AES 57TH INTERNATIONAL CONFERENCE

The Future of Audio Entertainment Technology: Cinema, Television, and the Internet

Hollywood, CA, USA
6–8 March, 2015

Organized in association with the TC on Sound for Digital Cinema & Television

Friday, March 6 — 8:00 am

OPENING REMARKS AND KEYNOTE ADDRESSES

[Welcome] Brian McCarty, Chair, AES Technical Committee Sound for Digital Cinema and Television, Conference Co-Chair

[Keynote] Acoustics for the Theater and Home—Moving Forward on a Foundation of Common Acoustical Science—*Floyd Toole*

PAPER SESSION 1 — 9.15 am

1-1 Predicting the In-Room Response of Cinemas from Anechoic Loudspeaker Data—*Linda A. Gedemer*, University of Salford, Manchester, UK; Harman International, Northridge, CA, USA

In recent years there has been a good deal of research undertaken on the in-room responses of cinemas, most notably on how they are measured and calibrated. Much of this research has been centered on measurement procedures such as the number of microphone positions to be spatially averaged, the measurement equipment itself, and the manipulation of the resultant room curve. This paper proposes that cinema measurement and calibration should instead place increased focus on the anechoic data of the loudspeaker being utilized. In this research, a well-defined loudspeaker was utilized and the playback system was held constant in each cinema so that the interaction between only the loudspeaker and the room could be observed.

1-2 The Audibility of Comb-Filtering Due to Cinema Screens—*David Elliott*,[1,2] *Keith Holland*,[2] *Philip Newell*[3]
[1]Sandy Brown Associates, UK
[2]ISVR, University of Southampton
[3]Acoustics Cosultant, Moaña, Spain

In most professional, commercial, and domestic cinemas the loudspeakers are mounted behind the screen so that the visual images on screen coincide with the position of the associated sound images. However, this requires the screen material to be both light-reflective and acoustically transparent, which is difficult to achieve. The resultant imperfect sound transmission gives rise to the sound from the loudspeakers being reflected back from the screen and then forward again as it reflects from the loudspeaker. The interference between the initial sound and the subsequent reflections gives rise to comb-filtering that depends upon the reflection coefficients of the screen and loudspeaker and the spacing between the screen and the loudspeaker. A change in separation will cause the respective phase difference between the direct and reflected waves to vary, thus yielding a different resultant wave at the point of the receiver. Little research has been carried out to date on the audibility of this comb filtering, and in particular, given a certain screen/loudspeaker setup, the effect of varying the spacing has not been investigated thoroughly. By modelling the system using MATLAB, the behavior of the perforated screen when subjected to an incident sound wave can be predicted. Once this is achieved, the effective filter for different separations between screen and loudspeaker can be used to create a database of simulated output signals that represent a portion of cinema audio being played through the screen/loudspeaker system in different geometric setups. These signals can then be used to determine the audibility of the comb-filtering effect at different separations via a comparative subjective testing technique. The results of these tests suggest that the audibility is, at the very least, barely audible at all separation distances.

1-3 Electroacoustic Measurements on Cinema B Chains in Australia—*David Murphy*, Krix Loudspeakers, Hackham, South Australia, Australia

Frequency responses and RTA measurement results of cinema screen systems are presented for a small selection of cinemas sizes and shapes in Australia. Measurements have been made using time gated techniques (MLSSA) and the traditional pink noise Real Time Analyser (RTA) process with a microphone array. Results from these methods will be compared with the manufacturer's quasi anechoic measurements, and probable reasons for the differences will be discussed.

1-4 Equalizing Effects of Perforated Movie Screens and the Future—*Glenn Leembruggen*,[1] *Philip Newell*[2]
[1]Acoustic Directions Pty. Ltd., Petersham, NSW, Australia
[2]Acoustics Cosultant, Moaña, Spain

Perforated cinema screens are currently in widespread use in cinemas and dubbing stages because to their high light-reflectivity. However, in the acoustical domain, such screens form a low pass filter that attenuates the high frequency response of the cinema loudspeakers. Recent studies have shown that the X-curve, which has long been adopted by the SMPTE as standard responses curve, primarily reflects this low pass filter action. It is therefore of significance to explore the ramifications of equalizing the effects of this low pass filter, in order to provide a flat frequency response for listeners. This paper reports on measurements of the spectral content of several movies and assesses the impacts on the headroom in the playback chain and the long-term power requirements of the loudspeaker drivers that equalization of this type would impose with the spectral content of these movies.

BREAK 10:15–10.30 am

WORKSHOP 10:30 am

Low Frequency Issues: Cinema/Home/Internet

Chair: **Anthony Grimani**, President, MSR Inc.

Panelists: *Manny LaCarrubba*, Chief Engineer, Sausalito Audio
Glenn Leembruggen, Acoustic Directions Pty. Ltd., Sydney University

LUNCH 12:00–1.00 pm

PAPER SESSION 2 1.00 pm

2-1 Enhanced Wide-Area Low-Frequency Sound Reproduction in Cinemas: Effective and Practical Alternatives to Current Sub-Optimal Calibration Strategies—*Adam J. Hill*,[1] *Malcolm O. J. Hawksford*,[2] *Philip Newell*[3], Acoustic Consultant
[1]University of Derby, UK
[2]University of Essex, UK
[3]Acoustics Cosultant, Moaña, Spain

This paper explores strategies for achieving accurate wide-area low-frequency sound reproduction in cinemas. Current standards for B-Chain calibration call for single channel low-frequency equalization aided by either single-point or spatially-averaged response measurements, an approach only applicable to a reasonably spatially invariant low-frequency response. A holistic approach to low-frequency coverage optimization is presented exploiting subwoofer arrays, their positioning, and multi-point signal processing. Acoustic-field examples are presented using finite-difference time-domain (FDTD) modeling software that expose a potential for superior wide-area signal reconstruction over that achieved using the current standards and recommendations

2-2 Considerations for the Generation and Measurement of Low Frequency Effects in Cinema Rooms—*Julius Newell*,[1] *Philip Newell*,[2] *Keith Holland*[3]
[1]Newell Acoustic Engineering, Lisbon, Portugal
[2]Acoustics Cosultant, Moaña, Spain
[3]ISVR, University of Southampton, UK

Little has been published about the repercussions of different source locations and measuring positions for the Low-Frequency Effects (LFE) loudspeakers in cinemas and dubbing theaters. The aim of this study is to determine the effects of the number and position of the loudspeakers on the uniformity of the response over the listening area and to assess the effect of the measuring of those responses by the choices made regarding the positioning of the microphones within typically used arrays.

WORKSHOP 1.30 pm

Speech and Dialog Intelligibility in Audio Entertainment—"Great Picture – Pity about the Dialogue"

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

Panelists: *Lon Bender*, Hearing is Precious
Kurt Graffy, Arup Engineering
David L. Smith, Bose Corp.
Microphone Representative TBA

BREAK 2:45–3:00 pm

PANEL DISCUSSION 3.00 pm

KEYNOTE ADDRESS 4.00 pm

[Keynote] Production of Audio for the Future—*Louis Hernandez Jr.*, Chairman and CEO,Avid Technologies, Inc.

RECEPTION: HOLLYWOOD ROOSEVELT 5:30–7:30 pm

Saturday, March 7

WORKSHOP 8:00 am

Loudness Wars at the Cinema: Noise Induced Impairment and Getting Our Dynamic Range Back

Robert T. Sataloff, M.D., D.M.A., F.A.C.S., Professor and Chairman, Department of Otolaryngology - Head and Neck Surgery, Senior Associate Dean for Clinical Academic Specialties, Drexel University College of Medicine

Eelco Grimm, Grimm Audio, Utrecht, Netherlands

PAPER SESSION 3 9.45 am

3-1 Dynamics and Low Frequency Ratio in Popular Music Recordings Since 1965—*Michael Oehler*,[1] *Christoph Reuter*,[2] *Isabella Czedik-Eysenberg*[2]
[1]University of Applied Sciences Düsseldorf, Düsseldorf, Germany
[2]University of Vienna, Vienna, Austria

The loudness, dynamic range, and energy distribution in low-frequency bands of popular music are analyzed. One objective was to operationalize popular music and construct a robust, balanced sample that covers a specific but relevant music market regarding annual revenues. The sample consists of the

"German Top 40" year-end charts from 1965 to 2013. Furthermore, different methods of measurement, such as LKFS or dBFS RMS, are used and compared. It could be shown that there was a significant increase of loudness, a decrease of the dynamic range, and an increasing importance of the low-frequency bands over time. While our results correspond to most previous research, there is a major difference regarding the recent data. It is frequently mentioned in studies that the process of decreasing dynamic range peaked in 2004 and, after that, the opposite trend occurred, namely, an increase in dynamic range. In the German music market, however, this seems to be true only for the time span from 2004 to 2010. From 2011 to 2013 a significant decrease of the dynamic range and an increase in loudness were found.

BREAK **10:00–10:15 am**

WORKSHOP **10:15 am**

Opportunities and Challenges in The Transition to Streamed Delivery of Audio Content

Chair: **Roger Charlesworth**, DTV Audio Group

Panelists: *Tim Carroll*, Telos Alliance
Sean Richardson, Director/Principal Audio Engineer, Starz Entertainment
Jeff Riedmiller, Dolby Laboratories
Jeff Dean, Meridian Audio Ltd.

PAPER SESSION 4 **11:45 am**

4-1 Adaptive Audio Reproduction Using Personalized Compression—*Andrew Mason,[1] Nick Jillings,[2] Zheng Ma,[3] Joshua D. Reiss,[3] Frank Melchior[1]*
[1]British Broadcasting Corporation, London, UK
[2]Birmingham City University, Birmingham, UK
[3]Queen Mary University of London, London, UK

Audio quality is very important to broadcasters' audiences, and unwanted loudness variations do compromise the quality of experience for the listener. Dynamic range control applied by the broadcaster can go some way to avoiding problems but can never take the individual environment of the listener into account. The listening conditions are a significant factor to be taken into account when dynamic range control is applied. The web audio API provided by HTML5 offers the possibility of performing dynamic range control under the control of the listener, tailoring it optimally for their individual situation. We have developed a system that demonstrates that this is achievable in a modern web browser. The implementation controls the compressor based on environmental noise level measured using the microphone present in most mobile device audio players.

4-2 Auditory Distance Perception with Static and Dynamic Binaural Rendering—*Gavin Kearney,[1] Xujia Liu,[1] Andrew Manns,[1] Marcin Gorzel[2]*
[1]University of York, York, UK
[2]Trinity College Dublin, Dublin, Ireland

In this paper we investigate the perception of sound source distance in relation to static and dynamic binaural systems. Reference data for distance perception in the median plane is first presented, which shows through subjective evaluation that under the test conditions there is no perceived difference in distance perception with sound sources presented at different azimuthal angles to the head. This data forms the hypothesis that head movements do not play an important role in auditory perception of source distance in real rooms. This is verified in relation to binaural systems through comparison of distance perception of binaural recordings versus head-tracked binaural rendering. The results also demonstrate that non-individualized head-related transfer functions can be effective for distance perception when compared to real sources.

4-3 Using a Binaural Spatialization Tool in Current Sound Post-Production and Broadcast Workflow—*Pierre Bézard,[1] Matthieu Aussal,[2,3] Matthieu Parmentier[4]*
[1]Louis Lumiere School for Cinema,La Plaine Saint-Denis CEDEX, France
[2]Centre de Mathématiques Appliquées de l'Ecole Polytechnique, Palaiseau, CEDEX France
[3]Digital Media Solutions, Noisiel, France
[4]France Televisions, Paris, France

This paper describes an experiment designed to study the differences between 5.1 audio played through loudspeakers and headphones in binaural and between compressed and uncompressed audio files. Differences in terms of spatial impression and of overall quality of the sound have been studied. This experiment was made in the context of "NouvOson," a Radio France website launched in March 2013 (http://nouvoson.radiofrance.fr) where audio contents are available online both in native 5.1 and processed in binaural using SpherAudio software by Digital Media Solutions. It also concerned the BILI Project, dealing with BInaural LIstening, involving Radio France, France Televisions, and DMS. Binaural processing theoretically allows the reproduction of 3D sound when listening through headphones; however, this technology still faces issues. These are not only due to the actual limits of research and development but also to the way we listen to and localize sounds. This experiment has shown that spatial characteristics, as well as timbre of the sound are modified. Besides, no real difference in the listener's perception has been found between binaural uncompressed files and AAC 192 kbps as well as MP3 192 kbps files.

LUNCH **12:30–1.15 pm**

WORKSHOP **1:15 pm**

Headphones: Design, Subjective/Objective Performance and Entertainment Content Production

Chair: **Robert Schulein**, RBS Consultants, Immersav.com

Robert Schulein, RBS Consultants, Immersav.com:

Introduction to headphone-based spatial audio

Tools and techniques for natural acoustical creation of im mersive sound field recordings using binaural head capture

Tom Ammermann, New Audio Technology GmbH:

Creating immersive content for headphone applications using digital production tools

[Headphone Demonstrations Supported by Conference Sponsor Sennheiser Electronic Corp.]

Headphones: Design, Subjective/Objective Performance and Entertainment Content Production (continued)

Tom Ammermann:

Interactive spatial audio engine for computer games using immersive headphone virtualization

Headphone measurements, user performance preferences, compensation techniques and standardization efforts

Chair: **Robert Schulein**

Panelists: *Todd Welti*, Harman International
Martin Walsh, DTS
Tad Rollow, Sennheiser

BREAK 3:30–3:45 pm

WORKSHOP 3:45 pm

Applications and Challenges of Object-Based Audio

Frank Melchior, Lead Technologist, BBC Research & Development; Presenter & Session Chair

Nuno Fonseca, Polytechnic Institute of Leira, Portugal
Sound Design Using Particles Systems and Virtual Microphones

PAPER SESSION 5 5:00–5:30 pm

5-1 A Description of an Object-Based Audio Workflow for Media Productions—*Alejandro Gasull Ruiz, Christoph Sladeczek, Thomas Sporer*, Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

The object-based audio approach can provide the audience with a new immersive sound experience that so far could not be achieved with conventional channel-based productions. Moved by these benefits, the interest of the media industry in applying this concept has increased dramatically in the recent years. However, the object-based approach does not only imply an improved listening experience, but it also brings a much more flexible workflow and new possibilities to sound engineers, from the recording process to the playback systems. This paper presents the results of a research project in this regard, in order to provide the creative production process with a flexible and simple workflow using the advantages of the object-based audio technologies.

5-2 Using Audio Objects and Spatial Audio in Sports Broadcasting—*Ulli Scuda,*[1] *Hanne Stenzel,*[1] *Dennis Baxter*[2]
[1]Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2]DennisBaxterSound LLC, Atlanta, Georgia, USA

This paper gives an overview on how immersive sound and interactive audio objects can be used in sports TV broadcasting with regard to the characteristic sound elements of different sport families. New broadcast standards target at bringing object based audio and 3D-Audio to the home. Once the technical infrastructure has been established, content is needed to make use of the innovative features. Sports TV broadcasting is a promising field for new audio standards because of its wide range of sound scenarios and its high market value. Some of these are presented and investigated with the question in mind, how audio objects can be captured in real time and how spatial audio can be used to make sports TV more exciting.

OPTIONAL TOUR: UNIVERSAL STUDIOS 5:30 pm

Sunday, March 8 IMMERSIVE AUDIO DAY

KEYNOTE ADDRESS 8:00 am

Immersive Audio: Status and Challenges—
Francis Rumsey

PAPER SESSION 6 9:00 am

6-1 Properties of Large-Scale Sound Field Synthesis—*Jens Ahrens, Hagen Wierstorf*, University of Technology Berlin Berlin, Germany

Sound field synthesis has been pursued as a promising approach for spatial audio reproduction for large listening areas. Research is typically performed on small and mid-size systems. An increasing number of systems of cinema size and larger exist, which have shown to exhibit properties that cannot be observed with smaller setups. In particular, practical limitations lead to artifacts whose perceptual saliency increases with array size. Depending on the situation these artifacts are most prominent in time domain or in frequency domain. In this paper we review the current state of knowledge on the properties of sound field synthesis using large-size loudspeaker arrays regarding both direct sound and reverberation.

6-2 nouvOson Website: How a Public Radio Broadcaster Makes Immersive Audio Accessible to the General Public—*Herve Dejardin, Edwige Ronciere*, Radio France, Paris, France

In March 2013, Radio France launched a new part of its website, called nouvOson, to broadcast 5.1 and binaural sound. A binaural technique was initially chosen to reach people who do not have a home theater set for 5.1 productions as well as for mobile applications. At the same time, Radio France became a founding member of the collaborative research project in binaural listening BILI, whose aim is to find an accessible way to personalize HRTF (Head-Related Transfer Function) for the general public. This paper discusses the progress of the nouvOson player project since 2013 and its outlook.

6-3 Producing 3D Audio in Ambisonics—*Matthias Frank, Franz Zotter, Alois Sontacchi*, University of Music and Performing Arts Graz, Graz, Austria

Ambisonics is a 3D recording and playback method that is based on the representation of the sound field excitation as a decomposition into spherical harmonics. This representation facilitates spatial sound production that is independent of the playback system. The adaptation to a given playback system (loudspeakers or motion-tracked headphones) is achieved by a suitable decoder. This paper gives an overview of the current state-of-the-art in Ambisonics including content production using Ambisonic main microphone arrays or panning of virtual sources, spatial effects, and reproduction by loudspeakers and headphones. The software for the whole production chain is already available as a VST-plugin suite for digital audio workstations.

BREAK 9:45–10:00 am

WORKSHOP 10:00 am

Cinema Immersive Audio Delivery Standards

Chair: **Brian Vessa**, Sony Pictures Entertainment

Panelists: *Brian Claypool*, Barco
Ton Kalker, DTS
Charles Robinson, Dolby Laboratories
Bert Van Daele, Auro Technologies

Co-Sponsored by the SMPTE (Society of Motion Picture & Television Engineers)

IMMERSIVE AUDIO DEMONSTRATIONS 11:30 am

LUNCH 12:15–1:15 pm

PAPER SESSION 7 1:15 pm

7-1 Does Stereoscopy Change Our Perception of Soundtracks?—*Etienne Hendrickx, Mathieu Paquier, Vincent Keohl*, University of Brest, Brest, France

Few psychoacoustic studies have been carried out about sound related to stereoscopic movies, and even less have tried to compare the effects of 2D and 3D video playback on the perception of soundtracks. In the present study 8 audio-visual sequences were presented in a theater to 44 subjects, in their non-stereoscopic version (2D) and their stereoscopic version (s-3D). For each presentation, subjects had to judge to what extent the 5.1 sound mix sounded frontal or "surround," in order to verify whether stereoscopy could have an influence on the perception of the front/rear balance of ambience sound. Results showed that the influence of stereoscopy was weak—for two sequences out of eight, subjects perceived the soundtrack as more frontal when the sequence was projected in its s-3D version. For the other sequences, the influence of stereoscopy was not significant.

7-2 Localization of Audio Objects in Multichannel Reproduction Systems—*Thomas Sporer*[1], *Judith Liebetrau,*[1] *Stephan Werner*[2], *Sara Kepplinger*,[1,2] *Timo Gabb,*[12] *Theresa Siedler*[2]
[1]Fraunhofer Institute of Digital Media Technology IDMT, Ilmenau, Germany
[2]Technische Universität Ilmenau, Ilmenau, Germany

With spatial audio systems the illusion of being in a sound scene should be created, which means to reproduce a natural sound field enveloping the listener. Perceived audio quality is related to the position of the listener in the reproduction room. Former studies indicated that with increased number of loudspeakers for reproduction a larger listening area with high audio quality can be created. In these studies quality was mostly assessed in terms of overall quality or basic audio quality; different factors like distortion, coloration, envelopment and localization were not assessed separately. The new MPEG-H standard should enable high efficiency coding and media delivery in heterogeneous environments. Part 3 of MPEG-H deals with the coding and delivery of high-quality 3D audio content, whereas for the producer the number of loudspeaker channels used for reproduction is unknown and flexible. In MPEG the overall audio quality was used for selection of most suitable coding and rendering algorithm. In addition to the MPEG testing and selection process, we conducted a study on localization performance dependent on listener position, number of loudspeakers and rendering algorithms. A new test method with an innovative testing framework was applied, reducing the influence of visual cues during testing. This method was used to evaluate three different rendering schemes and different loudspeaker setups at three different listening positions. The results showed that reproduction via 22 loudspeakers gives a better localization accuracy compared to 10 and 5 channels. More interestingly, a clear shift of the perceived position of sound objects to the right was observed. In former studies, where the loudspeakers were visible, such a shift could not be observed. This might support our assumption that the vision highly influences the localization perception and therefore listening tests without visual cues are required.

7-3 Next Generation Surround Decoding and Upmixing for Consumer and Professional Applications—*Mark Vinton, David McGrath, Charles Robinson, Phillip Brown*, Dolby Laboratories, USA

This paper describes a new spatial audio algorithm that creates a channel-based three-dimensional sound scene from two or more input channels. The algorithm was designed to decode matrix encoded programs (Lt/Rt). It is also an effective stereo upmixer; the signal relationships that guide the decoding algorithm (e.g., cross correlation) also provide appropriate cues to the intended spatial scene for standard, unencoded programs—we decode the artist's intent. Input channel configurations with more than two channels are decomposed into channel pairs that are then processed independently. Improvements relative to existing surround decoding systems include improved selectivity and separation due to multi-band processing; increased listener envelopment through independent processing for direct and diffuse signal components, and user-adjustable decorrelation; and support for an arbitrary number of output channels at user specified locations including elevation. The system described has been recently deployed in consumer and professional products for home, mobile, and cinema applications.

WORKSHOP 2:00 pm

Integrating Object-, Scene-, and Channel-Based Immersive Audio for Delivery to the Home

Chair: **Francis Rumsey**

Presenters: *Wilfried van Baelen*, Auro Technologies
Brett Crockett, Dolby Laboratories
Jean-Marc Jot, DTS
Schuyler Quackenbush, MPEG-H Alliance

BREAK 3:15–3:30 pm

WORKSHOP 3:30 pm

How to Make Big Small—Can We Really Bring Immersive Sound to the Home?

Chair: **Francis Rumsey**

Arnaud Laborie: Engineering a Home Cinema Processor to Handle Immersive Audio

Frank Melchior: Where Should We Draw the Line? The Baseline Renderer and Why It Is Essential for the Success for Next Generation Audio

Brian Vessa: Translating Immersive Cinema Mixes to the Home: Challenges and Compromises for the Studio

SUMMARY AND FUTURE PLANS 4:30 pm

Brian McCarty
Sean Olive
Francis Rumsey