AES 113th Convention Program

2002 October 5–8

Los Angeles Convention Center
Los Angeles, California, USA

Session A Saturday, October 5 9:00 am–11:30 am
Room 404AB

TRANSDUCERS, PART 1

Chair: **Marshall Buck**, Gibson Labs, Redondo Beach, CA, USA; Psychotechnology Inc., Los Angeles, CA, USA

9:00 am

A-1 Which Loudspeaker Parameters Are Important to Create the Illusion of a Live Performance in the Living Room?— *Siegfried Linkwitz*, Linkwitz Lab, Corte Madera, CA, USA

The preference in loudspeaker product design is for a small size, while preserving maximum low-frequency extension and output volume. If the goal is to create a realistic reproduction of a live event, then certain speaker parameters must be adequately controlled, such as volume displacement, intermodulation distortion, stored energy, and off-axis frequency response. Components must be carefully selected for low distortion performance. Parameters like phase linearity and cabinet diffraction are sometimes overrated. Multichannel loudspeaker setups require propagation delay correction and bass management if not all of the loudspeakers cover the full frequency range. These issues are reviewed at the advent of high resolution surround sound. The new technology can only fulfill its promise and expand into more than a niche market if capable loudspeakers are widely available.
Convention Paper 5637

9:30 am

A-2 Characterizing the Amplitude Response of Loudspeaker Systems—*Allan Devantier*, Harman International Industries, Inc., Northridge, CA, USA

The amplitude response of a loudspeaker system is characterized by a series of spatially averaged measurements. The proposed approach recognizes that the listener hears three acoustical events in a typical domestic environment: the direct sound, the early arrivals, and the reverberant sound field. A survey of fifteen domestic multichannel installations was used to determine the typical angle of the direct sound and the early arrivals. The reflected sound that arrives at the listener after encountering only one room boundary is used to approximate the early arrivals, and the total sound power is used to approximate the reverberant sound field. Two unique directivity indices are also defined, and the in-room response of the loudspeaker is predicted from anechoic data.
Convention Paper 5638

10:00 am

A-3 Graphing, Interpretation, and Comparison of Results of Loudspeaker Nonlinearity Measurement—*Alexander Voishvillo, Alex Terekhov, Gene Czerwinski, Sergei Alexandrov*, Cerwin-Vega Inc., Simi Valley, CA, USA

Harmonic distortion and THD do not convey sufficient information about nonlinearity in loudspeakers and horn drivers. Multitone stimulus and Gaussian noise produce more informative nonlinear response. Reaction to Gaussian noise can be transformed into a coherence or incoherence function. They provide information about nonlinearity in the form of easy-to-grasp frequency-dependent curves. Alternatively, the results of multitone measurement are difficult to interpret, compare, and overlay. A new method of depicting the results of multitone measurements has been developed. The distortion products are averaged in a moving frequency window. The result of the measurement is a single, continuous, frequency-dependent curve that takes into account the level of distortion products and their density. The curves can be easily overlaid and compared. Future development of a new method may lead to correlation between the level of distortion curves and the audibility of nonlinear distortion.
Convention Paper 5639

10:30 am

A-4 The Effects of Voice-Coil Axial Rest Position on Amplitude Modulation Distortion in Loudspeakers—*Ryan J. Mihelich*, Harman/Becker Automotive Systems, Martinsville, IN, USA

The magnetic field in the air gap of a conventional loudspeaker motor is often an asymmetric nonlinear function of axial position. Placement of the voice-coil into this asymmetrical field yields an asymmetric nonlinear force-factor, Bl, which is a primary cause of amplitude modulation distortion in loudspeakers. Adjustment of the rest position of the voice-coil in this field can alter the nature of this modulation distortion. Common practice is to nominally set the voice-coil at the geometric center of

the gap or at the position generating maximum Bl. A time-domain nonlinear simulator has been used to investigate effects of voice-coil placement in an asymmetric flux field on amplitude modulation distortion.
Convention Paper 5640

11:00 am

A-5 Nonlinearity in Horn Drivers—Where the Distortion Comes From?—*Alexander Voishvillo*, Cerwin-Vega Inc., Simi Valley, CA, USA

Compared to other components of professional sound systems (omitting free propagation distortion), horn drivers have the worst nonlinear distortion. Some of the driver's distortions can be mitigated by proper mechanical measures. However, distortions caused by nonlinear air compression and propagation are inherent to any horn driver. In this paper the comparison of nonlinear distortions caused by different sources is carried through measurements and modeling. The new dynamic model of the compression driver is based on the system of nonlinear differential and algebraic equations. Complex impedance of an arbitrary horn is considered by turning the impedance into a system of differential equations describing the pressure and velocity at the horn's throat. The comparison is carried out using harmonic distortion and the reaction to multitone stimulus.
Convention Paper 5641

Session B Saturday, October 5 9:00 am–11:00 am Room 406AB

SIGNAL PROCESSING, PART 1

Chair: **Rob Maher,** Montana State University, Bozeman, MT, USA

9:00 am

B-1 A Talker-Tracking Microphone Array for Teleconferencing Systems—*Kazunori Kobayashi, Ken'ichi Furuya, Akitoshi Kataoka*, NTT, Musashino-shi, Tokyo, Japan

We propose a beamforming method that is applicable to near sound fields where a filter-and-sum microphone array maintains better quality for the target sound than the conventional delay-and-sum array. We also describe a real-time implementation that includes steering of the beam to detected talker locations. With the use of a microphone array, our system also cuts levels of noise to achieve high-quality sound acquisition. Furthermore, it allows the talker to be in any position. Computer simulation and experiments show that our method is effective in teleconferencing systems.
Convention Paper 5642

9:30 am

B-2 An Alternative Model for Sound Signals Encountered in Reverberant Environments; Robust Maximum Likelihood Localization and Parameter Estimation Based on a Sub-Gaussian Model—*Panayiotis G. Georgiou, Chris Kyriakakis*, University of Southern California, Los Angeles, CA, USA

In this paper we investigate an alternative to the Gaussian density for modeling signals encountered in audio environments. The observation that sound signals are impulsive in nature, combined with the reverberation effects commonly encountered in audio, motivates the use of the sub-Gaussian density. The new sub-Gaussian statistical model and the separable solution of its maximum likelihood estimator are derived. These are used in an array scenario to demonstrate with both simulations and two different microphone arrays the achievable performance gains. The simulations exhibit the robustness of the sub-Gaussian-based method while the real world experiments reveal a significant performance gain, supporting the claim that the sub-Gaussian model is better suited for sound signals.
Convention Paper 5643

10:00 am

B-3 Imperceptible Echo for Robust Audio Watermarking—*Hyen-O Oh*[1]*, Jong Won Seok*[2]*, Jin Woo Hong*[2]*, Dae-Hee Youn*[1]

[1]Yonsei University, Seoul, Korea
[2]Electronics & Telecommunications Research Institute (ETRI), Daejon, Korea

In echo watermarking, the effort to improve robustness often conflicts with the requirement of imperceptibility. There have been inherent trade-offs in general audio watermarking techniques. In this paper we challenge the development of imperceptible but detectable echo kernels being directly embedded into the high-quality audio signal. Mathematical and perceptual characteristics of echo kernels are analyzed in a frequency domain. Finally, we can obtain a greater flat frequency response in perceptually significant bands by combining closely located positive and negative echoes. The proposed echo makes it possible to improve the robustness of an echo watermark without breaking the imperceptibility.
(Paper not presented at convention, but Convention Paper 5644 is available.)

10:30 am

B-4 New High Data Rate Audio Watermarking Based on SCS (Scalar Costa Scheme)—*Frank Siebenhaar*[1]*, Christian Neubauer*[1]*, Robert Bäuml*[2]*, Jürgen Herre*[1]

[1]Fraunhofer Institute for Integrated Circuits, Erlangen, Germany
[2]Friedrich-Alexander University, Erlangen, Germany

Presently, distribution of audio material is no longer limited to physical media. Instead, distribution via the Internet is of increasing importance. In order to attach additional information to the audio content, either for forensic or digital rights management purposes or for annotation purposes, watermarking is a promising technique since it is independent of the audio format and transmission technology.

State-of-the-art spread spectrum watermarking systems can offer high robustness against unintentional and intentional signal modifications. However, their data rate is typically comparatively low, often below 100 bit/s. This paper describes the adaptation of a new watermarking scheme called Scalar Costa Scheme (SCS), which is based on dithered quantization of audio signals. In order to fulfill the demands of high quality audio signal processing, modifications of the basic SCS, such as the introduction of a psychoacoustic model and new algorithms to determine quantization intervals, are required. Simulation figures and results of a sample implementation, which show the potential of this new watermarking scheme, are presented in this paper along with a short theoretical introduction to the SCS watermarking scheme.
Convention Paper 5645

Workshop 1 Saturday, October 5 9:00 am–11:00 am
Room 408A

STEREO AND SURROUND MICROPHONE TECHNIQUES (TUTORIAL)

Chair: **Geoff Martin,** McGill University, Montreal, Quebec, Canada, and Bang & Olufsen a/s, Struer, Denmark

Presenters: *Michael Bishop,* Telarc International
Doug Botnick, Independent Engineer/Producer
John Eargle, JME/JBL Professional
Richard King, Sony Music Studios
Mick Sawaguchi, NHK Broadcasting Center, Programme Production Operations, Sound Division

This tutorial workshop includes a panel of leading industry professionals from the areas of pop, classical and film music, as well as radio drama. Issues to be discussed include the characteristics of various microphone configurations as well as upward-and-downward compatibility considerations. This workshop, which has been organized by the AES Education Committee, will be of benefit to audio engineers of all backgrounds, including students.

11:10 am

Technical Committee Meeting on Microphones

Special Event
WHEN VINYL RULED!

Saturday, October 5	12:00 noon–6:00 pm
Sunday, October 6	12:00 noon–6:00 pm
Monday, October 7	10:00 am–6:00 pm
Tuesday, October 8	10:00 am–4:00 pm

Room 308A

Organizers: **Wes Dooley,** Audio Engineering Associates, Pasadena, CA, USA
Dale Manquen, Consultant, Thousand Oaks, CA, USA

Vinyl records ushered in an age of consumer high-fidelity. Magnetic tape recording was the technology that made practical the production of long-playing records. Assembling a component high-fidelity system became a widespread hobby for many of us, and for decades fueled the development of loudspeakers, electronics, and microphones. Listening to recorded music became a part of many people’s daily life.

Our historical exhibit will offer attendees an overview of production technology during the age of vinyl and spotlight its relevance to current music production.

The analog audio presentation will take the visitor from Ampex’s first machine, the 30 in/s, one-quarter-inch, full-track recordings to the leading edge technology of Michael Spitz’s contemporary 30 in/s, one-inch, two-track mastering decks. Capitol Records will have the spotlight on Saturday afternoon when Carson Taylor talks about and plays examples of Tower and location recordings of the fifties, sixties, and seventies. Other key highlights will include Jim Webb’s presentation of “12 Landmark Microphones” that made history; Kevin Gray’s and Stan Ricker’s presentation of vinyl disc mastering and record manufacturing; Paul McManus on the development and history of powered monitors from the fifties; and Ken Hirsch and David Gordon on the development of mixing console technology.

Saturday, Sunday, Monday, and Tuesday come up to Demo Room Row and hear good sounds, vintage and contemporary. Be sure to check the daily schedule outside Room 308A for exact presentation times.

Special Event
FREE HEARING SCREENINGS CO-SPONSORED BY THE AES AND HOUSE EAR INSTITUTE

Saturday, October 5	12:00 noon–6:00 pm
Sunday, October 6	10:00 am–6:00 pm
Monday, October 7	10:00 am–6:00 pm
Tuesday, October 8	10:00 am–4:00 pm

Booth 1881

Attendees are invited to take advantage of a free hearing screening co-sponsored by the AES and House Ear Institute. Four people can be screened simultaneously in the mobile audiological screening unit located on the exhibit floor. A daily sign-up sheet at the unit will allow individuals to reserve a screening time for that day. This hearing screening service has been developed in response to a growing interest in hearing conservation and to heighten awareness of the need for hearing protection and the safe management of sound.

Special Event
SONGWRITERS SHOWCASE: “IT’S ALL ABOUT THE SONG”

October 5–October 7	1:00 pm–8:00 pm
October 8	12:00 noon–3:00 pm

South Hall Lobby

Produced by: **Claudia Koal**

The Audio Engineering Society will present the Songwriters Showcase, “It’s All About the Song.” This event marks the third time the AES will incorporate live original music into its convention schedule. The Songwriters Showcase is a means of acknowledging the song as a key element in the motivation for technological advancements in audio. The roster of performers includes international composers and songwriters. The event features a wide spectrum of musical genres. Members of performing rights organizations ASCAP, BMI, and SESAC will perform original works at daily performances. A complete schedule of performers will be posted in the registration, press, and exhibition areas throughout the four-day convention. Enjoy special performances during the AES Mixers each evening.

Education Event
STUDENT DELEGATE ASSEMBLY 1
Saturday, October 5, 1:00 pm–2:30 pm
Room 402A

Chair: **Scott Cannon,** Stanford University Student Section, Stanford, CA, USA

Vice Chair: **Dell Harris**, Hampton University Student Section, Hampton, VA, USA

All students and educators are invited to participate in a discussion of opportunities in the audio field and issues of importance to audio education. A descriptive overview of conference events for students will also be given. This opening meeting of the Student Delegate Assembly will introduce the candidates for the coming year's election for chair and vice chair of the North and Latin America Regions. Each AES regional vice president may present two candidates for the election to be held at the closing meeting of the SDA. Election results and ➥

Recording Competition and Poster Awards will be given at the Student Delegate Assembly 2 on Tuesday, October 8, at 10:00 am, in Room 402A.

Session C Saturday, October 5 2:00 pm–6:00 pm Room 404AB

TRANSDUCERS, PART 2

Chair: **Steve Hutt**, Harman/Becker Automotive Systems, Martinsville, IN, USA

2:00 pm

C-1 The Bidirectional Microphone: A Forgotten Patriarch—*Ron Streicher*[1], *Wes Dooley*[2]

[1]Pacific Audio-Visual Enterprises, Pasadena, CA, USA
[2]Audio Engineering Associates, Pasadena, CA, USA

Despite being one of the progenitors of all modern microphones and recording techniques, the bidirectional pattern is still not very well understood. Its proper and effective use remains somewhat of a mystery to many recording and sound reinforcement engineers. In this paper the bidirectional microphone is examined from historical, technical, and operational perspectives. We review how it developed and exists as a fundamental element of almost all other single-order microphone patterns. In the course of describing how this unique pattern responds to sound waves arriving from different angles of incidence, we show that it very often can be successfully employed where other more commonly-used microphones cannot.
Convention Paper 5646

2:30 pm

C-2 Gaussian Mixture Model-Based Methods for Virtual Microphone Signal Synthesis—
Athanasios Mouchtaris, Shrikanth S. Narayanan, Chris Kyriakakis, University of Southern California, Los Angeles, CA, USA

Multichannel audio can immerse a group of listeners in a seamless aural environment. However, several issues must be addressed such as the excessive transmission requirements of multichannel audio, as well as the fact that to date only a handful of music recordings have been made with multiple channels. Previously, we proposed a system capable of synthesizing the multiple channels of a virtual multichannel recording from a smaller set of reference recordings. In this paper these methods are extended to provide a more general coverage of the problem. The emphasis here is on time-varying filtering techniques that can be used to enhance particular instruments in the recording, which is desired in order to simulate virtual microphones in several locations close to and around the sound source.
Convention Paper 5647

3:00 pm

C-3 Driver Directivity Control by Sound Redistribution—*Jan Abildgaard Pedersen, Gert Munch*, Bang & Olufsen a/s, Struer, Denmark

The directivity of a single loudspeaker driver is controlled by adding an acoustic reflector to an ordinary driver. The driver radiates upward and the sound is redistributed by being reflected off the acoustic reflector. The shape of the acoustic reflector is nontrivial and yields an interesting and useful directivity both in the vertical and horizontal plane. Two-dimensional FEM simulations and 3-D BEM simulations are compared to free field measurements performed on a loudspeaker using the acoustic reflector. The resulting directivity is related to results of previously reported psychoacoustic experiments.
Convention Paper 5648

3:30 pm

C-4 Pressure Response of Line Sources—*Mark S. Ureda*, JBL Professional, Northridge, CA, USA

The on-axis pressure response of a vertical line source is known to decrease at 3 dB per doubling of distance in the near field and at 6 dB in the far field. This paper shows that the conventional mathematics used to achieve this result understates the distance at which the -3 dB to -6 dB transition occurs. An examination of the pressure field of a line source reveals that the near field extends to a greater distance at positions laterally displaced from the centerline, normal to the source. The paper introduces the endpoint convention for the pressure response and compares the on-axis response of straight and hybrid line sources.
Convention Paper 5649

4:00 pm

C-5 High-Frequency Components for High-Output Articulated Line Arrays—*Doug Button*, JBL Professional, Northridge, CA, USA

The narrow vertical pattern achieved by line arrays has prompted much interest in the method for many forms of sound reinforcement in recent years. The live sound segment of the audio community has used horns and compression drivers for sound reinforcement for several decades. To adopt a line array philosophy to meet the demands of high level sound reinforcement, requires an approach that allows for the creation of a line source from the output of compression drivers. Additionally, it is desired that the line array take on different vertical patterns dependent upon use. This requires the solution to allow for the array to be articulated. Outlined in this paper is a waveguide/compression driver combination that is compact and simple in approach and highly suited for articulated arrays.
Convention Paper 5650

4:30 pm

C-6 High-Efficiency Direct-Radiator Loudspeaker Systems—*John Vanderkooy*[1], *Paul M. Boers*[2]

[1]University of Waterloo, Waterloo, Ontario, Canada
[2]Philips Research Labs, Eindhoven, The Netherlands

Direct-radiator loudspeakers become more efficient as the total magnetic flux is increased, but the accompanying equalization and amplifier modify the gains thus made. We study the combination of an efficient high-Bl driver with several amplifier types, including a highly efficient class-D amplifier. Comparison is made of a typical simulated driver, excited with a few different amplifier types, using various audio signals. The comparison is quite striking as the Bl value of the driver increases, significantly favoring the class-D amplifier.
Convention Paper 5651

5:00 pm

C-7 Audio Application of the Parametric Array—Implementation through a Numerical Model—
Wontak Kim[1], *Victor W. Sparrow*[2]

[1]Bose Corporation, Framingham, MA, USA
[2]Pennsylvania State University, University Park, PA, USA

Implementing the parametric array for audio applications is examined through numerical modeling and analytical approximation. The analytical solution of the nonlinear wave equation is used to provide guidelines on the design parameters of the parametric array. The solution relates the source size, input pressure level, and the carrier frequency to the audible signal response including the output level, beam width, and length of the interaction region. A time domain finite difference code that accurately solves the KZK nonlinear parabolic wave equation is used to predict the response of the parametric array. The accuracy of the numerical model is established by a simple parametric array experiment. In considering the implementation issues for audio applications of the parametric array, the emphasis is given to the poor frequency response and the harmonic distortion. Signal processing techniques to improve the frequency response and the harmonic distortion are suggested and implemented through the numerical simulation.
Convention Paper 5652

5:30 pm

C-8 Implementation of Straight-Line and Flat-Panel Constant Beamwidth Transducer (CBT) Loudspeaker Arrays Using Signal Delays—*D. B. (Don) Keele, Jr.*, Harman/Becker Automotive Systems, Martinsville, IN, USA

Conventional CBT arrays require a driver configuration that conforms to either a spherical cap-curved surface or a circular arc. CBT arrays can also be implemented in flat-panel or straight-line array configurations using signal delays and Legendre function shading of the driver amplitudes. Conventional CBT arrays do not require any signal processing except for simple frequency-independent shifts in loudspeaker level. However, the signal processing for the delay-derived CBT configurations, although more complex, is still frequency independent. This is in contrast with conventional constant-beamwidth flat-panel and straight-line designs which require strongly frequency-dependent signal processing. Additionally, the power response roll-off of the delay-derived CBT arrays is one-half the roll-off rate of the conventional designs, i.e., 3- or 6-dB/octave (line or flat) for the CBT array versus 6- or 12-dB/octave for the conventional designs. Delay-derived straight-line CBT arrays also provide superior horizontal off-axis response because they do not exhibit the ±90 degree right-left off-axis sound pressure buildup or bulge as compared to conventional circular-arc CBT arrays. In comparison to conventional CBT arrays, the two main disadvantages of delay-derived straight-line or flat-panel CBT arrays are 1) the more complicated processing required, which includes multiple power amplifiers and delay elements; and 2) the widening of the polar response at extreme off-axis angles particularly for arrays that provide wide coverage with beamwidths greater than 60 degrees. This paper illustrates its findings using numerical simulation and modeling.
Convention Paper 5653

Session D Saturday, October 5 2:00 pm–4:30 pm
Room 406AB

SIGNAL PROCESSING, PART 2

Chair: **Robert Bristow-Johnson,** Wave Mechanics, Burlington, VT, USA

2:00 pm

D-1 Automatic Design of Sound Synthesis Techniques by Means of Genetic Programming—*Ricardo A. Garcia*, Chaoticom, Hampton Falls, NH, USA

Design of sound synthesis techniques (SST) is a difficult problem. It is usually assumed that it requires human ingenuity to find a suitable solution. Many of the SSTs commonly used are the fruit of experimentation and long refinement processes. An automated approach for design of SSTs is proposed. The problem is stated as a search in the multidimensional SST space. It uses genetic programming (GP) to suggest valid functional forms and standard optimization techniques to fit their internal parameters. A psychoacoustic auditory model is used to compute the perceptual distance between the target and test sounds. The developed AGeSS (automatic generator of sound synthesizers) system is introduced, and a simple example of the evolved SSTs is shown.
Convention Paper 5654

2:30 pm

D-2 A Consumer-Adjustable Dynamic Range Control System—*Keith A. McMillen*, Octiv, Inc., Berkeley, CA, USA

Advances in technology have afforded listeners an available dynamic range in excess of 120 dB. While impressive in proper concert halls and listening rooms, large dynamic ranges are not always realistic for all environments and musical styles. This paper describes a practical multiband dynamics processor software object that can reside in low cost consumer products and allow the user to adjust dynamic range to fit his or her taste and listening environment.
Convention Paper 5655

3:00 pm

D-3 A Simple, Efficient Algorithm for Reduction of Hiss Amplification Under High Dynamics Compression—*Guillermo García*, CCRMA, Stanford, CA, USA

We present a very simple, effective, and computationally efficient algorithm to reduce the typical hiss amplification (or breathing) artifact of dynamics compressors working under high compression ratios. The algorithm works in the time domain, is very easy to implement, has a very low computational cost, and requires little program memory, therefore being of special interest for consumer-audio applications.
(Paper not presented at convention, but Convention Paper 5656 is available.)

3:30 pm

D-4 Adaptive Predistortion Filter for Linearization of Digital PWM Power Amplifier Using Neural Networks—Minki Yang[1], Jong-Hoon Oh[2]

[1]Pulsus Technologies, Inc., Seoul, Korea
[2]Pohang University of Science and Technology, Pohang, Korea

The paper presents a method to compensate for nonlinear distortion of digital pulse width modulation (PWM) power amplifiers by prefiltering the input signals using artificial neural networks. We first construct a model of the digital amplifier using artificial neural networks. Using this model, the artificial neural network model of a predistortion filter is trained such that the combined sys-

tem, the digital amplifier, and predistortion filter produces an output that is linearly proportional to the input. The simulation results show that the artificial neural networks of the predistortion filter can effectively correct the nonlinear distortion of the digital amplifier.
Convention Paper 5657

4:00 pm

D-5 Objective Measures of the Quality of Speech Transmission in a Real Mobile Network—Measuring, Estimate, and Prediction Method— *Dragana Sagovnovic*, Telekom Srbija, Belgrade, Yugoslavia

Contemporary means of communication (e.g., mobile telephony) have brought new limitations that telecom operators have to take into consideration. One of them is the fact that the types of deterioration of speech quality, perceived in mobile telephony, are different from the degradations noted in fixed telephony. This paper discusses, during the process of estimating the quality of speech transmission, the comparative (objective) methods for a formed database of degradations of a real mobile communications system. The consideration of the results of the objective-method tests is based on the development of a new objective method of speech quality. The results gathered during the comparison tests have been displayed and interpreted for different types of Serbian vowels: front vowels, /e/; mid vowels, /a/; and back vowels, /u/.
(Paper not presented at convention, but Convention Paper 5658 is available.)

4:40 pm

Technical Committee Meeting on Signal Processing

Workshop 2 Saturday, October 5 2:00 pm–5:00 pm
Room 408A

NEW MEDIA FOR MUSIC—ADAPTIVE RESPONSE TO TECHNOLOGY

Chair: **Dave Davis,** QCA Mastering, Cincinnati, OH, USA

Panelists: *Bob Ludwig*, Gateway Mastering and DVD, Portland, ME
Bill McQuay, Radio Expeditions: National Public Radio and the National Geographic Society, Baltimore, MD
Bobby Owsinski, Surround Associates, Studio City, CA
Mike Sokol, JMS Productions & Fits and Starts Productions, Hagerstown, MD

In 2002 the audio industry is at a crossroad. The CD is seemingly vulnerable to piracy by file trading and easy replication. Studios and engineers are seeing their markets shrink and rates decline. New formats are available to consumers, but there are no budgets for creating content. While technology gets the blame, any solution requires new technologies to succeed. We will examine product-oriented solutions that are available today and how to create new kinds of products that expand both artistic expression and markets.

This workshop focuses on using existing technology and adapting existing models to expand the music industry with new products. It is not an esoteric romp through dot.bomb schemes or a complicated strategy requiring consumers to replace their stereo equipment and record collections. The current crisis requires us to adapt and evolve incrementally, so this workshop is about identifying successful models and applying them to what we collectively do. Many of the technologies and models that can save us have already matured and found their way into the homes of music fans. 2002 is our last, best hope to restore the health and vitality to the music industry through our products, rather than legislation or litigation.

Dave Davis (QCA Mastering/UltraInteractive) will introduce the qualities of new media and how they can be used to add value to music products. He will explain why new media strategies are essential to our industry's adaptation to current cultural and technological challenges and demonstrate a number of existing products that provide fans with a richer experience. Finally, he will present some works that were conceived for rich media environments and require a DVD or computer to experience; for many new artists these technologies not only address economic challenges, but broaden their vision and expand their palette of tools.

Bob Ludwig (Gateway Mastering and DVD) will discuss the role of the mastering studio regarding SACD, DVD-V, DVD-A and enhanced CD. He will discuss how Gateway uses it's ftp site for projects and various pro and consumer file transfer technologies (Liquid Audio, DigiProNet, Rocket etc) and explain how an integrated facility can support an artist's vision from soup to nuts in one the building.

Bill McQuay (Radio Expeditions: National Public Radio and the National Geographic Society) is going to discuss the National Geographic *Expeditions* series as a forward looking model to support rich, challenging production for programming in a radio format. The Radio Expeditions feature runs once a month on NPR’s flagship news magazine Morning Edition. It lives on in that form for fans on NPR’s web site, where it can be streamed at will. The material was expanded into a lecture/presentation series with surround sound for a live venue. A DVD is in the works, presenting an even richer version of the program. These additional modes of presentation collectively support the use of audio in a deeper, richer way than is traditionally possible on broadcast radio.

Bobby Owsinski (Surround AssociatesA) is deeply involved in the development and exploration of surround mixing and will address the role of the creative engineer in production for multiple mix targets, as well as what studios and tracking engineers can do to facilitate surround work downstream.

Mike Sokol (JMS Productions & Fits and Starts Productions) will address the opportunities created by surround for audio professionals. Mike will be discussing a specific project in radio drama that was staged live in surround, and is working to broadcast it to home theaters via Digital cable using Dolby Digital. This represents both a reborn market, and the beginnings of surround-driven audio-only content in a digital broadcast environment. Mike will also discuss other alternative 5.1 audio markets, such as Powerpoint 5.1, as well as ways to do large music concerts with 5.1 surround elements. All of these markets offer additional production and revenue streams for studios trying to justify 5.1 production gear.

5:10 pm

Technical Committee Meeting on Archiving, Restoration, and Digital Libraries

Workshop 3 Saturday, October 5 2:00 pm–5:00 pm
Room 408B

WHAT AUDIO ENGINEERS SHOULD KNOW ABOUT HUMAN SOUND PERCEPTION

Presenters: *Durand Begault*, NASA Ames Research Center, Mountain View, CA, USA
William Martens, University of Aizu, Aizuwakamatsu-shi, Japan

Audio engineering professionals are intimately familiar with the art of producing acoustical phenomena to achieve specific psychoacoustic goals for applications such as music, multimedia, speech reinforcement, etc. However, they may be less familiar with the underlying mechanisms of hearing and perception that influence their everyday decisions.

This tutorial will review fundamental concepts of human sound perception, first from a monaural hearing perspective and then from a spatial hearing perspective. Topics include pitch, loudness, timbre, sensitivity to phase; temporal resolution and temporal integration; masking; spatial perception; perceptual effects of rooms; and differences between headphone and loudspeaker reproduction. Questions and comments from the audience will be encouraged.

5:10 pm

Technical Committee Meeting on Perception and Subjective Evaluation of Audio Signals

Education Event
MENTORING PANEL: STEPS TO THE FUTURE—EFFECTIVELY USING MENTORS TO HELP BUILD A CAREER
Saturday, October 5, 2:30 pm–4:00 pm
Room 402A

Chair: **Theresa Leonard,** The Banff Centre, Banff, Alberta, Canada

Panelists: *Jim Anderson,* Independent Recording Engineer, AES Regional VP
Edwin Dolinsky, Electronic Arts
Bill Dooley, SPARS, The Village Recorder
Lynn Fuston, Skywalker Sound
Richard King, Sony Music
David Moulton, Moulton Laboratories
Julie Perez, NBC TV, Music Mixer, "Late Night with Conan O'Brien"

Thriving in the audio industry takes more than technical aptitude—it helps to have guidance from industry professionals. Learn about the benefits of mentor relationships and how to develop your own network of industry connections via mentoring. Students have a chance to learn firsthand from industry veterans by attending a mentoring panel and then signing up for one-on-one time with one of the distinguished participants.

Special Event
14TH ANNUAL GRAMMY® RECORDING SOUNDTABLE: 21ST CENTURY REALITIES
Saturday, October 5, 4:00 pm–6:00 pm
Room 411 Theater

Moderator: **Howard Massey,** On The Right Wavelength Consulting

Panelists: *Brian Garten*
Ken Jordan
George Massenburg
Jack Joseph Puig
Elliot Scheiner
Al Schmitt

The National Academy of Recording Arts & Sciences, Inc. will present the 14th Annual GRAMMY® Recording SoundTable, moderated by Howard Massey. It's the 21st century: Recording budgets are shrinking, technologies are advancing, and public taste in music is changing. How does today's producer cope with these realities? Do 24 bits truly matter when most people are listening to music in MP3 format? Is there still room for analog in today's digital world? And, with entire hit records being made on laptop computers, is the professional recording studio becoming an endangered species? Don't miss the lively discussion as our panelists attack these topics head on.

Brian Garten is engineer for the Neptunes, and the man behind the console for two of 2002's biggest hits: Nelly's "Hot in Here" and "Dilemma," featuring Kelly of Destiny's Child. He has also recorded by Beyonce Knowles, Britney Spears, Mary J. Blige, Alicia Keys, Usher, No doubt, Busta Rhymes, and Justin Timberlake.

Beatmaster Ken Jordan of the alternative-electronic duo The Crystal Method brings a variety of influences to their dance-based productions, merging hip-hop, rock, and breakbeats into a unique blend that have made them an integral part of club culture. Their debut LP, *Vegas*, was released to critical acclaim in 1997, and their latest album, *Tweekend*, was released in 2001. The two albums have sold over a million and a half copies. For the very latest on this duo, check out www.thecrystalmethod.com.

George Massenburg is known as an engineer's engineer. His pristine work with artists like Little Feat, Toto, Linda Ronstadt, Lyle Lovett, Emmylou Harris, Carly Simon, Mary Chapin Carpenter, and Aaron Neville constantly sets new standards in audio excellence, and he is a technical pioneer as well. Recipient of a Technical GRAMMY® Award for his development of innovative recording tools such as the parametric equalizer, Massenburg is actively involved in surround sound mixing and the standardization of archiving materials.

Jack Joseph Puig is one of the most sought-after engineer/producers crafting music today. He owns an astonishing collection of vintage and modern recording equipment that he uses to create a unique engineering and production sound that merges the styles of the last four decades. You can hear his work on releases by Sheryl Crow, Vanessa Carlton, John Mayer, No Doubt, Counting Crows, Green Day, Goo Goo Dolls, Hole, Shelby Lynne, Weezer, Stone Temple Pilots, Tonic, Stevie Nicks, and Robbie Williams.

Engineer extraordinaire Elliot Scheiner has been nominated for more than a dozen GRAMMY® Awards and has won five of them, including the 2000 GRAMMY® Award for Best Engineered Album, Non-Classical. He has worked with a wide variety of multiplatinum artists, including Faith Hill, Beck, Steely Dan, Fleetwood Mac, The Eagles, Van Morrison, The Doobie Brothers, America, John Fogerty and Jimmy Buffett. He recently completed a surround sound remix of the classic Queen album *A Night at the Opera*, which won the 2002 DVD award for Best Audio.

There are few people truly deserving of the term legendary, but Al Schmitt is one of them. Not only has he won an astonishing nine GRAMMY® Awards, which include the 2001 Best Engineered Album, Non-Classical Award for Diana Krall's *The Look of Love*, he has also won two Latin GRAMMY® Awards. He has produced, engineered, and/or mixed more than 150 gold and platinum records for a diverse range of artists, including Frank Sinatra, Jefferson Airplane, Steely Dan, Barbra Streisand, Henry Mancini, and Duane Eddy.

Moderator Howard Massey is a noted industry consultant and veteran audio journalist. He has written for *Home Recording*, *Surround Professional*, *EQ*, *Musician*, *Billboard*, and many other publications, and is the author of eleven books, including *Behind the Glass*, a collection of interviews with the world's top record producers. Massey has also worked extensively as an audio engineer, producer, songwriter, and touring/session musician.

Special Event
AES MIXER
October 5–October 7 6:00 pm–8:00 pm
South Hall Lobby

A series of informal Mixers will be held each evening in the South Lobby entrance to the Los Angeles Convention Center, to enable convention attendees to meet in a social atmosphere after the day's activities and catch up with friends and colleagues from the industry. There will be music featuring performances from the "Songwriter's Showcase," a cash bar, and snacks.

Session E Sunday, October 6 9:00 am–12:00 noon
Room 406AB

SIGNAL PROCESSING, PART 3

Chair: **Jayant Datta**, Discrete Labs, Syracuse, NY, USA

9:00 am

E-1 Computationally Efficient Inversion of Mixed Phase Plants with IIR Filters—*Timoleon Papadopoulos, Philip A. Nelson*, University of Southampton, Southampton, UK

Inverse filtering in a single or in multiple channels arises as a problem in a number of applications in the areas of communications, active control, sound reproduction, and virtual acoustic imaging. In the single-channel case, when the plant $C(z^{-1})$ sought to be inverted has zeros outside the unit circle in the z-plane, an approximation to the inverse $1/C(z^{-1})$ can be realized with an FIR filter if an appropriate amount of modeling delay is introduced to the system. But the closer the zeros of $C(z^{-1})$ are to the unit circle (either inside or outside it), the longer the FIR inverse has to be, typically several tens of times longer than the plant. An off-line implementation utilizing a variant of the backward-in-time filtering technique usually associated with zero-phase FIR filtering is presented. This forms the basis on which a single-channel mixed phase plant can be inverted with an IIR filter of order roughly double than that of $C(z^{-1})$, thus decimating the processing time required for the inverse filtering computation.
Convention Paper 5659

9:30 am

E-2 Optimal Filter Partition for Efficient Convolution with Short Input/Output Delay—*Guillermo García*, CCRMA, Stanford, CA, USA

A new algorithm to find an optimal filter partition for efficient long convolution with low input/output delay is presented. For a specified input/output delay and filter length, our algorithm finds the nonuniform filter partition that minimizes computational cost of the convolution. We perform a detailed cost analysis of different block convolution schemes and show that our optimal-partition finder algorithm allows for significant performance improvement. Furthermore, when several long convolutions are computed in parallel and their outputs are mixed down (as is the case in multiple-source 3-D audio rendering), the algorithm finds an optimal partition (common to all channels) that allows for further performance optimization.
Convention Paper 5660

10:00 am

E-3 Filter Morphing—Topologies, Signals, and Sampling Rates—*Rob Clark*[1, 2], *Emmanuel Ifeachor*[2], *Glenn Rogers*[1]

[1]Allen & Heath Limited, Penryn, Cornwall, UK
[2]University of Plymouth, Plymouth, UK

Digital filter morphing techniques exist to reduce audible transient distortion during filter frequency response change. However, such distortions are heavily dependent on signal content, frequency response settings, filter topology, interpolation scheme, and sampling rates. This paper presents an investigation into these issues, implementing various filter topologies using different input stimuli and filter state change scenarios. The paper identifies the mechanisms causing these distortions, specifying worst case filter state change scenarios. The effects of existing interpolator schemes, finite word length, and system sampling rates on signal distortion are presented. The paper provides an understanding of filter state change, critical in the design of filter morphing algorithms.
Convention Paper 5661

10:30 am

E-4 Evaluation of Inverse Filtering Techniques for Room/Speaker Equalization—*Scott G. Norcross, Gilbert A. Soulodre, Michel C. Lavoie*, Communications Research Centre, Ottawa, Ontario, Canada

Inverse filtering has been proposed for numerous applications in audio and telecommunications, such as speaker equalization, virtual source creation, and room deconvolution. When an impulse response (IR) is at nonminimum phase, its corresponding inverse can produce artifacts that become distinctly audible. These artifacts produced by the inverse filtering can actually degrade the signal rather than improve it. The severity of these artifacts is affected by the characteristics of the filter and the method (time or frequency domain) used to compute its inverse. In this paper objective and subjective tests were conducted to investigate and highlight the potential limitations associated with several different inverse-filtering techniques. The subjective tests were conducted in compliance with the ITU-R MUSHRA method.
Convention Paper 5662

11:00 am

E-5 Using Subband Filters to Reduce the Complexity of Real-Time Signal Processing—
J. Michael Peterson, Chris Kyriakakis, University of Southern California, Los Angeles, CA, USA

Several high quality audio applications require the use of long finite impulse response (FIR) filters to model the acoustical properties of a room. Various structures for subband filtering are examined. These structures have the ability to divide long FIR filters into smaller FIR filters that are easier to use. Two structures will be discussed to process the signals in a real-time manner, time-convolution of spectrograms, and generalized filterbanks. Also filter estimation will be discussed.
Convention Paper 5663

11:30 am

E-6 Noise Shaping in Digital Test-Signal Generation—*Stanley P. Lipshitz*[1], *John Vanderkooy*[1], *Edward V. Semyonov*[2]

[1]University of Waterloo, Waterloo, Ontario, Canada
[2]Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russia

In an earlier paper we put forth an idea to use noise-shaping techniques in the generation of digital test signals. The previous paper proposed using noise shaping around an undithered quantizer to generate sinusoidal digital test signals with spectra having error nulls at the harmonics of the signal frequency, thus making digital distortion measurements of very great dynamic range possible. We extend this idea in this present paper in a number of ways. We show a) that dither is necessary in order to suppress spurious artifacts caused by the nonlinearity of an undithered noise shaper; b) that wider and deeper nulls at the harmonic frequencies can be achieved ➥

by using higher-order noise-shaper designs; c) that IIR filter designs can moderate the increased noise power that accompanies an increased FIR filter order; and d) some other novel uses of noise shaping in digital signal generation.
Convention Paper 5664

12:10 pm

Technical Committee Meeting on Multichannel and Binaural Audio Technologies

Session F Sunday, October 6 9:00 am–12:00 noon Room 404ABl

ROOM ACOUSTICS AND SOUND REINFORCEMENT

Chair: **Eric Benjamin,** Dolby Laboratories, San Francisco, CA, USA

9:00 am

F-1 Factors Affecting Accuracy of Loudspeaker Measurements for Computational Predictions—*Roger Schwenke*, Meyer Sound Laboratories, Berkeley, CA, USA

The delay of a signal from the input terminals of a loudspeaker amplifier to the output terminals of a microphone can be represented as two parts: one from the electrical input to acoustical transmission and an acoustical propagation delay from some point on the loudspeaker to the microphone. For computational models of mixtures of loudspeakers to be correct, these delays must be measured accurately. It will be shown that temperature differences as small as 1 degree Celsius between measurements of two models of loudspeakers can cause significant differences in the predicted sound field. Though sound speed is much less sensitive to changes in humidity, the difference between assuming a typical humidity and assuming zero humidity (which is the norm) can be significant.
Convention Paper 5665

9:30 am

F-2 Systems for Stereophonic Sound Reinforcement: Performance Criteria, Design Techniques, and Practical Examples—*Jim Brown,* Audio Systems Group, Inc., Chicago, IL, USA

Although stereo systems for large rooms were pioneered in a well documented work at Bell Labs in the 1930s, most modern practitioners appear to be ignorant of the most important aspect of that work as applied to modern sound reinforcement. This paper draws on the author's experience of over twenty years with both portable and permanent systems using two and three front-referenced channels. Design criteria and examples are presented to illustrate both good and bad design practices.
Convention Paper 5666

10:00 am

F-3 Cable Impedance and Digital Audio—*Stephen H. Lampen, David A. DeSmidt*, Belden Electronics Division, San Francisco, CA, USA

One of the key differences between cable designed for analog signals and cable designed for digital signals is the impedance of the cable. Why is impedance important for digital but not for analog? What effect do impedance variations or mismatching have on digital signals? Can you use Category 5e or Category 6 computer cable to run digital audio? Can you use coaxial cable to carry digital audio? This paper addresses all these questions and also outlines the limitations of digital cable designs.
Convention Paper 5667

10:30 am

F-4 Limitations of Current Sound System Intelligibility Verification Techniques—*Peter Mapp*, Peter Mapp Associates, Colchester, Essex, UK

The role of emergency sound and voice alarm systems in life safety management has never been so important. However, to be effective, it is essential that such systems are adequately intelligible. Verification of system intelligibility is therefore assuming ever-greater importance. While a number of verification techniques are available, each is not without its drawbacks. The paper reviews the available methods and, using the results of new research, highlights areas of weakness of the current techniques.
Convention Paper 5668

11:00 am

F-5 Robustness of Multiple Listener Equalization with Magnitude Response Averaging—*Sunil Bharitkar, Philip Hilmes, Chris Kyriakakis*, University of Southern California, Los Angeles, CA, USA

Traditionally, room response equalization is performed to improve sound quality at a given listener. However, room responses vary with source and listener positions. Hence, in a multiple listener environment, equalization may be performed through spatial averaging of magnitude responses at locations of interest. However, the performance of averaging-based equalization, at the listeners, may be affected when listener positions change. In this paper, we present a statistical approach to map variations in listener positions to a performance metric of equalization for magnitude response averaging. The results indicate that, for the analyzed listener configurations, the zone of equalization depends on the distance of microphones from a source and the frequencies in the sound.
Convention Paper 5669

11:30 am

F-6 Coax and Digital Audio—*Stephen H. Lampen, David A. DeSmidt*, Belden Electronics Division, San Francisco, CA, USA

Coaxial cables have been used to run digital audio signals for many years, and have been added to the AES specifications (AES3-id). How is coax different from twisted pairs? What are the distance limitations? What trade-offs are made going from digital twisted pairs to digital coax? These questions are all answered in this paper including a discussion of baluns, which are used to convert from one format to the other.
Convention Paper 5670

12:10 pm

Technical Committee Meeting on Acoustics and Sound Reinforcement

Workshop 4 Sunday, October 6 9:00 am–12:00 noon Room 408B

LARGE SIGNAL LOUDSPEAKER PARAMETER MEASUREMENT

Chair: **John Stewart,** Harman/Becker Automotive Systems, Martinsville, IN, USA

Panelists: *Doug Button,* JBL Professional, Northridge, CA, USA
David Clark, DLC Design, Wixom, MI, USA
Earl Geddes, GedLee Associates, Northville, MI, USA
Steve Hutt, Harman/Becker Automotive Systems, Martinsville, IN, USA
Wolfgang Klippel, Klippel, GmbH, Dresden, Germany
Steve Temme, Listen, Inc., Boston, MA, USA
Alexander Voishvillo, Cerwin-Vega Inc., Simi Valley, CA, USA

This workshop will present measurement methods used to characterize loudspeaker performance at large signal levels. A brief discussion of each technique—how it evolved, what benefit it brings, etc.—will be followed by a demonstration of the technique when possible. The discussion will include: high-power impedance measurement, X_{max}, model parameter nonlinearities, flux modulation, harmonic distortion, maximum SPL, and power compression.

12:10 pm

Technical Committee Meeting on Audio for Telecommunication

Workshop 5 Sunday, October 6 9:00 am–12:00 noon Room 408A

MIXING AND MASTERING IN MULTICHANNEL SURROUND (TUTORIAL)

Chair: **Michael Bishop,** Telarc International Corporation, Cleveland, OH, USA

Panelists: *John Eargle,* JME/JBL Professional
Frank Filipetti, Independent Engineer/Producer
Bob Ludwig, Gateway Mastering & DVD
George Massenburg, GML, LLC
Elliott Scheiner, Independent Engineer/Producer

This tutorial workshop includes a panel of leading industry professionals who will discuss various formats for mixing and mastering in surround. Topics to be covered include:

- Mixing and mastering for DTS/Dolby Digital Surround vs. discrete multichannel surround
- Mixing for 5.1 vs. 6.0, 7.1 or 10.2
- Multichannel mix recording formats in most common use and compatibility with mastering needs
- Mixing and mastering monitor loudspeaker setup/formats and compatibility with common home theater setups in the home
- Essential tools for mixing and mastering in multichannel surround
- Mix approaches and preferences: ambient surround vs. direct surround
- Challenges of mastering for DVD-A
- Challenges of mastering for SACD

Audio and music examples will be played. This workshop, which has been sponsored by the AES Education Committee, will be of benefit to audio engineers of all backgrounds, including students.

Regrettably, Dennis Purcel was scheduled to be a panelist in this workshop but has since passed away.

12:10 pm

Technical Committee Meeting on Transmission and Broadcasting

Special Event
PLATINUM PRODUCERS PANEL 1
PRODUCER, ENGINEER, STUDIO TECHNICIAN—BLURRING OF ROLES
Sunday, October 6, 12:30 pm–2:30 pm
Room 403A

Moderator: **Mitch Gallagher,** EQ Magazine

Panelists: *Rob Cavallo*
Mike Elizondo
Ron Fair
Ben Grosse
Taz Herzberg

It is no secret that the demands of creating audio and music in today's marketplace require a broad range of skills and flexibility. It is increasingly common for one person to handle what once were multiple dedicated jobs: producing, engineering, performing . . . and to do them all at the same time. We've gathered five of today's hottest producers and engineers; join as they discuss the tricks, techniques, mindset, and technologies that allow them to cover multiple roles at once—and how you can apply this information to your own situation.

Mitch Gallagher, the Editor in Chief of *EQ* magazine, began working in music professionally over 20 years ago. His background includes a degree in music, as well as graduate studies in electronic music composition and classical guitar. He is an author, journalist, teacher, touring and studio musician, recording engineer, project studio/multimedia production company owner, and award-winning composer. His first book, *Make Music Now!*, was recently released by Backbeat Books.

Mike Elizondo has shared production credit with Dr. Dre on the Rolling Stones “I Miss You” for the *Austin Powers* soundtrack. His songs and musicianship can be found on Eve’s Grammy winning “Let Me Blow Ya Mind,” Mary J. Blige’s “Family Affair,” Eminem’s “My Dad’s Gone Crazy.” Elizondo a production company “Tone Down Productions” with its first release, *Rebel through Columbia* due in early 2003.

Ron Fair is a veteran record man who is a rare combination of producer, A&R executive, musician, arranger, and engineer. He signed and produced multi-Grammy winner Christina Aguilera and also signed Lit. He is currently President of A&M Records and produced the vocal performances of Christina Aguilera, Mya, Pink, and Lil’ Kim.

Ben Grosse is a producer/engineer/mixer whose credits include Vertical Horizon, Six Pence None The Richer, Fuel, Filter, Guster, Ben Folds, Live, BT, Megadeth, 3rd Eye Blind, Red Hot Chili Peppers, Madonna and Crystal Method. He owns The Mix Room, Burbank, CA, which features two SSL-equipped rooms.

Taz Herzberg is a Los Angeles-based programmer and engineer. His credits include Counting Crows, Vanessa Carlton, Christina Aguilera, and the Grammy-awarded “Lady Marmalade” remake.

Workshop 7 Sunday, October 6 1:00 pm–3:00 pm Room 406AB

AES42-2001: AES STANDARD FOR ACOUSTIC-DIGITAL INTERFACE FOR MICROPHONES

Chair: **Stephan Peus,** Georg Neumann GmbH, Berlin, Germany

Panelists: *Manfred Bleichwehl*, *Steven Harris*, *Juergen Wahl*, Sennheiser Electronic Corporation, Van Nuys, CA, USA

This workshop introduces the new *AES42-2001 Standard for Acoustic-Digital Interface for Microphones*. Among its most interesting potentials are the many options for controlling remotely a wide variety of microphone features, which are not possible with traditional analog microphones. Microphones that are built according to this new standard will not only have an internal analog/digital converter but also extended digital signal processing capabilities (DSP). It can incorporate, for example, the equivalent of a high-gain, high-precision traditional microphone preamplifier, inside the digital microphone. Presenters will demonstrate some of these features and advantages and show potential applications using an operational microphone system, designed according to the new AES42-2001 Standard.

Session G Sunday, October 6 2:00 pm–6:00 pm Room 408A

MULTICHANNEL SOUND

Chair: **Durand Begault,** NASA Ames Research Center, Mountain View, CA, USA

2:00 pm

G-1 Interchannel Interference at the Listening Position in a Five-Channel Loudspeaker Configuration—*Geoff Martin*, McGill University, Montreal, Quebec, Canada, and Bang & Olufsen a/s, Struer, Denmark

It is commonly-accepted thinking that the use of a five-channel surround sound reproduction system increases the size of the listening area over that for two-channel stereophonic systems. In actual fact, for many types of program material, the area of this so-called sweet spot is reduced due to interference between the channels at the listener's ears. This effect is described and analyzed through theoretical evaluation and psychoacoustic listening tests.
Convention Paper 5677

2:30 pm

G-2 Breaking and Making the Rules: Sound Design in 6.1—*Ambrose Field*, University of York, York, UK

It is often colloquially said that there are no rules for the spatialization of multichannel content. This paper seeks to identify creative ways in which the multichannel systems of today and tomorrow can be harnessed to provide aesthetically convincing surround environments. The suitability of production techniques based on Ambisonic methods is re-evaluated for this task, and the discussion focuses on the role of multichannel systems in sound design for presentation in the movie theater.
Convention Paper 5672

3:00 pm

G-3 Localization of Lateral Phantom Images in a 5-Channel System With and Without Simulated Early Reflections—*Jason Corey, Wieslaw Woszczyk*, McGill University, Montreal, Quebec, Canada

Phantom images that rely on interchannel level differences can be produced easily for two-channel stereo. Yet one of the most difficult challenges in production for a five-channel environment is the creation of stable phantom images to the side of the listening position. The addition of simulated early reflection patterns from all five loudspeakers influences the localization of lateral phantom sources. Listening tests were conducted to compare participants' abilities to localize lateral sources under three conditions: power-panned sources alone, sources with simulated early reflection patterns, and simulated early reflection patterns alone (without direct sound). Our results compare localization error for the three conditions at different locations and suggest that early reflection patterns alone can be sufficient for source localization.
Convention Paper 5673

3:30 pm

G-4 The Minimum Number of Loudspeakers and its Arrangement for Reproducing the Spatial Impression of Diffuse Sound Field—*Koichiro Hiyama, Setsu Komiyama, Kimio Hamasaki*, NHK, Setagaya, Tokyo, Japan

It is important to find out how many loudspeakers are necessary for multichannel sound systems to reproduce the spatial impression of a diffuse sound field. This paper discusses the issue in the case where loudspeakers are placed symmetrically along a concentric circle around the listener and at the same height as the listener's ear. It becomes clear that the spatial impression of the diffuse sound field can be obtained by only two symmetrical pairs of loudspeakers (that is, four loudspeakers in all). In this arrangement, one pair of loudspeakers should be placed in the frontal area around the listener at an angle of about 60 degrees, and the other pair should be in the rear area with an angle of 120 degrees to 180 degrees.
Convention Paper 5674

4:00 pm

G-5 A Fuzzy Cerebellar Model Approach for Synthesizing Multichannel Recordings—*Ching-Shun Lin, Chris Kyriakakis*, University of Southern California, Los Angeles, CA, USA

New high capacity optical discs and high bandwidth networks provide the capability for delivering multichannel audio. Although there are many one- and two-channel recordings in existence, only a handful of multichannel recordings exist. In this paper we propose a neural network approach that can synthesize microphone signals with the correct acoustical characteristics of specific venues that have been characterized in advance. These signals can be used to generate a multichannel recording with the acoustical characteristics of the original venue. The complex semi-cepstrum technique is employed to extract features from musical signals recorded in a venue and these signals are sent into the fuzzy cerebellar model articulation controller (FCMAC) for training.
Convention Paper 5675

4:30 pm

G-6 Investigation of Listener Envelopment in Multichannel Surround Systems—*Gilbert A. Soulodre, Michel C. Lavoie, Scott G. Norcross*, Communications Research Centre, Ottawa, Ontario, Canada

It is now well understood that listener envelopment (LEV) is an essential component of good concert hall acoustics. An objective measure based on the late-lateral energy has been shown to perform well at predicting LEV. One goal of multichannel surround systems is to improve the re-creation of the concert hall experience in a home listening environment. By varying the amount of

late-lateral energy, such systems should allow the perception of LEV to be enhanced and controlled. In this paper the loudspeaker/listening room interactions are shown to limit the range of acoustical conditions that can be re-created. A series of formal subjective tests were conducted to determine if objective measures of late-lateral energy are suitable for predicting LEV in multichannel surround systems.
Convention Paper 5676

5:00 pm

G-7 The Significance of Interchannel Correlation, Phase, and Amplitude Differences on Multichannel Microphone Techniques—*Geoff Martin*, McGill University, Montreal, Quebec, Canada, and Bang & Olufsen a/s, Struer, Denmark

There is a measurable interference between correlated signals produced by multiple loudspeakers in a standard five-channel loudspeaker configuration, resulting in an audible comb filter effect. This is due to small individual differences in distances between the ears of the listener and the various loudspeakers. Although this effect is caused by the dimensions and characteristics of the monitoring environment, it can be minimized in the recording process, particularly through the relative placement of microphones and choice of their directional characteristics. In order to analyze this effect, the correlation of microphone signals and their amplitude differences in a recording environment are evaluated using theoretical models. This procedure is applied to coincident and spaced pairs of transducers for direct and reverberant sounds.
Convention Paper 5671

5:30 pm

G-8 A Multichannel Surround Audio Format Applying 3-D Sound Technology—*Myung-Soo Kang, Kyoo-Nyun Kim*, University of Ulsan, Ulsan, Korea

As systems employ multichannel audio more and more, it is necessary to consider the surround sound capability in the process of sound recording and playing. This paper presents a structured audio format and its application to design a more efficient surround sound system. Three-dimensional sound technology is used for localization of the sound source. We define the reusable sound object to clarify the audio format. Sound object is a unit of recorded sound samples that can be changed by various effect properties. Filter and 3-D properties are applied to change the sound objects in each track.
(Paper not presented at convention, but Convention Paper 5678 is available.)

Session H Sunday, October 6 2:00 pm–4:30 pm
Room 404AB

LOW BIT-RATE CODING, PART 1

Chair: **Gary Brown**, Tensilica, Santa Clara, CA, USA

2:00 pm

H-1 Scalable Lossless Audio Coding Based on MPEG-4 BSAC—*Doh-Hyung Kim, Jung-Hoe Kim, Sang-Wook Kim*, Samsung Advanced Institute of Technology, Suwon, Korea

In this paper a new hybrid type of scalable lossless audio coding scheme based on MPEG-4 BSAC (bit sliced arithmetic coding) is proposed. This method introduces two residual error signals, lossy coding error signal and prediction error signal, and utilizes the rice coding as a lossless coding tool. These kinds of processes enable an increase in the compression ratio. As a result of experiment, average total file size can be reduced about 50 to 60 percent of the original size. Consequently, a slight modification of the conventional MPEG-4 general audio coding scheme can give a scalable lossless audio coding functionality between lossy and lossless bitstream.
Convention Paper 5679

2:30 pm

H-2 Lossless Audio Coding Using Adaptive Multichannel Prediction—*Tilman Liebchen*, Technical University of Berlin, Berlin, Germany

Lossless audio coding enables the compression of digital audio data without any loss in quality due to a perfect reconstruction of the original signal. The compression is achieved by means of decorrelation methods such as linear prediction. However, since audio signals usually consist of at least two channels, which are often highly correlated with each other, it is worthwhile to make use of interchannel correlations as well. This paper shows how conventional (mono) prediction can be extended to stereo and multichannel prediction in order to improve compression efficiency. Results for stereo and multichannel recordings are given.
Convention Paper 5680

3:00 pm

H-3 Design of an Error-Resilient Layered Audio Codec—*Dai Yang, Hongmei Ai, Chris Kyriakakis, C.-C. Jay Kuo*, University of Southern California, Los Angeles, CA, USA

Current high quality audio coding techniques mainly focus on coding efficiency, which makes them extremely sensitive to channel noise, especially in high error rate wireless channels. In this paper we propose an error-resilient layered audio codec (ERLAC) which provides functionalities of both fine-grain scalability and error-resiliency. A progressive quantization, a dynamic segmentation scheme, a frequency interleaving technique, and an unequal error protection scheme are adopted in the proposed algorithm to construct the final error robust layered audio bitstream. The performance of the proposed algorithm is tested under different error patterns of WCDMA channels with several test audio materials. Our experimental results show that the proposed approach achieves excellent error resilience at a regular user bit rate of 64 kb/s.
Convention Paper 5681

3:30 pm

H-4 Application of a Concatenated Coding System with Convolutional Codes and Reed-Solomon Codes to MPEG Advanced Audio Coding—*Dong Yan Huang*[1]*, Say Wei Foo*[2]*, Weisi Lin*[3]*, Ju-Nia Al Lee*[4]

[1]Institute of Microelectronics, Singapore, Singapore
[2]Nanyang Technological University, Singapore, Singapore

[3]Laboratories of Information Technology, Singapore, Singapore
[4]National University of Singapore, Singapore

Reliable delivery of audio bitstream is vital to ensure the acceptable audio quality perceived by 3G network customers. In other words, an audio coding scheme that is employed must be fairly robust over the error-prone channels. Various error-resilience techniques can be utilized for the purpose. Due to the fact that some parts of the audio bitstream are less sensitive to transmission errors than others, the unequal error protection (UEP) is used to reduce the redundancy introduced by error resilience requirements. The current UEP scheme with convolutional codes and multistage interleaving has an unfortunate tendency to generate burst errors at the decoder output as the noise level is increased. A concatenated system combining Reed-Solomon codes with convolutional codes in the UEP scheme is investigated for MPEG advanced audio coding (AAC). Under severe channel conditions with random bit error rates of up to 5x10-02, the proposed scheme achieved more than 50 percent improvement in residual bit error rate over the original scheme at a bit rate of 64 kb/s and sampling frequency of 48 kHz. Under burst error conditions with burst error length of up to 4 ms, the proposed scheme achieved more than 65 percent improvement in bit error rate over the original scheme. The average percentage overhead incurred by using the concatenated system is about 3.5 percent of the original UEP scheme. Further improvements are made by decreasing the puncturing rate of convolutional codes. However, this can only be adopted when high protection is needed in extremely noisy conditions (e.g., channel BER significantly exceeds 1.00e-02) since it incurs increased overheads.
Convention Paper 5682

4:00 pm

H-5 A Simple Method for Reproducing High Frequency Components at Low Bit-Rate Audio Coding—*Jeongil Seo, Daeyoung Jang, Jinwoo Hong, Kyeoungok Kang,* Electronics & Telecommunications Research Institute (ETRI), Yuseong-Gu, Deajon, Korea

In this paper we describe a simple method for reproducing high frequency components at low bit-rate audio coding. To compress an audio signal at low bit rates (below 16-kb/s per channel) we can use a lower sampling frequency (below 16 kHz) or high performance audio coding technology. When an audio signal is sampled at a low frequency and coded at a low bit rate, high frequency components and reverberant sound are lost because of quantization noise between pitch pulses. In a short-term period, the harmonic characteristic of audio signals is stationary, so the replication of high-frequency bands with low-frequency bands can extend the frequency range of resulting sound and enhance the sound quality. In addition, for reducing the number of bands to be reproduced we adapted this algorithm at the Bark scale domain. For compatibility with a conventional audio decoder, the additional bitstream is added at the end of each frame, which is generated by a conventional audio coder. We adapted this proposed algorithm to MPEG-2 AAC and increased the quality of audio in comparison with the conventional MPEG-2 AAC coded audio at the same rate. The computational cost of the proposed algorithm is similar to or a little more than a conventional MPEG-2 AAC decoder.
Convention Paper 5683

4:40 pm

Technical Committee Meeting on Coding of Audio Signals

Workshop 6 Sunday, October 6 2:00 pm–5:00 pm
Room 408B

GAME AUDIO

Chair: **Aaron Marks,** On Your Mark Music Productions, CA, USA

Speakers: *Keith Arem,* PCB Productions (CD-ROM audio)
Jack Buser, Dolby Labs, San Francisco, CA, USA (Use of Surround Sound in Video Games)
Brian Schmidt, Microsoft (XBox audio)
Martin Wilde, Game Audio Programmer (Game audio programming issues)

This game audio workshop is designed to answer one of the most often asked questions: What does it take to make great game audio? Not only will we provide the answers, but we also will come armed with many specific audio examples designed to showcase the specific talents, techniques, and technologies used to help make the games business the billion dollar industry that it is today.

5:10 pm

Technical Committee Meeting on Audio for Games

Education Event
ONE-ON-ONE MENTORING SESSION 1
Sunday, October 6, 2:30 pm–4:30 pm
Room 402A

Students are invited to sign up for an individual meeting with audio mentors from the industry. The sign-up sheet will be located near the student center of the convention, and all students are invited to participate in this exciting and rewarding opportunity for focused discussion.

Education Event
EDUCATION FAIR
Sunday, October 6, 2:30 pm–4:30 pm
Room 403B

Institutions offering studies in audio—from short courses to graduate degrees—will be represented in a tabletop session. Information on each school's respective programs will be made available through the display of literature and academic guidance sessions with representatives. There is no charge for schools to participate and admission is free and open to everyone.

Education Event
STUDENT POSTER SESSION
Sunday, October 6, 2:30 pm–4:30 pm
Room 403B

The event will display the scholarly/research/creative works from AES student members. Since many institutions are engaged in both research and applied sciences of audio, this session will provide an opportunity to display and discuss these accomplishments with professionals, educators, and other students.

Special Event
AN AFTERNOON WITH . . . TOM HOLMAN
Sunday, October 6, 2:30 pm–5:00 pm
Room 411 Theater

Moderator: **George Petersen**, Mix Magazine, Emeryville, CA, USA

Tom Holman, whose distinguished career in audio, video, and film spans more than three decades, will participate in an interactive discussion with the audience. President of the TMH Corporation, and Professor of Film Sound at the University of Southern California School of Cinema-Television, Tom Holman has the rare ability to push the state of the art and challenge entire industries to achieve new quality standards. He creates new markets and redefines exiting markets by introducing new quality standards and designing products that perform to them.

Workshop 8 Sunday, October 6 3:00 pm–5:40 pm
Room 406AB

PROTECTING YOUR HEARING AGAINST LOSS: ASSESSMENT OF FUNCTIONAL HEARING ABILITY IN NOISE

Co-chairs: **Dilys Jones, Sigfrid Soli**, House Ear Institute, Los Angeles, CA, USA

These experts in the field of hearing will discuss the methods in which hearing can be protected from loss and will assess functional hearing ability in noise. The workshop will feature a tutorial on the anatomy of the ear, a discussion of the factors that contribute to hearing loss, and hearing evaluations.

Sigfrid Soli will also discuss: audiometric testing; use and interpretation of an audiogram; extended high-frequency testing; and how to protect and conserve one's hearing by using the Occupational Safety and Health Administration’s (OSHA) guidelines, hearing protection devices, in-ear monitors, and safe behavioral practices.

Education Event
EDUCATION FORUM
Sunday, October 6, 4:30 pm–6:00 pm
Room 402A

Co-hosts: **Theresa Leonard,** The Banff Centre, Banff, Alberta, Canada
Don Puluse

This event is a meeting of the AES Education Committee, teachers, authors, students, and members interested in the issues of primary and continuing education of the audio industry. It is an opportunity to discuss the programs of the Education Committee and to provide input for future projects of this committee.

Special Event
THE JAZZ SINGER
Sunday, October 6, 6:30 pm–7:30 pm
Room 403A

An old time radio re-creation of the Lux Radio Theatre production of “The Jazz Singer,” the first successful talking picture starring Al Jolson, will be staged on the 75th anniversary of the film’s original debut. Several veteran radio actors have been assembled to commemorate both the film and the original 1947 broadcast. This radio re-creation will feature Richard Halpern in the starring role, with Herb Ellis directing.

Session I Monday, October 7 9:00 am–12:00 noon
Room 404AB

LOW BIT-RATE CODING, PART 2

Chair: **Marina Bosi,** MPEG LA, Denver, Colorado

9:00 am

I-1 Perceptually-Based Joint-Program Audio Coding—*Christof Faller*[1], *Raziel Haimi-Cohen*[2], *Peter Kroon*[1], *Joseph Rothweiler*[1]

[1]Agere Systems, Murray Hill, NJ, USA
[2]Lucent Technologies, Murray Hill, NY, USA

Some digital audio broadcasting systems, such as Satellite Digital Audio Radio Services (SDARS), transmit many audio programs over the same transmission channel. Instead of splitting up the channel into fixed bit-rate subchannels, each carrying one audio program, one can dynamically distribute the channel capacity among the audio programs. We describe an algorithm which implements this concept taking into account statistics of the bit rate variation of audio coders and perception. The result is a dynamic distribution of the channel capacity among the coders depending on the perceptual entropy of the individual programs. This solution provides improved audio quality compared with fixed bit-rate subchannels for the same total transmission capacity. The proposed scheme is non-iterative and has a low computational complexity.
Convention Paper 5684

9:30 am

I-2 Incorporation of Inharmonicity Effects into Auditory Masking Models—*Hossein Najaf-Zadeh, Hassan Lahdili, Louis Thibault*, Communications Research Centre, Ottawa, Ontario, Canada

In this paper the effect of inharmonic structure of audio maskers on the produced masking pattern is addressed. In most auditory models, the tonal structure of the masker is analyzed to determine the masking threshold. Based on psychoacoustic data, masking thresholds caused by tonal signals are lower compared to those produced by noise-like maskers. However, the relationship between spectral components has not been considered. It has been found that for two different multitonal maskers with the same power, the one with a harmonic structure produces a lower masking threshold. This paper proposes a modification to the MPEG psychoacoustic model 2 in order to take into account the inharmonic structure of the input signal. Informal listening tests have shown that the bit rate required for transparent coding of inharmonic (multitonal) audio material can be reduced by 10 percent if the new modified psychoacoustic model 2 is used in the MPEG 1 Layer II encoder.
Convention Paper 5685

10:00 am

I-3 Binaural Cue Coding Applied to Audio Compression with Flexible Rendering—*Christof Faller, Frank Baumgarte*, Agere Systems, Murray Hill, NJ, USA

In this paper we describe an efficient scheme for compression and flexible spatial rendering of audio signals. The method is based on binaural cue coding (BCC) which was recently introduced for efficient compression of multichannel audio signals. The encoder input consists of separate signals without directional spatial cues, such as separate sound source signals, i.e., several monophon- ➥

ic signals. The signal transmitted to the decoder consists of the mono sum-signal of all input signals plus a low bit rate (e.g., 2 kb/s) set of BCC parameters. The mono signal can be encoded with any conventional audio or speech coder. Using the BCC parameters and the mono signal, the BCC synthesizer can flexibly render a spatial image by determining the perceived direction of the audio content of each of the encoder input signals. We provide the results of an audio quality assessment using headphones, which is a more critical scenario than loudspeaker playback.
Convention Paper 5686

10:30 am

I-4 Two-Pass Encoding of Audio Material Using MP3 Compression—*Martin Weishart*[1], *Ralf Göbel*[2], *Jürgen Herre*[1]

[1]Fraunhofer Institute for Integrated Circuits, Erlangen, Germany
[2]Fachhochschule Koblenz, Koblenz, Germany

Perceptual audio coding has become a widely-used technique for economic transmission and storage of high-quality audio signals. Audio compression schemes, as known from MPEG-1, 2, and 4 allow encoding with either a constant or a variable bit rate over time. While many applications demand a constant bit rate due to channel characteristics, the use of variable bit-rate encoding becomes increasingly attractive, e.g., for Internet audio and portable audio players. Using such an approach can lead to significant improvements in audio quality compared to traditional constant bit-rate encoding, but the consumed average bit rate will generally depend on the compressed audio material. This paper presents investigations into two-pass encoding which combines the flexibility of variable bit rate encoding and a predictable target bit consumption.
Convention Paper 5687

11:00 am

I-5 Technical Aspects of Digital Rights Management Systems—*Christian Neubauer*[1], *Karlheinz Brandenburg*[2], *Frank Siebenhaar*[1]

[1]Fraunhofer Institute for Integrated Circuits IIS-A, Erlangen, Germany
[2]Fraunhofer Arbeitsgruppe Elektronische Medientechnologie AEMT and Ilmenau University, Ilmenau, Germany

In today's multimedia world digital content is easily available to and widely used by end consumers. On the one hand high quality, the ability to be copied without loss of quality, and the existence of portable players make digital content, in particular digital music, very attractive to consumers. On the other hand the music industry is facing increasing revenue loss due to illegal copying. To cope with this problem so-called Digital Rights Management (DRM) systems have been developed in order to control the usage of content. However, currently no vendor and no DRM system is widely accepted by the market. This is due to the incompatibility of different systems, the lack of open standards, and other reasons. This paper analyzes the current situation of DRM systems, derives requirements for DRM systems, and presents technological building blocks to meet these requirements. Finally, an alternative approach for a DRM system is presented that better respects the rights of the consumers.
Convention Paper 5688

11:30 am

I-6 Configurable Microprocessor Implementation of Low Bit-Rate Audio Decoding—*Gary S. Brown*, Tensilica, Inc., Santa Clara, CA, USA

Using a configurable microprocessor to implement low bit-rate audio applications by tailoring the instruction set reduces algorithm complexity and implementation cost. As an example, this paper describes a Dolby digital (AC-3) decoder implementation that uses a commercially-available configurable microprocessor to achieve 32-bit floating-point precision while minimizing the required processor clock rate and die size. This paper focuses on how the audio quality and features of the reference decoder algorithm dictate the customization of the microprocessor. This paper shows examples of audio specific extensions to the processor's instruction set to create a family of AC-3 decoder implementations that meet multiple performance and cost points. How this approach benefits other audio applications is also discussed.
Convention Paper 5689

12:10 pm

Technical Committee Meeting on Network Audio Systems

Session J Monday, October 7 9:00 am–11:00 am
Room 406AB

HIGH-RESOLUTION AUDIO

Chair: **Rhonda Wilson,** Meridian Audio Limited, Huntingdon, Cambridgeshire, UK

9:00 am

J-1 DSD Processing Modules—Architecture and Application—*Nathan Bentall, Gary Cook, Peter Eastty, Eamon Hughes, Michael Page, Christopher Sleight, Mike Smith*, Sony Broadcast & Professional Research Labs, Oxford, UK

As demand continues to grow for production equipment targeting the high resolution, multichannel capabilities of SACD (super audio CD), there is increasing interest in adding DSD capability to both new and existing systems. The prospect of researching and implementing the necessary algorithms from scratch can be daunting. The high data rates, coupled with the asymmetric multipliers often required by the algorithms, make conventional von-Neumann-type DSP platforms (where most developers traditionally have their DSP expertise) seem sub-optimal. Based on custom DSD audio processing engines and packaged in the very compact SODIMM form factor, the modules described in this paper can add high quality, real time, low latency DSD audio processing functionality to a system with a minimum of development time.
Convention Paper 5690

9:30 am

J-2 Multichannel Audio Connection for Direct Stream Digital—*Michael Page, Nathan Bentall, Gary Cook, Peter Eastty, Eamon Hughes, Christopher Sleight*, Sony Broadcast & Professional Research Labs, Oxford, UK

The development of large-scale multitrack production equipment for direct stream digital (DSD) audio will be

substantially aided by the availability of a convenient multichannel interface. DSD is the high-resolution one-bit audio coding system for the super audio CD consumer disc format. Existing multichannel audio interconnection and networking protocols are not easily able to support the high-frequency sample clock required by DSD. A new interconnection type is introduced to provide reliable, low-latency, full-duplex transfer of 24 channels of DSD, using a single conventional office networking cable. The interconnection transfers audio data using Ethernet physical layer technology, while conveying a DSD sample clock over the same cable.
Convention Paper 5691

10:00 am

J-3 The Effect of Idle Tone Structure on Effective Dither in Delta-Sigma Modulation Systems: Part 2—*James A. S. Angus*, University of Salford, Salford, UK

This paper clarifies some of the confusion that has arisen over the efficacy of dither in PCM and sigma-delta modulation (SDM) systems. It presents a means of analyzing in-band, idle tone structure using chaos theory and describes a fair means of comparison between PCM and SDM. It presents results that show that dither can be effective in SDM systems.
Convention Paper 5692

10:30 am

J-4 DC Analysis of High Order Sigma-Delta Modulators—*Derk Reefman, Erwin Janssen*, Philips Research, Eindhoven, The Netherlands

A new method for the DC analysis of a sigma-delta modulator (SDM) is presented. The model used for the description of an SDM is adopted from Candy's model for a first order SDM. However, where Candy's model is exact for a first order SDM, it fails to be so in a higher order case. In our model, we deal with this by the introduction of stochastic behavior of the SDM and obtain the probability density distribution function of some variables which determine many of the characteristics of the SDM in the time domain. Comparison with simulation results shows that the assumption of stochastic behavior is rather good for SDM orders greater than 3, which display significant noise shaping. For lower orders (or less aggressive noise shaping) the approximation is less good. As an aside, the new model of sigma-delta modulation also clarifies why the time-quantized dither approach presented by Hawksford is much better compared to standard quantizer dithering.
Convention Paper 5693

11:10 am

Technical Committee Meeting on High-Resolution Audio

Workshop 9 Monday, October 7 9:00 am–12:00 noon Room 408A

STUDIO PRODUCTION AND PRACTICES

Chair: **George Massenburg,** GML, North Hollywood, CA, USA
with **David Smith**, Sony Music Studios, New York, NY, USA

This workshop will be an open discussion led by esteemed recording engineers George Massenburg and David Smith on delivery standards in digital and analog audio in the modern studio.

Mr. Massenburg and Mr. Smith, who are respectively the Chair and a member of the AES Technical Committee on Studio Practices, will detail their proven techniques for establishing high quality, archive-ready audio, using today's state of the art professional tools. They will concentrate on delivery formats, media requirements, and proper documentation, and will cover the myriad of formats embedded in current digital audio workstations. In addition, they will share their practical secrets for everyday success in the studio environment.

12:10 pm

Technical Committee Meeting on Studio Practices and Production

Workshop 10 Monday, October 7 9:00 am–12:00 noon Room 408B

LOUDSPEAKER LINE ARRAYS, PART 1: THEORY AND HISTORY

Chair: **Jim Brown,** Audio Systems Group, Inc., Chicago, IL, USA

Panelists: *Wolfgang Ahnert*, Acoustic Design Ahnert, Berlin, Germany
Stefan Feistel, Software Design Ahnert, Berlin, Germany
Kurt Graffy, Arup Acoustics, Los Angeles, CA, USA
David Gunness, Eastern Acoustic Works, Inc., Whitinsville, MA, USA
Don Keele, Harman/Becker Automotive Systems, Martinsville, IN, USA
Mike Klasco, Menlo Scientific, El Sobrante, CA, USA

This tutorial workshop is targeted at sound reinforcement professionals. The session includes development of basic concepts underlying line arrays, both simple and complex, an historical survey of commercial line array products, and a variety of line array design techniques. Methods of predicting line array performance are described, both at the conceptual and practical level. This workshop is a highly useful precursor to Workshop12, Loudspeaker Line Arrays, Part 2: Practice and Applications, at 1:30 pm this afternoon.

Presentations:

Historical Review of Line Array Products, by *Mike Klasco*

Basic Line Array Concepts," *Don Keele*

Electronically-Steered Arrays, by *David Gunness*

Measurement and Modeling of Multi-Element Loudspeakers and Modeling of Waveguide Elements, by *David Gunness*

Line Array System Design Considerations, by *Wolfgang Ahnert*

3-D Modeling of Line Arrays in Software, by *Stefan Feistel*

Sort-of Line Arrays? by *Jim Brown*

12:10 pm
Technical Committee Meeting on Loudspeakers

Education Event
ONE-ON-ONE MENTORING SESSION 2
Monday, October 7, 10:00 am–12:00 noon
Room 402A

Students are invited to sign up for an individual meeting with ➥

audio mentors from the industry. The sign-up sheet will be located near the student center of the convention, and all students are invited to participate in this exciting and rewarding opportunity for focused discussion.

Special Event
PLATINUM PRODUCERS PANEL 2:
PAST, PRESENT, AND FUTURE OF RECORDING
Monday, October 7, 12:30 pm–2:30 pm
Room 403A

Moderator: **Howard Massey**, On The Right Wavelength Consulting

Panelists: *Michael Bradford*
Bob Ezrin
Patrick Leonard
Larry Levine
Phil Ramone

How has the art of music recording changed over the past four decades? What is the state-of-the-art today? Where is recording going? Join us for a lively, fascinating discussion as four of the world's top record producers talk about how far we have come and make predictions for the future.

Multi-instrumentalist Michael Bradford (a.k.a Chunky Style) is one of today's hottest producers. Best known for his co-productions of Kid Rock's controversial *The History of Rock* and Uncle Kracker's hit single *Follow Me*, Bradford has also worked with Run-DMC, Terence Trent D'Arby, and Gregg Alexander, and has collaborated with renowned arranger Paul Buckmaster on a number of movie soundtracks.

Bob Ezrin has produced multiplatinum albums and live events for some of the world's leading artists, including Pink Floyd (*The Wall*), Alice Cooper, KISS, Rod Stewart, Peter Gabriel, Lou Reed, Roger Daltrey, and Nine Inch Nails. He has also established several successful production and publishing companies, a recording and mastering studio, and an independent label. He is also active in community service, and currently serves as Co-Chairman of the Board of Directors of The Los Angeles Mentoring Partnership, and is VP of the Board of Directors of the Mr. Holland's Opus Foundation. He is currently producing the new Jane's Addiction album.

Keyboardist Patrick Leonard produced Elton John's critically acclaimed 2001 album *Songs from the West Coast* and recently completed the new album for Duncan Sheik. He has also produced Madonna, Jewel, Bryan Ferry, Roger Waters, and Rod Stewart, and has enjoyed fame as both a solo artist and as half of the nineties rock duo Toy Matinee.

Larry Levine is responsible for some of the greatest classic rock recordings of all time. As Phil Spector's longtime engineer, he helped craft the wall of sound that made records by The Ronettes and other Spector artists sonically unique. Levine also manned the board for numerous Beach Boys albums, including the legendary *Pet Sounds*, as well as Ike and Tina Turner's famed *River Deep, Mountain High*. He has also recorded Herb Alpert, The Righteous Brothers, Billy Eckstine, and The Ramones.

Phil Ramone is one of the most revered names in the music business. A former child prodigy and classically trained violinist, the Pope of Pop has produced an astonishing list of platinum-selling artists, including Billy Joel, Frank Sinatra, Paul Simon, Ray Charles, Elton John, Bono, Quincy Jones, Madonna, Tony Bennett, Carly Simon, Gloria Estefan, Luciano Pavarotti, Natalie Cole, B.B. King, Paul McCartney, Sinead O'Connor, George Michael, James Taylor, and Jon Secada. He has also produced numerous theater, television, and film soundtracks and is currently working on album projects with Rod Stewart and Liza Minnelli.

Moderator Howard Massey is a noted industry consultant and veteran audio journalist. He has written for *Home Recording*, *Surround Professional*, *EQ*, *Musician*, *Billboard*, and many other publications, and is the author of 11 books, including *Behind the Glass*, a collection of interviews with the world's top record producers. Massey has also worked extensively as an audio engineer, producer, songwriter, and touring/session musician.

Education Event
STUDENT RECORDING COMPETITION
Monday, October 7, 1:00 pm–6:00 pm
Room 309

Hosts: **Theresa Leonard,** The Banff Centre, Banff, Alberta, Canada
Don Puluse

Finalists selected by an elite panel of judges will give brief descriptions and play recordings in the Classical and Jazz/Pop categories. One submission per category per student per student section or school is allowed. Meritorious awards will be presented at the closing Student Delegate Assembly meeting on Tuesday, October 8, at 10:00 am in Room 402AB.

1:00 pm–2:00 pm	Classical Category[1]
2:00 pm–3:00 pm	Surround Classical Category[1]
3:00 pm–4:00 pm	Jazz/Folk Category[2]
4:00 pm–5:00 pm	Pop/Rock Category[3]
5:00 pm–6:00 pm	Surround Non-Classical Category[3]

[1]**Judges:** *Michael Bishop, John Eargle, Richard King*
[2]**Judges:** *Jim Anderson, Lynn Fuston, Bill Schnee*
[3]**Judges:** *Frank Filipetti, Bob Ludwig, Elliot Scheiner*

Workshop 12 Monday, October 7 1:30 pm–5:30 pm
Room 408B

LOUDSPEAKER LINE ARRAYS, PART 2: PRACTICE AND APPLICATION

Chair: **John Murray,** Live Sound International Magazine!, San Francisco, CA, USA

Panelists: *François Deffarges*, Nexo, San Rafael, CA, USA
Mark Engebretson, JBL Professional, Northridge, CA, USA
Christian Heil, L'Acoustics, Oxnard, CA, USA
Tom McCauley, McCauley Sound, Inc., Puyallup, WA, USA
Jeff Rocha, Eastern Acoustic Works, Whitinsville, MA, USA
Evert Start, Duran Audio, Zaltbommel, The Netherlands

This workshop focuses on commercially available modular line array systems for performance audio use. Presenters who represent engineering departments of several competitive manufacturers and are actively working in the development of loudspeaker line arrays and for manufacturers of such systems will be featured.

Presentations:

Fresnel Zone Analysis, Wave Sculpturing Technology (WST), and HF DOSC Wave Guide, by *Christian Heil*

Arithmetic Spiral Arrays, MF Diffraction-Slot Band-Pass Section, and High-Efficiency Driver Design, by *Mark Engebretson*

Divergence Shading, MF Parabolic Separator, and MF & LF Band-Pass Technology, by *Jeff Rocha*

Hyperboloid Reflective Wave Source, MF Phase Plug, and Hypercardioid Subbass, by *François Deffarges*

Adaptive Density Inverse Flat Lens, MF Driver Design, and Inter-Cell Summation Aperture, by *Tom McCauley*.

Logarithmic Driver Spacing, Bessel FIR Filtering, and Steerable Lobes/2-Lobe Arrays, by *Evert Start*

Session K Monday, October 7 2:00 pm–4:00 pm
Room 404AB

RECORDING AND REPRODUCTION OF AUDIO

Chair: **Bob Moses,** Island Digital Media Group, Vashon, WA, USA

2:00 pm

K-1 Power Supply Regulation in Audio Power Amplifiers—*Eric Mendenhall*, Gibson Labs, Redondo Beach, CA, USA

Audio power amplifiers have typically been supplied power by the simplest possible means, usually an off-line supply with no line or load regulation, most commonly based on a line frequency transformer. Even modern amplifiers utilizing switchmode power supplies are usually designed without line or load regulation. The exception has been made for high-end audiophile amplifiers. The pros and cons of a regulated power supply are investigated.
Convention Paper 5694

2:30 pm

K-2 Audio Power Amplifier Output Stage Protection—*Eric Mendenhall*, Gibson Labs, Redondo Beach, CA, USA

This paper reviews a progression of circuits used for protecting bipolar power transistors in the output stages of audio power amplifiers. Design oriented methods of determining the protection locus are shown in a mathematical and graphical procedure. The circuits are then expanded from their standard configurations to allow for transient excursion beyond steady state limits, and thermally dependent protection limits, to better match the protection limits to the actual output stage capability. This allows the protection scheme to prevent output stage failure in the least restrictive way. A new method is shown for achieving a junction temperature estimation system without the use of a multiplier.
Convention Paper 5695

3:00 pm

K-3 Archiving Audio—*Jim Wheeler*, Tape Restoration & Forensics Company, Oceano, CA, USA

As hundreds of millions of tapes age, they begin to deteriorate. This paper describes how to recover these unplayable tapes as well as how to store them properly. This paper will also cover all of the issues of archiving audio, including high-capacity and inexpensive hard disk drives, as well as equipment obsolescence and new media.
Convention Paper 5696

3:30 pm

K-4 Finding a Recording Audio Education Program that Suits Your Career Choice—*Laurel Cash-Jones*, Burbank, CA, USA

This paper discusses the past, present, and future of recording audio education. It describes how the job market and educational requirements have changed, and takes a look at how to plan for a successful career. It also provides valuable information on getting and keeping a job in today's fast-paced world of professional audio.
Convention Paper 5697

4:10 pm

Technical Committee Meeting on Audio Recording and Storage Systems

Session L Monday, October 7 2:00 pm–4:30 pm
Room 406AB

AUDIO NETWORKING AND AUTOMOTIVE AUDIO

Chair: **John Strawn,** S Systems, Larkspur, CA, USA

2:00 pm

L-1 Studio Exploring Using Universal Plug and Play—*Rob Laubscher*[1], *Richard Foss*[2]

[1]Seattle, WA, USA
[2]Rhodes University, Grahamstown, South Africa

This paper explores the use of universal plug and play (UPnP) as a studio control technology. The architecture of a possible studio control technology is introduced. The elements of this studio control architecture are related to the architecture of UPnP. A sample implementation demonstrates the key aspects of using UPnP as a studio control technology.
Convention Paper 5698

2:30 pm

L-2 mLAN: Current Status and Future Directions— *Jun-chi Fujimori*[1], *Richard Foss*[2]

[1]Yamaha Corporation, Hamamatsu, Japan
[2]Rhodes University, Grahamstown, South Africa

"mLAN" describes a network that allows for the transmission and receipt of audio and music control data by audio devices. IEEE 1394 was chosen as the specification upon which to implement mLAN. mLAN has built on IEEE 1394 and related standards, introducing formats, structures, and procedures that enable the deployment of IEEE 1394 within a music studio context. This paper discusses these standards, their implementations, and provides pointers to the future evolution of mLAN.
Convention Paper 5699

3:00 pm

L-3 Mutually-Immersive Audio Telepresence— *Norman P. Jouppi*[1], *Michael J. Pan*[2]

[1]HP Labs, Palo Alto, CA, USA
[2]UCLA, Los Angeles, CA, USA

Mutually-immersive audio telepresence attempts to create for the user the audio perception of being in a remote location, as well as simultaneously to create the perception for people in the remote location that the user of the system is present there. The system provides bidirectional multichannel audio with relatively good fidelity, isolation from local sounds, and a reduction of local reflections. The system includes software for reducing unwanted feedback and joystick control of the audio environment. We are investigating the use of this telepresence system as a substitute for business travel. Initial ➥

user response has been positive, and future improvements in networking quality of service (QOS) will improve the interactivity of the system.
Convention Paper 5700

3:30 pm

L-4 Assessment of the Sound Field in a Car—
Chulmin Choi[1], *Lae-Hoon Kim*[1], *Sejin Doo*[2], *Yangki Oh*[3], *Koeng-Mo Sung*[1]

[1]Seoul National University, Seoul, Korea
[2]Dong-Ah Broadcasting College, Ansung, Korea
[3]Mokpo National University, Mokpo, Korea

This paper describes the measurement and assessment method for the sound field in a car cabin, which is so small that conventional room acoustical parameters cannot be employed directly. First, we measured the sound field using a multichannel microphone system and calculated some room acoustical parameters for judging their validity in the car cabin. As a result, we concluded that many of the conventional parameters do not have useful meaning in a car and its audio system. By analyzing the impulse responses from many cars, we developed some parameters for a more profound assessment of the sound field in a car.
Convention Paper 5701

4:00 pm

L-5 Measurement of the Speech Intelligibility Inside Cars—*Angelo Farina, Fabio Bozzoli,* University of Parma, Parma, Italy

The paper describes a measurement system developed for assessing speech intelligibility inside car compartments. This relates directly to the understandability of the voices being reproduced through the radio receiving system (news, traffic information, etc.), but in the future it will also be used for assessing the direct speech communications between the passengers and the performance of hands-free communication devices. The system is based on the use of two head-and-torso simulators, one equipped with an artificial mouth, the second equipped with binaural microphones. Only the second is used when the sound is being reproduced through the car's sound system. The MLS-based version of the STI method is used for performing the measurements, taking into account the effect of the background noise and the electro-acoustic propagation inside the compartment.
Convention Paper 5702

4:40 pm

Technical Committee Meeting on Automotive Audio

Workshop 11 Monday, October 7 2:00 pm–5:00 pm
Room 408A

RECENT DEVELOPMENTS IN MPEG-4 AUDIO

Chair: **Jürgen Herre,** Fraunhofer Institute for Integrated Circuits, Erlangen, Germany

Panelists: *Martin Dietz*, Coding Technologies, Nürnberg, Germany
Ralf Geiger, Fraunhofer Institute for Integrated Circuits IIS-A, Erlangen, Germany
Leon van de Kerkhof, Philips Digital Systems Laboratory, Eindhoven, The Netherlands

For more than a decade MPEG audio standards have been defining the state of the art in perceptual coding of audio signals. Several phases of standardization (MPEG-1, MPEG-2, MPEG-4) have been pushing borders beyond everyone's initial expectations. This workshop reports on three new extensions to the existing MPEG-4 audio standard which are currently under development. The work on compatible bandwidth extension of audio signals provides an additional performance boost for AAC coding at very low bit rates, permitting considerable signal quality at bit rates around 24 kbit/s per channel. A second extension augments the existing MPEG parametric audio coder with modes for representing high-quality audio signals. Finally, methods for lossless audio coding are under consideration, extending current MPEG-4 audio coders toward perfect representation at word lengths and sampling rates typically associated with high-definition audio. The workshop will provide background and demonstrations on these technologies.

Special Event
SPARS BUSINESS PANEL: WHAT I'VE HAD TO DO TO SURVIVE SINCE 9/11, ENRON, "DESKTOP" AUDIO, AND STIFFER COMPETITION
Monday, October 7, 3:30 pm–5:30 pm
Room 403A

Moderator: **David Amlen**, SPARS President, Sound on Sound (NYC)

Panelists: *Russ Berger*, Russ Berger Design Group, Addison, TX, USA
Steve Davis, Crawford Audio, Atlanta, GA, USA
Kevin Mills, Larrabee Studios, Los Angeles, CA, USA
Joe Montarello, Capitol Region Insurance
David Porter, Annex Digital, San Francisco, CA, USA

The Society of Professional Audio Recording Services (SPARS) is a 23-year-old professional organization dedicated to sharing practical, hands-on business information about audio facility ownership, management, and operations. This event, which is co-hosted by the AES, features an elite panel of studio owners and managers who will explore strategies for adapting your business to changing times. Panelists will illustrate, with stories and solutions, innovative ways they have handled unique business challenges faced in successfully running today's audio production facility.

Special Event
OPEN HOUSE OF THE TECHNICAL COUNCIL AND THE RICHARD C. HEYSER MEMORIAL LECTURE
Monday, October 7, 6:00 pm–8:00 pm
Room 403A

Lecturer: **James E. West**

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. Mr. Heyser was also an AES governor and AES Silver Medal recipient.

The distinguished Richard C. Heyser lecturer is Professor James West of Avaya Labs and Johns Hopkins University, coinventor of the electret condenser microphone. His lecture is titled, "Modern Electret Microphones and Their Applications."

It is well known that condenser microphones are the transducer of choice when accuracy, stability, frequency characteristics, dynamic range, and phase are important. But conventional condenser microphones require critical and costly construction as well as the need for a high dc bias for linearity. These disadvantages ruled out advanced practical microphone designs such as multi- element arrays and the use of linear microphones in telephony. The combination of our discovery of stable charge storage in thin polymers and the need for improved linearity in communications encouraged the development of modern electret microphones in the early 1960s at Bell Labs.

Others had suggested the use of electrets in transducers (electrical analog of a permanent magnet) in the late 1920s, but these and subsequent efforts all suffered from the insufficient stability of wax electrets under normal environmental conditions. Water molecules from atmospheric humidity were the main depolarizing factor in Carnauba and other wax electrets. The first broad application of wax electret microphones was discovered in captured World War II Japanese field equipment. Because of the decay of the polarization of the electret, these microphones had a lifetime of about six months.

Modern polymer electret transducers can be constructed in various sizes and shapes mainly because of the simplicity of the transducer. All that is needed in the mechanical system is a thin (25 microns) charged polymer, a small (irregular) air gap and a back plate. An impedance converter is necessary and is provided by a FET transistor to better match conventional electronic equipment. Applications of electret microphones range from very small hearing aid microphones (a few square millimeters) to very large single element units (20 cm diameter) for underwater and airborne reception of very low frequencies. Because the frequency and phase response of electret microphones are relatively constant from unit to unit, multiple element two-dimensional arrays can be constructed. We have constructed a two- dimensional, 400-element array for Arnold Auditorium at Bell Labs with electret elements that are available for under $1.00 each.

Telephone bandwidth and frequency characteristics have remained constant for the past 30 years while entertainment has brought high fidelity including surround sound into most homes throughout the world. People are accustomed to good quality sound and expect it in communication systems. The Internet Protocol (IP) offers the needed bandwidth to improve audio quality for telephony, but it will require broadband microphones and loudspeakers to provide customers with voice presence and clarity. Directional microphones for both handheld and hands-free modes are necessary to improve signal-to-noise ratios and to enable automatic speech recognition. Arrays with dynamic beam forming properties are required especially for conference rooms. Signal processing has made possible stereo acoustic echo cancellers and many other signal enhancements that improve audio quality. Dr. West will discuss some of the current work on broadband communications at Avaya Labs.

James E. West is a Bell Laboratories Fellow and former member of the Acoustics and Speech Research Department at Lucent Technologies, specializing in electroacoustics, physical acoustics, and architectural acoustics. He is now research scientist in the Multimedia Technologies Research Lab of Avaya. His pioneering research on charge storage and transport in polymers led to the development of electret transducers for sound recording and voice communication. Almost 90 percent of all microphones built today are based on the principles first published by West in the early 1960s. This simple but rugged transducer is the heart of most new telephones manufactured by Lucent and other producers of communication equipment. He is the recipient of the Callinan Award (1970), sponsored by the Electrochemical Society of America, the Senior Award (1970), sponsored by the IEEE Group on Acoustics, the Lewis Howard Latimer Light Switch and Socket Award (1989), sponsored by the National Patent Law Association, the George R. Stibitz Trophy, sponsored by the Third Annual AT&T Patent Award (1993), New Jersey Inventor of the Year for 1995, The Acoustical Society of America's Silver Medal in Engineering Acoustics (1995), an honorary Doctor of Science degree from New Jersey Institute of Technology (1997), the Golden Torch Award (1998) sponsored by the National Society of Black Engineers, the Industrial Research Institutes 1998 Achievement Award, and The Ronald H. Brown American Innovator Award (1999). The Acoustical Society of America has chosen one of his early papers on Electret Microphones as a Benchmark publication.

The lecture will be followed by a reception hosted by the Technical Council.

Special Event
TOUR OF CATHEDRAL OF OUR LADY OF THE ANGELS AND ORGAN RECITAL
Tuesday, October 8, 8:30 am–3:00 pm
Cathedral of Our Lady of the Angels
555 W. Temple Street
Los Angeles

Join with the American Institute of Organbuilders for a tour and series of lectures in the newly-completed Cathedral of Our Lady of the Angels. The prime focus of the tour will be "Opus 75"—the Cathedral's new 105 rank, 6019 pipe organ by Dobson Pipe Organ Builders (Lake City, Iowa). The tour also will feature a recital by the AES' "resident organist" Graham Blyth.

Included in the series of lectures will be "The Cathedral Organ: The Truth Revealed," by Manuel Rosales, consultant to the Cathedral for the organ project; "Design Adventures: Collaboration with Architects," by Lynn Dobson, President, Dobson Pipe Organ Builders; "How to Hold Up 55 Tons of Organ," by Jon Thieszen, a member of the organ design team; "Entering the Tonal Unknown," by John Panning and Samuel Soria, Cathedral Organists (this will be a lecture and demonstration of the features of "Opus 75").

At noon, lunch will be provided in the Archdiocesan Conference Center, accompanied by a lecture, "Organ Structure: Preparing for the Big One!" John Seest will offer a look into the support of the Dobson Organ in the massive new concrete Cathedral and the attention given to the high earthquake design loads that make this a unique organ structure. Following the lunch and lecture, Graham Blyth will perform a recital on the Dobson Organ.

To conclude the tour, Dennis Paoletti and John Prohs, acoustical consultants for the Cathedral, will discuss some of the unique design features and challenges they faced in realizing this massive and complex project.

Cost for the tour, including lunch, is $45 per person. Reservations are required and must be made at the Special Events Desk no later than 4:00 pm Sunday, October 6th. Convention attendees wanting to attend just the organ recital by Graham Blyth, but not the entire tour, are encouraged to check the poster in the Convention Center lobby for the starting time and additional details.

Attendance at the recital is free.

Session M Tuesday, October 8 9:00 am–12:00 noon
Room 404AB

PSYCHOACOUSTICS, PART 1

Chair: **Christopher Struck,** Dolby Laboratories, San Francisco, CA, USA

9:00 am

M-1 Comparison of Objective Measures of Loudness Using Audio Program Material—*Eric Benjamin*, Dolby Laboratories, San Francisco, CA, USA

Level measurements do not necessarily correlate well with subjective loudness. Several methods are described for making objective measurements that are designed to correlate better with actual loudness. Some of these measures are: A-weighting and B-weighting, the methods of Stevens and Zwicker as described in ISO 532A and B, and the method described by Moore and Glasberg. All of these measures are intended to describe (measure) the loudness of sounds with constant spectra. How well do these measures work with typical audio signals distributed via broadcast or recording?
Convention Paper 5703

9:30 am

M-2 Descriptive Analysis and Ideal Point Modelling of Speech Quality in Mobile Communication— *Ville-Veikko Mattila*, Nokia Research Center, Tampere, Finland

Descriptive analysis of processed speech quality was carried out by semantic differentiation, and external preference mapping was used to relate the attributes to overall quality judgments. Clean and noisy speech samples from different speakers were processed by various processing chains representing mobile communications, resulting in a total of 170 samples. The perceptual characteristics of the test samples were described by 18 screened subjects, and the final descriptive language with 21 attributes were developed in panel discussions. The scaled attributes were mapped to overall quality evaluations collected from 30 screened subjects by partial least square regression. The Phase II ideal point modeling was used to predict the quality with an average error of about 6 percent and to study the linearity of the attributes.
Convention Paper 5704

10:00 am

M-3 Relating Multilingual Semantic Scales to a Common Timbre Space—*William L. Martens, Charith N. W. Giragama*, University of Aizu, Aizu-Wakamatsu, Fukushima-ken, Japan

A single, common perceptual space for a small set of processed guitar timbres was derived for two groups of listeners, one group comprising native speakers of the Japanese language, the other group comprising native speakers of Sinhala, a language of Sri Lanka. Subsets of these groups made ratings on 13 bipolar adjective scales for the same set of sounds, each of the two groups using anchoring adjectives taken from their native language. The adjectives were those freely chosen most often in a preliminary elicitation. The results showed that the Japanese and Sinhalese semantic scales related differently to the dimensions of their shared timbre space that was derived using MDS analysis of the combined dissimilarity ratings of listeners from both groups.
Convention Paper 5705

10:30 am

M-4 Design and Evaluation of Binaural Cue Coding Schemes—*Frank Baumgarte, Christof Faller*, Agere Systems, Murray Hill, NJ, USA

Binaural cue coding (BCC) offers a compact parametric representation of auditory spatial information such as localization cues inherent in multichannel audio signals. BCC allows reconstruction of the spatial image given a mono signal and spatial cues that require a very low rate of a few kbit/s. This paper reviews relevant auditory perception phenomena exploited by BCC. The BCC core processing scheme design is discussed from a psychoacoustic point of view. This approach leads to a BCC implementation based on binaural perception models. The audio quality of this implementation is compared with low-complexity FFT-based BCC schemes presented earlier. Furthermore, spatial equalization schemes are introduced to optimize the auditory spatial image of loudspeaker or headphone presentation.
Convention Paper 5706

11:00 am

M-5 Evaluating Influences of a Central Automotive Loudspeaker on Perceived Spatial Attributes Using a Graphical Assessment Language— *Natanya Ford*[1], *Francis Rumsey*[1], *Tim Nind*[2]

[1]University of Surrey, Guildford, Surrey, UK
[2]Harman/Becker Automotive Systems, Martinsville, IN, USA

An investigation is described which further develops a graphical assessment language (GAL) for subjectively evaluating spatial attributes of audio reproductions. Two groups of listeners, those with previous experience of using a GAL and listeners new to graphical elicitation, were involved in the study which considered the influence of a central automotive loudspeaker on listeners' perception of ensemble width, instrument focus, and image skew. Listeners represented these attributes from both driver's and passenger's seats using their own graphical descriptors. Source material for the study consisted of simple instrumental and vocal sources chosen for their spectral and temporal characteristics. Sources ranged from a sustained cello melody to percussion and speech extracts. When analyzed using conventional statistical methods, responses highlighted differences in listeners' perceptions of width, focus, and skew for the various experimental conditions.
Convention Paper 5707

11:30 am

M-6 Multidimensional Perceptual Calibration for Distortion Effects Processing Software—*Marui Atsushi, William L. Martens*, University of Aizu, Aizuwakamatsu-shi, Fukushima-ken, Japan

Controlled nonlinear distortion effects processing produces a wide range of musically useful outputs, especially in the production of popular guitar sounds. But systematic control of distortion effects has been difficult to attain, due to the complex interaction of input gain, drive level, and tone controls. Rather than attempting to calibrate the output of commercial effects processing hardware, which typically employs proprietary distortion algorithms, a real-time software-based distortion effects processor was implemented and tested. Three distortion effect types were modeled using both waveshaping and a second order filter to provide more complete control over the parameters typically manipulated in controlling effects for electric guitars. The motivation was to relate perceptual differences between effects processing outputs and the mathematical functions describing the nonlinear waveshaping producing variation in distortion. Perceptual calibration entailed listening sessions where listeners adjusted the tone of each of nine test outputs, and then made both

pairwise dissimilarity ratings and attribute ratings for those nine stimuli. The results provide a basis for an effects-processing interface that is perceptually-calibrated for system users.
Convention Paper 5708

Workshop 13 Tuesday, October 8 9:00 am–12:00 noon
Room 408B

PERCEPTUAL ISSUES RELATED TO CASCADED AUDIO CODECS

Chair: **Thomas Sporer**, Fraunhofer Institute for Integrated Circuits, Ilmenau, Germany

Panelists: *Louis Fielder*, Dolby Laboratories, Inc., San Francisco, USA
Jürgen Herre, Fraunhofer IIS, Erlangen, Germany
Representitives from IRT, Munich, Germany; BBC, UK

Digital processing seemed to solve the problem of reduced audio quality due to copying forever, but today several stages of the transmission chain include perceptual audio coding schemes such as MPEG-2 Layer-2 and Layer-3, MPEG-4 AAC, AC-3, and ATRAC. This mixture of different coding formats leads to significant accumulation of coder distortions in each encoding step (so-called tandem coding) and associated serious loss of subjective audio quality. Additional problems can be caused by processing such as equalization and compression.

In this workshop typical usage scenarios will be presented and discussed. Results from listening tests conducted by ITU and EBU over the last few years will be summarized. Different solutions to reduce tandem artifacts and to increase robustness will be presented.

Workshop 14 Tuesday, October 8 9:00 am–12:00 noon
Room 408A

THE APPLICATION OF MULTICHANNEL SOUND FORMATS IN VEHICLES

Chair: **Richard Stroud**, Stroud Audio, Inc., Kokomo, IN, USA

Panelists: *David Clark*
Neal House
Mark Ziemba

Multichannel audio has made its way into the car. Many more applications of this technology will doubtless appear in new vehicles. Workshop presenters will discuss vehicular multichannel audio installation design goals, system development, available source material, recording considerations, and evaluation methodologies. Demonstration vehicles for multichannel listening will be present.

Special Event
AUDIO POST PRODUCTION IN 24p HDTV AND RELATED FORMATS
Tuesday, October 8, 10:00 am–11:30 am
Room 403A

Moderator: **Dennis Weinreich**, Videosonics Cinema Sound, London, UK

Panelists: *Colin Broad*, CB Electronics
Doug Ford, Skywalker Sound
Robert Predovich, Soundmaster Group
Scott Wood, Digidesign/Avid Technologies

The challenge of new technologies has significant impact on our professional workflow. As cutting-edge sound people we must be able to understand how best to utilize these technologies to benefit most from them—both technically and creatively. We need to understand what the technology is attempting to address and how we should use it to our best advantage. Some of us will try to find where it fits into our existing work flow. Others will see new creative opportunities. Many advances require complete rethinking of how we approach our job in order to benefit from them.

HD 24p is one such advance that will have significant impact on how we do our jobs. Currently, the lighter and seemingly less complex production tools cause filmmakers to think that HD will give a better look and sound to production with less effort. In time, after the advantages are weighed against the disadvantages, we will probably find ourselves working in ways we currently cannot foresee.

Our panel will discuss the present position on working with HD and 24p, with particular emphasis on Audio Post Production. We will look at issues from the production floor to sound editorial to the dubbing stage. The panel will be made up of contributors from Post facilities, manufacturers' representatives, and professionals in the field who can share their HD 24p experiences and enlighten us about this new and interesting technology.

This event is being presented by the APRS (UK) and co-hosted by the AES.

Education Event
STUDENT DELEGATE ASSEMBLY 2
Tuesday, October 8, 10:00 am–11:30 pm
Room 402A

Chair: **Scott Cannon,** Stanford University Student Section, Stanford, CA, USA

Vice Chair: **Dell Harris**, Hampton University Student Section, Hampton, VA, USA

At this meeting the SDA will elect new officers. One vote will be cast by the designated representative from each recognized AES student section in the North and Latin America Regions. Judges' comments and awards will be presented for the Recording Competitions and the Student Poster Session. Plans for future student activities at local, regional, and international levels will be summarized.

Special Event
ROAD WARRIORS PANEL: TRUE STORIES FROM THE FRONT LINES OF SOUND REINFORCEMENT
Tuesday, October 8, 12:30 pm – 2:00 pm
Room 403A

Introduction: **Paul Gallo**, founder of Pro Sound News, New York, NY, USA

Moderators: **Steve Harvey**, **Clive Young**, Pro Sound News, New York, NY, USA

Participants: *Greg Dean*
Kirk Kelsey
David Morgan

Introduced by Paul Gallo, publishing veteran and long-time friend to the live sound industry, this freewheeling panel of ➥

touring professionals will cover the latest trends, techniques, and tools that shape modern sound reinforcement. Moderators Steve Harvey and Clive Young will guide the all-star panel across subjects ranging from gear to gossip, in what promises to be an entertaining and educational 90 minutes, with the engineers on the business side of the microphone, saying something besides “testing” and “check” for a change!

Special Event
ORGAN RECITAL
Tuesday, October 8, 1:00 pm–2:00 pm
Cathedral of Our Lady of the Angels
555 W. Temple Street
Los Angeles

Organist: **Graham Blyth**

Visit the newly completed Dobson Organ at the new Cathedral of Our Lady of the Angels in downtown Los Angeles. The specially organized event will feature a lunchtime performance by Graham Blyth, the AES’ resident organ recitalist. His program will include: Antonio Soler’s “The Emperor's Fanfare,” Cesar Franck’s “Chorale No. 3 in A Minor,” Olivier Messiaen’s “Dieu Parmi Nous,” plus works by J. S. Bach and Louis Vierne. Mr. Blyth’s performance will be preceded by a series of presentations from the design and installation team for the Dobson Organ, and will be followed by a brief lecture from Dennis Paoletti and John Prohs, the Cathedral’s acoustical consultants.

Graham Blyth 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 the Trinity College of Music. In the late 1980s he renewed his studies with Sulemita Aronowsky for piano and with Robert Munns for organ.

Mr. Blyth 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, Amsterdam, Copenhagen, Munich, and Paris (Madeleine Church). He gives numerous concerts each year, principally as an organist and a pianist, but also as a conductor and a harpsichord player.

Mr. Blyth is founder and technical director of Soundcraft. He divides his time between his main career as a designer of professional audio equipment and organ-related activities. He has lived in Wantage, Oxfordshire, U.K., since 1984, where he is currently artistic director of the Wantage Chamber Concerts and director of the Wantage Festival of Arts. He is also founder and conductor of the Challow Chamber Singers & Players. He is involved with Musicom Ltd., a British company at the leading edge of the pipe organ control system and digital pipe synthesis design. He also acts as tonal consultant to the Saville Organ Company and is recognized as one of the leading voicers of digital pipe modeling systems.

Session N Tuesday, October 8 1:00 pm–4:00 pm
Room 404AB

PSYCHOACOUSTICS, PART 2

Chair: **Nick Zacharov,** Nokia Research Center, Tampere, Finland

1:00 pm

N-1 Industry Evaluation of In-Band On-Channel Digital Audio Broadcast Systems—*David Wilson,* Consumer Electronics Association, Arlington, VA, USA

The National Radio Systems Committee’s testing and evaluation program for in-band on-channel digital audio broadcast systems is described. The results of laboratory and field tests performed during 2001 on iBiquity Digital Corporation’s AM-band and FM-band IBOC DAB systems are reported. The conclusions drawn from the laboratory and field test results are also reported, and implications for the future are discussed.
Convention Paper 5709

1:30 pm

N-2 Comparisons of De Facto and MPEG Standard Audio Codecs in Sound Quality—*Eunmi L. Oh, JungHoe Kim*, Samsung Advanced Institute of Technology, Suwon, Korea

The current paper is concerned with assessing the sound quality of various audio codecs including ubiquitous de facto standards. Formal listening tests were conducted based on the ITU-R Recommendation BS.1116 in order to provide an objective measure of sound quality. Codecs tested included de facto standards that were commercially and noncommercially available and the MPEG general audio. In addition, our recently updated codec was tested. Test items consisted of usual MPEG test sequences and other sensitive sound excerpts at the bit rate of 64- and 96-kb/s stereo. Experimental results show that the sound quality of our newest codec out paces that of most of other codecs.
Convention Paper 5710

2:00 pm

N-3 Evaluating Digital Audio Artifacts with PEAQ— *Eric Benjamin*, Dolby Laboratories, San Francisco, CA, USA

Portions of the digital audio chain have been incrementally improved by development, such that objective specifications indicate a very high level of performance. Subjective reviews of these components often claim to observe substantial differences between products. This investigation uses the tool PEAQ (perceptual evaluation of audio quality) to measure the audio degradation caused by analog to digital converters, digital to analog converters, and sample rate conversion, and also to measure the minute incremental changes of codec audio quality that accompany very small changes in data rate.
Convention Paper 5711

2:30 pm

N-4 The Use of Head-and-Torso Models for Improved Spatial Sound Synthesis—*Ralph Algazi, Richard O. Duda, Dennis M. Thompson*, University of California, Davis, CA, USA

This paper concerns the use of a simple head-and-torso model to correct deficiencies in the low-frequency behavior of experimentally measured head-related transfer functions (HRTFs). This so-called snowman model consists of a spherical head located above a spherical torso. In addition to providing improved low-frequency response for music reproduction, the model provides the major low-frequency localization cues, including cues for low-elevation as well as high-elevation sources. The model HRTF and the measured HRTF can be easily combined by using the phase response of the model at all frequencies and by cross-fading between the dB magni-

tude responses of the model and the measurements. For efficient implementation, the exact snowman HRTF is approximated by two time delays and two first-order IIR filters. Because the poles are independent of the location of the virtual source, this supports a simple real-time implementation that allows for arbitrarily rapid head and source motion.
Convention Paper 5712

3:00 pm

N-5 Three-Dimensional Headphone Sound Reproduction Based on Active Noise Cancellation—*Daniël Schobben, Ronald Aarts*, Philips Research Labortories, Eindhoven, The Netherlands

Headphone signal processing systems that are commercially available today are not optimized for the individual listener. This results in large localization errors for most listeners. In this paper a system is introduced that requires a one-time calibration procedure, which can be carried out conveniently by the listener. This system consists of conventional headphones into which small microphones have been mounted. An active noise cancellation method is used to achieve a sound reproduction via headphones, which is as close as possible to a reference loudspeaker setup. The active noise cancellation system is based on adaptive filters that are implemented in the frequency domain.
Convention Paper 5713

3:30 pm

N-6 On the Design of Canonical Sound Localization Environments—*Eric J. Angel, Ralph Algazi, Richard O. Duda*, University of California, Davis, CA, USA

This paper addresses the design of virtual auditory spaces that optimize the localization of sound sources under engineering constraints. Such a design incorporates some critical cues commonly provided by rooms and by head motion. Different designs are evaluated by psychoacoustics tests with several subjects. Localization accuracy is measured by the azimuth and elevation errors and the front/back confusion rate. We present a methodology and results for some simple canonical environments that optimize the localization of sounds.
Convention Paper 5714

Workshop 15
Tuesday, October 8 — 1:00 pm–4:00 pm
Room 408A

CODING OF SPATIAL AUDIO—YESTERDAY, TODAY, TOMORROW

Chair: **Christof Faller,** Agere Systems, Murray Hill, NJ, USA

➥

Panelists: *Frank Baumgarte*, Agere Systems, Murray Hill, NJ, USA
Mark Davis, Dolby, San Francisco, CA, USA
Martin Dietz, Coding Technologies, Nürnberg, Germany
Gerald Schuller, Thomas Sporer, Fraunhofer Arbeitsgruppe Elektronische Medientechnologie AEMT; and Ilmenau University, Ilmenau, Germany

Low bit-rate audio coding has become ubiquitous in many of today's audio systems, most of which are able to handle stereo or multichannel audio signals. A closer examination of issues related to the compression of spatial audio reveals a number of complex perceptual and coding issues that need to be considered in order to achieve optimum coder performance. While a wealth of different approaches have evolved over the recent decade, there is no single technique serving all purposes equally well. This workshop will provide a review of the principles and practical approaches for coding of spatial audio, and a discussion of the different dimensions of the tradeoff bit rate vs. spatial quality, and characterizes commonly used coders w.r.t. these aspects.

Introduction, by *Christof Faller*

Spatial Perception, by *Thomas Sporer*

This talk will review spatial perception that is relevant for reproduction and coding of spatial audio.

Historical Development of Spatial Audio Reproduction and Transmission, by *Mark Davis*

The historical development of stereophonic and multichannel audio reproduction will be presented. It includes a description of early experiments for multiloudspeaker and two-loudspeaker stereophony, and the early techniques for pseudo stereophony, etc. For the historical development of spatial audio transmission, sum/difference transmission for FM radio stereophony will be described, and new techniques based on perceptual audio coding will be briefly mentioned.

Coding of Stereophonic Signals, by *Gerald Schuller*

This talk will focus on traditional ways of encoding of stereo audio signals such as separate coding (fails for certain signals), L/R coding and S/D coding and BMLD consideration, redundancy reduction/irrelevancy reduction, intensity stereo, and matrices as used in popular coders.

Parametric Stereo for Very Low Bit-Rate Stereo Coding, by *Martin Dietz*

This talk will describe the idea of parametic stereo coding: a mono signal, downmixed from the stereo source and compressed by powerful modern coding algorithms, is converted back into stereophonic sound by means of a coded parametric description of the spatial properties of the original signal.

Future Directions for Coding of Spatial Audio, by *Frank Baumgarte*

The technique and philosophy behind binaural cue coding (BCC) and other potential forward-looking technologies for coding of spatial audio will be described.

Education Event
STUDENT PAPER SESSION
Tuesday, October 8, 1:00 pm–3:00 pm
Room 406AB

The paper session provides opportunity for students to present their research work on all subjects covered by the AES. Professionals from the audio industry will be invited to discuss the students' projects.

Special Event
THE VIRTUAL STUDIO: DAW-BASED RECORDING
Tuesday, October 8, 2:30 pm–5:30 pm
Room 403A

Moderator: **Frank Wells**, Pro Sound News, New York, NY, USA

Panelists: *David Channing*,Independent Engineer
Lynn Fuston, 3D Audio
Jim Kaiser, Mastermix
Nathaniel Kunkel,Independent Engineer
Ralf Schluezen, TC Works
Rich Tozzoli,Independent Engineer

Everyday, more of the recording and post production studio infrastructure collapses into the frame of a computer. This two-part event will focus on technical concerns unique to DAW-based recording, as well as present power user application techniques.

Bits and Bytes and Bottlenecks

Inside a DAW, music exists as data, but a special kind of data that is sensitive to issues like word length, processor resolution, headroom and delay compensation. These aren't the sexy parts of the feature set, to be sure, but attention to the details can noticeably and measurably improve the final product. Our presentations will focus on some of the core issues that require attention.

Beyond the Basics

While a DAW can function as a simple recorder or editor, with plug-ins and other DSP tricks at their disposal, a DAW can also manipulate audio in sophisticated ways that were hitherto impossible. Power users will present advanced, cross platform techniques for the DAW user.

Moderator Frank Wells came to the recording industry after years of work as a radio broadcast engineer. Following nearly a decade as Chief of Technical Services for Masterfonics, Nashville, he defected to the world of trade journalism, first as the founding editor of *Audio Media, USA* and currently editor of *Pro Sound News* and executive editor of *Surround Professional*. He is also the current chair of the AES Nashville section.

David Channing has engineered and/or edited two Elton John records, the new Duncan Sheik record, Jewel's "Spirit" and projects with Enrique Iglesias, Melissa Etheridge, Shawn Colvin, and many more.

Lynn Fuston is owner/chief engineer of 3D Audio, with engineering credits that include Amy Grant, Mark O'Connor, DC Talk, Lee Greenwood, and the Newsboys. He has also produced a series of microphone and microphone pre-amplifier comparison CDs, and will soon release an A/D comparison disc.

Jim Kaiser is VP Central Region for the AES, and Director of Technology at Mastermix, Nashville. His background includes session engineering and technical consulting, including full facility technical design.

Nathaniel Kunkel is known for a broad body of recording work that includes CSNY, Little Feat, Linda Ronstadt and Lyle Lovett, including extensive work in 5.1 for DVD-A and Super Audio CD (the Graham Nash project, *Songs for Survivors*, and the 5.1 remix of James Taylor's *J.T.*).

Ralf Schluezen is CEO of TC Works, the software division of TC Electronic, based in Hamburg, Germany. He has extensive experience in plug-in design, including a stint with Steinberg before joining TC.

Rich Tozzoli is a New York-based engineer whose long list of credits includes extensive surround work with the likes of David Bowie, Foghat and Average White Band, along with numerous live classical recordings. He is also a contributing editor with *Surround Professional*.