijn*, Philips Research Europe, Eindhoven, The Netherlands

A widely distributed but multichannel audio reproduction system can be used to create dynamic spatial effects for various entertainment and communication applications. In this paper we focus on the follow-me audio effect where the sound source appears moving with the observer who is walking through a hallway or going from one room to another in the home environment. We give an overview of the array theory for the sparse distributed loudspeaker systems, study the binaural properties of the sound field rendered with a sparse line array, and compare two different dynamic rendering techniques in a new type of a listening test.
Convention Paper 7056

15:00

P11-6 Magic Arrays—Multichannel Microphone Array Design Applied to Multiformat Compatibility— *Michael Williams*, Sounds of Scotland, Paris, France

This paper describes the principles and design procedure of multiformat-compatible microphone arrays for a range of different segment coverage angles and for omnidirectional, hypocardioid, cardioid, and supercardioid microphones. At present the only practical solution available for the main microphone array for a multiple format recording is to use different microphone arrays for each of the required formats. This paper shows how this jungle of main microphone arrays can be replaced by a single 5-channel microphone array that will supply signals that are directly compatible with five standard formats: mono, 2- and 3-channel “stereo,” 4-channel “quadraphony,” and “multichannel” with the full five channels. The specific reproduction format can be chosen either during the production process as a function of the desired support media, or by the consumer from a multichannel media product according to their own particular listening configuration.
Convention Paper 7057

Workshop 8 — Sunday, May 6
12:30 – 14:30 — Room P

ADVANCED TECHNIQUES FOR BUILDING INTERACTIVE ENVIRONMENTS

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

Panelists: *Stefan Bilbao*, University of Edinburgh, Edinburgh, Scotland, UK
Renato Pellegrini, sonic emotion, Oberglatt (Zurich), Switzerland
Ville Pulkki, Helsinki University of Technology, Helsinki, Finland
Lauri Savioja, Helsinki University of Technology, Helsinki, Finland

Interactive environments are defined by the unpredictable actions of their users and by the real-time processing required to generate them. Until recently, many audio processing techniques that could be used in interactive environments require computational or bandwidth resources that prohibit their implementation in real time. Such techniques include methods to create audio content, to model the environment, to communicate audio data between remote participants, and to render the spatial audio output. The workshop will introduce state-of-the-art techniques for each of these four areas and discuss how suitable each is to particular applications in terms of practicality as well as computational and financial cost. The example techniques are physical modeling, interactive virtual reality environment modeling, wave field synthesis, and directional audio coding for each of the four areas respectively.

Live Sound Seminar 3 — Sunday, May 6
12:30 – 14:00 — Room E

d & b

Presenters: **Holger Blum**
Ralf Zuleeg

This session is comprised of two parts. Part one covers the dimensioning of line arrays and answers the question “Is more really more?“

Part two covers the advantage of digital signal wiring and how to achieve optimal sonic quality and high reliability.

Exhibitor Seminar — Sunday, May 6
12:30 — Room 357

MICROTECH GEFELL

Presenter: **Ulrich Apel**

“Surround Microphones Not Only for Concert Halls”

Microtech-Gefell presents microphones with special rainproof windscreens that are applicable not only in closed premises but for recordings of atmos on locations; even for film and television. Herewith Microtech-Gefell translated their outdoor-and environmental experiences into their studio-microphones. An application of this wind- and rainshield and an operation of new switchable microphone UM930 will be shown.

Session P12 Sunday, May 6 13:00 –14:30
Foyer IK

POSTERS: MICROPHONES AND LOUDSPEAKERS AND AUDIO IN COMPUTERS (GAMES, INTERNET, DESKTOP COMPUTER AUDIO)

13:00

P12-1 The Relation between Active Radiating Factor and Pressure Responses of Loudspeaker Line Arrays—*Yong Shen, Dayi Ou, Kang An*, Nanjing University, Nanjing, China

Active Radiating Factor (ARF) is an important parameter to analyze the loudspeaker line array when considering the gaps between each of two radiating transducers. The relation between ARF of the loudspeaker line array and the differential chart of its pressure responses in two distances (PRD) is analyzed. Some valuable conclusions about ARF and PRD are found. A method to estimate ARF by measuring pressure responses comes out.
Convention Paper 7058
[Paper was not presented, but Convention Paper 7058 is available for purchase.]

13:00

P12-2 Alternative Encoding Techniques for Digital Loudspeaker Arrays—*Fotios Kontomichos, Nicolas–Alexander Tatlas, John Mourjopoulos*, University of Patras, Patras, Greece

Recent developments in digital loudspeakers have resulted in the introduction of digital transducer arrays (DTA). In most implementations, DTA loudspeakers are driven by PCM encoded audio signals, usually resampled and requantized to an appropriate number of bits, in accordance to the number of the transducers constituting the DTA topology. However, given that DTAs generally increase harmonic distortion, especially for off-axis listening positions, optimization in signal encoding and bit-to-transducer assignment, is necessary. Here, a number of novel, alternative strategies are examined, concerning the input signal encoding via PCM-to-PWM conversion, as well as techniques for bit-assignment on the transducers of a DTA. These tests are supported by simulation results and comparisons, for different operating parameters.
Convention Paper 7059

13:00

P12-3 Online Identification of Linear Loudspeaker Parameters—*Bo Rohde Pedersen*,[1] *Per Rubak*[2]
[1]Aalborg University, Esbjerg, Denmark
[2]Aalborg University, Aalborg, Denmark

Feed forward nonlinear error correction of loudspeakers can improve sound quality. For creating an efficient feed forward strategy identification of the loudspeaker parameters is needed. The strategy of the compensator is that the nonlinear behavior of the loudspeakers has relatively small drift and only the linear loudspeaker parameters must be identified. In music systems this can be done with online transducer-less system identification using the voice coil current as feedback from the loudspeaker (plant). This is investigated in a simulation study for finding useful system identification algorithms. Two different identification techniques (ARMA and FIR) are compared. The stability of the nonlinearities is tested in a measurement series.
Convention Paper 7060

13:00

P12-4 Finite Element Analysis of Near Field Beam Forming in Safety Relevant Work Spaces—*Roman Beigelbeck*,[1] *Heinrich Pichler*[2]
[1]Austrian Academy of Sciences, Wiener Neustadt, Austria
[2]Consultant, Vienna, Austria

Due to their unique features, loudspeaker arrays are an interesting alternative to standard loudspeaker setups or headphone-based solutions in safety relevant workspaces such as air traffic control rooms. Consequentially, near field beam forming in small spaces plays an important role for this field of application. In this paper the sound design based on a set of loudspeaker arrays featuring their interaction with a typical air traffic control room infrastructure is investigated by means of finite element modeling. Guided by these results, optimized array parameters can be determined. Representative three-dimensional near field directional diagrams in front of the arrays are shown to visualize the sound field in different cases. Finally, these theoretical values are compared with practical results.
Convention Paper 7061

13:00

P12-5 Creating Directed Microphones from Undirected Microphones—*Emil Milanov, Elena Milanova*, Acoustical Engineers, Sofia, Bulgaria

In this paper we examine the possibility of creating directed microphones from undirected microphones. The result is achieved only by using acoustical elements and is valid for all types of microphones, regardless of their way of work (electro-dynamical, condenser, optical, electromechanical, etc.). By using alternations of the membrane usage, a force is obtained, which is equivalent to the effect of simultaneously operating undirected and bidirected (eight) microphones. The result is a microphone with a space characteristic equivalent to the Pascal curve (i.e., directed microphone with the traditional shapes of the space characteristic—cardioid, super cardioid and hyper cardioid). The shape of the space characteristic curve is near to theoretical and does not depend on the acoustic elements of the microphone. The microphone is directed, but does not have a proximity effect.
Convention Paper 7062

13:00

P12-6 Transducer with the Direct D/A Conversion Using the Optoacoustic Principle—*Libor Husník*, Czech Technical University in Prague, Prague, Czech Republic

Transducers with the direct D/A conversion,

sometimes called digital transducers, either loudspeakers or earphones, are searching their ways into being. There have been several attempts to design such devices, but none of them left research laboratories and made its way to commercial use as yet. Most of them use "classical" electroacoustic transduction principles, i.e., electrodynamic or electrostatic. In this paper the possibility to use optoacoustic transduction principles is explored. First, the principles of physical phenomena used in this transducer are revised. Then, some construction details in light of their usage in the digital earphone are described.
Convention Paper 7064

13:00

P12-7 Demystifying the Measurement of Impulse Response in Condenser Microphones—Part I—*Christian Langen*, Schoeps Mikrofone, Karlsruhe, Germany

Good impulse response is an important reason for preferring condenser microphones in audio applications that require high quality. However, it is difficult to characterize the impulse response of a microphone precisely. We cannot create an acoustic impulse that approximates the Dirac delta function closely enough that a microphone will emit only its own impulse response. Electrical spark discharges, pistol shots, and pressure-step methods all approximate the Dirac distribution, but due to their limitations one must still deconvolve the impulse responses of the excitation signal and that of the microphone itself. Since every known method for performing such deconvolution has further pitfalls of its own, a novel time-domain method of deconvolution is introduced.
Convention Paper 7065

13:00

P12-8 Toward Multimodal Interfaces for Intrusion Detection—*Miguel Garcia-Ruiz*,[1] *Miguel Vargas Martin*,[2] *Bill Kapralos*[2]
[1]University of Colima, Colima, Mexico
[2]University of Ontario Institute of Technology, Oshawa, Ontario, Canada

Network intrusion detection has generally been dealt with using sophisticated software and statistical analysis tools. However, occasionally network intrusion detection must be performed manually by administrators, either by detecting the intruders in real-time or by revising network logs, making this a tedious and time consuming labor. To support this, intrusion detection analysis has been carried out using visual, auditory or tactile sensory information in computer interfaces. However, little is known about how to best integrate the sensory channels for analyzing intrusion detection. We propose a multimodal human-computer interface to analyze malicious attacks during forensic examination of network logs. We describe a sonification prototype that generates different sounds according to a number of well-known network attacks.
Convention Paper 7066

13:00

P12-10 Steganographic Approach to Copyright Protection of Audio—*Suthikshn Kumar*, PES Institute of Technology, Bangalore, India

Steganography is the technique of hiding data in images and music. It is one of the powerful mechanisms by which useful copyright information is hidden in the audio. In this paper we propose the use of steganography and public key cryptography to store the copyright information and authenticate the original audio. A tool called Steger is being developed that automatically determines the original copyright holders of the audio content. This tool is useful in Digital Rights Management (DRM) enabling end-user systems such as PDAs, mobile phones, PCs, handheld devices, consumer electronics, etc.
Convention Paper 7067

Workshop 9 — **Sunday, May 6**
13:00 – 15:30 — **Room H**

THE HOWS AND WHYS OF SIGMA DELTA CONVERTERS

Chair: **Joshua Reiss**, Queen Mary, University of London, London, UK

Panelists: *Jamie Angus*, University of Salford, Salford, Greater Manchester, Salford, UK
Lars Risbo, Texas Instruments, Denmark, Lyngby, Denmark
Brian Trotter, Cirrus Logic, Inc., Austin, TX, USA

Sigma delta modulation is the most popular form of analog-to-digital conversion used in audio applications. They are also commonly used in D/A converters, sample rate converters, and digital power amplifiers. This workshop will discuss methods of operation, design, and use of sigma delta modulators. The theory behind their operation will be introduced and explained. We will discuss the issues with their use and how they can be resolved. We'll explain how their performance is assessed and how best to navigate through the specifications on a sigma delta modulator's data sheet. Finally, practical examples will be given to illustrate the concepts presented.

Tutorial 6 — **Sunday, May 6**
13:00 – 15:30 — **Room G**

BROADCAST CASE STUDIES

Presenters: **Dennis Baxter**, Audio for the Olympics
Akira Fukada, NHK tokyo
Gaute Nistov, NRK Oslo

Three experts will relate stories of their individual experiences in broadcasting.

Dennis Baxter (at 14:00) will tell of his experience broadcasting the Olympics. The Olympics uses production teams from all over the world including the host country, which is currently China. Most of these production teams are considered to be the best at their particular sports coverage. For example YLE (Finland) and NRK (Norway) produce Cross Country Skiing and Winter Biathlon, the BBC has produced Tennis and New Zealand covers sailing. The host country, 2004–Greece, 2006–Italy, 2008– China, is favored by the Host Broadcaster to participate as much as possible but often these broadcaster lack the experience. One of the greatest challenges in the broadcast production of the Olympics is a consistency in production. With sound mixers from over 30 different countries they bring with them various levels of technical skills and personal experience that influence the way the sound is produced. Additionally sound mixing is very personal and subjective and not easily definable as to what is right or wrong. The following factors influence sound mixing: (1) Cultural interpretation of television production; (2) Psychological – Television has been dominated by video and engineers and technicians who do not understand audio. Often there has been a lack of resources and support, and sound engineers sometimes just give up! (3) Personal prejudices – Most North American sound mixers disapprove of live sound sweetening; (4) Ego; (5) Experience. The presentation will explore these areas and the subjectivity of sound production.

Gaute Nistov's topic (at 13:00) is location recording from the bottom of the North Sea to the Pyramids of Egypt. On the 2nd of October 2006 singer Katie Melua performed a concert over 300 meters below sea level inside one of the concrete shafts that anchors the Troll gas rig to the sea bed. In addition to being Europe's highest selling European female artist last year this gig in the North Sea also secured Katie a world record for the deepest underwater concert. The special acoustic properties of the shaft and the very strict security measures on the platform were among the challenges for this extraordinary production.

Only weeks later in Cairo a performance of Norwegian playwright Henrik Ibsen's "Peer Gynt" was staged in front of the Great Pyramids of Giza. The combined effort of more than 30 actors and singers, the Cairo Symphony Orchestra, and a 60-plus strong choir at the outdoor arena, posed a very different set of requirements for a sound production that had to accommodate both a live transmission locally on the night as well as recording for postproduction. Nistov was in charge of the TV-sound production on both occasions and will discuss the technical solutions used with an emphasis on production planning.

Akira Fukada (at 14:45) talks of his challenges broadcasting two concerts in Japan. Two special concerts will be presented having taken place in demanding places in Japan: One of them is the concert at which the "field of summer" was performed at the city center of Hiroshima. This is a concert which looks back upon the 60th anniversary of the atomic bombing. The piece was performed at the exact place where the atomic bomb of Hiroshima was dropped. For Japan, that is a holy place. Therefore, there are many regulations and the concert was held in the severe environment of hot summer. It was not only broadcast live in 5.1 surround sound, but offered simultaneously all over the world on the Internet. The second concert took place inside of a mountain. This mountain has a huge base rock and the sound performed there produces characteristic reverberation.

Composer Isao Tomita and Fukada planned the concert using the sound properties of this space. First, music was performed by allotting a player to various places of a mountain. And those sounds are projected to the base rock using PA. The reflective sound is recorded with the surround microphone installed in the space. The performance had a distinctive sound due to this mountain's effect. However, during this concert it, unfortunately, rained; nevertheless the sound of the rain made for an exceptional sound effect caught by the surround microphone.

TECHNICAL TOUR 5
ORF Radio Broadcasting Studios
Sunday, May 6, 13:00 – 16:00

ORF Radio is the National Public Broadcaster for Austria, with three nationwide, nine regional, and two MW/SW stations. The tour to the Radio Broadcasting House gives access to the music recording studios, continuity studios for three different radio stations, and insight into the practical 5.1-Surround sound operation of the satellite radio channel OE1DD.

Sunday, May 6 13:00 Room 633
Technical Committee Meeting on Audio Forensics

STUDENT EVENT
Design Competition
Sunday, May 6, 13:30 – 16:00 Room 631/632

The design competition is a competition for audio projects made by students at any university or recording school, which will challenge students with an opportunity to showcase their technical skills. This is not for recording projects or theoretical papers. Designs will be judged by a panel of industry experts in design and manufacturing. Multiple prizes will be awarded.

Judges: Christoph Musialik, Juha Backman, Dan Lavry

Exhibitor Seminar Sunday, May 6
13:30 Room 357

D.A.V.I.D.

Presenter: **Andreas Hildebrand**

"Regionalization in Distributed Radio Broadcast"

Today's daily challenges in distributed PlayOut are the driving considerations in workflow management for many broadcasters. For this, D.A.V.I.D. provides a distributed broadcast server concept, including background processes aligning media and scheduling content automatically and just in time, ready to be broadcasted when needed at multiple operational sites.

Sunday, May 6 13:30 Room 142/143
Standards Committee Meeting on SC-02-08 Audio-File Transfer and Exchange

Workshop 10 Sunday, May 6
14:00 – 15:30 Room 560/561

ACOUSTIC CONCEPTS AND TIME MANAGEMENT FOR SURROUND RECORDINGS—JAZZ

Chair: **Ulrich Vette**, University of Music and Performing Arts Vienna, Austria

In recordings of jazz combos and big bands it is a common concept to record instruments acoustically separated by screens or even in multiple rooms. In a live recording situation this is not possible in the same way, so often close microphone positioning is the solution. By using path length delay compensation a different recording approach is possible: The use of main and ambience microphones leads to a more realistic spatial image and allows more natural sound colors of instruments than a mix with only spot microphones.

This acoustical concept will be demonstrated with a big band concert recorded at the Porgy and Bess Jazz Club in Vienna and a studio recording of a Jazz Quartet.

Sunday, May 6 14:00 Room 633
Technical Committee Meeting on Transmission and Broadcasting

Live Sound Seminar 4 Sunday, May 6
14:30 – 16:00 Room E

YAMAHA

Presenter: **Ruben van der Goor**

This presentation will give an overview, explanation, and the pros and cons of existing networks. Integration into Yamaha-based applications will also be demonstrated.

Exhibitor Seminar Sunday, May 6
14:30 Room 357

CUBE-TEC INTERNATIONAL

Presenter: **Stefan Zachau**

"Quadriga: The Archive Solution"

Cube-Tec presents the latest version of QUADRIGA, the world standard in automated quality controlled audio archival systems. QUADRIGA workstations are capable of simultaneously capturing multiple mono, stereo, or multichannel sources. The presentation shows the ingest process, Cube-Workflow/DOBBIN integration, and explains the new RF64/MBWF capabilities.

Workshop 11 Sunday, May 6
15:00 – 18:00 Room P

MUSIC 2.0: MUSIC AND THE SEMANTIC WEB—CONSUMING AND PRODUCING MUSIC IN THE 21ST CENTURY

Chair: **Mark Sandler**, Queen Mary, University of London, London, UK

Panelists: *Oscar Celma*, Universitat Pompeu Fabra
Lucas Gonze, Yahoo! Music
Yves Raimond, Queen Mary, University of London

Music and audio have evolved to encompass the Internet and all it offers for purchase and consumption of music; this is often called digital music. But the Internet is evolving, too. With Web 2.0 products already on offer, how will this affect digital music? Come to think of it, what exactly is Web 2.0?

This workshop will provide some answers to these questions. Topics to be covered by leading researchers include: Semantic Web and Web 2.0; Semantic Audio basics; current Music 2.0 projects; future Music 2.0 projects.

Exhibitor Seminar Sunday, May 6
15:00 Room 411

GEORG NEUMANN GMBH

Presenter: **Stephan Peus**

"Digital Microphones—Purpose, Meaning, and Aspects"

This seminar presents: the advantages of a fully digital signal chain; digital microphone technology in today's digital studio environment; the AES 42 Interface: The "chicken or egg" question; status quo concerning the implementation of the AES 42 standard; synchronization—the basic for any digital recordings; gain setting with digital microphones; how is the workflow in the recording changed by using digital microphones?; several examples of application.

Sunday, May 6 15:00 Room 633
Technical Committee Meeting on Loudspeakers and Headphones

Exhibitor Seminar Sunday, May 6
15:30 Room 357

DPA MICROPHONES

Presenters: **Eddy Bøgh Brixen, Mikkel Nymand**

"DPA Decca Tree & Surround Mount"

The DPA Decca Tree (D3)/Surround Mount (S5) is a highly versatile and stylish microphone mount for 2 to 5+ microphones. The unique building block design with an integrated cable option provides extreme flexibility allowing the possibility of numerous configurations. At this seminar the system is demonstrated by experienced audio engineers.

Sunday, May 6 15:30 Room 142/143
Standards Committee Meeting on SC-03-02 Transfer Technologies

Workshop 12 Sunday, May 6
16:00 – 17:30 Room 560/561

ACOUSTIC CONCEPTS AND TIME MANAGEMENT FOR SURROUND RECORDINGS—JAZZ

Chair: **Ulrich Vette**, University of Music and Performing Arts Vienna,

In recordings of jazz combos and big bands it is a common concept to record instruments acoustically separated by screens or even in multiple rooms. In a live recording situation this is not possible in the same way, so often close microphone positioning is the solution. By using path length delay compensation a different recording approach is possible: The use of main and ambience microphones leads to a more realistic spatial image and allows more natural sound colors of instruments than a mix with only spot microphones.

This acoustical concept will be demonstrated with a big band concert recorded at the Porgy and Bess Jazz Club in Vienna and a studio recording of a jazz quartet.

Tutorial 7 Sunday, May 6
16:00 – 18:00 Room H

CHAMPAGNE LIFESTYLE—BEER MONEY

Presenters: **Simon Bishop**, Freelance Sound Recordist, UK
Richard Merrick, Freelance Sound Recordist, UK

➥

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

STUDENT EVENT
Career/Job Fair
Sunday, May 6, 16:00 – 18:00 Foyer D

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

Sunday, May 6 16:00 Room 633
Technical Committee Meeting on Audio for Games

Session P13 Sunday, May 6 16:30 – 18:00
Foyer IK

POSTERS: MULTICHANNEL SOUND

16:30

P13-1 Headphones Technology for Surround Sound Monitoring—A Virtual 5.1 Listening Room—*Renato Pellegrini,*[1] *Clemens Kuhn,*[1] *Mario Gebhardt*[2]
[1]sonic emotion ag, Obergltt (Zurich), Switzerland
[2]Beyerdynamic GmbH & Co. KG, Heilbronn, Germany

This paper presents a headphone technology for professional surround monitoring with virtual 5.1 reproduction. Using perceptually motivated binaural signal processing and ultra sonic head tracking, this system enables the simulation of a loudspeaker set-up with correct localization and room impression. As a professional recording and mixing tool it provides the advantages of a portable headphone solution but avoids the known drawbacks such as inside-head localization, limited room perception, and turning of the sonic image with the listener's head. The combination of three technologies—binaural reproduction, room simulation, and head tracking—enables the reproduction of a virtual reference listening room for applications in studios, recording trucks, and mobile recording set-ups.
Convention Paper 7068

16:30

P13-2 Hybrid Sound Field Processing for Wave Field Synthesis System—*Hyunjoo Chung,*[1] *Hwan Shim,*[1] *JunSeok Lim,*[2] *Jae Hyoun Yoo,*[3] *Koeng-Mo Sung*[1]
[1]Seoul National University, Seoul, Korea
[2]Sejong University, Seoul, Korea
[3]Electronics and Telecommunications Research Institute (ETRI), Yusung-gu Daejeon, Korea

Using the wave field synthesis (WFS) method, the sound of a primary source was reproduced by plane waves. Although having some shortcomings, such as spatial aliasing, these plane waves enlarged the sweet spot of the listening area and decreased the localization error of the sound source. Also, we suggested a grouped reflections algorithm (GRA) for reproducing early reflections. This sequence of early reflections increased the spaciousness of the listening room environment. The result, obtained by applying this method, was implemented by linear arrays of 32 loudspeakers constructed in an anechoic room. For backward compatibility with standard five-channel surround titles, a new hybrid sound field processing algorithm using WFS and GRA method was implemented.
Convention Paper 7069

16:30

P13-3 Reproduction of Virtual Reality with Multichannel Microphone Techniques—*Timo Hiekkanen, Tero Lempiäinen, Martti Mattila, Ville Veijanen, Ville Pulkki*, Helsinki University of Technology, Espoo, Finland

The perceptual differences between virtual reality and its reproduction with different simulated multichannel microphone techniques were measured using listening tests. The virtual reality was generated using the image-source method and 16 loudspeakers in a 3-D arrangement in an anechoic chamber. Two spaced and two coincident microphone techniques were tested, namely Fukada tree, Decca tree, 1st order Ambisonics, and 2nd order Ambisonics. The spaced techniques utilized the 5.0 setup, and Ambisonics techniques utilized the quadraphonic setup. The perceptual difference was measured with ITU impairment scale.
Convention Paper 7070

Exhibitor Seminar Sunday, May 6
16:30 Room 357

APT

Presenter: **Jon McClintock**

"Surround Sound Broadcast Projects in Europe"

Sunday, May 6 17:00 Room 633
Technical Committee Meeting on Signal Processing

Sunday, May 6 17:00 Room 142/143
Standards Committee Meeting on SC-03-04 Storage and Handling of Media

SPECIAL EVENT
Open House of the Technical Council and the Richard C. Heyser Memorial Lecture
Sunday, May 6, 18:15 – 19:15 Room K

Lecturer: **Gerhard Steinke**

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 122nd AES Convention is Gerhard Steinke. He began his career at Radio Dresden as a sound engineer in 1947. In 1953 he moved to Berlin's Radio and Television Research Centre (RFZ), where he established a laboratory for acoustical-musical boundary problems in broadcasting. In 1956, Steinke set up the first subjective listening test group to assess sound recordings, studios, and impairments in the broadcasting chain. This concept and the associated findings are included in various international standards (OIRT, ITU-R, and EBU) and documents (SSF, AES) on listening tests and test rooms. He was also responsible for the introduction of stereophonic broadcasting in East Germany and established an experimental electronic music studio with the new Subharchord synthesizer in 1962. In 1971 he became the director of the Research and Development Department of Sound and Video System Technology of RFZ. Together with co-inventors he developed the "Delta Stereophony" sound reinforcement system and a home processor for multichannel sound. He moved to Deutsche Telekom in 1990 where he set up the research and development group for new sound transmission systems.

Gerhard Steinke lectured sound technology and electronic music at Berlin's University of Music in the Tonmeister discipline for 27 years. Since his retirement he published further numerous papers and lectures, and contributed documents to the Surround Sound Forum of the Tonmeister Society (VDT) and to the AES.

For his work in the field of standards he received the Honorary Golden Medal of the OIRT and was awarded the Bèkesy Medal for his contributions to audio by the Hungarian Acoustical Society.

Steinke is a life member and fellow of the AES and served as vice president, Europe Region of the AES from 1991 – 1993, where he initiated the inauguration of new AES sections in the Eastern European countries. He is also member of the VDT.

The title of his lecture is, "What Is Needed to Have the Audio-Eldorado at Home?"

Considering the various legitimate demands for high audio quality by the recipient at home, especially of the music enthusiasts, his presentation will review to what degree the requirements in standards for the whole transmission chain can be observed—from microphone/studio up the individual listening conditions at home—and which important problems are disturbing the compliance. On one side, economical constraints cannot allow that highest production quality can be brought to the listener via the networks in each case. On the other side, it can be shown that a lot of necessary improvements could be realized within the transmission process for the reconstruction of the original perception that a listener has had of the original sound source and sound aesthetic in the sense of the best possible approximation.

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

SPECIAL EVENT
Banquet
Sunday, May 6, 20:00 – 23:30
Wappensaal at the Wiener Rathaus (Vienna Town Hall)

This year's banquet will be held in the Wappensaal at the Wiener Rathaus (Vienna Town Hall), which is a beautiful room in a very famous landmark building. Fine food and wine will be complemented by discreet music. As always, one of the main pleasures of attending the banquet is to meet old friends and colleagues and make new ones.

STUDENT EVENT
Student Party
Sunday, May 6, 20:00 –22:000
Club Ost
Schwindgasse 1, 1040 Wien

This year the student's traditional get-together takes place in one of Vienna's renowned music locations—Club Ost. The Band "Vienna Funk Agency" performs funk and disco classics and for those who still need more there will be DJ's heating up the party. But don't miss the beginning of the Recording Competition Surround the next morning :).

Of course everybody is welcome to this party—you don't have to be a student to party like a student.

TECHNICAL TOUR 6
Wiener Musikverein—The Golden Hall and the 4 New Halls
Monday, May 7, 08:30 – 11:00

Creating the four new halls in Musikverein Vienna has been the largest building project since the famous "Golden Hal" was established in 1870. The original plan was the creation of four new rehearsal halls, since Musikverein had already met capacity problems. The excavation for a new Vienna Metro line next to the Musikverein Building triggered the construction of the project, reaching down to as much as 16 meters below ground. In 2000, the architects Holzbauer and Irresberger upgraded the project by "materializing" each hall: The Glass Hall (Magna Auditorium), the Metal Hall, The Stone Hall, and The Wooden Hall. This materialization provided a high representative character as well as a big challenge for the acoustical consultant. The opening of The Four New Halls took place in March 2004. The technical tour is guided by K. Bernd Quiring who designed the acoustics of these halls. Acoustical measures are presented and a life chamber music example will be presented in all halls in order to provide an opportunity for personal judgments of the participants.

Session P14 Monday, May 7 09:00 – 13:00
Room I

MICROPHONES AND LOUDSPEAKERS, PART 1

Chair: **Thomas Gmeiner**, AKG Acoustics GmbH, Vienna, Austria

09:00

P14-1 Refinements of Transmission Line Loudspeaker Models—*Juha Backman*, Nokia Corporation, Espoo, Finland

Simple waveguide models of loudspeaker enclosures describe well enclosures with simple interior geometry, but their accuracy is limited if used with more complex internal structures. A refinement of a transmission line loudspeaker model is discussed, presenting one-dimensional waveguide approximations for bends and corners. Bends and corners are represented as area changes in the line, approximated as one-dimensional line segments with parameters adjusted ➥

to match the exact solutions for sharp (rectangular) corners in a waveguide. Besides modeling, the paper discusses the sound transmission characteristics of commonly used bend types and the applicability of the results to folded horns.
Convention Paper 7071

09:30

P14-2 The Use of Negative Source Impedance with Moving Coil Loudspeaker Drive Units: An Analysis and Review—*Michael Turner*,[1,2] *David Wilson*[1]
[1]University of Leeds, W. Yorkshire, UK
[2]Switched Reluctance Drives Ltd., Harrogate, N. Yorkshire, UK

The effect of negative source impedance on the frequency response and pole-zero pattern of a moving coil loudspeaker drive unit is explored from first principles, and closed-form expressions for the transfer function and system poles are developed. Direct control of motor velocity via the substantial cancellation of voice coil impedance is discussed. Implementation using positive current feedback is analyzed, considering loop gain, damping, and stability from a control theory perspective. Pole placement techniques are shown to be effective in controlling theoretical system behavior at high frequencies. Modeled and measured results are presented. A selection of previous papers and applications concerned with operation of loudspeakers from negative source impedances is briefly reviewed. Practical issues and some possible applications are discussed.
Convention Paper 7072

10:00

P14-3 Effects of Acoustical Damping on Current-Driven Loudspeakers—*Rosalfonso Bortoni*,[1] *Sidnei Noceti Filho*,[2] *Homero Sette Silva*[3]
[1]THAT Corporation, Milford, MA, USA
[2]UFSC - Federal University of Santa Catarina, Santa Catarina, Brazil
[3]Selenium Loudspeakers, Nova Santa Rita, Brazil

Previous works show the benefits of exciting loudspeakers with current sources, but they do not present a study showing the behavior of this technique when acoustical damping is applied to the diverse types of loudspeakers. This paper presents theoretical and practical analysis of the frequency response of acoustically damped current-driven loudspeakers installed in closed box, vented box, and 4th and 6th order band-pass systems. Also, it presents a subjective analysis comparing a closed box system excited by voltage and current sources.
Convention Paper 7073

10:30

P14-4 Development of a Highly Directive Endfire Loudspeaker Array—*Marinus Boone*,[1] *Wan-Ho Cho,*[2] *Jeong-Guon Ih*[2]
[1]Delft University of Technology, Delft, The Netherlands
[2]Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea

Control of the directivity of loudspeaker systems is important in applications of sound reproduction with public address systems. The use of loudspeaker arrays shows great advantages to bundle the sound in specific directions. Usually the loudspeakers are placed on a vertical line and the directivity is mainly in a plane perpendicular to that line although the radiation direction can be adapted with filter techniques, called beamforming. In this paper we present results on the applicability of a loudspeaker line array where the main directivity is in the direction of that line, using so-called endfire beamforming, resulting in a "spotlight" of sound in a preferred direction. Optimized beamforming techniques were used, which were developed for the reciprocal problem of directional microphone arrays. Effects of the design parameters of the loudspeaker array system were investigated, and we found that the stability factor can be a useful parameter to control the directional characteristics. A prototype constant beamwidth array system was tested by simulation, and measurement and the results supported our findings.
Convention Paper 7074

11:00

P14-5 Mass Nonlinearity and Intrinsic Friction of the Loudspeaker Membrane—*Ivan Djurek*,[1] *Antonio Petosic*,[2] *Danijel Djurek*[2]
[1]University of Zagreb, Zagreb, Croatia
[2]AVAC – Alessandro Volta Applied Ceramics, Zagreb, Croatia

Vibration of the loudspeaker's membrane was analyzed in the regime of comparatively low driving currents (I_0 < 100 mA) in terms of mass nonlinearity M_{eff} and intrinsic friction R_M. The latter contributes to the damping term of the differential equation of motion and depends on the elongation of vibration. R_M is the sum of intrinsic friction R_i of the membrane and friction R_v coming from air viscosity on its surface. Independent measurements of flexural strength of the membrane were performed and correlated to experimental observations of the vibrating system. Experiments were also performed with membranes additionally reinforced by application of materials with higher Young modules.
Convention Paper 7075

11:30

P14-6 Modeling of an Electrodynamic Loudspeaker Using Runge-Kutta ODE Solver—*Antonio Petosic,*[1] *Ivan Djurek*,[1] *Danijel Djurek*[2]
[1]University of Zagreb, Zagreb, Croatia
[2]AVAC – Alessandro Volta Applied Ceramics, Zagreb, Croatia

The modeling of low frequency (<100Hz) electrodynamic loudspeaker is presented as one degree of freedom nonlinear damped oscillator described by an ordinary differential equation of motion. The model has been compared to an equivalent LRC circuit model, and it was shown that differential the equation approach is more suitable for calculations that include nonlinearities occurring in an electrodynamic loudspeaker, as well as couplings of different vibration modes, particularly those coming from vibrating air and

the loudspeaker itself. The nonlinear differential equation of periodically driven anharmonic oscillator was solved numerically, and calculated amplitude frequency dependence and electric impedance have been compared to the experimental data. Calculations included different working regimes of the loudspeaker being operated in an evacuated space and air.
Convention Paper 7076

12:00

P14-7 Chaotic State in an Electrodynamic Loudspeaker—*Danijel Djurek,*[1] *Ivan Djurek,*[2] *Antonio Petosic*[2]
[1]AVAC-Alessandro Volta Applied Ceramics, Zagreb, Croatia
[2]University of Zagreb, Zagreb, Croatia

An electrodynamic loudspeaker has been operated in a nonlinear regime when the *k*-factor strongly increases with displacements. For driving AC currents up to 2 A the vibration spectrum contains high frequency harmonics of the classic von Kármán type; for currents in the range 2.6 to 4 A doubling of driving period appears; and for currents in the range from 4 to 4.8 A multiple sequences of subharmonic vibrations begin with 1/4*f* and 3/4*f*. An application of currents higher than 4.8 A results in a white noise spectrum, which is a characteristic of chaotic state.
Convention Paper 7077

12:30

P14-8 An Improved Electrical Equivalent Circuit Model for Dynamic Moving Coil Transducers—*Knud Thorborg,*[1] *Andrew Unruh,*[2] *Christopher J. Struck*[3]
[1]Tymphany A/S, Taasrup, Denmark
[2]Tymphany Corporation, Cupertino, CA, USA
[3]Independent Consultant, San Francisco, CA, USA

A series combination of inductor and resistor is traditionally used to model the blocked electrical impedance of a dynamic moving coil transducer, such as a loudspeaker driver. In practice, semi-inductive behavior due to eddy currents and "skin effect" in the pole structure as well as transformer coupling between the voice coil and pole piece can be observed, but are not well represented by this simple model. An improved model using only a few additional elements is introduced to overcome these limitations. This improved model is easily incorporated into existing equivalent circuit models. The development of the model is explained and its use is demonstrated. Examples yielding more accurate box response simulations are also shown.
Convention Paper 7063

Session P15 Monday, May 7 09:00 – 13:00
Room K

LOW BIT-RATE AUDIO CODING

Chair: **Karlheinz Brandenburg**, Technical University Ilmenau, Ilmenau, Germany

09:00

P15-1 A Biologically-Inspired Low-Bit-Rate Universal Audio Coder—*Ramin Pichevar, Hossein Najaf-Zadeh, Louis Thibault*, Communications Research Centre, Ottawa, Ontario, Canada

We propose a new biologically-inspired paradigm for universal audio coding based on neural spikes. Our proposed approach is based on the generation of sparse 2-D representations of audio signals, dubbed as spikegrams. The spikegrams are generated by projecting the signal onto a set of over-complete adaptive gammachirp (gammatones with additional tuning parameters) kernels. A masking model is applied to the spikegrams to remove inaudible spikes and to increase the coding efficiency. The paradigm proposed in this paper is a first step toward the implementation of a high-quality audio encoder by further processing acoustical events generated in the spikegrams. Upon necessary optimization and fine-tuning our coding system, operating at 1 bit/sample for sound sampled at 44.1 kHz, is expected to deliver high quality audio for broadcast applications and other applications such as archiving and audio recording.
Convention Paper 7078

09:30

P15-2 The Relationship Between Basic Audio Quality and Selected Artifacts in Perceptual Audio Codecs —Part II: Validation Experiment—*Paulo Marins, Francis Rumsey, Slawomir Zielinski*, University of Surrey, Guildford, Surrey, UK

A pilot study was conducted to investigate the perceptual importance of selected audio coding artifacts and their relationship with basic audio quality. An additional experiment was undertaken to validate the results obtained in the pilot calibration experiment. A listening test was designed that required a panel of expert subjects to evaluate the selected artifacts used in the initial study. In this second experiment, however, certain experimental parameters were modified; these included different levels of degradation and program material. The outcomes of the validation experiment are presented in this paper along with a detailed evaluation of the impact of the chosen experimental artifacts on basic audio quality assessments for perceptual audio codecs.
Convention Paper 7079

10:00

P15-3 New Enhancements to Immersive Sound Field Rendition (ISR) System—*Chandresh Dubey,*[1] *Raghuram Annadana,*[1] *Deepen Sinha,*[1] *Anibal Ferreira*[1,2]
[1]ATC Labs, Chatham, NJ, USA
[2]University of Porto, Porto, Portugal

Consumer audio applications such as satellite radio broadcasts, multichannel audio streaming, and playback systems coupled with the need to meet stringent bandwidth requirements are eliciting newer challenges in parametric multichan-

➥

nel audio coding schemes. This paper describes the continuation of our research concerning the Immersive Soundfield Rendition (ISR) system and the different enhancements in various algorithmic components. The need to maintain a constant bit rate for many applications requires a rate control mechanism. The various strategies utilized in the rate loop mechanism are presented. In addition, an innovative phase compensated downmixing scheme has been incorporated in the ISR system so as to generate a high quality carrier signal. Enhancements have been made to the blind up-mixing scheme and considerable gains have been made in terms of acoustic diversity. The various enhancements of the ISR system and its performance are detailed. Audio demonstrations are available at http://www.atc-labs.com/isr.
Convention Paper 7080

10:30

P15-4 Aspects of Scalable Audio Coding—*Chris Dunn*, Independent Consultant, London, UK

Banded weight data is transmitted as side information within coded audio bit streams in order to achieve psychoacoustically-appropriate shaping of quantization noise. Methods of reducing the information overhead corresponding to weight data are discussed in the context of scalable bitplane coding. Two approaches to coding band weight data are compared in terms of coding efficiency and error resilience. In the first approach, weights are coded as a block of data at the beginning of each frame, using a predictor and Golomb coding of weight prediction residuals to achieve high coding efficiency. This approach is compared to coding weights for bands as they become significant, with weight data distributed across each coded bit stream frame.
Convention Paper 7081

11:00

P15-5 Source-Controlled Variable Bit Rate Extension for the AMR-WB+ Audio Codec—*Amélie Marty, Roch Lefebvre*, Université de Sherbrooke, Sherbrooke, Quebec, Canada

This paper presents a source-controlled, variable bit rate extension to the AMR-WB+ standard audio codec. AMR-WB+ allows multirate operation and, in particular, rate switching at every frame. However, the standard does not support source-controlled rate determination since it does not include a signal classifier. The proposed extension includes a signal classifier and rate mapping function for each signal class. Classification is performed at a lower frame rate compared to AMR-WB+, with typically one classification decision every second. Significant rate savings can be achieved by encoding speech at lower rates than other signals such as music. Applications include audio broadcasting over packet networks and storage of multimedia signals with mixed signals in the audio track.
Convention Paper 7082

11:30

P15-6 Multiple Description Coding for Audio Transmission Using Conjugate Vector Quantization—*Mylene Kwong, Roch Lefebvre, Soumaya Cherkaoui*, Université de Sherbrooke, Sherbrooke, Quebec, Canada

This paper explores robustness issues for real-time audio transmission over perturbed networks where multiple paths can be considered. Conjugate vector quantization (CVQ), a form of multiple description coding, can improve the resilience to packet losses. This work presents a generalized CVQ structure, where K>2 different conjugate codebooks are trained to create the best resulting codebook. Experiments show that four-description CVQ performs very closely to unconstrained VQ in clear channel conditions, while providing significant improvements in lossy channels. We also present a fast search algorithm that allows trade-offs between computational complexity and memory storage at the encoder. This robust quantization scheme can encode sensitive information such as spectral coefficients in a speech coder or a perceptual audio coder.
Convention Paper 7083

12:00

P15-7 MPEG Surround—The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding—*Jürgen Herre*,[1] *Kristofer Kjörling*,[2] *Jeroen Breebaart*,[3] *Christof Faller*,[4] *Sascha Disch*,[1] *Heiko Purnhagen*,[2] *Jeroen Koppens*,[5] *Johannes Hilpert*,[1] *Jonas Rödén*,[2] *Werner Oomen*,[5] *Karsten Linzmeier*,[1] *Kok Seng Chong*[6]
[1]Fraunhofer IIS, Erlangen, Germany
[2]Coding Technologies, Stockholm, Sweden
[3]Philips Research, Eindhoven, The Netherlands
[4]Agere Systems, Allentown, PA, USA
[5]Philips Applied Technologies, Eindhoven, The Netherlands
[6]Panasonic Singapore Laboratories Pte. Ltd., Singappore

In 2004 the ISO/MPEG Audio standardization group started a new work item on efficient and backward compatible coding of high-quality multichannel sound using parametric coding techniques. Finalized in fall of 2006, the resulting MPEG Surround specification allows to carry surround sound at bit rates that have been commonly used for coding of mono or stereo sound. This paper summarizes the results of the standardization process by describing the underlying ideas and providing an overview of the MPEG Surround technology. The performance of the scheme is characterized by the results of the recent verification tests. These tests include several operation modes as they would be used in typical application scenarios to introduce multichannel audio into existing audio services.
Convention Paper 7084

12:30

P15-8 Adaptive Design of the Preprocessing Stage for Stereo Lossless Audio Compression—*Florin Ghido, Ioan Tabus*, Tampere University of Technology, Tampere, Finland

We propose a novel lossless audio compression

scheme, which combines stereo preprocessing with stereo prediction. We show that such a scheme provides improved asymmetrical compression at almost no complexity increase for decoder (compared with stereo prediction alone), or the same compression for lower decoder complexity. The stage of stereo prediction is preceded by a rotation-like channel transformation, which improves compression by requiring smaller inter-channel optimal prediction orders and by obtaining smaller magnitudes for prediction coefficients. On a corpus consisting of 84 audio files (in CD-A format), for the OptimFROG-AS (asymmetric) lossless audio compressor using stereo prediction with orders 8/4, we obtained, on an audio corpus (in CD-audio format) of size 51.6 GB, compression improvements up to 5.10 percent on average 0.23 percent.
Convention Paper 7085

Workshop 13 **Monday, May 7**
09:00 – 11:30 **Room H**

SURROUND RECORDING AND REPRODUCTION WITH HEIGHT

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

Panelists: *Arnaud Laborie*, Trinnov Audio, France
Jeff Levison, DTS, CA, USA
Ville Pulkki, Helsinki University, Finland
Wilfried Van Baelen, Galaxy Studio, The Netherlands
Helmut Wittek, Schoeps Microphones, Germany
Wieslaw Woszczyk, McGill University, Canada

Two-dimensional multichannel sound systems such as 5.1, 6.1, and 7.1 are widely applied to movie theaters, home audio, and broadcasting. While new video formats such as 4k and 8k for the digital cinema is currently discussed, next-generation multichannel audio systems for the movie theater such as 10.2 and 22.2 are also discussed. New features on next-generation multichannel audio include the sensation of height and elevation of sound sources. A three-dimensional audio system is not only discussed among the sound professionals for movie productions but also among the broadcasters and home audio developers. This workshop will review the latest proposals of three-dimensional audio systems and will discuss the surround sound recording and reproduction to provide the sensation of height.

Panelists will give separate demonstrations in another room following this workshop. Details will be announced.

Workshop 14 **Monday, May 7**
09:00 – 11:00 **Room P**

PROCESS MODELING FOR BROADCAST PROFESSIONALS

Cochairs: **Chris Chambers**, BBC
Karl Petermichl, ORF

Managing workflow processes is an increasingly challenging task due to the complexity of modern, multimedia-oriented program departments. Process modeling is a way to graphically draw and analyze workflows, as well as integrating business parameters from the "Balanced Score Card" model into those processes.

Process modeling with the method ADONIS (used at ORF) and modeling with BPMN and UML in various applications will be explained and demonstrated.

Workshop 15 **Monday, May 7**
09:00 – 10:30 **Room 560/561**

ACOUSTIC CONCEPTS AND TIME MANAGEMENT FOR SURROUND RECORDINGS—CLASSICAL

Chair: **Ulrich Vette**, University of Music and Performing Arts Vienna, Austria

Experiences with acoustical surround recordings for the ITU 5.1 loudspeaker setup often lead to the conclusion to combine different ambience microphones to reproduce ambient sound. In this workshop a concept for time relations between ambience microphones in different distances to the sound source will be presented and discussed.

Moreover there will be a comparison between surround reproduction using the ITU 5.1 loudspeaker setup and advanced ideas using a second pair of lateral or elevated rear speakers. Multitrack recordings will be presented from research about different surround microphone setups at the University of Music and performing Arts, Vienna and examples of concert and studio-recordings with more than one pair of microphones for surround reproduction.

STUDENT EVENT
Recording Competition—Surround
Monday, May 7, 09:00 – 12:00 Room G

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

Schedule:
09:00 – Classical Surround
10:00 – Non-Classical Surround
11:00 – Sound For Picture

Judges:
Classical Surround: Werner Dabringhaus, Josef Schütz, Jürg Jecklin
Non-Classical Surround: Ronald Prent, Akira Fukada, TBA
Picture: Florian Camerer, Bernhard Maisch, Wilfried van Baelen

We would like to acknowledge the generosity of our sponsors for the student recording competition: AKG, Audio-Technica, DTS, Magix, PMC, RME, Schoeps, Sennheiser.

Monday, May 7 **09:00** **Room 633**
Technical Committee Meeting on Human Factors in Audio Systems

Monday, May 7 09:00 Room 142/143
Standards Committee Meeting on SC-03-12 Forensic Audio

Session P16 Monday, May 7 09:30 – 11:00
Foyer IK

POSTERS: ROOM AND ARCHITECTURAL ACOUSTICS AND SOUND REINFORCEMENT

09:30

P16-1 Acoustic Treatment of the Regional Flight Control Center Hall in Zagreb, Croatia—*Marko Horvat, Hrvoje Domitrovic, Sanja Grubesa*, University of Zagreb, Zagreb, Croatia

Acoustic treatment has been realized in the hall of Regional Flight Control Center in Zagreb, Croatia, upon complaints made by the flight control operators working in the mentioned hall. The primary complaint was that the operators could hear each other too well across the hall, due to unwanted reflections, so the main task was to reduce those reflections. The emphasis was also made on reducing the reverberation time of the hall, proven to be too long for the size and intended purpose of the hall, thereby reducing the background noise level in the hall as well.
Convention Paper 7086

09:30

P16-2 Investigating Classroom Acoustics by Means of Advanced Reproduction Techniques—*Nicola Prodi*,[1] *Andrea Farnetani*,[1] *Yuliya Smyrnova*,[1,2] *Janina Fels*[3]
[1]University of Ferrara, Ferrara, Italy
[2]Polish Academy of Sciences, Poland
[3]RWTH Aachen University, Aachen, Germany

A research was undertaken to investigate the loss of Italian language word intelligibility in classrooms caused by low signal to noise ratio and too high reverberation. In the first part of the paper, impulse responses and background noises were measured in two primary schools using different mono, binaural, and B-format probes. A dummy head with child morphology was also used for the first time in this context. It was thus possible to compare the performance of a child head to the conventional adult one. Then the restitution of the recorded sound fields in a dedicated listening room was accomplished, using stereo dipole and ambisonics technologies.
Convention Paper 7087

09:30

P16-3 Perception of Concert Hall Acoustics in Seats Where the Reflected Energy Is Stronger than the Direct Energy—*David Griesinger*, Harman Specialty Group, Bedford, MA, USA

This paper describes a series of experiments into sound perception when the direct/reverberant ratio (d/r) is low. Sound source localization and the perception of being adequately close to the musicians are improved when the direct sound dominates the total reflected energy for about 40 ms, during which time the direct sound can be separately perceived. For such a hall the impressions of loudness, clarity, and localization are satisfactory and nearly unchanged over a 6 dB range of d/r. As the time period of direct sound dominance decreases, the d/r ratio must be higher for equal subjective clarity.
Convention Paper 7088

09:30

P16-4 Relation between Correlation Characteristics of Sound Field and Width of Listening Zone—*Elena Prokofieva*, Linn Products Ltd., Glasgow, Scotland, UK

The principal features of an even sound field are large directivity and diffusive nature of the radiated field. Directivity pattern of stereo loudspeakers was analyzed to determine the degree of coherence of the sound field in the room. Measurements of the sound field in close to a standard listening room conditions were conducted with the loudspeakers placed in ITU-recommended and specially selected positions. The results showed that the radius of correlation corresponds to the size of the sweet spot. Relocating loudspeakers in the room and taking into consideration the room environment influence can help to enlarge the listening zone within the room. These conclusions were confirmed by listening tests and recommendations on sound field correlation can be established.
Convention Paper 7089
[Paper not presented but Convention Paper 7089 is available for purchase.]

09:30

P16-5 On the Implementation of a Room Acoustics Modeling Software Using Finite-Differences Time-Domain Method—*José Lopez*,[1] *José Escolano*,[2] *Basilio Pueo*[3]
[1]Technical University of Valencia, Valencia, Spain
[2]Technical University of Jaen, Linares (Jaen), Spain
[3]University of Alicante, Alicante, Spain

The Finite-Difference Time-Domain (FDTD) approximation method has been introduced into acoustics in the last years to solve field problems numerically. However, the huge amount of computer power needed to be used in the modeling of large rooms has delayed the launch of commercial applications, being the major part based on ray-tracing. This paper analyzes the viability of a FDTD implementation for this task in today's personal computers and presents the resulting application. All simulation stages from the architectural model, the generation of the mesh, implementation of the recursion, parallelization, and, finally, the result in the form of impulse response are discussed.
Convention Paper 7090

TECHNICAL TOUR 7
Vienna Symphonic Library
Monday, May 7, 09:30 – 13:30

Have a sneak peek into the Silent Stage of the Vienna Symphonic Library, one of the biggest producers of virtual instruments world-wide. Here, recording musicians as well as sound technicians and musical directors find optimal conditions for capturing perfect samples—ultimate control over the recording conditions and the absolute silence required for sampling sessions—in a studio solely designed for that purpose. See a presentation of the Vienna Instruments, the most powerful virtual instruments currently on the market, tailor-made for the massive sample database of over 1.5 million samples recorded in the Silent Stage. Meet the creator of MIR, the Multi Impulse Mixing and Reverberation Engine which adds space to the alchemy of digital orchestral music production. Based upon the principle of multi impulse response convolution, the Vienna Symphonic Library digitizes ("samples") the characteristics of spaces like the Great Hall of the Wiener Konzerthaus with unprecedented detail and methodology, taking this technology to the next level and capturing the great concert stages of the world in their full glory—wall to wall, floor to ceiling, in three dimensions!

Monday, May 7 10:00 Room 633
Technical Committee Meeting on Hearing and Hearing Loss Prevention

Exhibitor Seminar Monday, May 7
10:30 Room 357

LEAR CORPORATION/DIRAC RESEARCH

Presenters: **Nilo Casimiro Ericsson, Mathias Johansson, Armin Prommersberger**

"Dirac Live™—A New Approach to Room Correction Algorithms"

A method for improving car audio systems through digital correction of the impulse/frequency responses of individual loudspeakers is presented. The ability to simultaneously optimize different time and frequency domain criteria and flexible spatial averaging are the main benefits of the new approach. The solution is featured in the new BMW M5 Individual High End Sound System, recently reviewed by Auto, Motor und Sport (number 8/2007 pp. 146–148) stating: "the most impressive car stereo sound in the world." An A/B comparison to conventional filtering is available in LEAR's demo car in our common exhibition booth.

Workshop 16 Monday, May 7
11:00 – 12:30 Room 560/561

ACOUSTIC CONCEPTS AND TIME MANAGEMENT FOR SURROUND RECORDINGS—CLASSICAL

Chair: **Ulrich Vette**, University of Music and Performing Arts Vienna, Austria

Experiences with acoustical surround recordings for the ITU 5.1 loudspeaker setup often lead to the conclusion to combine different ambience microphones to reproduce ambient sound. In this workshop a concept for time relations between ambience microphones in different distances to the sound source will be presented and discussed.

Moreover there will be a comparison between surround reproduction using the ITU 5.1 loudspeaker setup and advanced ideas using a second pair of lateral or elevated rear speakers. Multitrack recordings will be presented from research about different surround microphone setups at the University of Music and performing Arts, Vienna and examples of concert and studio-recordings with more than one pair of microphones for surround reproduction.

Workshop 17 Monday, May 7
11:00 – 13:00 Room 631/632

HANDS-ON DEMONSTRATION: SUBJECTIVE EVALUATION BY REFERENCE RECORDINGS WITH SURROUND MAIN MIKINGS FOR CLASSICAL ORCHESTRA

Presenters: **Hideo Irimajiri**, Mainichi Broadcasting Corp.
Toru Kamekawa, Tokyo National University of Fine Arts and Music
Masayuki Mimura, Yomiuri Telecasting Corp.
Hideaki Nishida, Asahi Broadcasting Corp.
Koichi Ono, Kansai Telecasting Corp.

There are many different setups for surround main microphones for classical music and orchestra. But it is very difficult to research and study them practically and academically under identical conditions and judge their performance. Consequently the AES Japan Surround Study Project has been organized and put into practice after one year of preparation. It was organized around 10 broadcasters and 2 universities; 12 manufacturers supported by HBF provided financial support. There were 15 different combinations of main and ambience microphone setups that were recorded on 96 channels independently in Pro Tools HD at 24 bit / 96-kHz. The musical examples were performed by the Osaka Philharmonic Orchestra on September 24-27, 2006.

In this workshop each individual setup will be played back. Participants will have the opportunity for feedback in a listening test environment, and the data will be collected for subjective evaluation.

Workshop 17 will be repeated four times throughout the convention: W-17, W-20, W-28, and W-29.

Exhibitor Seminar Monday, May 7
11:00 Room 441

SENNHEISER

Presenter: **Gregor Zielinsky**

"High Quality Surround Recording with the New MKH 8000 Microphone"

In this seminar Sennheiser presents the brand-new MKH 8000 Studio Microphone Series for the first time ever. On the examples of classical symphony and big band multitrack recordings, this "hands-on" seminar will show several surround microphone techniques using the MKH 8000. The different recordings will be explained and can be compared and judged by the listeners.

Monday, May 7 11:00 Room 633
Technical Committee Meeting on Perception and Subjective Evaluation of Audio

Monday, May 7 11:00 Room 142/143
Standards Committee Meeting on SC-03-06 Digital Library and Archive Systems

Session P17 Monday, May 7 11:30 – 13:00
Foyer IK

POSTERS: SPATIAL AUDIO PERCEPTION AND PROCESSING

11:30

P17-1 Sound Source Localization and B-Format Enhancement Using Sound Field Microphone Sets—*Charalampos Dimoulas, Kostantinos Avdelidis, George Kalliris, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The current paper focuses on the implementation of sound-field microphone arrays for sound source localization purposes and B-format enhancement. There are many applications where spatial audio information is very important, while reverberant sound-field and ambient noise deteriorate the recording conditions. As examples we may refer to sound recordings during movie production, virtual reality environments, teleconference and distance learning applications using 3-D audio capabilities. B-format components, provided from a single sound field microphone, are adequate to estimate sound source direction of arrival, while the combination of two sound field microphones allows estimating the exact source location. In addition, the eight (or more) available signal components can be used to apply delay and sum techniques, enabling SNR improvements and virtual positioning of a signal B-format microphone to any desired place. Simplicity, reduced computational load, and effectiveness are some of the advantages of the proposed methodology, which is evaluated via software simulations.
Convention Paper 7091

11:30

P17-2 Research on Widening the Virtual Listening Space in the Automotive Environment—*Jeong-Hun Seo, Lae-Hoon Kim, Hwan Shim, Koeng-Mo Sung*, Seoul National University, Seoul, Korea

This paper represents the research about a way to widen the virtual space in cars. Generally, the interior of cars contains small volume compared to normal listening environments. This makes listeners feel a little stuffy. Therefore, the way to widen a virtual space in cars is needed. One of the most important cues for spaciousness is the lateral reflections in accordance with room acoustics, so we will widen virtual space in cars using artificial lateral reflections in automotive environments.
Convention Paper 7092

11:30

P17-3 Perceptual Distortion Maps for Room Reverberation—*Thomas Zarouchas, John Mourjopoulos*, University of Patras, Patras, Greece

From reverberated audio signals and using as reference the input (anechoic) audio, a number of distortion maps are extracted indicating how room reverberation distorts in time-frequency scales, perceived features in the received signal. These maps are simplified to describe the monaural time-frequency/level distortions and the distortion of the spatial cues (i.e., interchannel cues and coherence), which are very important for sound localization in a reverberant environment. Such maps are studied here as functions of room parameters (size, acoustics, distance, etc.), as well as due to input signal properties. Overall perceptual distortion ratings are produced and reverberation-resilient signal features are extracted.
Convention Paper 7093

11:30

P17-4 A New Structure for Stereo Acoustic Echo Cancellation Based on Binaural Cue Coding —*Yoomi Hur, Young-Choel Park, Dae Hee Youn*, Yonsei University, Seoul, Korea

Most stereo teleconferencing systems involve an acoustic echo canceller to remove undesired echoes. However, it is difficult for the stereo echo cancellers to converge to the true echo path since the cross-correlation between the stereo input signals is high. To solve the problem, we propose a new structure that is combined with Binaural Cue Coding (BCC). BCC is a method for multichannel spatial rendering based on one down-mixed audio channel and side information. Based on the BCC, we propose a new single channel adaptive filter for the stereo echo cancellation, which takes the down-mixed monaural signal as the reference. Efficient voice conference systems can be implemented by using the proposed structure, since the BCC scheme to transmit stereo signals as mono signals with a number of side information, enables low-bit-rate transmission. Simulation results confirm that the convergence speed is increased and the misalignment problem is solved. In addition, the proposed structure has better tracking capability.
Convention Paper 7094

11:30

P17-5 Efficient Binaural Filtering in QMF Domain for BRIR—*David Virette,*[1] *Pierrick Philippe,*[2] *Gregory Pallone,*[1] *Rozenn Nicol,*[1] *Julien Faure,*[1] *Marc Emerit,*[2] *Alexandre Guerin*[1]
[1]France Telecom, Lannion, France
[2]France Telecom, Cesson-Sévigné, France

The MPEG Surround standard includes two "native" binaural processing modules for reproducing 3-D audio content over headphones. In this paper we present a novel and efficient binaural room impulse response (BRIR) modeling algorithm extending their possibilities. It is based on a parametric decomposition of the BRIR and is integrated within the subband domain implementation of the MPEG Surround binaural decoder. First we show that for impulse responses with room effects, our approach offers a significant reduction in terms of computational requirements compared to the native methods. Second, we report results from listening tests comparing different trade-offs between complexity and quality. We show that using our method,

the complexity can be reduced by a factor of two while preserving the optimum quality.
Convention Paper 7095

11:30

P17-6 A Parametric Model of Head-Related Transfer Functions for Sound Source Localization—*Youngtae Kim, Jungho Kim, Sangchul Ko*, Samsung Advanced Institute of Technology, Gyeonggi-do, Korea

A simple and effective parametric model of head-related transfer functions is presented for synthesizing binaural sound for practical 3-D sound reproduction systems. The suggested model is based on a simplified time-domain description of the physics of wave propagation and diffraction, and the components of the model have a one-to-one correspondence with the main characteristics of the measured head-related transfer functions such as sound diffraction, delay, and reflection. Their parameters are derived from some sets of the measurements, and thus enable the model to fit significant perceptual impact hidden in head-related transfer functions. Finally simple subjective listening tests verify the perceptual effectiveness of the model. This will show some promise of permitting efficient implementation in real-world applications.
Convention Paper 7096

11:30

P17-7 Binaural Response Synthesis from Center-of-Head Position Measurements for Stereo Applications—*Sunil Bharitkar, Chris Kyriakakis*, University of Southern California/Audyssey Labs., Los Angeles, CA, USA

In two-channel or stereo applications, such as for televisions, automotive infotainment, and hi-fi systems, the loudspeakers are typically placed substantially close to each other. The sound field generated from such a setup creates an image that is perceived as monophonic while lacking sufficient spatial "presence." Due to this limitation, a virtual sound technique may be utilized to widen the soundstage to give the perception to listener(s) that sound is originating from a wider angle using head-related-transfer functions (HRTFs). In this paper we present a method to synthesize responses at a listener's ears given simply two room-response measurements, where each measurement is obtained between a loudspeaker, in a stereo loudspeaker setup, and an assumed center-of-head position (where the listener is assumed to be seated). The binaural response synthesis approach uses head-shadowing models (for inter-aural intensity differences) and the Woodworth-Schlosberg delay model. This approach is useful when dummy heads are not readily available for HRTF measurements as well as to generalize the approach to reflect measurements that would have been obtained over a large corpus of data (viz., human subjects). We also compare the responses obtained from this approach with measurements done with a dummy-head.
Convention Paper 7097

11:30

P17-8 Physical and Filter Pinna Models Based on Anthropometry—*Patrick Satarzadeh, V. Ralph Algazi, Richard O. Duda*, University of California at Davis, Davis, CA, USA

This paper addresses the fundamental problem of relating the anthropometry of the pinna to the localization cues it creates. The HRTFs for isolated pinnae (which are called PRTFs) are analyzed and modeled for sound sources directly in front of the listener. It is shown that a low-order filter model, with parameters suggested by or derived from pinna anthropometry, provides a good fit to the data. Methods are reported for adjusting the model parameters to fit the PRTF data. It is often possible to estimate the model parameters from a few geometrical measurements of the pinna. However, direct estimation from pinna anthropometry in general remains an unsolved problem, and the nature of this problem is discussed.
Convention Paper 7098

11:30

P17-9 A Novel Approach to Robotic Monaural Sound Localization—*Fakheredine Keyrouz, Abdallah Bou Saleh, Klaus Diepold*, Technische Universität München, Munich, Germany

This paper presents a novel monaural 3-D sound localization technique that robustly estimates a sound source within a 2.5-degree azimuth deviation and a 5-degree elevation deviation. The proposed system, an upgrade of monaural-based localization techniques, uses two microphones, one inserted within the ear canal of a humanoid head equipped with an artificial ear, and the second held outside the ear, 5 cm away from the inner microphone. The outer microphone is small enough so that minimal reflections that might contribute to localization errors are introduced. The system exploits the spectral information of the signals from the two microphones in such a way that a simple correlation mechanism, using a generic set of Head Related Transfer Functions (HRTFs), is used to localize the sound sources. The low computational requirement provides a basis for robotic real-time applications. The technique was tested through extensive simulations of a noisy reverberant room and further through an experimental setup. Both results demonstrated the capability of the monaural system to localize, with high accuracy, sound sources in a three-dimensional environment even in presence of strong noise and distortion.
Convention Paper 7099

11:30

P17-10 Optimized Binaural Modeling for Immersive Audio Applications—*Christos Tsakostas,*[1] *Andreas Floros*[2]
[1]Holistiks Engineering Systems, Athens, Greece
[2]Ionian University, Corfu, Greece

Recent developments related to immersive audio systems mainly originate from binaural audio processing technology. In this paper a

➥

novel high-quality binaural modeling engine is presented suitable for supporting a wide range of applications in the area of virtual reality, mobile playback, and computer games. Based on a set of optimized algorithms for Head-Related Transfer Functions (HRTF) equalization, acoustic environment modeling, and cross-talk cancellation, it is shown that the proposed binaural engine can achieve significantly improved authenticity for 3-D audio representation in real-time. A complete binaural synthesis application is also presented that demonstrates the efficiency of the proposed algorithms.
Convention Paper 7100

11:30

P17-11 Head-Related Transfer Function Calculation Using Boundary Element Method—*Przemyslaw Plaskota, Andrzej B. Dobrucki*, Wroclaw University of Technology, Wroclaw, Poland

Measuring the head-related transfer function (HRTF) is an efficient method in taking the influence of human body on sound spectrum into consideration. The database used in reproduction of the sound source position is built using the measurement results. The base is individual for each human, which makes it impossible to make a versatile base for all listeners. In this paper a numerical model of an artificial head is presented. The model allows the determination of the value of HRTF without the measurements. The model includes both geometrical and acoustical parameters. A method that is often used to determine the acoustical field parameters is the boundary element method, which was used to calculate the values of HRTF in this paper.
Convention Paper 7101

11:30

P17-12 Binaural Room Synthesis and Binaural Sky—Flexible Virtual Monitoring for Multichannel Audio also with Height Loudspeakers—*Klaus Laumann,[1] Jörg Hör,[1] Gerd Spikofski,[1] Roman Stumpner,[1] Günther Theile,[1] Helmut Wittek[2]*
[1]Institut für Rundfunktechnik (IRT), Munich, Germany
[2]Schoeps Mikrofone, Karlsruhe, Germany

This is a presentation only; no Convention Paper will be available for purchase.

Workshop 18 **Monday, May 7**
11:30 – 13:30 **Room H**

THE REPLAY OF HISTORICAL MAGNETIC TAPE—MORE THAN PRESSING THE PLAY BUTTON

Chair: **Dietrich Schüller**, Phonogrammarchiv Vienna

Panelists: *Friedrich Engel*, Magnetic Recording Historian
Alexander Füller, Replay Practitioner, Phonogrammarchiv Vienna
Bernhard Graf, Replay Practitioner, Phonogrammarchiv Vienna
Nadja Wallaszkovits, Chief Technician Audio, Phonogrammarchiv Vienna

Since analog magnetic tape technology is no longer a part of the audio production process, specific knowledge is endangered to decline. The workshop discusses the various problems occurring in the transfer process of historical magnetic audio tapes. Basing on a definition of historical tape brands and an overview of early magnetic tape developments, the practical handling of critical tapes is outlined.

Starting with the analysis of the physical and chemical preservation status of the individual tape, the choice and adjustment of the replay equipment and parameters are discussed. Problems of carrier handling and physical restoration are demonstrated, as well as possible signal enhancement on the playback process only. The workshop focuses on a practical demonstration of handling and reproduction of original tapes from the early ages of magnetic recording, stored under irregular conditions.

Tutorial 8 **Monday, May 7**
11:30 – 13:00 **Room P**

SYNC AND TIMECODE—BRICKS NOT TRICKS

Presenter: **Chris Woolf**, Broadcast Engineering Systems Ltd.

Synchronization and timecode—intimately connected but not to be confused with each other—form the time-axis shells of buildings within which most sound practitioners must house their work. Rules-of-thumb, tricks-that-seem-to-work, and even blind faith often support rather shaky structures so this tutorial provides some underpinning: a foundation of solid bricks.

The session presumes very little and will be useful to those with limited experience. However, it will also appeal to those with gnarled hands and a lot of dust under their fingernails but who harbor secret doubts about the security of their techniques —dark glasses and a false moustache may be worn.

Exhibitor Seminar **Monday, May 7**
11:30 **Room 357**

AUDIO EXPORT GEORGE NEUMANN

Presenters: **Haimo Godler, ORF; Christoph Keller, Audio Export**

"KoKo—The ORF Radio Archive"

The software for archiving, interfacing, and retrieval was developed especially for the needs of ORF. The solution is a perfect example for a seamless integration between the digital production, the on-air tools based on the automation-system Radiomax, and the sound archive. This exhibitor seminar presents the main ideas of the ORF radio archive to the visitors.

Exhibitor Seminar **Monday, May 7**
12:30 **Room 357**

MINNETONKA

Presenter: **John Schur**

"Minnetonka's Audio Tools Workflow Architecture"

Rapidly increasing numbers of audio files and formats represent the largest challenge in media production, raising numerous productivity and quality issues. Minnetonka will present the AudioTools Workflow Architecture and demonstrate how the first products in the AudioTools family—Batch Pro and Version Pro—will help studios to be productive, efficient, and organized.

Monday, May 7 12:30 Room 142/143
Standards Committee Meeting on SC-03-07 Audio Metadata

Workshop 19 Monday, May 7
13:00 – 15:00 Room G

PLATINUM ENGINEERS / 5.1 MIXING TECHNIQUES

Chair: **George Massenburg**, GML, TN, USA

Panelists: *Jeff Wolpert*, Desert Fish Inc.
Ronald Prent, Galaxy Studios

Even with the increasing confusion over the future of surround music mixes there is an increasing demand for efficiently produced surround products for film and television, as well as for music-only release. At the same time, it is economically tempting for film and television producers to accept surround mixing simulacrums ("faux 5.1" processing such as Unwrap or Logic 7) even when multichannel originals are available for remix.

Methods and techniques will be presented in this workshop to demonstrate how good, modern surround mixes are being made by successful professional practitioners. The workshop will cover subjects such as: Different approaches to “spreading out” multichannel sources; Strategies for different media; Use of the center channel; Bass management; Reverb, delay, and other effects; and Monitoring.

In addition to the panel presentation, there will be several “break-out” presentations the next day in the same room where each of the panelists will present one or more mixes, discuss and present techniques, and interact with the audience. Seats for these sessions will be limited; tickets will be available after the workshop.

Monday, May 7 13:00 Room 633
Technical Committee Meeting on Semantic Audio Analysis

STUDENT EVENT
Education Forum Panel
Monday, May 7, 13:30– 15:30 Room P

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

Panelists: *Thomas Goerne*, Detmold University of Music
Douglas McKinnie, Middle Tennessee State University
Hans Vesterberg, School of Music in Piteå

Educators will discuss pedagogical issues related to teaching audio. Success in field of audio requires an understanding of not only audio theory and concepts, but also practical experience with audio. In fact some concepts are best assimilated through practical experience. Conversely some practical experiences require prior knowledge of the underlying theory to achieve the best results. The discussion will focus on how educators can best balance the presentation of theory and practice when teaching audio.

Exhibitor Seminar Monday, May 7
13:30 Room 357

D.A.V.I.D.

Presenter: **Klaus Hellmich**

“Global Interfaces Optimized for Exchanging Large Media Files”

Combining video content with radio workflows, increasingly large multimedia files have to be exchanged among multiple locations. As part of its “Moves Media” product line, D.A.V.I.D. provides the DigaMailbox-IP System, offering simple, secure and reliable connections over the Internet that are optimized for exchanging extensive media files between bureaus scattered all over the world.

Session P18 Monday, May 7 14:00 – 17:00
Room I

MICROPHONES AND LOUDSPEAKERS, PART 2

Chair: **Martin Opitz**, AKG Acoustics GmbH, Vienna, Austria

14:00

P18-1 Applications of the Acoustic Center—*John Vanderkooy*, BW Group, Ltd., Steyning, UK and University of Waterloo, Waterloo, Ontario, Canada

This paper focuses on uses for the acoustic center concept, which in this paper represents a particular point for a transducer that acts as the origin of its low-frequency radiation or reception. The concept, although new to loudspeakers, has long been employed for microphones when accurate acoustic pressure calibration is required. A theoretical justification of the concept is presented and several calculation methods are discussed. We first apply the concept to subwoofers, for which the acoustic center is essentially a cabinet dimension away from the center of the cabinet. This has an influence on its radiation pattern in a normal room with reflecting walls. A second application that we consider is the effective position of a microphone, which is necessary if it is to be used for accurate calibration of acoustic pressure. A final application that we consider is the effective position of the ears on the head at lower frequencies. Calculations show that the acoustic centers of the ears are well away from the head, and the effective ear separation is larger than expected. This has implications for the human localization mechanism. Measurements on a Kemar mannequin show that the separation is even larger than expected from the calculations, and most of this can be understood, but the measurements at the lowest frequencies are somewhat uncertain.
Convention Paper 7102

14:30

P18-2 Development of a Finite Element Headphone Model —*Dominik Biba, Martin Opitz*, AKG Acoustics GmbH, Vienna, Austria

For the development of high-end headphones, ➥

numerical simulation of the acoustic behavior is an efficient tool. While lumped-element models are valid for frequencies up to a few kilohertz, finite-element models are valid for higher frequencies too. A headphone model using finite elements and boundary elements was built in three phases. In parallel to building the numerical model a real-world sample was built and measured. The validity of the model was verified by comparison of both the radiated sound field and the membrane modal behavior. Agreement between measured and computed amplitude frequency responses was achieved.
Convention Paper 7103

15:00

P18-3 Development of Camera Mountable 5.0 Surround Microphone and Method of 3-Channel to 5-Channel Signal Recomposing System—*Minoru Kobayashi,*[1] *Setsu Komiyama,*[2] *Satoshi Kikkawa,*[2] *Takako Kawashima,*[2] *Takeshi Ishii*[1]
[1]Sanken Microphone Co., Ltd., Tokyo, Japan
[2]NHK Japan Broadcasting Corporation, Tokyo, Japan

This paper describes the development of small 5.0 surround microphone and 3-channel to 5-channel signal re-composing system. The principal reason we started this study was broadcasters' recent demand for a small, light-weight camera-mounted 5.0 surround microphone for documentaries, dramas, and sports shooting outdoors. We have produced a 5.0 surround microphone; but if we want to use it as a camera-mounted microphone, there is a limitation we could not ignore. The limitation is the number of audio tracks. The commonly available HD cameras in the market have only 4 audio tracks. In order to overcome the limitation, we developed a 3-channel to 5-channel signal re-composing system.
Convention Paper 7104

15:30

P18-4 Noise-Robust Recognition System Making Use of Body-Conducted Speech Microphone—*Shunsuke Ishimitsu,*[1] *Masashi Nakayama,*[2] *Toshikazu Yoshimi,*[3] *Hirofumi Yanagawa*[4]
[1]University of Hyogo, Himeji, Hyogo, Japan
[2]Toyohashi University of Technology, Toyohashi, Aichi, Japan
[3]Pioneer Corp., Kawagoe, Saitama, Japan
[4]Chiba Institute of Technology, Chiba, Japan

In recent years, speech recognition systems have been introduced in a wide variety of environments such, as vehicle instrumentation. Speech recognition plays an important role in ships' chief engineer systems. In such a system, speech recognition supports engine room controls, and lower than 0-dB signal-to-noise ratio (SNR) operability is required. In such a low SNR environment, a noise signal can be misjudged as speech, dramatically decreasing the recognition rate. Hence, speech recognition systems operating in low SNR environments have not received much attention. Therefore, this paper focuses on a recognition system that uses body-conducted signals. Such signals are seldom affected by background noise, and thus a high recognition rate can be expected in low SNR environments such as an engine room. Since noise is not introduced within body-conducted signals that are conducted in solids, even within sites such as engine rooms, which are low SNR environments, construction of a system with a high recognition rate can be expected. However, within the construction of such systems, in order to create models specialized for body-conducted speech, learning data consisting of sentences that must be read in numerous times is required. Therefore, in the present paper we applied a method in which the specific nature of body-conducted speech is reflected within an existing speech recognition system with only small numbers of vocalizations. Simultaneously, the measure by pretreatment was also worked on.
Convention Paper 7105

16:00

P18-5 Revisiting Proximity Effect Using Broadband Signals—*Laurent Millot,*[1,2] *Mohammed Elliq,*[2] *Manuel Lopes,*[2] *Gérard Pelé,*[1,2] *Dominique Lambert*[1,2]
[1]Université Paris, Paris, France
[2]ENS Louis-Lumiere, Noisy Le Grand, France

Experiments mainly studying the proximity effect are presented. Pink noise and music were used as stimuli and a combo guitar amplifier as source to test several microphones: omnidirectionnal and directional. We plotted in-axis levels and spectral balances as functions of x, the distance to the source. Proximity effect was found for omnidirectionnal microphones. In-axis level curves show that $1/x$ law seems poorly valid. Spectral balance evolutions depend on microphones and, moreover, on stimuli: bigger decreases of low frequencies with pink noise; larger increases of other frequencies with music. For a naked loudspeaker, we found similar in-axis level curves under and above the cut-off frequency and propose an explanation. Listening to equalized music recordings demonstrates will help to demonstrate proximity effect for tested microphones.
Convention Paper 7106

16:30

P18-6 Anechoic Measurements of Particle-Velocity Probes Compared to Pressure Gradient and Pressure Microphones—*Wieslaw Woszczyk,*[1] *Masakazu Iwaki,*[2] *Takehiro Sugimoto,*[2] *Kazuho Ono,*[2] *Hans-Elias de Bree*[3]
[1]McGill University, Montreal, Quebec, Canada
[2]NHK Science & Technical Research Laboratories, Setagaya-ku, Tokyo, Japan
[3]Microflown Technologies, Zevenaar, The Netherlands

A number of anechoic measurements of Microflown™ particle velocity probes are compared to measurements of pressure-gradient and pressure microphones made under identical acoustical conditions at varying distances from a point source having a wide frequency range. Detailed measurements show specific response changes affected by the distance to the source, and focus on the importance of transducer cali-

bration with respect to distance.
Convention Paper 7107

Session P19 Monday, May 8 14:00 – 17:00
Room K

SIGNAL PROCESSING, SOUND QUALITY DESIGN

Chair: **Alfred Kraker**

14:00

P19-1 Idle Tone Behavior in Sigma Delta Modulation
—Enrique Perez Gonzalez, Josh Reiss, Queen Mary, University of London, London, UK

This paper examines the relationship between various unwanted phenomena that plague audio engineers in the design of sigma delta modulators. This work aims to clarify the difference and relationship between DC idle tones, long limit cycles, short limit cycles, and "periodic" short limit cycles, while extending the current knowledge in idle tone behavior. A relationship between the periodic input to the quantizer of a 1-bit delta sigma modulator and the appearance of idle tones is shown. It is shown that for a large range of input signal magnitudes, the fundamental frequency of idle tones is proportional to the DC input. This finding has also been used to examine idle tone aliasing. Numerous simulations are reported which confirm these findings.
Convention Paper 7108

14:30

P19-2 Low Distortion Sound Reproduction Using 8-Bit uC and ZePoC-Algorithms—*Jan Wellmann, Olaf Schnick, Wolfgang Mathis*, Leibniz University Hannover, Hannover, Germany

The ZePoC-encoding algorithm for Class-D amplification allows the complete separation of the signal-baseband from all higher-frequency switching artifacts. Real-Time-ZePoC-Encoding demands a lot of computational power, but in applications where recorded signals should be reproduced, they can be encoded by a software-defined ZePoC-System in advance. Reproducing this pre-encoded signal has very low hardware requirements: no digital-analog-converter or linear amplifier is needed; the playback device must only contain memory; a counter for forming the rectangular output-signal; and, if higher output-power is required, an additional switching power-stage and filter. A simple system made up of an 8-bit micro-controller at 16-Mhz clock-rate could reach a signal-to-noise-ratio of 80 dB and a usable frequency range of up to 15 kHz. A test-system made up of an 8-bit RISC-Processor, external memory, and a single-transistor, single power-supply switching-stage will be presented.
Convention Paper 7109

15:00

P19-3 Efficient, High-Quality Equalization Using a Multirate Filterbank and FIR Filters—*Riitta Väänänen, Jarmo Hiipakka*, Nokia Research Center, Helsinki, Finland

This paper presents a digital signal processing algorithm for efficient and high-quality audio equalization. In this approach, the original full-band audio signal is first down-sampled and separated into two or more subband signals using a multirate filterbank, after which the equalization is performed in the down-sampled domains. After the equalization, the signal is up-sampled and combined back to a full-band audio signal. Linear phase FIR filters, designed based on the user-controlled parameters, are used to implement the actual equalization. The method presented in this paper helps in designing an implementation that results in computational savings, while still preserving optimal sound quality with any equalization parameter setting.
Convention Paper 7110

15:30

P19-4 Correction of Crossover Phase Distortion Using Reversed Time All-Pass IIR Filter—*Veronique Adam, Sebastien Benz*, Goldmund (Audio Networks SA), Geneva, Switzerland

The purpose of this paper is to describe a correction implementation of group delay distortion arising from a two-way loudspeaker system crossover. Having determined an IIR all-pass filter having a group delay response corresponding to that of the system crossover to be corrected, we have validated under Matlab and implemented in DSP the time reversal solution proposed by S. A. Azizi,. S. R. Powell, and P. M. Chau, enabling an IIR filter to be inversed, while retaining stability and causality. In addition to theory and calculation validation, we have also carried out preliminary listening tests, supporting the evaluation of timber modification, sound clarity, and space localization due to the group delay distortion correction.
Convention Paper 7111

16:00

P19-5 Natural Timbre in Room Correction Systems
—Jan Abildgaard Pedersen, Henrik Green Mortensen, Lyngdorf Audio, Skive, Denmark

Room correction systems are often found to provide a timbre, which is described to be artificial or unnatural. This paper presents a new approach to this problem, which is based on the finding that part of the influence of a listening room is natural to the human ear and should not be removed by a room correction system. More specifically the smooth increase of level toward lower frequencies, also referred to as room gain, must be preserved after applying a room correction system. In the described system this is done as an integral part of the automatic target calculator, which also takes into account the main characteristics of the used loudspeaker, e.g., lower cut-off frequency and directivity index.
Convention Paper 7112

16:30

P19-6 Multi Core/Multi Thread Processing in Object-Based Real Time Audio Rendering: Approaches and Solutions for an Optimization Problem
—Ulrich Reiter, Andreas Partzsch, Technische Universität Ilmenau, Ilmenau, Germany

This paper presents considerations, approaches, and solutions to the problem of optimization of thread distribution for multi core processing in real time audio rendering environments. It explains some basic problems, describes the constraints, and finally suggests an approach based on solving an optimization problem by analyzing a directed graph representing the signal processing flow. The suggested approach can handle an arbitrary number of CPU cores and is therefore well primed for future processor developments.
Convention Paper 7159

Session P20 Monday, May 7 14:00 – 15:30
Foyer IK

POSTERS: PSYCHOACOUSTICS, PERCEPTION, AND LISTENING TESTS

14:00

P20-1 Detection and Lateralization of Sinusoidal Signals in the Presence of Dichotic Pink Noise—*Tuomo Raitio, Heidi-Maria Lehtonen, Petteri Laine, Ville Pulkki,* Helsinki University of Technology, Espoo, Finland

This paper investigates the ability to lateralize low-frequency sound in the presence of interfering dichotic noise. This is addressed by measuring the detection and lateralization thresholds of four sinusoidal signals (62.5, 125, 250, and 500 Hz) in the presence of uncorrelated pink noise in headphone listening. In lateralization tests the signals were positioned to left or right by delaying either of the headphone channels by 0.5 ms. The results show that the lateralization threshold does not depart from detection threshold at frequencies 250 and 500 Hz. Interestingly, below 250 Hz the lateralization threshold rises fast, and at 62.5 Hz, the signal has to be amplified 18 dB from the detection level before being lateralized correctly. This suggests that low-frequency ITD decoding mechanisms are easily distracted by random changes in a signal phase. This explains, at least partly, why the direction of a subwoofer cannot be detected easily in surround sound listening of broad-band signals.
Convention Paper 7113

14:00

P20-2 The Practice and Study of Ear Training on Discrimination of Sound Attributes—*Zhi Liu, Fan Wu, Qing Yang*, Beijing Union University, Beijing, China

In order to improve subjects' discrimination of sound attributes, an ear training course has been designed. The training includes: discrimination of a pure tone's frequency, the frequency changes, sound level changes, musical instrument timbre, irregularity of frequency response. etc. All the items were carried on in interlaced order to avoid listening fatigue. Meanwhile, some explanation of the psychoacoustic principles and some tests were also conducted. In all, 57 subjects, divided into 2 groups, took part in the training course for about 15 weeks. After the special ear training, most subjects made great progress with a nearly 85 percent average correctness rate.
Convention Paper 7114

14:00

P20-3 The Training and Analysis on Listening Discrimination of Pure Tone Frequency—*Zhi Liu, Fan Wu, Qing Yang*, Beijing Union University, Beijing, China

After special and scientific ear training, more than 90 percent of people will get great progress on the discrimination of pure tone frequency. This has been proven by long-term ear training for two groups of subjects. Several pure tones were selected in an octave step for the training. Fifty-seven subjects took part in the training. The training was conducted once a week, for a total of 15 weeks for one group. The average correctness rate increased from around 60 percent to above 90 percent. The test results also show that human ears have poor discrimination with middle frequencies, while strong ability with high frequency and low frequencies. The relationship between the improvement and training time indicates that the training has the similar effect as the physical training.
Convention Paper 7115

14:00

P20-4 Virtual Hearing Aid—A Computer Application for Simulating Hearing Aid Performance—
Andrzej Czyzewski,[1] *Bozena Kostek*,[1,2]
Lukasz Kosikowski[1]
[1]Gdansk University of Technology, Gdansk, Poland
[2]Institute of Physiology and Pathology of Hearing, Warsaw, Poland

The virtual hearing aid is a computer application allowing an approximate simulation of hearing aid performance. The computer application implements algorithms simulating band-pass filters, compressors, and the perceptual masking strategies for audio signal processing. Individual persons' hearing characteristics were taken into account for this purpose. The experimental part comprises verification of engineered algorithms implemented to virtual hearing prosthesis. The paper also contains results of examinations of patients aimed at verifying the applicability of the proposed signal processing strategy to the domain of hearing prostheses.
Convention Paper 7116

14:00

P20-5 Training Versus Practice in Spatial Audio Attribute Evaluation Tasks—*Rafael Kassier, Tim Brookes, Francis Rumsey*, University of Surrey, Guildford, Surrey, UK

Listener training in published studies has tended to focus on simple repetitive practice of experimental tasks without feedback. Time savings in listening panel selection and training could be accomplished if a more general training system could be used and applied to a variety of tasks. In order for a training system for spatial audio listening skills to prove effective, it must demon-

strate that learned skills are transferable, and it must compare favorably with repetitive practice on specific tasks. A novel study to compare a training system with repetitive practice has been extended to include a total of 48 subjects. Transfer is assessed and practice and training are compared against a control group for tasks involving transfer of spatial audio training.
Convention Paper 7117

14:00

P20-6 Subjective Assessment of Quality of Multimedia Signals by Means of A-B Test—*Stefan Brachmanski*, Wroclaw University of Technology, Wroclaw, Poland

In this paper an automated method of subjective assessment of speech, music, image, and video quality has been described. In the method the sound, image or video samples were randomized and paired in A-B sets and then presented to a group of listeners. On the base A-B results a preference matrix was calculated. The conversion from the preference matrix to a numerical scale was performed with accordance to Thurstone's V model of paired comparisons. The method was applied to evaluate an influence of various coding techniques on a quality of video signals and natural speech.
Convention Paper 7118

14:00

P20-7 Influence of Visual Stimuli on the Sound Quality Evaluation of Loudspeaker Systems —*Alex Karandreas, Flemming Christensen*, Aalborg University, Aalborg, Denmark

Product sound quality evaluation aims to identify relevant attributes and assess their influence on the overall auditory impression. Extending this sound-specific rationale, the present paper evaluates overall impression in relation to hearing and vision, specifically for loudspeakers. In order to quantify the bias that the image of a loudspeaker has on the sound quality evaluation of a naive listening panel, loudspeaker sounds of varied degradation are coupled with positively or negatively biasing visual input of actual loudspeakers, and in a separate experiment by pictures of the same loudspeakers.
Convention Paper 7119

14:00

P20-8 The Study of Audio Equipment Evaluations Using the Sound of Music—*Shunsuke Ishimitsu,*[1] *Koji Sakamoto,*[1] *Keitaro Sugawara,*[2] *Toshikazu Yoshimi,*[2] *Atusushi Makino,*[2] *Katsuhiro Sasaki,*[3] *Hirofumi Yanagawa*[4]
[1]University of Hyogo, Himeji, Hyogo, Japan
[2]Pioneer Corp., Kawagoe, Saitama, Japan
[3]Tohoku Pioneer Corp., Tendo, Yamagata, Japan
[4]Chiba Institute of Technology, Narashino, Chiba, Japan

In this paper we considered audio equipment evaluation using musical sounds. Audio amplifiers were set up as the evaluation targets, and sound quality differences between them were visualized by a wavelet analysis using an actual musical sound signal. We considered the cause of these differences and then tried to connect the sound impression to an analysis result. WT of the sound of music was carried out to evaluate two amplifiers. The sound quality of amplifier A related to the esthetic factor of the “depth” and has been analyzed as the high region of WT, and that of amplifier B related to the force factor and has been analyzed as the low region of WT. Thus, we were able to visualize the impression of listening to the music by correlation of the auditory experiment and wavelet analysis.
Convention Paper 7120

Tutorial 9 **Monday, May 7**
14:00 – 16:00 **Room H**

BALANCED INTERFACES

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

One goal in the design of audio equipment is to maintain a high signal-to-noise ratio. But audio equipment most often operates on utility AC power, which, even under ideal conditions, normally creates ground voltage differences, magnetic fields, and electric fields. RF energy is increasingly omnipresent, too. Balanced interfaces are capable of conveying wide dynamic range analog audio signals while giving them unrivaled immunity to interference. Realizing this full capability in real-world, mass-produced equipment is not necessarily costly but requires some understanding of several common mistakes made by equipment designers. The telephone company pioneered the widespread use of balanced lines and for 50 years virtually all audio equipment used transformers at its balanced inputs and outputs—their high noise rejection was taken for granted.

When solid-state differential amplifiers began replacing transformers, most designers failed to recognize the importance of common-mode impedances—which are solely responsible for noise rejection. Instead, most believed that “balance” meant equal and opposite signal swings—which is a myth. As a result, most modern audio equipment has poor noise rejection when operating in real-world systems, even though it may have impressive rejection in a laboratory test. The IEC recognized this dichotomy when they revised their CMRR test standards in 2000 (at the urging of this author). A new IC uses bootstrap techniques to raise its common-mode impedances, and real-world noise rejection, to levels comparable to the finest transformers.

The three basic types of balanced output circuits, each with a peculiar set of trade-offs, must be accommodated by balanced input circuits. Further, certain cable constructions and shield connections can degrade noise rejection of an otherwise perfect interface. A very common equipment design error, the “pin 1 problem,” causes shield connections to behave as low-impedance audio inputs, allowing power-line noise and RF interference to enter the signal path.

TECHNICAL TOUR 8
Mediathek
Monday, May 7, 14:00 – 17:00

The Österreichische Mediathek in Vienna is the national archive for audio and video recordings. Over 1.2 million audio recordings are stored on various carriers like tape, records, 78s, CDs, DATs, or audio cassettes. Since the year 2000 our staff developed and implemented half automated workflows for digitizing all these different formats. Meanwhile there exists a complete system for digitizing over cataloging to long term preservation in a mass storage system. The tour starts with a short presentation of this concept, followed by a visit of the storage vault and various premises. Also interesting is the digital presentation platform of the content: the catalog and a growing collection of galleries on the Web.

Monday, May 7 14:00 Room 633
Technical Committee Meeting on Audio Recording and Mastering Systems

Exhibitor Seminar Monday, May 7
14:30 Room 357

ODEON

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

"ODEON Room Acoustic Simulations"

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

Monday, May 7 14:30 Room 142/143
Standards Committee Meeting on SC-02-12 Audio Applications of IEEE 1394

Workshop 20 Monday, May 7
15:00 – 17:00 Room 631/632

HANDS-ON DEMONSTRATION SUBJECTIVE EVALUATION BY REFERENCE RECORDINGS WITH SURROUND MAIN MIKINGS FOR CLASSICAL ORCHESTRA

Presenters: **Hideo Irimajiri**, Mainichi Broadcasting Corp.
Toru Kamekawa, Tokyo National University of Fine Arts and Music
Masayuki Mimura, Yomiuri Telecasting Corp.
Hideaki Nishida, Asahi Broadcasting Corp.
Koichi Ono, Kansai Telecasting Corp.

There are many different setups for surround main microphones for classical music and orchestra. But it is very difficult to research and study them practically and academically under identical conditions and judge their performance. Consequently the AES Japan Surround Study Project has been organized and put into practice after one year of preparation. It was organized around 10 broadcasters and 2 universities; 12 manufacturers supported by HBF provided financial support. There were 15 different combinations of main and ambience microphone setups that were recorded on 96 channels independently in Pro Tools HD at 24 bit / 96-kHz. The musical examples were performed by the Osaka Philharmonic Orchestra on September 24–27, 2006.

In this workshop each individual setup will be played back. Participants will have the opportunity for feedback in a listening test environment, and the data will be collected for subjective evaluation.

Exhibitor Seminar Monday, May 7
15:00 Room 441

SENNHEISER

Presenter: **Gregor Zielinsky**

"High Quality Surround Recording with the New MKH 8000 Microphone"

In this seminar Sennheiser presents the brand-new MKH 8000 Studio Microphone Series for the first time ever. On the examples of classical symphony and big band multitrack recordings, this "hands-on" seminar will show several surround microphone techniques using the MKH 8000. The different recordings will be explained and can be compared and judged by the listeners.

Monday, May 7 15:00 Room 633
Technical Committee Meeting on Network Audio Systems

Workshop 21 Monday, May 7
15:30 – 18:00 Room G

RADIO MUSIC RECORDING

Panelists: *Martin Leitner*, ORF
Anton Reininger, ORF
Josef Schütz, ORF
Ulrike Schwarz, BR

This workshop will present three senior sound balancing engineers from ORF Radio, sharing their own individual working procedures toward music recordings in 5.1 and stereo for live and deferred broadcasting. The musical genres will encompass classical, jazz, and world music. A method used for surround pickup, the "Karreth-Disc," will be introduced by the BR engineer.

Exhibitor Seminar Monday, May 7
15:30 Room 357

SCHOEPS

Presenter: **Helmut Wittek**

"Double M/S Recording in Stereo and Surround"

"Double M/S" is a technique for two-channel or multichannel stereo recording. It uses a compact microphone arrangement that requires only three recorder channels, while allowing full postproduction processing. During the seminar, recorded examples from a range of scenarios will be presented. The new Schoeps VST Plug-In will also be introduced and its capabilities demonstrated.

Workshop 22 Monday, May 7
16:00 – 18:00 Room H

CURRENT ISSUES IN BINAURAL LOUDNESS ASSESSMENT

Chair: **Ville P. Sivonen**, University Hospital of Oulu, Oulun Yliopisto, Finland

Panelists: *Pavel Zahorik*, University of Louisville, Lousiville, KY, USATBA

The most commonly-used standardized model for predicting the loudness of sound fields (ISO-532B, 1975) relies on the use of monophonic acoustical measurements and does not take into account the signals that actually reach a listener's ears. That is, the acoustical input to the model is obtained in the absence of a listener, using a microphone placed in the position where the center of the listener's head would be. Binaural sound fields encountered in typical measurements are more complex than those assumed by the standard model, most often consisting of a combination of spatially localized sound source(s) and spatially distributed indirect sound. Workshop panelists will discuss current issues in binaural loudness modeling and report the results of recent studies on directional loudness perception.

TECHNICAL TOUR 9
Organ Concert Recording
Monday, May 7, 16:00 – 21:00

This tour will center on the setup and recording of Graham Blyth's organ concert at the Jesuitenkirche. This will be a hands-on session with direct involvement of anybody attending. The concert will be recorded in 2+2+2+2 surround with height and in conventional 5.1 surround. The following aspects of the recording process will be explained during the setup: microphone selection and placement; equipment configuration, recording chain; and on-location monitoring via headphones.

Monday, May 7 16:00 Room 633
Technical Committee Meeting on Microphones and Applications

Workshop 23 Monday, May 7
16:30 – 17:30 Room P

BC 9001: QUALITY MANAGEMENT SYSTEMS FOR BROADCASTERS

Chair: **Guillaume Cheneviere**, CERTIMEDIA

A group of international experts has compiled a comprehensive certification program that is compatible with ISO 9001:2000. This program is targeted directly to the needs and organizational procedures of national public and private radio and television broadcasters and should be of high interest to anyone looking for a framework to analyze and certify the practical operation of their stations.

Exhibitor Seminar Monday, May 7
16:30 Room 357

MICROTECH GEFELL

Presenter: **Ulrich Apel**

"Surround Microphones Not Only for Concert Halls"

Microtech-Gefell presents microphones with special rainproof windscreens that are applicable not only in closed premises but for recordings of atmos on locations; even for film and television. Herewith Microtech-Gefell translated their outdoor-and environmental experiences into their studio-microphones. An application of this wind-and rainshield and an operation of new switchable microphone UM930 will be shown.

Monday, May 7 17:00 Room 633
Technical Committee Meeting Advisory Group on Regulations

Exhibitor Seminar Monday, May 7
17:30 Room 357

FLEX ACOUSTICS

Presenter: **Niels Werner Adelman-Larsen**

"Acoustics for Rock—The Flexible Bass Absorber"

Flex Acoustics made what seems to be the first ever scientific investigation into the recommended room acoustics for rock and pop concert halls. The results showed the necessity of additional bass absorption and led to the invention of a flexible bass absorber. The investigation and the absorber will be presented.

SPECIAL EVENT
Organ Recital by Graham Blyth
Monday, May 7, 20:00 – 22:00
Jesuitenkirche
Dr. Ignaz Seipel-Platz, Vienna

Graham Blythe's traditional organ concert will be given at the Jesuitenkirche (Jesuit Church), also known as the Universitätskirche (University Church), an ornate church on Dr. Ignaz Seipel-Platz, immediately adjacent to the old University of Vienna buildings. It was built between 1623 and 1627 on the site of an earlier chapel. In 1703, Brother Andrea Pozzo, SJ, an architect, painter and sculptor, and a master in the quadratura, was invited by the emperor Leopold I to redecorate the church. He added twin towers and reworked the facade in an early Baroque style with narrow horizontal and vertical sections. Despite its relatively austere exterior, the interior is remarkably opulent with ersatz marble pillars, gilding, and a number of allegorical ceiling frescoes. Immediately adjacent is the Aula, where Haydn's oratorio *The Creation* had its premiere, as did Beethoven's Seventh Symphony.

The new organ of the Jesuitenkirche is one of the most famous organs in Vienna. Its control elements are rebuilt from a typical Cavaillé-Coll console.

The program features works by Bach, Franck (Three Pieces), Vierne, and Widor (selections from 6th 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 (Liebfrauen Dom), Paris (Madeleine and St. Etienne du Mont) and Berlin. He has lived in Wantage, Oxfordshire, since 1984 where he is currently Artistic Director of Wantage Chamber Concerts and Director of the Wantage Festival of Arts.

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

➥

Session P21 Tuesday, May 8 09:00 – 12:30
Room K

INSTRUMENTATION AND MEASUREMENT

Chair: **Heinrich Pichler**, Technical University Vienna, Austria

09:00

P21-1 Advancements in Impulse Response Measurements by Sine Sweeps—*Angelo Farina*, University of Parma, Parma, Italy

Sine sweeps have been employed for long time for audio and acoustics measurements, but in recent years (2000 and later) their usage became much larger, thanks to the computational capabilities of modern computers. Recent research results now allow for a further step in sine sweep measurements, particularly when dealing with the problem of measuring impulse responses, distortion, and when working with systems that are neither time invariant nor linear. The paper present some of these advancements and provide experimental results aimed to quantify the improvement in signal-to-noise ratio, the suppression of pre-ringing, and the techniques employed for performing these measurements cheaply using a standard PC and a good-quality sound interface, and currently available loudspeakers and microphones.
Convention Paper 7121

09:30

P21-2 The Challenges of MP3 Player Testing—*Steve Temme, Pascal Brunet, Zachary Rimkunas*, Listen, Inc., Boston, MA, USA

MP3 player audio performance is discussed including measurements of frequency response, phase response, crosstalk, distortion, sampling rate errors, jitter, and maximum sound pressure level with headphones. In order to make these measurements, several measurement techniques and algorithms are presented to overcome some of the challenges of testing MP3 players. We discuss test equipment requirements, selection of test signals, and the effects of the encoding on these test signals. A new method for measuring noncoherent distortion using any test signal including music is also presented.
Convention Paper 7122

10:00

P21-3 Tracking Harmonics and Artifacts in Spectra Using Sinusoidal and Spiral Maps—*Palmyra Catravas*, Union College, Schenectady, NY, USA

A technique for tracking a harmonic series in a spectrum using a combination of sinusoidal and spiral maps is described. The spiral map enhances patterns that appear when a sinusoid is sampled near the Nyquist rate. The correspondence between the maps facilitates derivation of properties and motivates the use of curves that cut across the sinusoid or spiral. As an application, the spatial separation of a specific musical pitch from an artifact is demonstrated.
Convention Paper 7123

10:30

P21-4 Spatial Distribution Meter: A New Method to Display Spatial Impression Over Time—*Joerg Bitzer*, University of Applied Science Oldenburg, Oldenburg, Germany

A wide-spread method for visualizing the spatial behavior of stereo signals is the vector-scope or goniometer, which shows the relation between the left and right channel. Well-trained eyes can see misbalance and mono compatibility problems in these very fast changing figures. However, this analysis tool contains no information over time and no details can be seen. In this paper we introduce a new method for analyzing stereo signals, which is based on the vector-scope but shows the behavior over time. The final graph looks like a spectrogram/sonogram, where the axes are time and angle. Useful applications of this new spatial distribution meter are the analysis of stereo impulse responses, the shift of stereo base over time, and the typical applications of a vector-scope.
Convention Paper 7124

11:00

P21-5 Low Level Audio Signal Transfer Through Transformer Conflicts with Permeability Behavior Inside Their Cores—*Menno van der Veen*, ir. bureau Vanderveen bv, Zwolle, The Netherlands

At the threshold of audibility, the signal and flux density levels in an amplifier with audio transformers are very small. At those levels the relative magnetic permeability of the iron transformer core collapses and the inductance of the transformer becomes very small. The impedances connected to the transformer plus its signal level and frequency-dependent inductance behave as a high pass filter that corner frequencies slip into the audio bandwidth, resulting in a nonlinear signal transfer through the transformer. This paper explains deviations in the reproduction of micro details at the threshold of audibility.
Convention Paper 7125

11:30

P21-6 New Techniques for Measuring Speech Privacy and Efficiency of Sound Masking Systems—*Peter Mapp*, Peter Mapp + Associates, Colchester, Essex, UK

Speech privacy is becoming an increasingly important aspect for many workplace and security environments as well as hospital and medical centers where patient confidentially is of critical importance. Traditionally, speech privacy has been measured by means of the Articulation Index, transposed to rate privacy rather than intelligibility (PI=1-AI). However, this is an indirect and cumbersome method that usually requires a spreadsheet calculation to yield the Privacy Index rating. The paper discusses the potential use of STI and STIPa as direct measures of

speech privacy. The benefits and limitations of the methods are highlighted together with the results from a number of case studies. It is concluded that while the method has potential merit, a number of the limiting factors require further research.
Convention Paper 7126

12:00

P21-7 Onset Detection Method in Piano Music: Sensibility to Threshold Values—*Luis Ortiz-Berenguer, Javier Casajus-Quiros, Marison Torres-Guijarro, Jon Beraceochea*, Technical University of Madrid, Madrid, Spain

Piano music transcription requires a stage of onset detection. Every time a new note or chord is played, a new analysis of the note or chord is needed. It is a critical issue to correctly detect if a new note or chord has been played. Onset detection should have a simple solution, but several problems arise when attempting to perform it. This paper present a study of the sensibility of a detection method depending on the adjustable parameters. It also compares some results with a simpler method based on the analysis of the derivative of the energy envelope. The methods have been tested with six piano compositions. As a conclusion, accurate automatic onset detection in piano music is not a simple task even in the case of notes played alone.
Convention Paper 7127

Session P22 Tuesday, May 8 09:00 – 10:30
Foyer IK

POSTERS: AUDIO NETWORKING

09:00

P22-1 Benefits of Using SIP for Audio Broadcasting Applications—*Serge de Jaham*, AETA Audio Systems, Le Plessis Robinson, France

The SIP protocol has gained popularity for setting up temporary audio links over IP networks for broadcast applications. This paper briefly describes SIP and discusses its distinctive advantages, especially in comparison with proprietary systems. The main and obvious benefit is standardization, which opens the way to interoperation between different makes of codecs. SIP, as a signaling protocol, readily provides efficient methods for link setups, while preserving ease of use. Now a mature technology in the VoIP field, it is supported by a wide range of network devices and includes provision for specific issues like firewall or NAT traversal. As a result, SIP should be the key to the transition from ISDN to IP networks, while providing at least as flexible operating modes.
Convention Paper 7128

09:00

P22-2 Managing the Leap from Synchronous to IP for Radio Broadcasters: A Look at Equipment, Network, and Compression Considerations—*Simon Daniels*, APT, Belfast, Ireland, UK

Increasingly radio broadcasters are looking at making the leap from synchronous networks to IP networks for their distribution and contribution links. The advantages of migrating away from synchronous networks to IP networks are numerous but often tempered by a number of concerns regarding the IP transport mechanism including latency, lost packets, packet size, protocol selection, jitter and algorithm selection. This paper will address concerns such as which IP Protocol is most suitable for real time audio delivery, which algorithm is most suited to IP to reduce the affects of the inherent latency on an IP network, how to protect against packet loss and how to deal with the inherent delay involved in packetizing audio for delivery over an IP network.
Convention Paper 7129

09:00

P22-3 An XML-Based Approach to Audio Connection Management—*Richard Foss, Brad Klinkradt*, Rhodes University, Grahamstown, South Africa

An XML-based approach to firewire audio connection management has been developed that allows for the creation of connection management applications using a range of implementation tools. The XML connection management requests flow between a client and server, where the client and server can reside on the same or separate workstations. The server maintains the state of the firewire audio device configuration as well as information about potential users. XML is also used to control user access of devices.
Convention Paper 7130

09:00

P22-4 Rhythm Based Error Correction Approach for Scalable Audio Streaming Over the Internet— *J. C. Cuevas-Martinez, P. Vera-Candeas, N. Ruiz-Reyes*, University of Jaén, Linares, Jaén, Spain

Multimedia is nowadays the most important kind of information over the internet due to the impressive growth of the Web and streaming technologies. Although there are faster lines, the amount of potential users can exceed the actual available band width. In that way, scalable audio streaming makes it possible. However, error correction is left in a second level of importance for multimedia, using in some cases TCP, FEC (forward error correction) that are useless in low bit rate coders or even nothing. Therefore, in this paper a rhythm-based error correction approach is presented. This solution can avoid important redundant information, leaving almost all the error processing at the decoder side, without any feedback to the sender.
Convention Paper 7131

➡

Workshop 24 **Tuesday, May 8**
09:00 – 11:00 **Room H**

DIGITAL STORAGE CARRIERS FOR ARCHIVES

Chair: **Stefani Renner**, Ingenieurbuero Renner

Panelists: *Wolfgang Draese*, Hitachi Data Systems
Magnus Widmer, IBM Switzerland, Storage Systems Group

Nearly all sectors of the professional audio industry, in particular large radio stations, face the same problem: masses of analog archives, waiting to be digitized. Because digitalization pervades all levels of production today, converting from analog archives to digital media has become a necessity, not only to improve quality and preservation but also to make those archives compatible with current production methods. So, it should be no surprise that digital archiving has become a major field of interest in broadcasting.

This workshop will focus on identifying the best storage carrier for archives. Since one storage carrier no longer fits every situation—as it did back when analog tape predominated—each archive should be able to utilize the mass storage system that best meets its needs. Today's archives generally use one of three types of mass storage carrier: tape-based, hard disk, and optical. In this workshop we will present an overview of these three types of carrier along with their specific features and conclude with an in-depth discussion on how to select the best storage system for your archiving needs. Aspects such as the future-proof of the storage and the total cost of ownership are examined as well as questions about the behavior in high humidity, in dusty environments or at high temperatures. In addition, real-life experiences with archive installations in broadcasting will be shared.

STUDENT EVENT
Student Delegate Assembly Meeting—Part 2
Tuesday, May 8, 09:00 – 10:30 Room P

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 Design and Recording Competitions. Plans for future student activities at local, regional, and international levels will be summarized.

TECHNICAL TOUR 10
Boesendorfer Grand Piano Factory
Tuesday, May 8, 09:00 – 13:00

Watch the birth of a famous Boesendorfer Grand Piano! See the initial preprocessing stages of the wooden and iron raw material, the treatment of the iron frame, the sound board, pin blocks, key blocks, the frame installation, string production, voicing, assembly, etc. You can get more information of the process at www.boesendorfer.com.

Tuesday, May 8 **09:00** **Room 142/143**
Standards Committee Meeting AESSC Plenary

Session P23 **Tuesday, May 8** **09:30 – 12:00**
Room I

ANALYSIS AND SYNTHESIS OF SOUND

Chair: **Gerhard Graber**, Technical University Graz, Austria

09:30

P23-1 Sound-Transformation and Remixing in Real-Time—*Hannes Raffaseder*, St. Pölten University of Applied Sciences, St. Pölten, Austria

Starting with a short overview on some very basic principles of sound perception, this paper acts on the assumption that recording, storage, editing, and reproduction of audio signals have compensated for at least some of these principles and, therefore, have significantly changed human listening habits. Reflecting on these changes, the idea of sound transformation and remixing in real-time is suggested as part of the performance and composition process. Some techniques are explored and a number of implementations are introduced.
Convention Paper 7132

10:00

P23-2 Hybrid Time-Scale Modification of Audio—*Patrick-André Savard, Philippe Gournay, Roch Lefebvre*, Université de Sherbrooke, Sherbrooke, Quebec, Canada

This paper presents a novel technique for time-scale modification (TSM), which integrates time-domain and frequency-domain processing. The method relies on frame-by-frame classification to choose between different techniques adapted to different signal types. Provisions are taken to seamlessly switch between techniques. The result is a more universal TSM algorithm that yields continuous high quality results on a wider range of audio signals. The method is tested on mixed-content signals and formal listening tests results are discussed.
Convention Paper 7133

10:30

P23-3 New Audio Editor Functionality Using Harmonic Sinusoids—*Wen Xue, Mark Sandler*, Queen Mary, University of London, London, UK

This paper introduces the application of harmonic sinusoid model in an audio editor. The harmonic sinusoid model is a parametric model for representing pitched audio events, allowing amplitude and pitch evolution. While standard audio editors enable the user to select a time or frequency range to edit, with the harmonic sinusoidal parameters estimated in phase, we are able to select a pitched event and edit it as if it were separated from the background. The user interface is designed as a simple one-click selection, while the user is given further options for better results.
Convention Paper 7134

11:00

P23-4 Measurement and Optimization of Acoustic Feedback of Control Elements in Cars—*Alexander Treiber, Gerhard Gruhler*, Heilbronn University, Heilbronn, Germany

Acoustical quality of control elements in cars is

increasingly important for manufacturers in order to improve the quality, appearance, and security of their products. This paper presents methods and tools used in an ongoing research project. The project's goal is to support the industry with the definition of suitable parameters and limits as well as to develop realizable proposals for measuring equipment. Jury tests are hereby used to create the scientific basis for the hearing-related benchmarking of signals.
Convention Paper 7135

11:30

P23-5 On the Training of Multilayer Perceptrons for Speech/Nonspeech Classification in Hearing Aids—*Lorena Álvarez, Enrique Alexandre, Lucas Cuadra, Manuel Rosa-Zurera*, Universidad de Alcalá, Alcalá de Henares, Spain

This paper explores the application of multilayer perceptrons (MLP) to the problem of speech/nonspeech classification in digital hearing aids. When properly designed and trained, MLPs are able to generate an arbitrary classification frontier with a relatively low computational complexity. The paper will focus on studying the key influence of the training process on the performance of the system. An appropriate election of the training algorithm will help to provide better classification with a lower number of neurons in the network, which leads to a lower computational complexity. The results obtained will be compared with those obtained from two reference algorithms (the Fisher linear discriminant and the k-Nearest Neighbour), along with some comments regarding the computational complexity.
Convention Paper 7136

Workshop 25 **Tuesday, May 8**
09:30 – 11:30 **Room 631/632**

"WANNA FEEL MY LFE?" AND 5.1 OTHER QUESTIONS TO CONFUSE YOUR GIRLFRIEND AND SCARE YOUR GRANDMA

Cochairs: **Florian Camerer**, Austrian TV
Bosse Ternstrom, Swedish Radio

This radical multichannel battle features two notorious surround-sound-nerds with an armoury to die for. From the loudest to the softest, the most brain-melting to the downright boring, a variety of mixes will provide the basis of a discussion of esthetics, philosophies and the quest for the ultimate place in town for a cold beer.

Exhibitor Seminar **Tuesday, May 8**
10:30 **Room 357**

LEAR CORPORATION/DIRAC RESEARCH

Presenters: **Nilo Casimiro Ericsson, Mathias Johansson, Armin Prommersberger**

"Dirac Live™—A New Approach to Room Correction Algorithms"

A method for improving car audio systems through digital correction of the impulse/frequency responses of individual loudspeakers is presented. The ability to simultaneously optimize different time and frequency domain criteria and flexible spatial averaging are the main benefits of the new approach. The solution is featured in the new BMW M5 Individual High End Sound System, recently reviewed by Auto, Motor und Sport (number 8/2007 pp. 146–148) stating: "the most impressive car stereo sound in the world." An A/B comparison to conventional filtering is available in LEAR's demo car in our common exhibition booth.

Workshop 26 **Tuesday, May 8**
11:00 – 12:30 **Room 560/561**

ACOUSTIC CONCEPTS AND TIME MANAGEMENT FOR SURROUND RECORDINGS—ORGAN

Cochairs: **Werner Dabringhaus**, MDG
Ulrich Vette, University of Music and Performing Arts, Vienna

This presentation concentrates on the comparison of different surround concepts with the outlook for upcoming technical possibilties. The organ recording from the recital of Graham Blyth in the Jesuitenkirche (see Technical Tour 9, page 15) will be reproduced in 2+2+2+2, ITU 5.1 and further loudspeaker setups.

Exhibitor Seminar **Tuesday, May 8**
11:00 **Room 411**

SENNHEISER

Presenter: **Gregor Zielinsky**

"High Quality Surround Recording with the New MKH 8000 Microphone"

In this seminar Sennheiser presents the brand-new MKH 8000 Studio Microphone Series for the first time ever. On the examples of classical symphony and big band multitrack recordings, this "hands-on" seminar will show several surround microphone techniques using the MKH 8000. The different recordings will be explained and can be compared and judged by the listeners.

Tutorial 10 **Tuesday, May 8**
11:30 – 13:30 **Room H**

GROUNDING AND SHIELDING

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

Many designers and installers of audio/video systems think of grounding and interfacing as a "black art." Do signal cables really "pick up" noise, presumably from the air like a radio receiver? Equipment manufacturers, installers, and users rarely understand the real sources of system noise and ground loop problems, routinely overlooking or ignoring basic laws of physics. Although myth and misinformation are epidemic, this tutorial brings insight and knowledge to the subject. Signals accumulate noise and interference as they flow through system equipment and cables. Both balanced and unbalanced interfaces transport signals but are also vulnerable to coupling of interference from the power line and other sources. The realities of ac power distribution and safety are such that some widely used noise reduction strategies are both illegal and dangerous. Properly wired, ➥

fully code-compliant systems always exhibit small but significant residual voltages between pieces of equipment as well as tiny leakage currents that flow in signal cables. The unbalanced interface has an intrinsic problem, common-impedance coupling, making it very vulnerable to noise problems. The balanced interface, because of a property called common-mode rejection, can theoretically nullify noise problems. Balanced interfaces are widely misunderstood and their common-mode rejections suffer severe degradation in most real-world systems. Many pieces of equipment, because of an innocent design error, have a built-in noise coupling mechanism dubbed the "pin 1 problem" by Neil Muncy. A simple troubleshooting method that uses no test equipment will be described. It can pinpoint the exact location and cause of system noise. Most often, devices known as ground isolators are the best way to eliminate noise coupling. Signal quality and other practical issues are discussed as well as how to properly connect unbalanced and balanced interfaces to each other. While immunity to RF interference is a part of good equipment design, it must often be provided externally. Finally, power line treatments such as technical power, balanced power, power isolation transformers, and surge suppression are discussed.

Tutorial 11 **Tuesday, May 8**
11:30 – 13:30 **Room P**

AUDIO CONTRIBUTION WITH IP OVER WAN NETWORKS

Presenters: **Mathias Coinchon**, EBU
Lars Jonsson, EBU/Swedish Radio
Gregory Massey, APT Ltd., Ireland

Audio over IP end units have become common in radio and tv operations for streaming of audio over IP networks, from remote sites or local offices into main studio centers. ISDN is gradually replaced by IP-circuits.

The IP networks used can be well-managed private networks with controlled quality of service. The Internet is increasingly also used for various cases of radio contribution, especially over longer distances. Radio correspondents will have the choice in their equipment to use either ISDN or Internet via ADSL to deliver their reports. In France even distribution to FM transmitters via IP over well managed MPLS-networks is planned to replace older circuits.

More than 15 manufacturers provide equipment for these applications.

With almost no exceptions, end units coming from one manufacturer today are not compatible with another company's unit. Based on an initiative coming from German vendors and broadcasters, the European Broadcasting Union, EBU, has started a project group, N/ACIP, Audio Contribution over IP, to suggest a method of create interoperability for audio over IP. A draft standard has been proposed by the EBU. Some manufacturers already have begun implementation as a minimum interoperability option.

The tutorial will cover the standardization process and give a basic overview of audio over IP.

Tuesday, May 8 **12:00 – 18:00**
Room G

PLATINUM BREAKOUT SESSIONS

The panelists of the workshop "Platinum Engineers – 5.1 Mixing Techniques" will do a hands-on demonstration of several of their mixes with audience interactivity. The mixes are presented with the original multitrack sessions, so they can be broken down into the individual elements, and mixing techniques can be experienced very intuitively.

Schedule:

12:00 – George Massenburg
13:00 – Ronald Prent
14:00 – Jeff Wolpert
15:00 – George Massenburg
16:00 – Ronald Prent
17:00 – Jeff Wolpert

Workshop 27 **Tuesday, May 8**
13:00 – 14:30 **Room 560/561**

ACOUSTIC CONCEPTS AND TIME MANAGEMENT FOR SURROUND RECORDINGS—ORGAN

Cochairs: **Werner Dabringhaus**, MDG
Ulrich Vette, University of Music and Performing Arts, Vienna

This presentation concentrates on the comparison of different surround concepts with the outlook for upcoming technical possibilties. The organ recording from the recital of Graham Blyth in the Jesuitenkirche (see Technical Tour 9, page 15) will be reproduced in 2+2+2+2, ITU 5.1 and further loudspeaker setups.

Workshop 28 **Tuesday, May 8**
13:00 – 15:00 **Room 631/632**

HANDS-ON DEMONSTRATION: SUBJECTIVE EVALUATION BY REFERENCE RECORDINGS WITH SURROUND MAIN MIKINGS FOR CLASSICAL ORCHESTRA

Presenters: **Hideo Irimajiri**, Mainichi Broadcasting Corp.
Toru Kamekawa, Tokyo National University of Fine Arts and Music
Masayuki Mimura, Yomiuri Telecasting Corp.
Hideaki Nishida, Asahi Broadcasting Corp.
Koichi Ono, Kansai Telecasting Corp.

There are many different setups for surround main microphones for classical music and orchestra. But it is very difficult to research and study them practically and academically under identical conditions and judge their performance. Consequently the AES Japan Surround Study Project has been organized and put into practice after one year of preparation. It was organized around 10 broadcasters and 2 universities; 12 manufacturers supported by HBF provided financial support. There were 15 different combinations of main and ambience microphone setups that were recorded on 96 channels independently in Pro Tools HD at 24 bit / 96-kHz. The musical examples were performed by the Osaka Philharmonic Orchestra on September 24-27, 2006.

In this workshop each individual setup will be played back. Participants will have the opportunity for feedback in a listening test environment, and the data will be collected for subjective evaluation.

Session P24 **Tuesday, May 8** **13:30 – 15:00**
Room I

AUDIO-VIDEO SYSTEMS

Chair: **Gregor Widholm**, Musikuniversität Wien, Austria

13:30

P24-1 Production and Live Transmission of 22.2 Multichannel Sound with Ultra-High Definition TV —*Toshiyuki Nishiguchi,*[1] *Yasushige Nakayama,*[1] *Reiko Okumura,*[1] *Takehiro Sugimoto,*[1] *Atsushi Imai,*[1] *Masakazu Iwaki,*[1] *Kimio Hamasaki,*[1] *Akio Ando,*[1] *Shoji Kitajima,*[2] *Yutaka Otsuka,*[2] *Satoko Shimaoka*[2]
[1]NHK Science & Technical Research Laboratories, Tokyo, Japan
[2]NHK Broadcast Engineering Department, Tokyo, Japan

A 22.2 multichannel sound system was developed for ultra-high definition TV. The improvement in the spatial quality created by this system as compared to that of two-dimensional sound was evaluated and reported in previous papers. The first experiment on large-scale live production and transmission of 22.2 multichannel sound with ultra-high definition video was carried out to show the possible application of this system to next-generation broadcasting. In Tokyo, 22.2 multichannel sound was live mixed and transmitted to Osaka using an IP optical network. This paper describes in detail this live production and its transmission using the 22.2 multichannel sound system. It also discusses various issues of sound design, capturing, and mixing for three-dimensional sound.
Convention Paper 7137

14:00

P24-2 Automated Audio Detection, Segmentation, and Indexing with Application to Postproduction Editing—*Charalampos Dimoulas, Christos Vegiris, Kostantinos Avdelidis, George Kalliris, George Papanikolaou*, Aristotle University of Thessaloniki, Thessaloniki, Greece

The current paper deals with audio event detection, segmentation, and characterization in order to be further utilized in postproduction. Browsing, selection, and characterization of audio-visual content is a tiresome task, especially in audio/video editing applications, where an enormous amount of recordings with different characteristics is usually involved. Automated detection, segmentation, and general audio classification are essential to deploy flexible and effective audio-visual content management. A multi-resolution scanning procedure, based mainly in wavelet processing, is currently proposed, where various energy-based comparators and signal-complexity metrics have been tested for detection purposes. A variety of audio features, including many MPEG-7 audio low level descriptors, have been considered for events' characterization and indexing purposes. Extraction of the detection/ characterization results via MPEG-7 description schemes or similar indexing files are considered.
Convention Paper 7138

14:30

P24-3 A New Method for Measuring Time Code Quality —*Michael Beckinger, René Rodigast, Florian Müller-Kähler*, Fraunhofer Institut IDMT, Ilmenau, Germany

A high-quality time code and word clock synchronization is essential to prevent audio drop outs and flutters in sound studios. A bad adjustment of time code generators, respectively word clock synchronizers, requires extensive error checks in synchronization networks. For this reason, a new measurement method is presented that enables sound engineers to measure longitudinal time code jitter and to check time code/word clock synchronization.
Convention Paper 7139

Session P25 Tuesday, May 8 13:30 – 17:30
Room K

ROOM AND ARCHITECTURAL ACOUSTICS AND SOUND REINFORCEMENT

Chair: **Franz Lechleitner**, Austrain Academy of Sciences, Vienna, Austria

13:30

P25-1 50 Years of Sound Control Room Design—*Jan Voetmann*, DELTA Acoustics, Hørsholm, Denmark

Sound control room design is an interesting corner of small room acoustics and represents most of the problems found here: frequency balanced reverberation time, proper distribution of room modes, low frequency reproduction, sound source and receiver positioning, etc. The function of the control room is twofold, which is often overlooked. On one hand, the control room together with the monitor loudspeakers should reproduce as faithfully as possible the efforts of the sound engineer and the producer in creating a new recording. On the other hand, the control room should mimic the perceived acoustics of an average living room when checking the final result of the recording. Simply because most musical productions are aimed at the listening environment of a living room.
Convention Paper 7140

14:00

P25-2 Acoustics in Rock and Pop Music Halls—*Niels W. Adelman-Larsen,*[1] *Eric R. Thompson,*[2] *Anders Christian Gade*[2]
[1]Flex Acoustics, Lyngby, Denmark
[2]Technical University of Denmark, Lyngby, Denmark

The existing body of literature regarding the acoustic design of concert halls has focused almost exclusively on classical music, although there are many more performances of rhythmic music, including rock and pop. Objective measurements were made of the acoustics of twenty rock music venues in Denmark and a questionnaire was used in a subjective assessment of

➥

those venues with professional rock musicians and sound engineers. Correlations between the measurements and the questionnaire answers lead, among others, to a recommendation for reverberation time as a function of hall volume. Since the bass frequency sounds are typically highly amplified, they play an important role in the subjective ratings, and the 63 Hz band must be included in objective measurements and recommendations.
Convention Paper 7141

14:30

P25-3 The Flexible Bass Absorber—*Niels W. Adelman-Larsen*,[1] *Eric R. Thompson*,[2] *Anders Christian Gade*[2]
[1]Flex Acoustics, Lyngby, Denmark
[2]Technical University of Denmark, Lyngby, Denmark

Multipurpose concert halls face a dilemma. They host different performance types that require significantly different acoustic conditions in order to provide the best sound quality to both performers, sound engineers, and audience. Pop and rock music often contain high levels of bass sound energy but still require high definition for good sound quality. The mid- and high-frequency absorption is easily regulated, but adjusting the low-frequency absorption has typically been too expensive or requires too much space to be practical for multipurpose halls. A practical solution to the dilemma has been developed. Measurements were made on a variable and mobile low-frequency absorber. The paper presents the results of prototype sound absorption measurements as well as elements of the design.
Convention Paper 7142

15:00

P25-4 Improvements to Binary Amplitude Diffusers
—*Elizabeth Payne-Johnson*,[1] *Gillian Gehring*,[1] *Jamie Angus*[2]
[1]University of Sheffield, Sheffield, South Yorkshire, UK
[2]University of Salford, Salford, Greater Manchester, UK

Improved forms of diffusion structures based on absorption reflection gratings are presented. The theory, design, advantages, and limitations of these structures are discussed and their performance presented. Two methods of improving performance are suggested. The first structure is based on diffusion limited aggregation, which models non-regular fractal growth. The second structure was a panel with square absorption patches of a variable size that was determined by an m-sequence. Of the two, the second structure performed best. This paper demonstrates that improved diffusing structures that take up less space than phase reflecting ones are possible.
Convention Paper 7143

15:30

P25-5 An MLS Method for Nonstationary and Outdoor Acoustic Paths—*Jamie Angus, David Waddington*, University of Salford, Salford, Greater Manchester, UK

The correlation properties of a directly carrier-modulated code sequence modulation signal are exploited to investigate sound propagation in turbulent air. An experiment is described in which the correlation properties of the spread spectrum signal are demonstrated and are used to calculate accurate times of flight that compare well with sonic anemometer measurements of speed of sound. The results illustrate that a directly carrier-modulated code sequence modulation system can provide significantly improved ways of measuring sound propagation outdoors. Moreover, the technique directly measures wind speed. This can be used to compensate the time of flight thus allowing the measurement of acoustic impulse responses in nonstationary media, for example outdoors, where reliable measurements have previously been difficult to obtain.
Convention Paper 7144

16:00

P25-6 Holographic Sound Field Analysis with a Scalable Spherical Microphone-Array—*Anton Schlesinger*,[1] *Giovanni Del Galdo*,[2] *Jörg Lotze*,[2] *Stephan Husung*,[2] *Bernhard Albrecht*[2]
[1]Delft University of Technology, Delft, The Netherlands
[2]Technical University of Ilmenau, Ilmenau, Germany

Room acoustic parameters vary greatly with the position of the receiver and of the source, so that we cannot extract exhaustive information on the room acoustics from independent single-point measurements. Using array measurements permits the prediction of the sound field with a high spatial resolution and leads to a more precise assessment of the room acoustic properties. We propose an array technique to investigate room acoustics by reconstructing the volumetric sound field from measurements taken on the surface of a sphere, by means of the methods of nearfield acoustical holography (NAH). A virtual spherical single-microphone-array was constructed and successfully tested in room acoustical modal analysis.
Convention Paper 7145

16:30

P25-7 A Comparison of Modeling Techniques for Small Acoustic Spaces such as Car Cabins—*Neil Harris*, New Transducers Ltd. (NXT), Huntingdon, Cambridgeshire, UK

This paper results from a case study comparing the relative cost effectiveness of three modeling techniques applied to a small acoustic space such as a car cabin. The techniques considered are finite element analysis, analytical solutions, and the quasi-analytical ray-trace or image method. A simple test-case is used to compare solution times and accuracy.
Convention Paper 7146

17:00

P25-8 Acoustical and Musical Design of the Sea Organ in Zadar—*Ivan Stamac*, Stims d.o.o., Zagreb, Croatia

The Sea Organ in Zadar, Croatia, is an awarded

urban architectural installation using sea wave random kinetic energy to produce quasi-musical sounds. It contains 35 flue pipes built into subterranean tunnels having outward-bound apertures for the sound to emanate. Each flue pipe is blown by a column of air pushed in turn by a column of moving water entering an immersed tube. The pipes are tuned to 9 tones of the diatonic major chords G and C6. The series of excited tones is a statistical function of time- and space-distributed wave energy to particular pipes. In this paper the acoustical and musical design propositions and solutions, as parts of the multidiscipline design process, will be presented.
Convention Paper 7147

Session P26 Tuesday, May 8 13:30 – 15:00
Foyer IK

POSTERS: INSTRUMENTATION AND MEASUREMENT

13:30

P26-1 A Wireless PDA-Based Acoustics Measurement Platform—*Petros Alexandridis, Nicolas-Alexander Tatlas, Panos Hatziantoniou, John Mourjopoulos*, University of Patras, Patras, Greece

The proposed platform allows acoustic measurements via a flexible, portable system, based on commercially available hardware, such as a personal digital assistant (PDA) equipped with a wireless adapter and a digital audio capture card as well as a personal computer (PC) interconnected with off-the-shelf wireless networking hardware. Using this hardware, three software applications were implemented: (i) a device driver that handles the communication of the digital audio capture card with the PDA; (ii) a PDA application that realizes the WiFi connection with the personal computer, also incorporating a recording function that captures data and presents the user with their analysis; and (iii) the personal computer application that initiates the playback sequence as dictated by the connected PDA. The system can assist the fast measurement of large spaces. Room Impulse Response (RIR) measurement tests were conducted in a laboratory room, in order to evaluate the effectiveness and functionality of the measurement system.
Convention Paper 7148

13:30

P26-2 Nonlinear Cross Talk in Personal Computer-Based Audio Systems—*R. Allan Belcher, Jonathon Chambers*, Cardiff University, Cardiff, Wales, UK

International Electrotechnical Commission (IEC) and AES standards provide comprehensive tests for the performance of audio analog to digital (ADC) and digital to analog (DAC) converters for both consumer and professional applications. It is usually assumed that the ADC is more likely to degrade audio sound quality than the DAC. Tests on two samples of a professional quality PC-based audio system are presented that show that a stereo DAC can introduce unexpected nonlinear effects. These results suggest that a future revision of the standards should include a measure of interchannel nonlinear cross talk in the stereo DAC. Results are presented and an intermodulation distortion (IMD) loop test proposed to enable this measurement to be made with precision.
Convention Paper 7149

13:30

P26-3 A Low-Distortion Fast-Settling Audio Oscillator: A Tribute to the Late Peter J. Baxandall, Analog Audio Expert—*John Vanderkooy*, University of Waterloo, Waterloo, Ontario, Canada

This paper is dedicated to the memory of Peter Baxandall, well-known for his work in audio and electronics. It is an exposition and analysis of a low-distortion fast-settling audio oscillator that he designed and built. Normal oscillators are shown to suffer from amplitude instability when the thermally-variable controlling resistance has a long time constant. The genius of the present two-integrator design is that it derives its amplitude stability from the cancellation of two square wave signals, of which one is fixed in amplitude, the other proportional to the oscillator output, with a threshold. A detailed analysis of the oscillator is presented. The result is an oscillator with a distortion below 0.01 percent and settling times of approximately 1 oscillation period. It is particularly useful in automated test equipment.
Convention Paper 7150

13:30

P26-4 Direct Current Offset and Balance for Audio Transformers Used with Paralleled Tubes or Solid State Devices—*Aristide Polisois, Pierre Touzelet*, S.E.R.E.M.E. S.A., Bondoufle Cedex, France

In May 2005 (118th AES Convention, Barcelona, Spain, Paper 6346), I described a self compensated transformer for SE audio amplifiers, designed as SC-OPT and based on the principle that an auxiliary winding (tertiary), crossed by the same current as the primary winding, opposes a magnetic flux that reduces the overall flux, produced by the direct current, to almost zero, thus leaving the whole magnetic headroom in the core, for alternating current purposes. The opposed alternating current built in the tertiary was short-circuited with a suitable capacitor. Subsequently, a dedicated auxiliary magnetic core was added to the tertiary, acting as a flux escape, to reduce the antagonism of the tertiary on the primary, which is responsible for the loss of primary inductance. An improved layout to obtain the same result was achieved with the SC-SCC-SET (Split Core-Stereo Common Circuit-Single Ended Transformer), invented by Polisois and Mariani (120th AES Convention, Paris, France, Paper 6831). This model allows significantly improving the bass range. However, to achieve the DC generated flux cancellation, it needs an external balancing device of the DC flowing in the two primaries (left and right channel), situated on the same magnetic core. The ➥

transformer described hereafter named 4x4 SC-SCC-SET (4x4 Self Compensated-Single Common Circuit-Single Ended Transformer), overcomes this requirement. It also has many novel features, proceeding from the adopted arrangement of the windings.
Convention Paper 7151

13:30

P26-5 Acoustical Issues and Proposed Improvements for NASA Spacesuits—*Durand R. Begault, James L. Hieronymus*, NASA Ames Research Center, Moffett Field, CA, USA

This paper reviews current acoustical issues relevant to the design of future NASA spacesuits, based on measurements conducted in the current Mark III advanced prototype surface suit and proposes solutions for improving voice communications. Methods for mitigating problems including noise from the air supply, structure-borne noise from the suit, and detrimental acoustical reflections are reviewed.
Convention Paper 7152

13:30

P26-6 Design, Construction, and Qualification of the New Anechoic Chamber at Laboratorio de Sonido, Universidad Politécnica de Madrid—*Juan José Gómez-Alfageme, José Luis Sánchez-Bote, Elena Blanco-Martin*, Universidad Politécnica de Madrid, Madrid, Spain

The year 2005 has seen the design and construction of a new anechoic chamber at Laboratorio de Sonido of the Universidad Politécnica de Madrid. This new chamber has a free volume of 70 cubic meters and is built with rock wool wedges covered with a porous cotton cloth. The chamber cutoff frequency is 150 Hz. The chamber has been qualified according to that established in the ISO 3745 standard for the determination of the maximum distance between the sound source and the measurement position where the inverse square law is observed, within some tolerance. For the qualification, different types of excitation signals have been used as pure tones, broadband noise, narrow band noise, and pseudorandom sequences MLS.
Convention Paper 7153

TECHNICAL TOUR 11
Music Studio Thomas Rabitsch
Tuesday, May 8, 13:30 – 16:30

Studio Rabitsch is the home of producer/arranger/musician Thomas Rabitsch who started as a keyboard player in the Viennese Underground Music scene, later palying with Austria's most popular and successful popstar, Falco. Among various pop-music projects for a vast number of Austrian artists, he works for TV and radio companies, as a music supervisor and composer as well as arranger. The main studio features an AMS/Neve Capricorn console with 5.1 Dynaudio Air20 Monitoring and an extensive ProTools and Nuendo rig. Visitors will be able to get a first impression of a current surround sound remix project of a Falco-Live-Concert.

Tutorial 12 **Tuesday, May 8**
14:00 – 17:00 **Room H**

LOUDSPEAKER NONLINEARITIES—CAUSES, PARAMETERS, SYMPTOMS

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

This tutorial addresses the relationship between nonlinear distortion measurements and nonlinearities that are the physical causes for signal distortion in loudspeakers, headphones, micro-speakers, and other transducers. Using simulation techniques characteristic symptoms are identified for each nonlinearity and presented systematically in a guide for loudspeaker diagnostics. This information is important for understanding the implications of nonlinear parameters and for performing measurements that describe the loudspeaker more comprehensively. The practical application of the new techniques are demonstrated on practical examples.

Tutorial 13 **Tuesday, May 8**
14:00 – 16:00 **Room P**

AUDIO QUALITY IN NEW LOW BIT-RATE DISTRIBUTION SYSTEMS

Presenters: **Lars Jonsson**, EBU/Swedish Radio
Gregory Massey, APT Ltd., Ireland, UK
Gerhard Stoll, IRT Munich, Germany

Audio quality in new low bit-rate distribution systems—with a broadcasters cascade perspective with many re-encodings at too low bit rates. What is the solution to this problem?

In all new digital distribution systems broadcasters are faced with the problem of using perceptual coding systems in all stages of the broadcasting chain, from early capturing to contribution over editing and on-air. Systems in the home are now also using recording with digital media with low bit rate systems. The resulting overall quality is often degraded by the cascading artifacts in more than five steps of coding and re-encoding.

This tutorial discusses state of the art methods and listening test results within the EBU to overcome these problems.

TECHNICAL TOUR 12
Vienna State Opera
Tuesday, May 8, 14:00 – 16:00

The Vienna State Opera is one of the most famous and important opera houses in the world. With about 50 operas and 20 ballet performances every season it is a true repertoire theater with a special bonus: the orchestra consists of members of the world-renowned Vienna Philharmonic! The tour through this prestigeous house will include a behind-the-scenes look of the stage and the most prominent mechanical installations and will continue to various well-known halls within the State Opera like the “Gobelin-Saal,” the “Marmorsaal,” and the “Schwindfoyer.” There is also a fully-featured organ to be seen in the “Orgelsaal”before the visitors will get an in-depth view of the elaborate electro-acoustical possibilities like the tailor-made TOA ix-3000 theater-console, the computer-controlled amplifier- and sequence switching as well as the enormous loudspeaker installation. Professor Wolfgang Fritz, the head of the electroacoustical departement, will be the guide on this exciting tour through one of the premier sights of Vienna and Austria.

Session P27 Tuesday, May 8 15:30 – 17:00
Foyer IK

POSTERS: ANALYSIS AND SYNTHESIS OF SOUND

15:30

P27-1 Time Signature Detection by Using a Multiresolution Audio Similarity Matrix—*Mikel Gainza, Eugene Coyle*, Dublin Institute of Technology, Dublin, Ireland

A method that estimates the time signature of a piece of music is presented. The approach exploits the repetitive structure of most of the music, where the same musical bar is repeated in different parts of a piece. The method utilizes a multiresolution audio similarity matrix approach, which allows comparisons between longer audio segments (bars) by combining comparisons of shorter segments (fraction of a note). The time signature method only depends on musical structure, and does not depend on the presence of percussive instruments or strong musical accents.
Convention Paper 7154

15:30

P27-2 Signal Processing Parameters for Tonality Estimation—*Katy Noland, Mark Sandler*, Queen Mary, University of London, London, UK

All musical audio feature extraction techniques require some form of signal processing as a first step. However, the choice of low level parameters such as window sizes is often disregarded, and arbitrary values are chosen. We present an investigation into the effects of low level parameter choice on different tonality estimation algorithms and show that the low level parameters can make a significant difference to the results. We also show that the choice of parameters is algorithm specific, so optimization is required for each different method.
Convention Paper 7155

15:30

P27-3 Audio Effects for Real-Time Performance Using Beat Tracking—*A. M. Stark, M. D. Plumbley, M. E. P. Davies*, Queen Mary, University of London, London, UK

We present a new class of digital audio effects that can automatically relate parameter values to the tempo of a musical input in real-time. Using a beat tracking system as the front end, we demonstrate a tempo-dependent delay effect and a set of beat-synchronous low frequency oscillator (LFO) effects including auto-wah, tremolo, and vibrato. The effects show better performance than might be expected as they are blind to certain beat tracker errors. All effects are implemented as VST-plugins that operate in real-time, enabling their use both in live musical performance and the off-line modification of studio recordings.
Convention Paper 7156

15:30

P27-4 JAVA Library for Automatic Musical Instruments Recognition—*Piotr Aniola, Ewa Lukasik*, Poznan University of Technology, Poznan, Poland

The paper presents an open source Java library intended for analysis and classification of musical instrument sounds. It consists of two main parts: one devoted for feature extraction and the second performing musical instruments recognition and similarity assessment. The project's plug-in based structure enables further extendibility of both modules. In the current version two separate sound modeling algorithms have been implemented: k-means and Gaussian Mixture Models. The software has been created for the purpose of recognition of different exemplars of the same type of instruments and validated for electric guitars, guitar-amplifiers, and violins. The Java project follows the latest trends in software engineering. It enables the developer to easily create highly usable, reliable, and extendable programs. The entire software discussed here is open source.
Convention Paper 7157

15:30

P27-5 Extraction of Long-Term Rhythmic Structures Using the Empirical Mode Decomposition—*Peyman Heydarian, Joshua D. Reiss*, Queen Mary, University of London, London, UK

Long-term musical structures provide information concerning rhythm, melody, and the composition. Although highly musically relevant, these structures are difficult to determine using standard signal processing techniques. In this paper a new technique based on the time-domain empirical mode decomposition is explained. It decomposes a given signal into its constituent oscillations that can be modified to produce a new version of the signal. It enables us to analyze the long-term metrical structures in musical signals and provides insight into perceived rhythms and their relationship to the signal. The technique is explained, and results are reported and discussed.
Convention Paper 7158

Session P28 Tuesday, May 8 15:30 – 16:30
Room I

AUDIO IN COMPUTERS (GAMES, INTERNET, AND DESKTOP COMPUTER AUDIO)

Chair: **Gregor Widholm**, Musikuniversität Wien, Austria

15:30

P28-1 A Distributed Real-Time Virtual Acoustic Rendering System for Dynamic Geometries—*Raine Kajastila,*[1] *Samuel Siltanen,*[1] *Peter Lundén,*[2] *Tapio Lokki,*[1] *Lauri Savioja*[1]
[1]Helsinki University of Technology, Espoo, Finland
[2]Interactive Institute, Stockholm, Sweden

A novel room acoustic simulation system capable of producing interactive sound environ- ➥

ments in dynamic and complex 3-D geometries is introduced. The system is distributed to several modules that share the same 3-D geometry. All changes made by one module are immediately updated in all other modules in real time. The auralization tools of the system include a geometry reduction tool, a beam tracing algorithm, and a sound rendering application. The geometry reduction simplifies 3-D models for a beam tracing module that forwards direct sound and early reflection paths for sound rendering. The sound rendering application contains an automatic estimation of late reverberation parameters, based on early reflections.
Convention Paper 7160

16:00

P28-2 To Create Spatial Auditory Events via More Channel Headphones Related on Portable 5.1 / 5.0 Surround Reproductions of Sound—*Florian Koenig*, ULTRASONE AG, Tutzing, Germany

In the future portable surround devices will be the successor of stereo applications in games, cell/mobile phone applications or mp3-players. This portable technique evolution needs worldwide compatible electro-acoustic headphones without any "mean" HRTF (head-related transfer function) or binaural DSP (digital signal processing). Such headphones offer natural and realistic 3-D images of sound with a minimum of elevation effects and a virtual distance perception front and back. The individualized HRTF, for instance, made by a near-field offset/de-centric headphone loudspeaker placement at the earcups offers measurable advantages in regards to a "mean" HRTF via DSP. Past AES conferences presented papers stating further problems such as the compatible downmix of 5.1 to 2.0 signals mainly due to DSP based headphones, but also to remind to the 4.0 downmix seeing 4-channel surround sound headphones. As well discussing the connections of headphones we need to remember that there should be available a standardized vario 3.5" jack for stereo and more channel signal supply. This paper presents the basics of how to realize this spatial auditory event via 4-channel headphones plus the mode of a direct audio signal supply that is loudspeaker compatible.
Convention Paper 7161

Workshop 29 **Tuesday, May 8**
16:00 – 18:00 **Room 631/632**

HANDS-ON DEMONSTRATION: SUBJECTIVE EVALUATION BY REFERENCE RECORDINGS WITH SURROUND MAIN MIKINGS FOR CLASSICAL ORCHESTRA

Presenters: **Hideo Irimajiri**, Mainichi Broadcasting Corp.
Toru Kamekawa, Tokyo National University of Fine Arts and Music
Masayuki Mimura, Yomiuri Telecasting Corp.
Hideaki Nishida, Asahi Broadcasting Corp.
Koichi Ono, Kansai Telecasting Corp.

There are many different setups for surround main microphones for classical music and orchestra. But it is very difficult to research and study them practically and academically under identical conditions and judge their performance. Consequently the AES Japan Surround Study Project has been organized and put into practice after one year of preparation. It was organized around 10 broadcasters and 2 universities; 12 manufacturers supported by HBF provided financial support. There were 15 different combinations of main and ambience microphone setups that were recorded on 96 channels independently in Pro Tools HD at 24 bit / 96-kHz. The musical examples were performed by the Osaka Philharmonic Orchestra on September 24-27, 2006.

In this workshop each individual setup will be played back. Participants will have the opportunity for feedback in a listening test environment, and the data will be collected for subjective evaluation.