Extended Program/Book of Abstracts of the

EAA Joint Symposium on Auralization and Ambisonics

2014, Berlin

Editors:

Prof. Dr. Stefan Weinzierl
Audio Communication Group, TU Berlin

Prof. Dr. rer. nat. Michael Vorländer
Institute of Technical Acoustics, RWTH Aachen

Dr. Franz Zotter
Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz

Dr. Hans-Joachim Maempel
State Institute for Music Research Stiftung Preußischer Kulturbesitz

Alexander Lindau
Audio Communication Group, TU Berlin

Printed: 5/5/2014 9:14:00 AM

Thursday, April 3rd 2014

Plenary Talk

Author: Lauri Savioja

Title: Trends in Room Acoustics Modeling

Abstract: None

Paper Session A

PA.A.1

Author: Stefan Drechsler

Title: An Algorithm for Automatic Geometry Simplification for Room Acoustical Simulation Based on Regression Planes

Abstract: The numerical simulation of room acoustics is based - besides other properties - on the geometry of the evaluated room. During the planning phase of a building, these geometry data are usually available as a closed polyhedron stored in a computer readable CAD file format that was prepared for other purposes as, e.g., visualization. These geometry data must be processed to extract a reasonable small number of acoustically appropriate polygons to reduce calculation times for the simulation of room acoustics. The size of the appropriate polygons is closely related to the sound wave length. This paper presents an algorithm to reduce the number of polygons by iteratively replacing a group of polygons by a bigger polygon. This polygon is defined by a regression plane minimizing the squared distances of the points in the polygons of the group. The group of polygons is a growing set containing polygons with vertices within a layer of limited thickness around the intermediate regression plane. This thickness can be adapted to the examined wavelength. Besides the presentation of the basic algorithm with examples, some hints for the practical application in room acoustics are discussed together with the possibility to use intermediate results from the calculation of the regression planes to estimate the additional scattering coefficients that are necessary to model the original roughness of the simplified surface.

PA.A.2

Author: Alexander Pohl, Uwe M. Stephenson

Title: Combining Higher Order Reflections with Diffractions without Explosion of Computation Time: The Sound Particle Radiosity Method

Abstract: The simulation of sound propagation in large rooms and urban environments is mainly performed by geometric simulation methods like ray tracing or the Sound Particle Simulation Method (SPSM). Hence, a severe deficiency is that wave effects are not included, especially if screening or diffraction effects are important. A method to introduce diffraction is the Uncertainty relation Based Diffraction (UBD) model, which has been successfully evaluated recently. To find close edges as sources of diffraction, a subdivision of the room into convex subspaces is performed by virtual walls. However, this causes a recursive split-up of Sound Particles (SPs) at each diffraction event. This effect should be compensated by a reunification of SPs. Therefore, the Sound Particle Radiosity (SPR) has been found that combines the SPSM with an advantage of the radiosity method: the re-unification of sound energy that uses a discretization of the surface into small patches. Now, SPR has been extended to 3D for the first time. To increase the available memory and to decrease the computation time, a parallelization has been implemented for the first time. First results indicate that the discretization of the virtual walls into patches yields additional but tolerable errors in the simulation of diffraction. However, even in 2D, SPR requires a huge memory. To solve this problem in 3D remains a great challenge, even more for more complex rooms. Also a method for a convex subdivision to 3D still has to be found.

Poster Session A

PO.A.1

Author/s: Matthias Kronlachner, Franz Zotter

Title: Warping and Directional Loudness Manipulation Tools for Ambisonics

Abstract: The production and re-production of sound scenes in the Ambisonic domain offers flexibility regarding the loudspeaker placement around the listening area. Correct decoding should result in a spatial audio perspective that is independent of the loudspeaker configuration. In case a modification of this perspective is needed, or directional alterations of the amplitude, applying such transformations in the Ambisonic domain is more challenging than directly changing the production or loudspeaker placement. Nevertheless, for achieving flexibility and consistency, finding such algorithms in the Ambisonic domain is feasible and the development of such is presented in this work. Apart from rotating the sound field in its entirety, only few transformation matrices have been described for Higher Order Ambisonics. One publications even claims a dominance transform works for first order only. This paper presents an implementation for finding appropriate transformation matrices for Higher Order Ambisonics. Moreover, available tools do not equip the audio engineer with necessary metering tools to monitor the directional loudness levels in Ambisonic recordings or productions. This paper presents a software for the visualization of Ambisonic signals which helps to comprehend and verify the developed transformations.

PO.A.2

Author: Sönke Pelzer, Bruno Masiero, Michael Vorländer

Title: 3D Reproduction of Room Auralizations by Combining Intensity Panning, Crosstalk Cancellation and Ambisonics

Abstract: When auditory selective attention started to be in the interest of acoustical and psychological research the dichotic reproduction was used due to limited technical capabilities, but still today many psychological experiments on the cocktail-party effect are performed using a dichotic reproduction. For more realistic scenes and a better spatial impression, auralization can be used and the presentation of stimuli should be binaural. The most intuitive but also most complex way of reproducing a realistic virtual scene is a loudspeaker-setup in an anechoic room. Binaural reproduction via headphones is more convenient, especially for psychologists without measurement facilities. In this study a psychologically established experiment has been conducted with real sources (loudspeakers in an anechoic chamber) and binaural reproduction via headphones using individual as well as nonindividual HRTFs. In contrast to usual acoustical localization tasks subjects are not asked to focus directly on the localization, the accuracy of the sources' positions, the externalization, etc. Instead, the localization quality is analyzed indirectly, since subjects mainly focus on the psychological attention task. Subjects listen to two competing speakers at different positions in space and have to report and categorize the target speaker's words guided by visual cues. Independent analyzed variables are reaction times and error rates for a total of 73 subjects. This way it is obtained an objective measure of the quality of methods for reproduction of auralized sounds.

PO.A.3

Author/s: Torben Wendt, Steven van de Par, Stephan D. Ewert

Title: Perceptual and Room Acoustical Evaluation of a Computational Efficient Binaural Room Impulse Response Simulation Method

Abstract: A fast and perceptively plausible method for synthesizing binaural room impulse responses (BRIR) is presented. The method is principally suited for application in dynamic and interactive evaluation environments (e. g., for hearing aid development), psychophysics with adaptively changing room reverberation, or simulation and computer games. In order to achieve a low computational cost, the proposed method is based on a hybrid approach. Using the image source model (ISM; Allen and Berkley [J.Acoust. Soc. Am. Vol. 66(4), 1979]), early reflections are computed in a geometrically exact way, taking into account source and listener positions as well as wall absorption and room geometry approximated by a "shoebox". The ISM is restricted to a low order and the reverberant tail is generated by a feedback delay network (FDN; Jot and Chaigne [Proc. 90th AES Conv., 1991]), which offers the advantages of a low computational complexity on the one hand and an explicit control of the frequency dependent decay characteristics on the other hand. The FDN approach was extended, taking spatial room properties into account such as room dimensions and different absorption characteristics of the walls. Moreover, the listener orientation and position in the room is considered to achieve a realistic spatial reverberant field. Technical and subjective evaluations were performed by comparing measured and synthesized BRIRs for various rooms. Mostly, a high accuracy both for some common room acoustical parameters and subjective sound properties was found. In addition, an analysis will be presented of several methods to include room geometry in the FDN.

PO.A.4

Author/s: David A. Dick, Michelle C. Vigeant

Title: A Comparison of Late Lateral Energy (GLL) and Lateral Energy Fraction (LF) Measurements Using a Spherical Microphone Array and Conventional Methods

Abstract: Late Lateral Energy Level (GLL) and Lateral Energy Fraction (LF) are two room acoustics measures that have both been shown to correlate with certain aspects of the spatial impression of a listening space. In order to obtain these quantities, the lateral energy must be measured, which is typically carried out using microphones with a figure-of-eight (figure-8) polar pickup pattern. However, most commercially available figure-8 microphones are intended for use in audio recording applications, and are not laboratory-grade or designed for room acoustics impulse response (IR) measurements. Such microphones may suffer from non-ideal frequency response and/or directivity patterns. This study compares measurements that were taken in a 2500 seat auditorium using an omni-directional and studio-grade figure-8 microphone pair versus the omni-directional and dipole components extracted from a 32 element spherical microphone array. The results show that the two measurement methods agree in the 500 Hz and higher octave bands, but differ at low frequencies due to differences in the directivity patterns. The difference of the LF average from 125 Hz to 1 kHz for the two methods was between 0.59 and 1.81 just noticeable differences (JNDs) at the six receiver locations. The difference of the GLL average from 125 Hz to 1 kHz for the two methods was between 0.02 and 0.48 JNDs (applying the JND for strength of 1 dB). It was also found that repeatability error was present at one of the six receiver locations for the LF measurements, but was very small for the GLL measurements.

PO.A.5

Author/s: Antti Kuusinen

Title: An Anechoic Audio Corpus for Room Acoustics and Related Studies

Abstract: Anechoic or semi-anechoic instrument recordings are readily available for academic purposes on a few different sites online. Anechoic recordings are commonly used in auralizations, which today practically means convolving recordings with simulated or measured room impulse responses. Besides the possibility of being used as such, these recordings offer other possibilities for the generation of test stimuli. Many studies, such as, studies on auditory distance perception or source separation, would benefit from available experimental materials which would not be strictly musical but could still be linked to the perception of musical stimuli. The goal of the current investigation is to develop a procedure for generating such materials, i.e., an anechoic audio corpus which can be used in the future investigations of room acoustics and in related fields. Moreover, the aim is to provide a framework for further development of processes where a large number of stimuli can be generated in a systematic way. In this study, the proposed framework is instantiated by producing two sets of stimuli by either directly segmenting anechoic music or randomly combining different segments of anechoic instrument tracks. Music information retrieval (MIR) approach is used to calculate 14 musical features of the generated sets of stimuli. Principal component analysis is used to analyse the sample spaces enabling the experimenter to select a small number stimuli with desired characteristics. The benefits and drawbacks of this stimuli generation approach including some important theoretical underpinnings of experimental design are also discussed.

PO.A.6

Author/s: Tim T. Ziemer, Rolf Bader

Title: Introducing a Method to Measure, Store, and Reconstruct the Spatial Sound Radiation Characteristics of Musical Instruments

Abstract: Musical instruments radiate sound in various spatial patterns. Several methods have established to measure this spatial sound Radiation Characteristic RC. Approaches exist to physically recreate the RC by means of loudspeaker arrays. Investigations on the psychoacoustic effects of physical parameters such as RC have been carried out and some correlations have been found between physical and psychoacoustic parameters. This paper introduces a concept to combine these three fields of research in a complete chain from measurement, analysis and storage to physical reproduction of the RC of musical instruments which allows for physical and psychoacoustic investigations. Based on this concept, a system including measurement, digital signal processing and sound field synthesis has been implemented and tested. This system is able to measure, analyze and store the RC of musical instruments and reconstruct it within an extended listening area with high precision.

PO.A.7

Author/s: Johannes Klein, Martin Pollow, Michael Vorländer

Title: Optimized Spherical Sound Source for Auralization with Arbitrary Source Directivity

Abstract: The auralization of measured room impulse responses (RIRs) is traditionally bound to the directivity of the source as well as of the receiver. For the comparability of room acoustical measurements ISO 3382 requires the source and the receiver to be of an omnidirectional directivity. Other source directivity patterns cannot be auralized using RIRs obtained this way. In order to include the spatial information the room impulse response has either to be measured with a sound source of the desired directivity or - assuming the room to be a linear time-invariant system - it can be generated by superposing a set of measurements with a source of known directivity. The advantage of the latter method is that it generates a set of RIRs that can be used to derive the RIR for an arbitrary directivity up to a certain spherical harmonic order in post processing. This article describes a superposition method and a specialized measurement source for the measurement of room impulse responses for arbitrary source directivity and discusses their capabilities and the limitations. The measurement source was developed using an analytical model. The directivity patterns used for the post processing originate from high-resolution measurements of the actual device. The deviation compared to the analytical model is analyzed regarding the radiation pattern and the achievable synthesis accuracy.

PO.A.8

Author/s: Martin Pollow, Johannes Klein, Stefan Zillekens, Janina Fels

Title: Analysis and Processing of Rapidly Measured Individual HRTFs for Auralization

Abstract: Head-related transfer functions (HRTFs) are commonly used for the auralization of binaural acoustic scenes. This can be done either by using the HRTFs in discrete directions or by transferring them to a spatially continuous representation. The latter method is more sensitive to measurement errors due to imperfections of the measurement setup or due to spatial undersampling, but has the advantage of a physically correct interpolation and range extrapolation of the perceived sources when the description is correct. With the new trend of using individualized HRTFs by measuring human listeners instead of generic dummy heads, the requirements regarding the measurement speed have risen, yielding potentially higher

deviations than with slowly measured generic HRTFs. This contribution uses a loudspeaker array system for the measurement of individual HRTFs and evaluates and possibly corrects the inaccuracies of the measurement results. One step is to account for the actual loudspeaker positions, which can be determined using the results of a specialized microphone array with an optimization algorithm. The scattering influence of the equipment itself can be minimized by a modification of the original setup. A comparison of the simultaneously measured left and right ear HRTFs gives hints about the actual position of a possibly mislocated subject's head. The goal is to derive an accurate description of individually measured HRTFs in order to successfully perform auralization using features of spherical acoustics such as interpolation and range extrapolation of the measured HRTFs.

Paper Session B

PA.B.1

Author/s: Jonathan A. Hargreaves, Yiu W. Lam

Title: An Energy Interpretation of the Kirchhoff-Helmholtz Boundary Integral Equation and its Application to Sound Field Synthesis

Abstract: Most spatial audio reproduction systems have the constraint that all loudspeakers must be equidistant from the listener, a property which is difficult to achieve in real rooms. In traditional Ambisonics this arises because the spherical harmonic functions, which are used to encode the spatial sound-field, are orthonormal over a sphere and because loudspeaker proximity is not fully addressed. Recently, significant progress to lift this restriction has been made through the theory of sound field synthesis, which formalizes various spatial audio systems in a mathematical framework based on the single layer potential. This approach has shown many benefits but the theory, which treats audio rendering as a sound-soft scattering problem, can appear one step removed from the physical reality and also possesses frequencies where the solution is non-unique. In the time-domain Boundary Element Method approaches to address such non-uniqueness amount to statements which test the flow of acoustic energy rather than considering pressure alone. This paper applies that notion to spatial audio rendering by re-examining the Kirchhoff-Helmholtz integral equation as a wave-matching metric, and suggests a physical interpretation of its kernel in terms of common energy flux between waves. It is shown that the spherical basis functions (spherical harmonics multiplied by spherical Bessel or Hankel functions) are orthogonal over any arbitrary surface with respect to this metric. Finally other applications are discussed, including design of high-order microphone arrays and the coupling of virtual acoustic models to auralization hardware.

PA.B.2

Author/s: Frank Schultz, Sascha Spors

Title: Comparing Approaches to the Spherical and Planar Single Layer Potentials for Interior Sound Field Synthesis

Abstract: Practical approaches to sound field synthesis are described by the single layer potential for which monopoles are driven by appropriate driving functions. The solution of the sound field synthesis problem is represented by the Helmholtz integral equation, postulating the superposition of a single and a double layer potential. In order to derive an implicit or explicit solution of the driving function the Helmholtz integral equation is simplified to a single layer potential. This is achieved by imposing boundary conditions on the involved scalar fields, i.e. the sound pressure or the Green's function. Basically three important approaches to an exact solution for the unknown driving function exist in literature. The first method uses a direct and explicit solution of the single layer potential by means of operator theory, which also can be interpreted as the solution of a convolution integral. A second approach imposes the homogeneous Dirichlet boundary condition on the pressure in the Helmholtz integral equation. This results in an explicit solution for the driving function which is equivalent to the first method. By imposing the homogeneous Neumann boundary condition on the freefield Green's function a third approach derives a suitable single layer potential driving function for planar problems only. The paper discusses the approaches in a consistent framework showing that the same principles hold for planar and spherical problems. We prove the equivalence of all three solution methods for planar problems, which extends the well-known derivation of the Rayleigh integral from the angular spectrum domain.

PA.B.3

Author/s: Jens Ahrens

Title: Challenges in the Creation of Artificial Reverberation for Sound Field Synthesis: Early Reflections and Room Modes

Abstract: Practical implementations of sound field synthesis evoke considerable artifacts that have to be considered in the creation of artificial reverberation. The most prominent artifact is spatial aliasing, which manifests itself as additional wave fronts that follow the desired synthetic wave front in time. These additional wave fronts propagate into different directions and occur at intervals that are similar to the intervals at which acoustic reflections occur in real rooms. It may be assumed that the human auditory system is not capable of differentiating aliasing artifacts and room reflections so that a synthetic reflection pattern should be designed such that it evokes a plausible pattern together with the aliased wave fronts. Two potential solutions are outlined. Finally, the capability of sound field synthesis of synthesizing room resonances (room modes) is analyzed and the promising results are illustrated based on numerical simulations.

PA.B.4

Author/s: Stephen Oxnard, Damian Murphy

Title: Achieving Realistic Auralisations Using an Efficient Hybrid 2D Multi-Plane FDTD Acoustic Model

Abstract: This research examines the validity of utilising a 2D multiplane FDTD acoustic model to simulate low frequency sound propagation as part of a hybrid room impulse response (RIR) synthesis system. Analytic results, pertaining to the comparison of simulated low frequency multiplane RIRs with both practical RIR measurements and 3D FDTD simulated RIRs, demonstrate that a good level of accuracy is attained through use of this hybrid modelling paradigm. This claim is further supported, in part, by comparative subjective test results. Furthermore, 2D multiplane simulations are shown to be far more efficient than full 3D FDTD modelling procedures as they achieve a 98% reduction in computation time.

Paper Session C

PA.C.1

Author/s: Frank Wefers, Jonas Stienen, Sönke Pelzer, Michael Vorländer

Title: Interactive Acoustic Virtual Environments Using Distributed Room Acoustic Simulations

Abstract: This publication presents how the computational resources of PC clusters can be used to realize low-latency room acoustic simulations for comprehensive virtual scenes. A benefit is not only to realize more complex scenes (including a multitude of sound sources, acoustically coupled-rooms with sound transmission), but also to minimize the system response times for prototyping applications (e.g. interactive change of materials or geometry) in simpler applications. PC clusters prove to be a suitable platform for room acoustic simulations, as the incorporated algorithms, the image source method and stochastic ray-tracing, are largely free of data interdependencies. For the computation in massive parallel systems the simulation of a room impulse response is separated into individual parts for the direct sound (DS), early reflections (ER) and diffuse late reverberation (LR). Additional decomposition concepts (e.g. individual image sources, frequency bands, sub volumes) are discussed. During user interaction (e.g. movement of the sources/listeners) the system is continuously issued new simulation tasks. A real-time scheduler decides on significant updates and assigns simulation tasks to available cluster nodes. Thereby the three simulation types are processed with different priorities. The multitude of (asynchronously) finished simulation tasks is transformed into room impulse responses. Convolution with the audio signals is realized by non-uniformly partitioned convolution in the frequency domain. The filter partitioning is adapted to the update rates of the individual impulse response parts (DS, ER, LR). Parallelization strategies, network protocols and performance figures are presented.

PA.C.2

Author/s: Matthew Azevedo, Jonah Sacks

Title: Auralization as an Architectural Design Tool

Abstract: Auralization provides a valuable tool that allows architects, building owners, and other decision-makers to directly experience the aural implications of design decisions and allows them to make more informed choices. Standard numerical metrics are difficult to relate to aural phenomena without significant practice and frequently fail to capture acoustical issues that are essential to the basic functionality of spaces. Consultants at Acentech have been using auralizations of full soundscapes including many independent sources as design and communication tools for a variety of projects including atria, lecture halls, theaters, and performance spaces. These auralizations have included natural speech and electro-acoustic reinforcement, crowd activity, interactions between PA systems and room acoustics, HVAC noise, wall and window transmission, and the subjective effects of

sound masking. In general, clients find the experience of listening to their as-yet unbuilt spaces to be exciting and useful. Though most are not trained listeners, they typically move quick-ly past the “wow” stage and into critical listening and candid discussion of the different acoustical treatments presented and of the overall sound of the space. This helps architects and project owners to feel connected to the acoustical aspect of the design, and it helps the team to agree on design decisions that may have significant implications regarding cost and aesthetics. This paper presents several case studies of projects where auralization was an integral part of the design process. Additionally, it describes a rapid auralization design and development process using a MaxMSP-based real-time ambisonic convolution platform.

Artistic contributions, demonstrations, and workshops on Thursday

Ambisonic Environment, Room 212 (Museum)

Author/s: Matthias Kronlachner

Title: Workshop: Production and Post-Production of Ambisonic Recordings

Abstract: This workshop presents cross-plattform audio plug-ins for creating Ambisonic recordings within the Digital Audio Workstations Reaper and Ardour. A step-by-step tutorial as well as several listening examples should demonstrate the possibilities of the software. The plug-in suite includes encoders, converter and modification tools, decoders for loudspeakers and headphones and a visual Ambisonic metering tool. The encoder plug-in may be connected to a central remote control, mapping the current position and audio level of every single track in the DAW onto the surface of a sphere. Recently filters have been designed at the Institute of Electronic Music and Acoustics Graz for converting the 32 microphone signals from the Eigenmike into 4th Order Ambisonics. Post production scenarios by applying the developed warping and directional loudness modification tools to Ambisonic field recordings are presented.

Ambisonic Environment, Room 212 (Museum)

Author/s: Sönke Pelzer

Title: Demo Presentation: 3D Reproduction of Room Auralizations by Combining Intensity Panning, Crosstalk Cancellation and Ambisonics

Abstract: This demonstration deals with the auralization of room acoustics using loudspeaker arrays. Three different 3D reproduction techniques were implemented into a room acoustics simulator: Ambisonics, Vector-base Amplitude Panning and Crosstalk Cancellation. The generation of spatial room impulse responses is split into direct sound, early reflections, and diffuse reverberation. Each part can be auralized with a different reproduction method, so that the effect of the method related to the part of the impulse response can be analyzed. Hybrid combinations reproduce the parts with different methods simultaneously. The mixed reproduction has the potential to benefit from advantages of each single method while avoiding their individual drawbacks. For example a method which allows good localization can be used to reproduce the direct sound and early reflections, while a method with higher immersion and envelopment is used for the diffuse decay. The workshop includes an introduction to the software framework and associated production work flow and shows helpful tools for the setup of loudspeaker-based 3D audio.

Demo Room (SIM)

Author/s: Frank Wefers

Title: Workshop: Interactive Acoustic Virtual Environments Using Room Acoustic Simulations

Abstract: This workshop presents an interactive virtual environment of a spanish medivial church, including the real-time simulation of room acoustics. Participants experience an audio-visual demo and can freely navigate through the building. The scene is visualized using a mobile, stereoscopic screen (Powerwall), equipped with optical tracking. Binaural 3D sound is reproduced using headphones and/or dynamic crosstalk cancellation. The room acoustic is simulated in real-time on a multi-core machine, using hybrid geometrical acoustics based on image sources and ray-tracing. The simulation of a room impulse response is separated into individual parts for the direct sound (DS), early reflections (ER) and diffuse late reverberation (LR). During user interaction (e.g. movement of the sources/listeners) the system continuously updates these parts. Real-time filtering is realized using partitioned frequency-domain convolution techniques, with adapted update rates for the individual impulse response parts (DS, ER, LR). The authors present the design of the system, it's software and hardware components and report on the creation process of the virtual scene.

Demo Room (SIM)

Author/s: Audio Communication Group, Berlin

Title: Demonstrations: Dynamic Binaural Recording and Reproduction

Abstract: The Audio Communication Group presents several demonstrations related to dynamic binaural recording and rendering. Hence, you may shortly borrow head and ears from our binaural measurement robot FABIAN a device for Fast and Automatic Binaural Impulse response AcquisitioN. Further, you may listen to the Audio Communication Group's Binaural Reference System demonstrating real-time binaural rendering and featuring the recently developed binaural headphones BK211 together with the DSP-controlled amplifier BKamp. As content we present a comparison of (1) a data-, (2) a model- and (3) several massive-multichannel-loudspeaker-based auralizations of a lecture hall. Finally, we will present our prototype of a Motion Tracked Binaural (MTB-) microphone array (Algazi, et al. 2004). You may listen to MTB recordings while changing the used crossfade algorithm and the number of microphones in real-time.

MediaLab (TU Berlin)

Author/s: Audio Communication Group, Berlin

Title: Demonstration: Virtual Concert Hall

Abstract: Have you ever asked yourself, how your concert experience would differ, if you could watch and listen to the same performance of a musical piece in different venues? Research in the field of audio-visual room perception requires experimental variations which can be realized in simulated environments only. Within the SEACEN framework, the audio communication group built an immersive virtual environment, aiming at recreating the opto-acoustical properties of a concert venue as exactly as possible. Data-based dynamic binaural synthesis allows for a plausible acoustic simulation, whereas a high resolution projection of stereoscopic photographs on a 180° cylindrical screen provides a three-dimensional semi-panoramic view. Within our demonstration you will be able to switch in between six divergent concert venues (theatre, opera, chamber music hall, concert hall, church, roman basilica) while enjoying the same opto-acoustic performance of a string quartet.

Wellenfeld H0104, TU Berlin

WFS Concerts

Artist: Hans Tutschku:

"Rituale" (rituals) works with recordings of human voices and instrumental sounds from different cultures and melts them into a sound ritual. "Rituale" uses the extraordinary possibilities to place and move sound sources within a wave field. As the location of sources within the audience space becomes available, the sounding entities are approaching the listener in a physical way.

Artist: TBA

"Tau" (2005/11) was originally written for the INA/GRM Acousmonium and adapted for WFS in 2011. Based on a previous piece "Studies for Thunder (2005)", a virtual closed world was created to imply an immersive sensation of a macroscopic world in which microscopic events are embedded. During the performance, individual layers are filtered, mixed and distributed in space. Real time parameters control the movements of the sounds fed to the WFS system.

Friday, April 4th 2014

Paper Session D

PA.D.1

Author/s: Matthias Frank

Title: Localization Using Different Amplitude-Panning Methods in the Frontal Horizontal Plane

Abstract: Amplitude panning is the simplest method to create phantom sources in the horizontal plane. The most commonly employed amplitude panning methods are Vector-Base Amplitude Panning (VBAP), Multiple-Direction Amplitude Panning (MDAP), and Ambisonics. This article investigates the localization of frontal phantom sources created by VBAP, MDAP, and Ambisonics (with and without max-rE weighting) at the central listening position in a listening experiment. The experiment was conducted under typical non-

anechoic studio listening conditions and utilized pink noise and a regular array of 8 loudspeakers for all methods. The experimental results are compared to different predictors: a binaural localization model using measured binaural room impulse responses, the direction of the measured sound intensity vector, and the directions of the simpler velocity and energy vectors. The article hereby addresses the questions of how close the actually localized directions of the different panning methods are compared to the desired directions, and how good the predictors match the experimental results.

PA.D.2

Author/s: Symeon Mattes, Philip Nelson, Filippo Fazi and Michael Capp

Title: Exploration of a Biologically Inspired Model for Sound Source Localization in 3D Space

Abstract: Sound localization in 3D space relies on a variety of auditory cues resulting from the encoding provided by the lower and higher regions of the auditory path. During the last 50 years different theories and models have been developed to describe psychoacoustic phenomena in sound localization inspired by the processing that is undertaken in the human auditory system. In this paper, a biologically inspired model of human sound localization is described and the encoding of the known auditory cues by the model is explored. In particular, the model takes as an input binaural and monaural stationary signals that carry information about the Interaural Time Difference (ITD), the Interaural Level Difference (ILD) and the spectral variation of the Head Related Transfer Function (HRTF). The model processes these cues through a series of linear and nonlinear units, that simulate the peripheral and the pre-processing stages of the auditory system. The encoded cues, which in the model are represented by excitation-inhibition (EI) and the time average (TA) activity patterns, are then decoded by a central processing unit to estimate the final location of the sound source.

PA.D.3

Author/s: Stefan Klockgether, Steven van de Par

Title: A Model for the Prediction of Room Acoustical Perception Based on the Just Noticeable Differences of Spatial Perception

Abstract: The accurate physical simulation of room acoustics is a very complex task. It is known, that the human auditory system is limited in its ability to distinguish between subtle differences in the acoustics of a room. This is for example demonstrated by the observation that certain room acoustical parameters can be highly variable on very short distances while the human auditory system seems to be insensitive to the variations in the associated perception of room acoustical attributes. Therefore it seems attractive, to better understand the human auditory perception of room acoustics and to better know what properties of room acoustics are perceptually relevant. This would help to determine how accurate a room acoustical simulation needs to be, for example to sound the same as the real room that is simulated. The influence of the interaural cross-correlation of the first part and the reverberant tail of a binaural room impulse response on spatial perception is surveyed by directly manipulating the impulse response. The data is compared to a modelling approach. A two-stage binaural psychoacoustic model is presented, to predict the perception of the acoustics of a room. The first stage extracts the binaural cues and is based on the limits of spatial perception. The just noticeable differences of the interaural cross correlation are used to tune the model. The second stage predicts several perceptual attributes which characterize the perception of room acoustical attributes. The model uses the information of binaural room impulse responses for the prediction of the apparent width of sources and the perceived listener envelopment.

Plenary Talk

Author: Alexander Raake and Jens Blauert

Title: Listening and Assessing with Binaural Models

Abstract: We report on a novel approach to modeling human binaural listening. Our EU-funded project TWO!EARS replaces current thinking about auditory modeling by a systemic approach in which human listeners are regarded as multi-modal agents that develop their concept of the world by exploratory interaction. The approach is based on a structural link from binaural perception to judgment and action, realized by interleaved signal-driven (bottom-up) and hypothesis-driven (top-down) processing within an innovative expert-system architecture. The system achieves object formation based on Gestalt principles, meaning assignment, knowledge acquisition and representation, learning, logic-based reasoning and reference-based judgment. More specifically, the system assigns meaning to aural events by combining signal- and symbol-based processing in a joint model structure, integrated with proprioceptive and visual percepts.

A system of this kind is suitable for approaching the problem of aural quality in a novel way, taking into account that aural quality is not an inherent property of sounds but happens in a complex process, specified by judgments where the properties of the sounds (their “character“) are compared to conceptual references that represent the expectations of the listeners. The references are task- and listener-specific. Establishing these references requires knowledge inherent in the modeling system, namely, a specific aural-world model.

To obtain ground-truth data for the world model, aural quality can be analyzed in terms of plausibility. It can be attempted to assess plausibility in an indirect fashion, observing listeners’ involvement and immersion into aural scenes and, thereby, implicitly considering the meaning associated with it. If, instead of mere form-related fidelity, the ability of sounds under test to convey meaning to listeners is considered, the communication quality can be honored as an important constituent of aural quality or, in other words, the function aspect is taken into account in addition to the form aspect.

Paper Session E

PA.E.1

Author/s: Fabian Brinkmann, Alexander Lindau, Martina Vrhovnik, Stefan Weinzierl

Title: Assessing the Authenticity of Individual Dynamic Binaural Synthesis

Abstract: Binaural technology allows to capture sound fields by recording the sound pressure arriving at the listener’s ear canal entrances. If these signals are reconstructed for the same

listener the simulation should be indistinguishable from the corresponding real sound field. A simulation fulfilling this premise could be termed as perceptually authentic. Authenticity has been assessed previously for static binaural resynthesis of sound sources in anechoic environments, i.e. for HRTF-based simulations not accounting for head movements of the listeners. Results indicated that simulations were still discernable from real sound fields, at least, if critical audio material was used. However, for dynamic binaural synthesis to our knowledge – and probably because this technology is even more demanding – no such study has been conducted so far. Thus, having developed a state-of-the-art system for individual dynamic auralization of anechoic and reverberant acoustical environments, we assessed its perceptual authenticity by letting subjects directly compare binaural simulations and real sound fields. To this end, individual binaural room impulses were acquired for two different source positions in a medium-sized recording studio, as well as individual headphone transfer functions. Listening tests were conducted for two different audio contents applying a most sensitive ABX test paradigm. Results showed that for speech signals many of the subjects failed to reliably detect the simulation. For pink noise pulses, however, all subjects could distinguish the simulation from reality. Results further provided evidence for future improvements.

PA.E.2

Author/s: Benjamin Bernschütz, Arnau Vázquez Giner, Christoph Pörschmann, Johannes Arend

Title: Perceptual Aspects Concerning the Binaural Reproduction of Plane Waves with Reduced Modal Order

Abstract: Modal descriptions of measured or simulated sound fields using spherical harmonics enjoy popularity, and the binaural reproduction of the respective datasets using headphones is of great interest. A common method to extract directional information in the space domain from an underlying modal description is using plane wave decomposition techniques. Usually a set of head related transfer functions (HRTFs) is involved in a next step in order to establish the typical binaural cues that can be evaluated by the human auditory system. Due to their nature, HRTFs carry substantial information in higher modal orders in proportion to the temporal frequency. Owing to different reasons that are not subject of discussion herein, measurement or simulation systems often deliver a comparatively low number of resolvable modes. This leads to plane wave descriptions of limited modal order that entail substantial adaptation problems to HRTFs. Even if the adaptationcan be optimized using appropriate spatial sampling schemes for the HRTFs or spatial downsampling techniques, considerable technical differences remain between a native far-field HRTF and its binaural order-reduced plane wave counterparts. Listening experiments were conducted in order to evaluate some of the respective perceptual differences. The paper presents and discusses the involved signal processing, the design of the listening experiments and a selection of the experiment's results.

Poster Session B

PO.B.1

Author/s: Michael Schoeffler, Susanne Westphal, Alexander Adami, Harald Bayerlein, Jürgen Herre

Title: Comparison of a 2D- and 3D-Based Graphical User Interface for Localization Listening Tests

Abstract: Recently, there is a trend in developing new multi-channel formats towards adding additional loudspeakers in elevated positions. While the common 5.1 surround sound system only has loudspeakers in the horizontal plane, more complex systems, such as 10.2 or 22.2, include two or more elevated loudspeakers. When listening to music using a multi-channel playback system, the audio material has often not been produced for the used system, e.g. listening to 10.2 material while using a 5.1 surround system. In such cases, the audio material has to be down- or upmixed. Compared with listening to the original audio material, down- or upmixing affects the listening experience. The localization of sound sources is one attribute that might be affected by down- or upmixing the audio material. In the past, some localization listening tests were conducted by using an user interface depicting a two-dimensional representation of the scene. When it comes to elevated loudspeakers, a third dimension also has to be depicted by the user interface. In this work, an experiment was conducted where participants had to locate sound sources by using two different graphical user interfaces (GUIs). The first GUI consisted of two static images of the scene: a top-view and a front-view. The other GUI had a fully adjustable 3D visualization of the scene. The main purpose of the experiment is to investigate the differences between both GUIs. This includes the time participants spend on each GUI and the difference in the responses. This work is a contribution to the development of new evaluation methods for new and existing multi-channel audio formats and renderers.

PO.B.2

Author/s: Julian Grosse, Steven van de Par

Title: Perceptual Optimization of Room-In-Room Reproduction with Spatially Distributed Loudspeakers

Abstract: It is often desirable to reproduce a specific room-acoustic scene, e.g. a concert hall in a playback room, in such a way that the listener has a plausible and authentic spatial impression of the original sound source including the room acoustical properties. In this study a perceptually motivated approach for spatial audio reproduction is developed. This approach optimizes the spatial and monaural cues of the direct and reverberant sound separately. More specifically, the (monaural) spectral cues responsible for the timbre and the (binaural) interaural cross correlation (IACC) cues, responsible for the listener envelopment, were optimized in the playback room to restore the auditory impression of the recording room. The direct sound recorded close to the source is processed with an auditory motivated gammatone filterbank such that the spectral cues, ITD's and ILD's are comparable to the direct sound in the recording room. Additionally, the reverberant sound, which was recorded at two distant locations from the source, is played back via dipole loudspeakers. Due to the arrangement of the two dipole loudspeakers, only the diffuse sound field in the playback

room is excited, therefore the spectral cues and the IACC of the reverberant sound field can be controlled independently to match the cues that were present in the recording room. As indicated by a preliminary listening test the applied optimization is perceptually similar to the reference signal and is generally preferred when compared to a conventional room-in-room reproduction.

PO.B.3

Editorial Remark: This contribution was withdrawn eventually.

Author/s: Josefa Oberem, Vera Lawo, Iring Koch, Janina Fels

Title: Evaluation of Individual and Nonindividual Binaural Reproduction Methods in an Experiment on Auditory Selective Attention

Abstract: When auditory selective attention started to be in the interest of acoustical and psychological research the dichotic reproduction was used due to limited technical capabilities, but still today many psychological experiments on the cocktail-party effect are performed using a dichotic reproduction. For more realistic scenes and a better spatial impression, auralization can be used and the presentation of stimuli should be binaural. The most intuitive but also most complex way of reproducing a realistic virtual scene is a loudspeaker-setup in an anechoic room. Binaural reproduction via headphones is more convenient, especially for psychologists without measurement facilities. In this study a psychologically established experiment has been conducted with real sources (loudspeakers in an anechoic chamber) and binaural reproduction via headphones using individual as well as nonindividual HRTFs. In contrast to usual acoustical localization tasks subjects are not asked to focus directly on the localization, the accuracy of the sources' positions, the externalization, etc. Instead, the localization quality is analyzed indirectly, since subjects mainly focus on the psychological attention task. Subjects listen to two competing speakers at different positions in space and have to report and categorize the target speaker's words guided by visual cues. Independent analyzed variables are reaction times and error rates for a total of 73 subjects. This way it is obtained an objective measure of the quality of methods for reproduction of auralized sounds.

PO.B.4

Author/s: Pawel Malecki

Title: Auralization of Several Churches and Listening Comparison Using Multidimensional Scaling Approach

Abstract: Modern auralization techniques allow to make better assessment of particular interior type for different destination and purposes. The quality and reality of acoustic recording and reproduction systems increase so the results of this kind of research are much more reliable. The article shows the psychoacoustic comparison of different reverberant interiors. The auralization was provided using 1st order ambisonics spatial impulse responses convoluted with anechoic choral music. Listening tests were conducted within the 16-channel sound system. The subjects were tested using the pair comparison method and the results were analyzed with the multidimensional scaling approach.

PO.B.5

Author: Lukas Aspöck, Sönke Pelzer, Frank Wefers, Michael Vorlaender

Title: A Real-Time Auralization Plugin for Architectural Design and Education

Abstract: The role of acoustics in architectural planning processes is often neglected if the designer lacks necessary experience in acoustics. Even if an acoustic consultant is involved he might be presented with limited options after the initial planning process. Some disadvantageous decisions might be hard to reverse then. To improve and facilitate the construction process permanent immediate feedback should be given to the designer. Planning cannot be imaged today without live 3D visual rendering. But also acoustics should be rendered in real-time to provide the same type of intuitive feedback. Therefore a real-time room acoustics auralization was implemented into a popular CAD-Modeling tool. Binaural room impulse responses are continuously updated using image sources and ray tracing algorithms and convolved in real-time with audio feed from recorded sounds or the user's microphone. The CAD model can be freely modified during the simulations including geometry, surface materials and source and receiver positions. Using streaming low-latency convolution, an immediate feedback is provided to the user.

PO.B.6

Author: Eugen Rasumow, Matthias Blau, Martin Hansen, Simon Doclo, Steven van de Par, Volker Mellert, Dirk Püschel

Title: The Impact of the White Noise Gain (WNG) of a Virtual Artificial Head on the Appraisal of Binaural Sound Reproduction

Abstract: As an individualized alternative to traditional artificial heads, individual head-related transfer functions (HRTFs) can be synthesized with a microphone array and digital filtering. This strategy is referred to as "virtual artificial head" (VAH). The VAH filter coefficients are calculated by incorporating regularization to account for small errors in the characteristics and/or the position of the microphones. A common way to increase robustness is to impose a so-called white noise gain (WNG) constraint. The higher the WNG, the more robust the HRTF synthesis will be. On the other hand, this comes at the cost of decreasing the synthesis accuracy for the given sample of the HRTF set in question. Thus, a compromise between robustness and accuracy must be found, which furthermore depends on the used setup (sensor noise, mechanical stability etc.). In this study, different WNG are evaluated perceptually by four expert listeners for two different microphone arrays. The aim of the study is to find microphone array-dependent WNG regions which result in appropriate perceptual performances. It turns out that the perceptually optimal WNG varies with the microphone array, depending on the sensor noise and mechanical stability but also on the individual HRTFs and preferences. These results may be used to optimize VAH regularization strategies with respect to microphone characteristics, in particular self-noise and stability.

PO.B.7

Author: Jian Zhang, Chundong Xu, Risheng Xia, Junfeng Li , Yonghong Yan

Title: Dependency of the Finite-Impulse-Response-Based Head-Related Impulse Response Model on Filter Order

Abstract: Various approaches have been reported on HRIR modeling to lighten the high computation cost of the 3-D audio systems without sacrificing the quality of the rendered sounds. The performance of these HRIR models have been widely evaluated usually in terms of the objective estimation errors between the original measured HRIRs and the modeled HRIRs. However, it is still unclear how much these objective evaluation results match the psychoacoustic evaluations. In this research, an efficient finite-impulse-response (FIR) model is studied as a case study which is essentially based on the concept of the minimum-phase modeling technique. The accuracy dependency of this modeling approach on the order of FIR filter is examined with the objective estimation errors and the psychoacoustic tests. In the psychoacoustic tests, the MIT HRIR database are exploited and evaluated in terms of sound source localization difference and sound quality difference by comparing the synthesized stimuli with the measured HRIRs and those with the FIR models of different orders. Results indicated that the measured hundred-sample-length HRIRs can be sufficiently modeled by the low-order FIR model from the perceptual point of view, and provided the relationship between perceptual sound localization/quality difference and the objective estimation results that should be useful for evaluating the other HRIR modeling approaches.

PO.B.8

Author: Iain Laird, Damian Murphy, Paul Chapman

Title: Comparison of Spatial Audio Techniques for Use in Stage Acoustic Laboratory Experiments

Abstract: Real-time auralisation systems are increasingly being used by researchers aiming to observe how particular stage and auditorium configurations affect a musician’s performance technique. These experiments typically take place in controlled laboratory conditions equipped with auralisation systems capable of reproducing the 3D acoustic conditions of a performance space in response to a performing musician in real-time. This paper compares the performance of First Order Ambisonics and Spatial Impulse Response Rendering in terms of both objective measurements and subjective listening tests. It was found that both techniques spatialised single reflections with similar accuracy when measured at the sweet spot. Informal listening tests found that the techniques produced very similar perceived results both for synthesised impulse responses and for measured stage acoustic impulse responses.

Paper Session F

PA.F.1

Author/s: Alexander Lindau, Vera Erbes, Steffen Lepa, Hans-Joachim Maempel, Fabian Brinkmann, Stefan Weinzierl

Title: A Spatial Audio Quality Inventory for Virtual Acoustic Environments (SAQI)

Abstract: The perceptual evaluation of virtual acoustic environments may be based on overall criteria such as plausibility and authenticity or by using catalogues of more detailed auditory qualities as, e.g., loudness, timbre, localization, etc. However, only the latter will be suitable to reveal specific shortcomings of a simulation under test and allow for a directed technical improvement. To this end a common vocabulary of relevant perceptual attributes appears desirable. Existing vocabularies for the evaluation of sound field synthesis, spatialization technologies and virtual environments were often generated ad hoc by the authors or have focused only on specific perceptual aspects. To overcome these limitations, we have developed a Spatial Audio Quality Inventory (SAQI) for the evaluation of virtual acoustic environments. It is a consensus vocabulary comprising 48 verbal descriptors of perceptual qualities assumed to be of practical relevance when comparing virtual environments to real or imagined references or amongst each other. The vocabulary was generated by a Focus Group of 20 German speaking experts for virtual acoustics. Five additional experts helped verifying the unambiguity of all descriptors and the related explanations. Moreover, an English translation was generated and verified by eight bilingual experts. The paper describes the methodology and the outcome, presenting the vocabulary in the English version.

PA.F.2

Author/s: Rozenn Nicol, Laetitia Gros, Cathy Colomes, Olivier Warusfel, Markus Noisternig, Hélène Bahu, Brian FG Katz, Laurent S. R. Simon

Title: A Roadmap for Assessing the Quality of Experience of 3D Audio Binaural Rendering

Abstract: Today there are 2 major evolutions in spatial audio. First, an enhanced 3D audio experience, where virtual sound sources can be accurately synthesized in any direction, is possible with technologies such as binaural, Wave Field Synthesis, Higher Order Ambisonics or Vector Base Amplitude Panning. Second, 3D audio is on the way to being democratized through binaural adaptation for headphone listening. These evolutions call for revisiting the methods and tools used to assess the perception of spatial sound reproduction. The first objective of this paper is to delineate the problem, by exploring the potential dimensions and the related attributes underlying the perception of spatial sound, mainly within the context of binaural reproduction. Secondly, assessment methods, including both standard and less conventional ones, are listed, and their relevance for the measure of the attributes previously identified is discussed.

Round Table

Participants: Stefan Weinzierl, Stefan Feistel, Claus-Lynge Christensen, Michael Vorländer

Title: Round Robin on Auralization

Abstract: This round table is intended to stimulate discussions on the prerequisites for a planned Round Robin on Auralization to be conducted by the SEACEN research consortium. Within this Round Robin, existing algorithms for numerical (model-based) sound field simulation shall be evaluated psychically and perceptually using a fair and consistent test design. Therefore, for instance, a suitable exchange format for simulated sound fields needs to be decided. Hereby, temporal or spectral representations, as, e.g., expansions into series of spherical harmonic functions or plane wave representations may be discussed with

respect to specific advantages and disadvantages. Further, issues of auralization of sound fields represented in such a way should be considered (choice of HRTF data sets, required HRTF pre-treatment, issues of dynamic auralization). Finally, suitable measures for physical and perceptual accuracy may be discussed, as, e.g., physical error measures, choice of room acoustical parameters, and evaluation according to integrative and qualitative differentiated measures for perceptual accuracy).

Paper Session G

PA.G.1

Author/s: Hannes Pomberger, Florian Pausch

Title: Design and Evaluation of a Spherical Segment Array with Double Cone

Abstract: This article discusses modal beamforming applied to our recently built panoramic microphone array prototype. Modal beamforming using conventional spherical microphone arrays is based on the decomposition of the sound pressure in spherical harmonics. The decomposition in spherical harmonics requires a distribution of microphones covering all directions, even if only a limited range of directions is of interest. For a panoramic array, the range of the zenith angle is bounded above and below which corresponds to a spherical segment. Solving the Helmholtz equation for a two-point Neumann boundary condition in the zenith angle yields orthogonal functions for a spherical segment, which we will refer to as spherical segment harmonics. This is physically equivalent to a sound field delimited by a rigid infinite double conical surface. Our prototype consists of microphones distributed on a spherical segment and a rigid double cone of finite length. We show that the new beamforming is applicable, even though the practical implementation does not meet the theoretical model of infinitely long boundaries. The pattern synthesis approach is verified by acoustic measurements of the array prototype.

PA.G.2

Author/s: Jonathan Sheaffer, Shahar Villeval, Boaz Rafaely

Title: Rendering Binaural Room Impulse Responses from Spherical Microphone Array Recordings Using Timbre Correction

Abstract: The technique of rendering binaural room impulse responses from spatial data captured by spherical microphone arrays has been recently proposed and investigated perceptually. The finite spatial resolution enforced by the microphone configuration restricts the available frequency bandwidth and, accordingly, modifies the perceived timbre of the played-back material. This paper presents a feasibility study investigating the use of filters to correct such spectral artifacts. Listening tests are employed to gain a better understanding of how equalization affects externalization, source focus and timbre. Preliminary results suggest that timbre correction filters improve both timbral and spatial perception.

Artistic contributions, demonstrations, and workshops on Friday

Ambisonic Environment, Room 212 (Museum)

Author/s: David Monacchi

Title: Fragments of Extinction (Borneo 2012) – A Periphonic Audio-Video Concert Based on Rainforest Ecosystems 3D Recordings

Abstract: Following the extensive data collection carried out with space-preservative recording methodologies during the last field recording trip to the remote equatorial primary rainforests of Brunei and Sarawak - Borneo, the author proposes a sequence of sonic experiences where pure unaltered recordings, in the 1st part of the concert, are then complemented with digital sound synthesis in the 2nd part. The spatial complexity and inter-species ecoacoustic order within the different sonic habitats (primary lowland dipterocarp forest, alluvial forest, pond and riverbank forest), manifesting the balanced interplay among hundreds of biophonies, have been recorded with the highest definition possible to be presented with periphonic playback systems. Compositionally, different levels of time-lapse, explorations of audible and inaudible sonic languages, and a real-time spectrogram analysis video projection, allow the audience to understand the ecosystems' internal configurations. Subtle sensor-driven live integrations ideally build then a powerful metaphor of our species collaborating with these extraordinary composite ecosystems. The aim of the international long-term project Fragments of Extinction is to bring the current biodiversity crisis to broad public's attention, through spatial audio ecosystems' reconstructions.

Demo room (SIM)

Author/s: Sönke Pelzer, Lukas Aspöck

Title: Workshop: Real-time Room Acoustics Simulation and Auralization in SketchUp

Abstract: This workshop demonstrates a real-time room acoustics simulation software and its integration as a plug-in into the popular 3D modeling software "SketchUp". All acoustics related controls and tools are integrated into the original software's GUI, which results in a very easy and intuitive handling of acoustics simulations. A high degree of interactivity is provided by continuously updating the simulation results while working on the 3D model. Any modification of the geometry, of materials or of source and receiver positions has immediate impact on the results. The user sees a visualization of room acoustics parameters directly inside the modeling window and listens to the auralization of the room characterized by its binaural room impulse response.

Related publications:

[1] Pelzer et al.: Interactive real-time simulation and auralization for modifiable rooms, Proc. of International Symposium on Room Acoustics (ISRA 2013)

[2] Aspöck et al.: A real-time auralization plugin for architectural design and education, Proc. of the EAA Joint Symposium on Auralization and Ambisonics (2014).

Demo room (SIM)

Author/s: Frank Wefers

Title: Workshop: Interactive Acoustic Virtual Environments Using Room Acoustic Simulations

Abstract: see program of Thursday (above)

Demo room (SIM)

Author/s: Audio Communication Group, Berlin

Title: Demonstrations: Dynamic Binaural Recording and Reproduction

Abstract: see program of Thursday (above)

Saturday, April 5th 2014

Paper Session H

PA.H.1

Author/s: Peter Stitt, Stéphanie Bertet, Maarten van Walstijn

Title: Off-Centre Localisation Performance of Ambisonics and HOA for Large and Small Loudspeaker Array Radii

Abstract: Ambisonics and Higher Order Ambisonics (HOA) are scalable spatial audio techniques that attempt to present a sound scene to listeners over as large an area as possible. A localisation experiment was carried out to investigate the performance of a first and third order system at three listening positions - one in the centre and two off-centre - using a 5 m radius loudspeaker array. The results are briefly presented and compared to those of an earlier experiment on a 2.2 m loudspeaker array. In both experiments the off-centre listeners were placed such that the ratio of distance from the centre to the array radius was constant in both experiments. The test used a reverse target-pointer adjustment method to determine the error, both signed and absolute, for each combination of listening position and system. The signed error was used to indicate the direction and magnitude of the shifts in panning angle introduced for the off-centre listening positions. The absolute error was used as a measure of the performance of the listening position and system combinations for a comparison of their overall performance.

PA.H.2

Author/s: Georgios Marentakis, Franz Zotter, Matthias Frank

Title: Vector-Base and Ambisonic Amplitude Panning: A Comparison Using Pop, Classical, and Contemporary Spatial Music

Abstract: Vector-Base Amplitude Panning (VBAP) and Ambisonics are commonly used in 3D audio reproduction via loudspeakers. While research has been investigating their properties using psychoacoustic test signals, there is only a small number of investigations employing musical material. Considering the musical application of these spatialization methods, we

present an experimental study characterizing quality aspects using excerpts that belong to three different musical genres (popular, classical, and contemporary spatial music). The study compares seven configurations of vector-base and Ambisonic amplitude panning in a hemispherical listening environment that is permanently installed in the IEM CUBE. Four configurations thereof use 24 loudspeakers, and the others use a subset of 12 loudspeakers. In pairwise comparisons, participants rated each configuration pair on a quasi-continuous scale in terms of preference, envelopment, spatial clarity, sound quality, and stability. Perceptual scales were then constructed which revealed how configurations ranked in terms of each attribute. The ranking of the tested configurations on the perceptual scales was dependent on the musical material. In the case of the popular and the classical music piece, results were relatively consistent and participants tended to prefer the configurations that used 12 loudspeakers. Results indicate that preference judgments are correlated to envelopment, sound quality, and spatial clarity.

PA.H.3

Author/s: Nicolas Epain, Craig Jin, Franz Zotter

Title: Ambisonic Decoding with Constant Angular Spread

Abstract: Designing Ambisonic decoders for irregular loudspeaker layouts has long been an issue. The quality of an Ambisonic decoder for a particular virtual source direction can be characterised via three measures: the energy error, the angular error and the angular spread. The energy error measures the mismatch between the energy of the speaker signals and the energy of the virtual source. The angular error measures the mismatch between the direction of the so-called energy vector and the direction of the virtual source. Lastly, the angular spread is determined by the norm of the energy vector and is indicative of the source width perceived by a listener. In recent works, decoders have been presented which yield minimal energy and angular errors for irregular speaker layouts. However, these decoders offer no control on the angular spread. In this paper we present a method for calculating Ambisonic decoders providing a nearly constant angular spread across source directions while maintaining very low energy and angular errors.

Plenary Talk

Author: Craig Jin, R. Zolfaghari, Nicolas Epain, and A. Tew

Title: From Super-Resolution Acoustic Imaging to Binaural Rendering

Abstract: In this plenary talk, we address two topics: super-resolution acoustic imaging and individualization of binaural rendering. We connect the two disparate topics by discussing the possible application of super-resolution analyses for HOA binaural rendering using microphone arrays. A detailed review is given of our research into super-resolution acoustic imaging in the Higher-Order Ambisonic (HOA) domain. The concept of performing a plane-wave decomposition using sparse recovery and a spatial dictionary is explained. Empirical examples are given of upscaling an HOA sound scene and super-resolution beamforming in a reverberant environment. We briefly describe direct-diffuse sound separation in the HOA domain and discuss sub-dividing the super-resolution acoustic imaging problem so it may be performed on a general purpose graphical processing unit (GPGPU). The possible application

of super-resolution analyses for HOA binaural rendering is discussed and a perceptually-motivated method for binaural decoding of HOA is presented. We then discuss methods for exploring the HRIR individualization problem using the Sydney York Morphological and Recording of Ears (SYMARE) database and a shape analysis method known as Large Deformation Diffeomorphic Metric Mapping.

Paper Session I

PA.I.1

Author/s: Markus Zaunschirm, Franz Zotter

Title: Measurement-Based Modal Beamforming Using Planar Circular Microphone Arrays

Abstract: This paper describes how to use a planar circular pressure-zone table-top microphone array for modal beamforming. Its goals are similar as for spherical arrays: higher-order resolution and a more-or-less steering-invariant beampattern design in the three-dimensional half space. As conventional circular arrays lack control of the beampattern in the vertical array plane, the proposed arrangement tries to fix this shortcoming to allow both horizontal and vertical control of beamforming. To provide a fully calibrated decomposition into the directional modes, the proposed beamforming approach is based on measurement data. From a MIMO (multiple-input-multiple-output) system description of the measurement data in the spherical harmonics domain, an inverse MIMO system of filters is designed for decomposing the microphone array signals into those spherical components eligible for modal beamforming. For an efficient measurement and robust set of decomposition filters, a reduced set of measurement positions and a regularisation strategy is suggested.

PA.I.2

Author/s: Franz Zotter, Matthias Frank, Matthias Kronlachner, Jung-Woo Choi

Title: Efficient Phantom Source Widening and Diffuseness in Ambisonics

Abstract: Object-based spatial audio considers virtual sound sources having a width/diffuseness parameter. This parameter aims at controlling the perceived width or diffuseness of the auditory object, or phantom source, created by the renderer. Width/diffuseness provides an important salience parameter that is independent of perceived direction and timbre. A highly efficient sparse filter structure for two-channel stereophony was described and tested recently, but it becomes ineffective for most parts of a large audience. This paper presents phantom source width/diffuseness control for Ambisonics. The new approach is a remarkably elegant application of the previously described stereo phantom source widening on Ambisonics. Compared with former experimental data, our experiments show a greater freedom of increasing the width and widening that works for a larger listening area.

Poster Session C

PO.C.1

Author/s: Giso Grimm, Torben Wendt, Volker Hohmann, Stephan D. Ewert

Title: Implementation and Perceptual Evaluation of a Simulation Method for Coupled Rooms in Higher Order Ambisonics

Abstract: A fast and perceptively plausible method for rendering acoustic scenarios with moving sources and moving listeners is presented. The method is principally suited for application in dynamic and interactive evaluation environments (e.g., for hearing aid development), psycho-physics with adaptively changing the spatial configuration, or simulation and computer games. The simulation distinguishes between the direct sound, sound reflected and diffracted by objects of limited size, diffuse sound surrounding the listener, e.g., diffuse background sounds and diffuse reverberation, and 'radiating holes' for simulation of coupled adjacent rooms. Instead of providing its own simulation of room reverberation, the proposed simulation method generates appropriate output signals for external room reverberation simulators (e.g., see contribution by Wendt et al.). The output of such room reverberation simulators is then taken either as diffuse surrounding sound if the listener position is within the simulated room, or as input into a 'radiating hole', if the listener is in an adjacent room. Subjective evaluations are performed by comparing measured and synthesized transitions between coupled rooms.

PO.C.2

Author/s: Diego Murillo, Filippo Fazi, Mincheol Shin

Title: Evaluation of Ambisonics Decoding Methods with Experimental Measurements

Abstract: Ambisonics is a sound reproduction technique based on the decomposition of the sound field using spherical harmonics. The truncation in the number of coefficients used to recreate the sound field leads to reproduction artifacts which depend on the frequency and the listener spatial location. In this work, the performance of three different decoding methods (Basic, Max-rE and In-Phase) has been studied and evaluated in the light of the results of experimental measurements. The latter were performed using a spherical array composed of 40 uniformly distributed loudspeakers and a translating 29-channel linear microphone array. An error analysis is presented based on the difference between the desired and synthesized sound pressure and acoustic intensity field. The results indicate that, as expected, the size of the region of accurate sound field reconstruction reduces as frequency increases, but with different trends depending on the type of decoder implemented.

PO.C.3

Author/s: César Salvador, Shuichi Sakamoto, Jorge Treviño, Yôiti Suzuki

Title: Embedding Distance Information in Binaural Renderings of Far Field Recordings

Abstract: Traditional representations of sound fields based on spherical harmonics expansions do not include the sound source distance information. As multipole expansions can accurately encode the distance of a sound source, they can be used for accurate sound field reproduction. The binaural reproduction of multipole encodings, though, requires head-related transfer functions (HRTFs) with distance information. However, the inclusion of distance information on available data sets of HRTFs, using acoustic propagators, requires

demanding regularization techniques. We alternatively propose a method to embed distance information in the spherical harmonics encodings of compact microphone array recordings. We call this method the Distance Editing Binaural Ambisonics (DEBA). DEBA is applied to the synthesis of binaural signals of arbitrary distances using only far-field HRTFs. We evaluated DEBA by synthesizing HRTFs for nearby sources from various samplings of far-field ones. Comparisons with numerically calculated HRTFs yielded mean spectral distortion values below 6 dB, and mean normalized spherical correlation values above 0.97.

PO.C.4

Author/s: Sam Clapp, Anne Guthrie, Jonas Braasch, Ning Xiang

Title: Evaluating the Accuracy of the Ambisonic Reproduction of Measured Soundfields

Abstract: A spherical microphone array can encode a measured soundfield into its spherical harmonic components. Such an array will be subject to limitations on the highest spherical harmonic order it can encode and encoding accuracy at different frequencies. Ambisonics is a system designed to reproduce the spherical harmonic components of a measured or virtual soundfield using multiple loudspeakers. In ambisonic systems, the size of the sweet spot is wavelength dependent, and thus decreases in size with an increase in frequency. This paper examines how to reconcile the limitations of the recording and playback stages to arrive at the optimum ambisonic decoding scheme for a given spherical array design. In addition, binaural models are used to evaluate these systems perceptually.

PO.C.5

Author/s: Khemapat Tontiwattanakul, Filippo Fazi, Philip Nelson

Title: Comparison of Analytical and BEM-based Beamforming Strategies for Spherical Microphone Arrays

Abstract: Most of the beamformer strategies applied to spherical microphone arrays are based on the analytical solution of the Helmholtz equation. In this paper an alternative technique is presented to design a beamformer that relies on the use of boundary element method (BEM) for the solution of the inverse scattering problem that arises with a rigid sphere microphone array. Initially, the problem is formulated as an integral equation providing a representation of the sound field by means of an infinite superposition of plane waves impinging on the microphone array. The kernel of the integral, which relates the plane wave amplitudes (source strengths) to the acoustic pressure captured by the microphones, is computed with the BEM. The problem is then discretized and formulated as a system of linear equations, the solution of which is used to design the beamformer. The results of numerical simulation and experimental measurement are used to compare the performance of the proposed beamformer design and of the traditional spherical harmonic beamforming.

PO.C.6

Author/s: Stephanie Bertet, Peter Stitt

Title: Investigation on the Influence of the Order on Ambisonics Reproduced Sound Scene over Headphones

Abstract: Ambisonics can be used to reproduce a sound scene over either loudspeakers or headphones using head related transfer functions, simulating the audio path between the loudspeaker position and the ears of the listener. Based on spherical harmonics decomposition, the intermediate B-format enables a large number of encoded sources for a fixed decoder and reproduction system. This study investigates binaurally reproduced sound scenes for 2D ambisonic orders 1 to 4 decoded on 2M + 2 virtual loudspeakers using two decoder options, basic and mixed basic and max rE. Two sound scenes are used, simulating an audio conference and a scene in a kitchen at home. Similarity ratings are obtained from pair- wise comparisons between all of the combinations of systems and a binaural reference. The results are analysed to define a perceptual space using a multidimensional scaling method and presented as three-dimensional spaces.

PO.C.7

Author/s: Brian F.G. Katz, Markus Noisternig, Olivier Delarozière

Presenter: Laurent Simon

Title: Scale Model Auralization for Art, Science, and Music: The Stupaphonic Experiment

Abstract: The use of acoustical scale models has been replaced for the most part by computational models and numerical simulations for room acoustic studies as well as artificial reverberation units. There remains however a number of acoustical phenomena which are difficult to address with computer simulations, such as coupled volumes, diffraction, and complex scattering, due to the computational complexity and/or calculation time necessary for addressing such acoustical wave phenomena on the scale of room acoustical problems, even small rooms. This paper presents a pilot study of a rather unique artistic architectural structure consisting of a self-supporting construction composed of small stacked linear elements. Acoustically, the structure combines modal behavior, concave forms, and very regular scattering patterns. An example scale model has been constructed and studied in order to separate different construction features and their associated acoustics effects. In an attempt to explore the interest of the specific acoustic for musical performance, a computational platform was created to utilize the scale model as a physical convolution reverberation unit for musical performance.

PO.C.8

Author: Daniel Protheroe

Title: Measurement and Visualisation of Room Responses in Level, Time and Direction

Abstract: The acoustical characteristics of an existing room are typically determined from single or dual channel impulse response measurements. The resulting data has limited or no directional information, therefore losing information of crucial importance in room acoustics. This contribution presents the performance of IRIS, a fully directional (i.e. 3-D) impulse response measurement system which utilises a compact Ambisonic microphone array. IRIS can analyse the sound field at a position in a fully objective way: sound reflections in terms of level, time and direction. This data can be easily visualised, e.g. as a “hedgehog pattern”. This is beneficial for observing the directional distribution of early and late sound energy, and for comparing different seating areas, variable acoustics settings etc, and identifying physical aspects of the room which are maybe causing undesirable reflections. Spatial parameters such as the early lateral energy fraction can also be determined from the 3-D information. The IRIS system has been validated by collecting 3-D impulse responses in varied rooms, and the results appear to be strongly correlated with the respective physical environments.