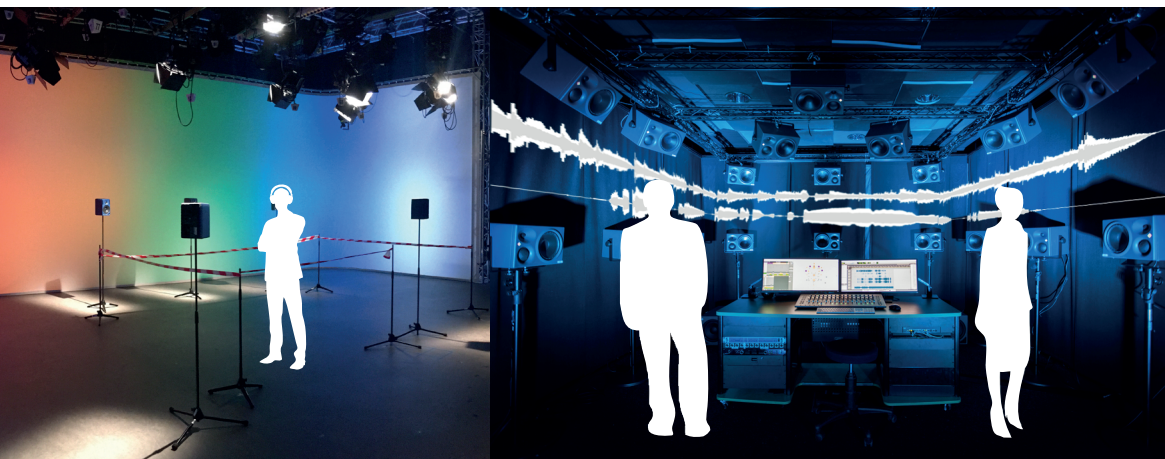# ICSA 2019

5th International Conference on Spatial Audio
26. - 28. September 2019 in Ilmenau, Germany

AUDIO AT ALL SITES

vdt Verband Deutscher
     Tonmeister e.V.

TECHNISCHE UNIVERSITÄT
ILMENAU

ICSA 2019

5th INTERNATIONAL CONFERENCE ON SPATIAL AUDIO

September 26th to 28th, 2019

Ilmenau, Germany

Audio for Virtual, Augmented and Mixed Realities

The ICSA 2019 focuses on a multidisciplinary bringing together of developers, scientists, users, and content creators of and for spatial audio systems and services. A special focus is on audio for so-called virtual, augmented, and mixed realities.

The fields of ICSA 2019 are:

▷ Development and scientific investigation of technical systems and services for spatial audio recording, processing and reproduction
▷ Creation of content for reproduction via spatial audio systems and services
▷ Use and application of spatial audio systems and content presentation services
▷ Media impact of content and spatial audio systems and services from the point of view of media science.

The ICSA 2019 is organized by VDT and TU Ilmenau with support of Fraunhofer Institute for Digital Media Technology IDMT.

Contents

Program Information

Organizational Issues

Check-In Times and Conference Info-Desk

> ▷ 26.09.2019 – 08:00-16:30 @ Foyer Helmholtz Building
> ▷ 27.09.2019 – 08:00-12:30 and 15:30-16:30 @ Hemholtz Building H1519
> ▷ 28.09.2019 – 08:00-15:00 @ Hemholtz Building H1519

Lunch

At Thursday 26.09. and Friday 27.09. there will be Lunch at Mensa Ehrenberg from 12:30 to 13:30. Each participant receives a voucher for one main menu and one drink for each day. Mensa 1 is reserved for the ICSA participants.

Address: Mensa Ehrenberg, Max-Planck-Ring 1, 98693 Ilmenau

On Saturday 28.09. there will be a Lunch snack in the Helmholtz Building from 12:30 to 13:30 at the catering space (in front of the listening lab).

Coffee Breaks and Drinks

There is coffee and tea available for self-service at the catering space (in front of the listening lab). Coffee break times are from 10:00 to 10:30 and from 15:30 to 16:30. Further conference drinks are available during the entire day. **Enjoy your food and drinks only in areas designated for this purpose. Please do not take them into the Helmholtz Hörsaal or the demo and workshop labs.**

Transport in and around Ilmenau

> ▷ **Train**: EIB – Erfurter Industriebahn, access schedule via www.bahn.com/en
> ▷ **Bus**: There are 3 bus routes available, A, B and C. Schedules can be found using the following link: https://www.iov-ilmenau.de/fahrplan/fahrplandruck/
> ▷ **Taxi**
>
> > ◇ Taxi Ilmenau - www.taxi-ilmenau.de | phone: +49 3677-666666
> > ◇ Taxi am Bahnhof – www.taxi-am-bahnhof.de | phone: +49 3677-667788

WiFi access

  ▷ Helmholtz Building: SSID: **icsa2019** | password: **spas0919**
  ▷ Fraunhofer IDMT: SSID: **IDMT-Guest** | password: **spas0919**

By using this service, you agree to the following terms and conditions:

  ▷ This service is provided free of charge and without obligation to our visitors. We cannot guarantee availability and security and we assume no responsibility for possible damages in conjunction with the use of the guestnet access.
  ▷ We do not guarantee permanent availability of your connection. We reserve the right to interrupt or terminate this service at any time and without prior notice.
  ▷ Data transmission in a WLAN is subject to eaves dropping. You are responsible for any necessary encryption of your data e.g. by using a VPN connection to your company network.
  ▷ We cannot prevent attacks from the internet or from other computers using this network. The use of a personal firewall, virus scanner, etc to protect your system, is your personal responsibility.
  ▷ You may use your username and password exclusively for your own purpose and only for the duration of your stay.
  ▷ You are not allowed to run a Server or to provide internet services over this network.
  ▷ Please note that you are of course required to abide by all laws.

Data privacy: Username and connection data (IP-, MAC-Address, times) will be saved and processed for up to 30 days, and used for monitoring errors, performance and possible misuse. This information will be handed over to law enforcement authorities as need be.

Social Events

25.09.2019 – 20:00 – Get together Jam Session

Since many of you will already arrive on Wednesday (or earlier), we organized a meeting point: We would warmly like to welcome you in the Baracke 5 e.V. for a meeting prior to the conference. There will be time and space to talk, take a beer and a snack, and to have a nice jam session together. Some instruments will be available. Beer and snacks can be purchased on site.

Address: Baracke 5 e.V., Neuhäuser Weg 10, 98693 Ilmenau

26.09.2019 – 19:00 - 22:30 – Welcome Reception

Join all participants, authors, presenters and organizers of ICSA 2019 for a welcome party. It takes place in the foyer of the Helmholtzbau, in front of the Helmoltz lecture hall, in the middle of the exhibition area. By the way: We thank all of our exhibitors and sponsors who rendered this ICSA possible! #Meeting Point# is the Helmholtz Hörsaal. At the Welcome Reception food and drinks are served from the Ilmenauer Studentenclub e.V. (ILSC) represented by the bc-Studentenclub and the bi-Studentenclub. Drinks and food are kindly **sponsored partly by Musikelectronic Geithain and mh acoustics**.

Address: Helmholtz Building, Helmholtz-Platz 2, 98693 Ilmenau

27.09.2019 – 19:00 - 20:00 – Organ Concert

Hans-Jürgen Freitag, Cantor of St. Jakobus church in Ilmenau, on September 27 is conducting compositions for choir, recorder, harpsichord, and the large Walcker organ. The ensemble will perform the contemplative piece "Golgotha" by contemporary composer Günther Berger alongside works by J. S. Bach. Berger links the events in the Passion of Jesus with a lament for the Shoah, a topic which he has revisited, almost obsessively, time and again. This musical rendition of the subject, particularly the instrumental, transcends conventional expectations of music for recorder and organ. The audience is sure to leave with an indelible memory of the concert. **The entrance is free, donations are welcome.**

Address: St.-Jakobus-Kirche, Kirchplatz 1, 98693 Ilmenau

27.09.2019 – 20:30 - 22:30 – Dinner at Hotel Tanne

ICSA invites all visitors with an ICSA full day ticket to a dinner at Friday evening. Please come and join us at the Hotel Tanne. **All Beverages at this evening are free of charge until the drinking contingent runs out.**

Address: Hotel Tanne, Lindenstraße 38, 98693 Ilmenau

Information for Presenters and Authors

Posters

Presenters are asked to hang out their posters upon arrival. Poster strips will be available at your poster slot. Time to discuss the posters is reserved on Thursday and on Friday from 15:30 to 16:40 during the coffee break. Presenters are encouraged to be present at their poster at this time on Friday but may also be present on Thursday. Please remove the poster before departure. Remaining posters will be disposed of.

Talks

Presenters are asked to be at the Helmholtz-Hörsaal about half an hour before the start of their session to have a technical test of their slides. Oral presentations should not exceed 15 minutes in length, in order to leave at last 5 minutes after each talk for questions and discussion.

Workshops and Demos

Presenters are asked to be at the presentation room 10 minutes before the session. Please make sure, that all demonstrations and workshop contents are ready to run. For further questions and consultations for technical preparation please contact the Workshops and Demo Chairs.

ICSA proceedings

The proceeding will be available for members at the homepage of the VDT. Please check https://vdt-icsa.de/ for further information.

Time Table

Please look at the online version of the time table here: https://vdt-icsa.de/program/

Rooms

All rooms in this booklet are shortened using an acronym. The following table lists the acronyms to the real name and building. Please see page 79 for a map of the campus.

Acronym	Name	Building	Plan on Page
H-Hs	Helmholtz-Hörsaal		
AV1	Audiovisual Lab 1		
AV2	Audiovisual Lab 2	Helmholtz Building	80
LL	Listening Lab		
H1554	Helmholtz-Room 1554		
TV Studio	Television Studio	Media Lab 2	79
TV Demo	Editorial Room		
3DP	3D Presentation Room		
SR-0.10	Seminar Room 0.10/0.12	Fraunhofer IDMT	84
WS	WaveStudio		

Data Privacy

We expressly point out that within the framework of the conference, images and sound material will be produced in the form of photos and video recordings. The recordings should document both the event itself and the participation of individual persons. By participating in the event, the participant declares his consent to the use and publication of such recordings for the purpose of public reporting on the event.

Keynotes

Gregor Zielinsky

Thursday, 26th of September, 09:10 - 10:00   Room: H-Hs

Gregor Zielinsky

Recording natural instruments for the purpose of object based recording and post processing is a huge problem. There are three important aspects, that usually interfere with each other. Which is: Sound-quality, mainly meaning naturally, original sound - Optimal crosstalk suppression - Recording all instruments together in there natural playing situation. These contradictions usually lead to unsatisfying result. Using a classical mic setup with common high quality spot mics will not result in cross talk suppression as it is needed. Using common contact microphones will lead to unsatisfying sound. Separating instruments and recording them one by one in anechoic chambers not only leads to very bad sound, but also to musical problems as extreme intonation and rhythmical effects. The workshop will show a compete new approach to recording instruments for object based recording. The Remic system allows a perfect crosstalk rejection as well as an optimal natural sound. Also, the Remics can be attached to the instruments in an easy way. These new recording options will lead to perfect results in all kind of situations as 3D recording, Games, VR, AR and any kind of situation where instruments are needed in a separated way. We will play a session that will show the sound of the Remics, the Crosstalk rejection as well as with way how to use them in an object based mix situation as for headphones or speakers.

David Griesinger: Learning to Listen to Headphones

Friday, 27th of September, 09:10 - 10:00   Room: H-Hs

David Griesinger

David Griesinger about his keynote: The author has been recording live music, testing audio equipment, and enjoying sound with headphones for more than sixty years. We have tried dynamic, piezoelectric, and electrostatic phones, but none of them could reproduce the timbre of a carefully equalized loudspeaker array. In this talk we will present some of the experiences that led to the realization that the resonances that the concha and ear canals create are individual in frequency, directionally dependent, and modified in individual ways by headphones. In-head localization and inaccurate timbre results. Measuring HRTF functions at the blocked ear canal ignores these resonances. To hear realistic sound, or to reproduce binaural recordings, we must individually match a headphone to our own ears. We have found a simple, non-invasive method to match the frequency response at the eardrum from a pair of headphones to the frequency response at the eardrum from a frontal loudspeaker. The result is frontal localization without head tracking and accurate timbre. A research version of our Windows app is available on our website, and there is a YouTube video showing how to use it. We have recently developed simple to use versions of the app for AAX and VST3 systems. We will demonstrate these apps informally at this conference and in an AVID booth at the AES conference in New York.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Session 1: Recording and Production

Thursday, 26th of September

Capturing 3D Audio: A pilot study on the spatial and timbral auditory perception of 3D recordings using main-array and front-rear separation in diffuse field conditions

Thursday, 26th of September, 10:30   Room: H-Hs

Janet Grab

This research is a preliminary pilot experiment into the subjectively perceived differences between the recordings resulting from three different 3D microphone arrays: A Bowles-Array with a vertical coincident height-channel microphone layer, a Fukada-Tree/Hamasaki-Cube configuration and a Hybrid-Array containing the signals from the Bowles-Array main layer and the Hamasaki-Cube height layer. It was hypothesised that the arrays in concern will produce recordings that shall lead each to an increased perception of specific attributes for all sources tested (cello, violin, handpan, djembe, guitar). In order to detect possible patterns in spatial and timbral auditory perception subjective listening tests included direct scale magnitude estimations for the attributes Naturalness, Presence, Preference, Width, Localisation Accuracy, Distance/Depth, Envelopment, Spatial Balance, Room Perception, Vertical Image Shift, Vertical Image Spread, and Vertical Frequency Separation, and category scaling for the assessment of timbral attributes. Results suggest that none of the arrays in concern conveyed an increased perception of any of the attributes for all sources, which disproves the hypothesis and indicates a source-dependent performance. Simultaneously patterns in the subject responses have been detected which could be explained through psychoacoustic findings focussing on the correlation of perception between the attributes in question. Furthermore, by trying to explain the obtained differences in auditory perception between the different arrays, some assumptions could be made upon what components of which array could have contributed to a specific perception. These findings could serve as a reference for future experiments in the fields of 3D recording techniques or psychoacoustics.

3D audio in Ambisonics and production tools

Thursday, 26th of September, 10:50   Room: H-Hs

Franz Zotter, Matthias Frank

In this lecture we will present some background and basics behind Ambisonics work as a production tool and intermediate audio format. We will show some available practical tools that are for free, how their working principles are, and in which way they can be used with channel based or (first/higher-order) Ambisonic microphone recordings, loudspeakers, or headphones.

icsa2019@tu-ilmenau.de

Field Test for Immersive and Interactive Audio Production with the Gewandhausorchester Leipzig using MPEG-H

Thursday, 26th of September, 11:10   Room: H-Hs

Jakob Mäsel, Christian Simon, Jin Choi, Andreas Schulz, Andreas Silzle

Next Generation Audio (NGA) codecs offer new features for consumers such as advanced user interactivity, immersive sound and optimized reproduction across different classes of playback devices. Yet, 'real world' experience in the application of NGA workflows in classical music production is limited: The realistic depiction of the original concert hall sound and the treatment of elevated sound sources impose challenges when recording and mixing. Besides producing immersive formats, sound engineers face new tasks like the integration of user interactivity, ensuring downmix compatibility and loudness consistency. This paper expounds our field test for an NGA production workflow found suitable for complex orchestral music. Using MPEG-H 3D Audio as an example, we recorded three full-length concerts with Gewandhausorchester Leipzig, among them the prestigious New Year's concert with Beethoven's 9th symphony conducted by Gewandhauskapellmeister Andris Nelsons. We developed a microphone setup for the immersive recording of classical music in Gewandhaus and achieved very convincing results with the recording and mixing strategies delineated in this paper. The reverberance in the recordings was found to be realistically relatable to the excellent perception in the original concert hall. Elevated sources, like the choir and soloists, were clearly depicted as such. For metadata authoring and mastering, free MPEG-H production tools were used to regulate loudness, dynamic range and downmix compatibility, targeting playback over loudspeakers and 3D-audio soundbars. Additionally, we tested two different interactivity approaches suitable for classical music listeners.

Basics for 3D audio for live events

Thursday, 26th of September, 11:30   Room: H-Hs

Lasse Nipkow

This talk is an introduction to the workshop ‚3D audio for live events' on the following day. Reproduction of 3D audio through a PA system means an upscaling the speaker setup to a much larger listening area. It requires much more speakers to avoid audible holes between the speakers. During the talk, Lasse Nipkow will explain how the speakers should best be arranged due to psychoacoustic phenomena. On the other hand, he will explain what kind of speaker signals should be used so that listeners have as impressive a sound experience as possible.

ADAM AUDIO

S3A
[1999–2009]

S3X-H
[2009–2017]

S3H
[2017–today]

TECHNOLOGIES EVOLVE,
VALUES REMAIN.

20 YEARS
ADAM AUDIO

Session 2: Tools

Friday, 27th of September

Ambisonics Decoder Description (.ADD) – developing a new file format for Ambisonics decoding matrices

Friday, 27th of September, 10:30   Room: H-Hs

Gabriel Arlauskas, Jonas Ohland, Henning Schaar

Different software solutions have been developed for the calculation and implementation of Ambisonics decoding matrices. The present paper presents and describes a new data file format which can be used as an intermediate between solutions. Currently available software solutions use particular data conventions causing difficult compatibility and exchangeability. In the present work an open-source toolkit is developed for storing, handling and using Ambisonics decoding matrices. The toolkit includes tools for conversion from common matrix data conventions to the ADD-format and back, calculating decoding matrices, decoding Ambisonics signals and extracting existing matrices from external decoding tools. The new ADD-format and toolkit enables increased flexibility in production workflows and eliminates the drawbacks and limitations regarding compatibility between software solutions.

A Systematic Performance Investigation of Convolution Algorithms for Synthetic Room Reverberation

Friday, 27th of September, 10:50   Room: H-Hs

Wolfgang Fohl, Stefan Heidtmann

This paper presents a systematic investigation of optimization strategies for the convolution algorithm. Special attention is given to features relevant for the creation of virtual room acoustics, where the source signal is convolved with a room impulse response signal which has a length of several seconds. Examined were optimizations for the discrete convolution in the time domain and for the partitioned fast convolution in the frequency domain. Applied technologies were usage of AVX instructions, and GPU computing with the OpenCL framework. The results of the various algorithms are evaluated in terms of sample throughput. Various influence factors on the measured performance were identified. It turned out, that even ambitious projects with more than 10 channels and filter response lengths of several seconds may be rendered in real time with the GPU version of the discrete convolution.

Auralization systems for simulation of augmented reality experiences in virtual environments

Friday, 27th of September, 11:10   Room: H-Hs

Peter Dodds, Sebastia V. Amengual Gari, W. Owen Brimijoin, Philip W. Robinson

Augmented reality has the potential to connect people anywhere, anytime, and provide them with interactive virtual objects that enhance their lives. To deliver contextually appropriate audio for these experiences, a much greater understanding of how users will interact with augmented content and each other is needed. This contribution presents a system for evaluating human behavior and augmented reality device performance in calibrated synthesized environments. The system consists of a spherical loudspeaker array capable of spatial audio reproduction in a noise isolated and acoustically dampened room. The space is equipped with motion capture systems that track listener position, orientation, and eye gaze direction in temporal synchrony with audio playback and capture to allow for interactive control over the acoustic environment. In addition to spatial audio content from the loudspeaker array, supplementary virtual objects can be presented to listeners using motion-tracked unoccluding headphones. The system facilitates a wide array of studies relating to augmented reality research including communication ecology, spatial hearing, room acoustics, and device performance. System applications and configuration, calibration, processing, and validation routines are presented.

The Distribution of Ambisonic and Point Source Rendering to Ethernet AVB Speakers

Friday, 27th of September, 11:30   Room: H-Hs

Sean Devonport, Richard Foss

Point source rendering is used by many object-based audio systems to mix audio objects to loudspeaker arrangements. Algorithms such as Distance-Based Amplitude Panning and Vector-Base Amplitude Panning allow for audio objects to have their locations rendered with high precision. It has been shown that in the context of loudspeaker rendering, point sources rendered with Ambisonics are often spatially blurred. However, Ambisonics does have the advantage of being able to create interesting spatial audio effects and ambient scenes can be recorded using Ambisonic microphones. This paper intends to highlight the advantages that may be gained by combining Ambisonics with virtual point source rendering. It is well known that the processing required for rendering both point source and Ambisonics can have a large overhead. To mitigate this, a distributed spatial audio system based on Ethernet AVB and distributed endpoint processors is modified to incorporate both point source rendering and Ambisonics. An example is given of how point source rendering can be integrated with Ambisonics using this system with existing software.

icsa2019@tu-ilmenau.de

Session 3: Listening Test 1

Friday, 27th of September

Towards the Perception of Sound Source Directivity Inside Six-Degrees-of-Freedom Virtual Reality

Friday, 27th of September, 16:40   Room: H-Hs

Thomas Robotham, Olli S. Rummukainen, Emanuël A. P. Habets

Sound source directivity is a measure of the distribution of sound, propagating from a source object. It is an essential component of how we perceive acoustic environments, interactions and events. For six-degrees-of-freedom (6-DoF) virtual reality (VR), the combination of binaural audio and complete freedom of movement introduces new influencing elements into how we perceive source directivity. This preliminary study aims to explore if factors attributed to 6-DoF VR have an impact on the way we perceive changes of simple sound source directivity. The study is divided into two parts. Part I comprises of a control experiment in a non-VR monaural listening environment. The task is to ascertain difference limen between reference and test signals using a method of adjustment test. Based on the findings in Part I, Part II implements maximum attenuation thresholds on the same sound source directivity patterns using the same stimuli in 6-DoF VR. Results indicate that for critical steady-state signals, factors introduced by 6-DoF VR potentially mask our ability to detect loudness differences. Further analysis of the behavioral data acquired during Part II provides more insight into how subjects assess sound source directivity in 6-DoF VR.

Listening Tests with Individual versus Generic Head-Related Transfer Functions in Six-Degrees-of-Freedom Virtual Reality

Friday, 27th of September, 17:00   Room: H-Hs

Olli S. Rummukainen, Thomas Robotham, Axel Plinge, Frank Wefers, Juergen Herre, Emanuël A. P. Habets

Individual head-related transfer functions (HRTFs) improve localization accuracy and externalization in binaural audio reproduction compared to generic HRTFs. Listening tests are often conducted using generic HRTFs due to the difficulty of obtaining individual HRTFs for all participants. This study explores the ramifications of the choice of HRTFs for critical listening in a six-degrees-of-freedom audio-visual virtual environment, when participants are presented with an overall audio quality evaluation task. The study consists of two sessions using either individual or generic HRTFs. A small effect between the sessions is observed in a condition where elevation cues are impaired. Other conditions are rated similarly between individual and generic HRTFs.

icsa2019@tu-ilmenau.de

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

How positioning inaccuracies influence the spatial upsampling of sparse head-related transfer function sets

Friday, 27th of September, 17:20   Room: H-Hs

Christoph Pörschmann, Johannes Mathias Arend

Determining full-spherical individual sets of head-related transfer functions (HRTFs) based on sparse measurements is a prerequisite for various applications in virtual acoustics. To obtain dense sets from sparse measurements, spatial upsampling of sparse HRTF sets in the spatially continuous spherical harmonics (SH) domain can be performed by an inverse SH transform. This limits an HRTF interpolation in the spatially continuous spherical harmonics (SH) domain can result in audible artifacts. In a previous publication we presented the SUpDEq method (Spatial Upsampling by Directional Equalization), which reduces these artifacts by a directional equalization prior to the SH transform. Generally, apart from the spatial resolution of the HRTF set, measurement inaccuracies, for example caused by displacements of the head during the measurement, can influence the spatial upsampling as well. By this direction-depending temporal and spectral deviations are added to the dataset, which in the process of spatial upsampling can cause artifacts comparable to spatial aliasing errors. To reduce the influence of the distance inaccuracies, we present a method for distance error compensation that performs an appropriate distance-shifting of the measured HRTFs. Determining the required values for the shift benefits from the directional equalization performed by SUpDEq and results in time-aligning the directionally equalized HRTFs. We analyze the influence of the angular and distance displacements on spectrum, on interaural cues and on modeled localization performance. While limited angular inaccuracies only have a low impact, already small random distance displacements cause strong impairments, which can be significantly reduced applying the proposed distance error compensation method.

Enhanced Immersion for Binaural Audio Reproduction of Ambisonics in Six-Degrees-of-Freedom: The Effect of Added Distance Information

Friday, 27th of September, 17:40   Room: H-Hs

Axel Plinge, Sebastian Schlecht, Olli Rummukainen, Emanuel Habets

The immersion of the user is of key interest in the reproduction of acoustic scenes in virtual reality. It is enhanced when movement is possible in six degrees-of-freedom, i.e., three rotational plus three translational degrees. Further enhancement of immersion can be achieved when the user is not only able to move between distant sound sources, but can also move towards and behind close sources. In this paper, we employ a reproduction method for Ambisonics recordings from a single position that uses meta information on the distance of the sound sources in the recorded acoustic scene. A subjective study investigates the benefit of said distance information. Different spatial audio reproduction methods are compared with a multi-stimulus test. Two synthetic scenes are contrasted, one with close sources the user can walk around, and one with far away sources that can not be reached. We found that for close or distant sources, loudness changing with the distance enhances the experience. In case of close sources, the use of correct distance information was found to be important.

Session 4: Methodology

Saturday, 28th of September

New potentials for portable spatial audio with MEMS based speakers

Saturday, 28th of September, 09:00   Room: H-Hs

Daniel Beer, Tobias Brocks

There is a high demand for portable spatial sound on the market. Manufactures of headphones and hearing aids have to deal with miniaturization of electroacoustic transducer (micro speakers) size by keeping sound quality and energy efficiency (battery life time) at the same time. This is especially important for 3D-Headphones with multiple loudspeakers. Different techniques are known to downsizing transducers. A very successful method is given by the semiconductor industry. The use of the so-called MEMS technology (Micro-Electro-Mechanical-System) leads to great success in applications for microphones and accelerometers. This success has triggered a high interest in the potential of MEMS technology for speaker manufacturing. Based on patents, the initial approaches of MEMS loudspeakers will be presented. An outlook of the arising new potentials for portable spatial audio with MEMS based speakers is given.

Deep Neural Network Approaches for Selective Hearing based on Spatial Data Simulation

Saturday, 28th of September, 09:20   Room: H-Hs

Simon Hestermann, Hanna Lukashevich, Christoph Sladeczek

Selective Hearing (SH) refers to the listener's attention to specific sound sources of interest in their auditory scene. Achieving SH through computational means involves detection, classification, separation, localization and enhancement of sound sources. Deep neural networks (DNNs) have been shown to perform these tasks in a robust and time-efficient manner. A promising application of SH are intelligent noise-cancelling headphones, where sound sources of interest, such as warning signals, sirens or speech, are extracted from a given auditory scene and conveyed to the user, whilst the rest of the auditory scene remains inaudible. For this purpose, existing DSP-based noise cancellation approaches need to be combined with machine learning techniques. In this context, we evaluate various DNN architectures for sound source detection and separation. In addition, we propose a data simulation approach for generating different sound environments for a virtual pair of headphones. The Fraunhofer SpatialSound Wave technology is used for the training data simulation and a realistic evaluation of the trained models. For the evaluation, a three-dimensional acoustic scene is simulated via the object based audio approach.

Exploring Audiovisual Support Systems for In-Car Multiparty Conferencing

Saturday, 28th of September, 09:40   Room: H-Hs

Florian Weidner, Bernhard Fiedler, Johannes Redlich, Wolfgang Broll

Calling while driving poses a severe safety risk. When more than two people are involved in a call - a conference call - this risk increases even more. Intelligent vehicles could offer support systems that ease the cognitive burden of such a multiparty calls. We explore the possibilities of such advanced driving assistant systems (ADAS) in two ways: first, we investigate object-based spatial audio where each remote caller is modelled as a distinct audio source. Second, we apply a non-intrusive ambient stereoscopic 3D (S3D) visualization that indicates the current speaker and its location. In a between-subject design driving simulator study (n=56), we assess workload, user experience and driving performance. Surprisingly, we found no positive effect of object-based audio. However, we present evidence how a supporting visualization might lower situational stress and increase the system's dependability. We conclude that a supportive and intelligent stereoscopic visualization is a promising candidate for enhancing multiparty conference calls while driving.

Session 5: Listening Test 2

Saturday, 28th of September

The Virtual Acoustic Spaces Unity Spatializer with custom head tracker

Saturday, 28th of September, 10:30   Room: H-Hs

Thomas Resch, Markus Hädrich

The Virtual Acoustic Spaces (VAS) Unity Spatializer is a plugin for dynamic binaural synthesis for Unity. It can handle Impulse Responses (IRs) of arbitrary length (limited only by hardware resources). Hence, it is possible to calculate the binaural synthesis not only with Head Related Transfer Functions (HRTFs), but also on the basis of Binaural Room Impulse Responses (BRIRs). The plugin can also virtualize reflections calculated by raytracing and it is possible to load an individual IR set for each instance. In addition to being compatible with off-the-shelf Cross Reality (XR) hardware it features a Bluetooth binding for an easily built, custom-made head tracker based on an ESP32 board. It is therefore predestined for audio augmented reality applications.

Room geometry inference using sources and receivers on a uniform linear array

Saturday, 28th of September, 10:50   Room: H-Hs

Youssef El Baba, Andreas Walther, Emanuel Habets

State-of-the-art room geometry inference algorithms estimate the shape of a room by analyzing peaks in room impulse responses. These algorithms typically require the position of the source wrt the receiver array, this position is often estimated with sound source localization, which is susceptible to high errors under common sampling frequencies. This paper proposes a new approach, namely using an array with a known geometry and consisting of both sources and receivers. When these transducers constitute a uniform linear array, new challenges and opportunities arise for performing room geometry inference. We propose solutions designed to address these challenges, but also designed to leverage the opportunities for better results.

Detection of Constant Phase Distortions in Filters for Sound Field Synthesis

Saturday, 28th of September, 11:10   Room: H-Hs

Frank Schultz, Nara Hahn, Sascha Spors

Filters with constant phase shift, often in conjunction with +-3/6 dB amplitude decay per octave, frequently occur in sound field synthesis and sound reinforcement applications. These ideal filters are known as (half) differentiators/integrators and exhibit zero group delay and +-45/90 degree phase shift. It is well known that certain group delay distortions in electro-acoustic systems are audible for trained listeners for critical audio stimuli, such as transient, impulse-like and square wave signals. It is of interest if linear distortion by a constant phase shift is audible as well. For that, we conducted a series of ABX listening tests, diotically presenting non-phase shifted references against their treatments with different phase shifts. The experiments revealed that (similar to group delay) for the critical square waves, this can be clearly detected, which generally depends on the amount of constant phase. Here, -90 degree (Hilbert transform) is comparably easier to detect than other phase shifts. For castanets and pink-noise shaped Dirac impulse, the detection rate tends to guessing for most listeners, although the polarity inversion (+-180) is detectable, thereby confirming earlier studies on this topic. Our results motivate to apply constant phase shift filters to ensure that also the most critical signals are technically reproduced as best as possible (cf. modern electronic dance music with typical crest factor of as low as 6 dB). In the paper, we furthermore give the analytical expressions for continuous and discrete-time infinite impulse responses for unit magnitude filter with arbitrary constant phase shift.

Plausibility versus Authenticity of binaural walk-through scenarios

Saturday, 28th of September, 11:30   Room: H-Hs

Annika Neidhardt, Anna-Maria Zerlik

In various studies the plausibility and authenticity of dynamic binaural reproduction systems was evaluated by randomly providing either the real audio scene or its binaural simulation. So far, these studies only focused on dynamic head rotation without changes of the head position. In this paper, we present a first attempt to evaluate the perception of binaural walk-throughs in a similar way. BRIRs were measured with a Kemar 45ba dummy head with open AKG K1000 headphones to generate a binaural simulation of a real approaching motion towards a loud-speaker. No individualization was applied. The participants had to identify, whether they were listening to the real scene or the binaural reproduction. In the first part, the real scene was not included to test whether its consideration in the test affects the subjects' inner reference. The results suggest, that this is the case. Trained listeners were able to identify simulation and real reproduction in most cases. The main goal of the study was the determination of the weaknesses of the current system in terms of perceptual aspects which reveal the simulation. These observations are discussed in detail in order to develop suggestions for improving the system.

Perceptual Evaluation of Spatial Resolution in Directivity Patterns 2

Saturday, 28th of September, 11:50   Room: H-Hs

Matthias Frank, Manuel Brandner

The incorporation of source directivity is important for a plausible and authentic auralization. While high-resolution measurement setups and data exist, it is yet not clear how detailed the directivity information has to be measured and reproduced with regard to perception. In particular, when source and listener are at the same location, resulting in a high direct-to-reverberant energy ratio, the precise shape of the directivity pattern might not yield perceptual differences. The paper approaches this question by a listening experiment in a virtual environment with generic directivity patterns and coincident position of listener and source. The experiment compares different spatial resolutions (spherical harmonic orders) of the directivity patterns for multiple virtual listener/source positions/orientations and levels of direct sound for speech and noise. The virtual environment employs a higher-order image-source model and binaural, dynamic Ambisonic playback. The results show that the exact shape of the directivity pattern is often perceptually irrelevant, while the preservation of the direct-to-reverberant energy ratio is more important.

Posters

Friday, 27th of September

Data set: BRIRs for position-dynamic binaural synthesis measured in two rooms

Friday, 27th of September, 15:30 - 16:40   Room: H-Hs

Annika Neidhardt

Binaural room impulse responses were measured with a KEMAR 45BA dummy head. It was placed at different positions located on a line with a length of 2m in a 25cm positional resolution. and an azimuth resolution of 4°. Two source positions were considered in the setup, one in front of the line, one at the side. The same arrangement of source and receiver positions was realized in two different rooms, a quite dry listening laboratory and a quite reverberant seminar room. The data set is valuable for realizing, testing and studying dynamic binaural walk-through scenarios in the two different rooms. It is provided online.

Evaluation of various neural network features for binaural localization of speech signals in the horizontal plane

Friday, 27th of September, 15:30 - 16:40   Room: H-Hs

Song Li, Mengsha Liu, Jürgen Peissig

Binaural localization of speech signals has been widely applied in human-computer interaction systems, communication devices, etc. Traditionally, the binaural cues, i.e., frequency-dependent interaural level difference (ILD) and interaural time difference (ITD), are often used to localize binaural signals in the horizontal plane. The spectral information, especially the positive gradient, is an important cue of the binaural sound localization in the sagittal plane. It is still unknown, whether this cue can be used for sound localization in the horizontal plane or not. The mel frequency cepstral coefficient (MFCC) is commonly applied in automatic speech and speaker recognition, which can also be used as an acoustic feature to localize speech signals. Furthermore, it is interesting to investigate whether the difference of MFCC between two ears (DMFCC) can be used for binaural sound localization or not. In the present study, the above-mentioned acoustic cues, i.e., ILD, ITD, spectral information, MFCC and DMFCC, are used as neural network features for binaural localization of speech signals in the horizontal plane, and the performance of these features in terms of localization accuracy is evaluated.

Mixed Realities: a live collaborative musical performance

Friday, 27th of September, 15:30 - 16:40   Room: H-Hs

Andrea Genovese, Marta Gospodarek, Agnieszka Roginska

The question of remote music collaboration is ever-relevant in the field of immersive audio. The proliferation of commercial immersive devices for virtual and mixed reality enables today's musicians to experience enhanced forms of virtual presence when remotely connected to their peers. In the presented work, a dancer and percussionists have been recorded with both microphones and an OptiTrack motion capture system. Their audiovisual presence is converted into a game character avatar able to be reproduced as a pre-recorded performer through VR headsets. During the exhibition, a live percussionist performer wearing a motion-capture suit enters the performance as an additional character in the virtual scene. Audience members are also brought in the experience by means of a VR headset and are able to observe and hear the live collaboration between the real musician and the pre-recorded virtual members of the ensemble. The goal is that of creating a compelling and cohesive immersive experience where the real and virtual audio layers blend seamlessly. To match the auditory expectations set by the presence of a real source, the virtual audio material is treated to acoustically match the characteristics that the same instruments would have in the intended exhibition space via measured impulse responses and dynamic binaural rendering. This paper illustrates an overview of the proposed method used to create this novel musical experience and a discussion about the impressions of participants, audience and musicians. Future technical enhancements for the involved elements are discussed along with proposed evaluation procedures and variations to the pipeline.

A Virtual Acoustic Environment for Psychoacoustic Assessment of Wind Turbine Noisescapes

Friday, 27th of September, 15:30 - 16:40   Room: H-Hs

Jakob Bergner, Stephan Preihs, Jürgen Peissig

The massive expansion of wind energy as part of the transition to renewable energy supply has gained increasing scepticism and negative emotions within the public especially in rural areas where wind farms exist or are planned. Among various social, economic and ecological arguments, objectors worry about noise exposure of the affected population. Beside the overall sound pressure level that might be increased by wind turbines, acoustic properties like tonality and amplitude modulation are widely perceived as increasingly annoying. Most research on these properties has been carried out in an isolated, singular and mostly physically motivated manner where overall auditory perception is rarely considered. To reduce this gap and to get a better understanding on how the interaction of the mentioned properties influences the perception of wind turbine noise, systematic psychoacoustic testing offers the potential to contribute to basic scientific findings. In order to carry out excessive, comparable and valid perceptual evaluations, a reproduction environment for wind turbine noise was designed, realized and validated. It consists of a recording and processing stage for capturing existing wind turbine noisescapes, a method to model and synthesize plausible sound generation and propagation, as well as a plausible audiovisual reproduction environment that respects ecological validity. This paper presents the underlying interdisciplinary and project-related research goal, the general methodology of development and validation, as well as an outlook for current and future psychoacoustic application.

icsa2019@tu-ilmenau.de

Investigation on spatial auditory perception using non-uniform spatial distribution of binaural room impulse responses

Friday, 27th of September, 15:30 - 16:40   Room: H-Hs

Stephan Werner, Florian Klein, Georg Götz

For spatial audio reproduction in the context of virtual and augmented reality, a position-dynamic binaural synthesis can be used to reproduce the ear signals for a moving listener. A set of binaural room impulse responses (BRIRs) is required for each possible position of the listener in the room. The required spatial resolution of the BRIR positions can be estimated by spatial auditory perception. If the resolution is too low, jumps in perception of direction and distance and coloration effects occur. This contribution presents an evaluation of spatial audio quality using different spatial resolutions of the position of the used BRIRs. The evaluation is performed with a moving listener. The test persons evaluate any abnormalities in the spatial audio quality. The result is a comparison of the quality and the spatial resolution of the various methods used.

Development of a method for measuring, displaying, and evaluating behavior in a position-dynamic binaural synthesis application

Friday, 27th of September, 15:30 - 16:40   Room: H-Hs

Dominik Zapf, Stephan Werner, Stephan Fremerey

As binaural systems become more and more important, understanding user behavior can be critical to the ease of use and efficiency of designing these applications. For this purpose, the test subjects are tested with a motion tracking system. Six degrees of freedom for the position and head orientation of a listener are recorded. The movement of 23 people is recorded in five different test scenarios and for an exploration time of approximately three minutes. Each of the first four scenarios contains a different specific task where the listener is asked to find a specific audio object. The last scenario consists of a piece of music in which the test persons can move freely. The explorable area is 4 m x 4 m, with a spatial grid of the used binaural filters of 0.25 m steps and 5° in head rotation. Several features are extracted from the data: Yaw angle, inclination angle, yaw angle speed, xyz position, walking speed, path and area walked, time for localization and reaction time. It turns out that most people behave similarly, especially in simpler scenarios. The walkable area and the exploration behaviour, however, seemed to be very individual.

Real-time Estimation of Reverberation Time for Selection of suitable binaural room impulse responses

Friday, 27th of September, 15:30 - 16:40   Room: H-Hs

Florian Klein, Annika Neidhardt, Marius Seipel

The aim of auditory augmented reality is to create a highly immersive and plausible auditory illusion combining virtual audio objects and scenarios with the real acoustic surrounding. For this use case it is necessary to estimate the acoustics of the current room. A mismatch between real and simulated acoustics will easily be detected by the listener and will probably lead to In-head localization or an unrealistic acoustic envelopment of the virtual sound sources. This publication investigates State-of-the-Art algorithms for blind reverberation time estimation which are commonly used for speech enhancement algorithms or speech dereverberation and applies them to binaural ear signals. The outcome of these algorithms can be used to select the most appropriate room out of a room database. A room database could for example include pre-measured or simulated binaural room impulse responses which could directly be used to realize a binaural reproduction. First results show promising results combined with low computational effort. Further strategies for enhancing the used method are proposed in order to create a more precise reverberation time estimation.

Sweet area size for the envelopment of a recursive and a non-recursive diffuseness rendering approach

Friday, 27th of September, 15:30 - 16:40   Room: H-Hs

Matthias Blochberger, Matthias Frank, Franz Zotter

We compare two algorithms that create diffuse sound fields in terms of the sweet area size they produce in a 10 m by 10 m playback room. One approach employs random frequency-dependent group delays to generate a set of minimally correlated impulse responses used as filters for multi-channel diffusion of a mono sound. Canfield-Dafilou presented a frequency-dependent maximum group delay value as constraint on the randomness of this method, ensuring minimal audible artifacts in studio environments. We relax those constraints to enable a stable, enveloping, and diffuse listening experience filling the targeted larger audience area size, which however, unavoidably yields the impression of spaciousness and reverberation. Consequently, the new FIR approach competes with the IIR response feedback-delay-network diffusion as alternative. We conduct listening experiments to reveal properties and effectiveness of both methods, in particular regarding sweet area size and sound quality.

Workshops

From a demo room to an arena: design of spatial sound reinforcement systems

Thursday, 26th of September, 13:30 - 14:20   Room: 3DP
Friday, 27th of September, 16:00 - 16:50   Room: LL

Javier Frutos Bonilla

All the different multichannel spatial algorithms have one thing in common: the loudspeakers used to represent the audio object must overlap themselves on the listening area. The nominal dispersion angle, maximum sound pressure level and frequency range are some key factors when chosing the right loudspeaker. However, is that all you need to consider? What happens with the coverage and the SPL decay when installing loudspeakers at different heights? Do we need the same loudspeaker type when using playback content as when installing them over a stage? Can we combine different types in a single installation? In small demo rooms, clubs or cinema-like venues we consider that the whole audience is on a single listening area. What happens when we install the system in an arena or a theatre where we need to divide the audience into different subspaces? This workshop addresses all these questions, gives an insight on how the loudspeaker types influence the results, and discusses different system design solutions using practical examples.

Remicing instruments. Object Based recording of instruments with optimized natural sound.

Thursday, 26th of September, 13:30 - 14:20   Room: SR-0.10

Gregor Zielinsky, Trine Thybo

Recording natural instruments for the purpose of object based recording and post processing is a huge problem. There are three important aspects, that usually interfere with each other. Which is: A) Sound-quality, mainly meaning naturally, original sound, B) Optimal crosstalk suppression, C) Recording all instruments together in there natural playing situation. These contradictions usually lead to unsatisfying result. Using a classical mic setup with common high quality spot mics will not result in cross talk suppression as it is needed. Using common contact microphones will lead to unsatisfying sound. Separating instruments and recording them one by one in anechoic chambers not only leads to very bad sound, but also to musical problems as extreme intonation and rhythmical effects. The workshop will show a compete new approach to recording instruments for object based recording. The Remic system allows a perfect crosstalk rejection as well as an optimal natural sound. Also, the Remics can be attached to the instruments in an easy way. These new recording options will lead to perfect results in all kind of situations as 3D recording, Games, VR, AR and any kind of situation where instruments are needed in a separated way. We will play a session that will show the sound of the Remics, the Crosstalk rejection as well as with way how to use them in an object based mix situation as for headphones or speakers.

icsa2019@tu-ilmenau.de

Live Concert-Recordings BigBand and Popbands in 9.1 - mixed in the box

Thursday, 26th of September, 14:30 - 15:20   Room: 3DP

Peter Brümmer

In this listening session I would like to demonstrate, how multitrack recordings, which were meant to be broadcasted in stereo, can come to a new life as an immersive audio event. The produced separation of the different sections allows the listener to be in the centre of the music – like a band leader. Live recordings from international established artists (WDR BigBand, Incognito, Level 42 etc.) will be taken as an example for 9.1 demonstrations. Furthermore the workflow as well as the used low budget equipment will be explained. Have fun!

Audio perception in the context of quality assurance for manufacturing and industry

Thursday, 26th of September, 14:30 - 15:20   Room: SR-0.10

Sara Kepplinger

The use of acoustic event detection in industrial quality control allows for contact less quality assurance. Thus, this might be useful in predictive maintenance, end-of-line-, or in-line-testing. Furthermore, it may support the use of available information (i.e., sound) to reduce rejects and production losses. In fact, faults such as broken ball bearings or defective engines are heard and recognized easily by experienced machine operators, but hardly or only costly recognized otherwise. Due to highly complex noise scenarios in fabrication plants it is challenging to detect audible errors in the human hearing range. However, industrial sound analysis solutions that are robust against unpredictable background noise in factories but sensitive to subtle differences in the observed sound are developed. Therefore, qualitative annotation of data is inevitable. Hence, the remaining question is about collecting adequate feedback and annotations from the experts, the machine operators recognizing their personal ear-catcher. In this two hours workshop we would like to address this challenge by providing a three-parted program. We will start with a formal introduction to the topic, then provide a hands-on activity to allow the experience of the way of "hearing quality", and end with a moderated panel discussion to talk about useful methods, (non-)existing standards and how-toes in order to support future annotation tasks and overall acoustic quality assurance processes in industry. The outcome of this workshop supports further evaluation activities or in-situ feedback collection towards finding the most adequate way of annotation for experienced machine operators using sound source description for quality assurance.

The Piece "Cosmic Space" for a Spatial Sound Dome and VR, AR or MR Installation

Thursday, 26th of September, 14:30 - 15:20   Room: WS
Friday, 27th of September, 13:30 - 14:20   Room: WS

Goda Marija Gužauskaite

The piece "Cosmic Space" was created two years ago and it took exactly the same time to arrange it into the spatial sound sphere. The piece is an algorithmic composition of 200 rhythmical lines which were spread into the sphere systematically in two different ways: as 200 sound sources and as 4 stereo files. In the presentation given by the composer and arranger, will be compared both arrangements as well as presented the algorithmic structure of the composition and the final realisation of the spatial sound and virtual, augmented or mixed reality's installation.

Comparison of the 3D-microphone-systems FOA, SOA, ORTF-3D and au3Dio

Friday, 27th of September, 10:00 - 10:50   Room: LL
Friday, 27th of September, 16:00 - 16:50   Room: TV Studio

Christian Vaida

How do the 3D-microphone-systems FOA, SOA, ORTF-3D and au3Dio differ tonally? The preference of the recordings, the depth and the localisation shall be compared. The work „Te Deum" by H. Berlioz is very suitable for this test, due to its wide lineup in the front with symphonic orchestra, soloists, three discrete choirs with 180 singers in total as well as an organ from far behind. The criteria for chosing the 3D-systems were that they should be as different as possible and available on the market. Techniques that require complicated setups such as OCT-3D, Hamasaki and others were omitted deliberately. In practice the majority of 3D-audio-recordings is done with small ambisonics microphones, therefore bigger setups with spaced microphones should at least be easy to assemble. The sound characteristics of different microphones play a role indeed, but cannot be considered in this test. The point is to compare the different setups in principle. FOA (4 capsules) and SOA (8 capsules) don't contain time differences, the ORTF-3D setup with its 8 cardioid microphones features already a small time difference component but is still compact, whereas the au3Dio system features bigger spacings and 10 micprophones. To ensure comparability, all systems are recorded simultaneously at the same position and afterwards converted to third order ambisonics. For ORTF-3D and au3Dio, the discrete microphone signals are panned to their respective position, FOA and SOA are upmixed with Harpex. The TOA-stream of all recordings can then be played back on any 3D-setup using the IEM AllRaDecoder.

icsa2019@tu-ilmenau.de

———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————
———————————————————————————————

3D Audio Recording Practice - Example of a Flexible Workflow

Friday, 27th of September, 11:00 - 11:50   Room: LL
Saturday, 28th of September, 11:30 - 12:20   Room: LL

Matthias Domke, Christian Birkner

Immersive 3D audio sounds great, offers fantastic design possibilities and comes very close to a natural listening experience. However, there are still only a few productions, as the workflow is often seen as complicated and incompatible with established processes. Nevertheless, this highly interesting topic was investigated at the HTWK Leipzig in cooperation with practical partners within the scope of final theses and projects. A very interesting and flexible 3D recording workflow was developed, which is introduced in the context of this lecture. Particularly remarkable in this context is the (seemingly) unconventional use of measuring microphones for 3D audio production. The background and advantages of this approach will also be presented in the lecture.

Real-life Experiences in Object-based Audio Production

Friday, 27th of September, 11:00 - 11:50   Room: TV Studio

Werner Bleisteiner, Christian Schimmöller, Helge Schwarz

Bayerischer Rundfunk (BR) was one of the partners in the EU Horizon 2020 research project OR-PHEUS (2015-2018) that pioneered the development of Next Generation Audio in broadcasting. BR's main role in this project was to create and produce suitable content for the project's pilots. Due to this objective, a variety of different pieces of typical broadcast genres and formats were created: music, documentary, general interest, sports, etc. Quite often, also different technologies for capturing and production were applied, from legacy 'dummy head binaural' to Ambisonics and real object-based audio, and all these at times even in parallel. In a practical 1-hour workshop at ICSA we are to demonstrate a selection of these productions in their different formats and give listeners the chance to experience and discuss both technical and creative questions of the different approaches with the producers and sound engineers involved.

icsa2019@tu-ilmenau.de

3D audio for live events

Friday, 27th of September, 13:30 - 14:20   Room: 3DP

Lasse Nipkow

Many applications for 3D audio are aimed at live events. Several manufacturers of PA systems now offer solutions. These are not optimized for native 3D audio, but achieve an immersive impression due to placement of objects and added reverberation. To operate native 3D audio for live events, there are two key questions to consider: — A) How should the loudspeaker layout be chosen so that the audience gets the most impressive listening experience at a reasonable cost? — B) How should 3D audio content be designed so that it reaches the audience as effectively as possible through such PA systems? Two phenomena can be used for 3D audio: Envelopment and projection of sound sources in the front. The most impressive case is provided by a loudspeaker setup using imaginary connecting lines between the loudspeakers, which results in a volume. Vertical phantom sound sources presented in the front sound more natural than horizontal ones. The number of vertical front loudspeakers determins the resolution of the image. The wedge-shaped loudspeaker setup, that means with height loudspeakers in the front, meets both requirements minimally. The greater the listening area, the more loudspeakers are needed. The greater the distances of the loudspeakers from each other, the more audible holes are perceived, depending on the acoustics of the playback room.

Beyond Ambisonics - 360 & VR Audio Content Creation

Friday, 27th of September, 13:40 - 14:20   Room: SR-0.10

Tom Ammermann

VR and 360 are mainly headphone applications. Thus, binaural headphone virtualization, well known for decades, becomes important as never before. Fortunately, current binaural virtualization technologies offers a fantastic new experience far away from creepy low quality 'surround simulations' in the past. The workshop will offer an overview, shows new approaches and how to deal with the content creation.

icsa2019@tu-ilmenau.de

Using audio objects in live and entertainment applications

Friday, 27th of September, 14:30 - 15:20   Room: 3DP

Christoph Sladeczek, Josua Hagedorn

Multi-channel loudspeaker systems are used for the generation of spatial audio. Especially for live and entertainment applications an intuitive interaction approach is needed to allow realtime generation of spatial audio scenes. The object-based audio approach is the most promising concept, which is about to revolutionize live and entertainment sound reproduction. An audio object is characterized by an input signal (signal of a voice or an instrument) and corresponding metadata representing the properties of the object like position or gain. The generation of loudspeaker signals is conducted by an audio renderer, who creates the speaker signal interactively. This workshop gives insides into the general concept of object-based audio, its background developments and focuses on usage in live and entertainment applications. The workshop will be conducted using Fraunhofer IDMT's SpatialSound Wave technology, which is used in entertainment facilities worldwide.

Sound for extreme 360° productions

Friday, 27th of September, 14:30 - 15:20   Room: SR-0.10

Martin Rieger

The workshop shows various examples of 360-degree video productions under challenging conditions, featuring location recordings and post-production. The purpose of the talk is to give practical insights of immersive VR-video and how sound on vision needs to be contemplated, which varies a lot from usual film formats and requires a lot of knowledge additional to audio as such. Different technologies and sometimes even custom solutions are needed on set and post. There is no use for a boom microphone and its operator, which gets replaced by an immersive microphone array which there is, just like for 360° cameras, no perfect setup for every occasion as people tend to claim that there is. Martin's experience ranges from a 360 documentary in Vietnam, Cambodia, and Thailand traveling with the team on motorbikes, to working in knee-high snow for a freeride experience in Austria, and to accompany the EU at aid operations in Belgium, Denmark and Bangladesh. Thus, he experienced a lot of different types of challenges, which classic production sound mixers and post-production facilities don't need to take care of. Therefore, he realized the importance of having a very specialized audio engineer who can, while on location recording, think about his own following 3D audio mix with sound design, restoration and re-recording. Considering the language used in the presentation and nature of the topic, Martin wants to share entertaining, yet informative stories and listening examples of the experiences he made with interesting virtual reality productions around the world.

Assessing Public Noisescapes By Means of Virtual Acoustic Environments

Friday, 27th of September, 17:00 - 17:50   Room: TV Studio

Jakob Bergner, Stephan Preihs, Jürgen Peissig

The qualitative human perception of complex auditory scenes is a multidimensional process which is strongly dependent on the context. This means for example that a single person can perceive the same sound very differently when listening conditions or time are changed. Thus, qualitative in-situ assessments of soundscapes and noisescapes are superimposed by the temporary context in most cases which leads to difficulties when it comes to reliability and thus further comparability of results. One possible approach to separate the contextual part from qualitative ratings of an auditory scene and to leaving only the auditory perception is to vary either property, context or acoustics, while the respective other stays constant. Hence, this workshop discusses the possibility of contributing to this issue by developing a virtual acoustic environment (VAE) where real and virtual noisescapes can be reproduced and be enriched by additional (visual) context. General conceptual and technical challenges of the proposed approach are introduced and subsequently illustrated by a current research project on wind turbine noise, including various immersive audio examples. Participants of the workshop are very welcome to engage in a concluding discussion.

A network-based array microphone for motion-tracked binaural applications

Saturday, 28th of September, 10:30 - 11:20   Room: LL

Thaden Cohrs, Lukas Treybig

The workshop provides an overview of the system design for array microphones and the required signal processing within an audio network. We present a spherical microphone array for 360° binaural motion-tracked applications in Ethernet-based AVB/TSN networks. An exemplary real time signal processing will be discussed. Moreover the management and controlling capability of network audio devices will be mentioned.

Spatial audio production for immersive fulldome projections

Saturday, 28th of September, 10:30 - 11:20   Room: TV Studio

Johannes Ott, Niklas Wienböker, Anca-Stefania Tutescu, Jan Rosenbauer, Thomas Görne

Fulldome is an immersive half-spherical video format designed for planetariums, often combined with 3D audio playback. This offers new ways of perceiving sound in space and enhancing fulldome productions with audiovisual synergy. However, the production auf audio for fulldome video poses some technical and artistic challenges. Limited time slots and resources seldom allow to work on sound productions inside a planetarium directly. Likewise, the various spatial audio technologies provide the user with fairly different approaches to create, position and move sounds in three-dimensional space. The workshop will discuss three different approaches to create spatial audio content for fulldome productions in remote studios and to present them in a planetarium. The object based proprietary Fraunhofer "Spatial Sound Wave" (SSW) system, scene based Higher Order Ambisonics (HOA), and channel based production will be compared. Technical challenges and the potential for storytelling in 3D audio will be discussed.

Holographic directivity measurement of line sources and sound panels

Saturday, 28th of September, 11:30 - 12:20   Room: TV Studio

Christian Bellmann

Compared to a standard Hifi stereo system, new 3D listening experiences have increased the complexity of sound systems a lot. To realize immersive audio reproduction, many loudspeakers distributed in the listening room may be placed. Other methods use more complex sound sources like line arrays and sound panels to shape distinct beams, including controlled reflections from the room boundaries. The complex control algorithms necessitate directivity data of each individual transducer with highly accurate phase information. The workshop will discuss the field of directivity measurement for multi-transducer sound systems and will compare the benefits and limits of a traditional far field measurement vs. new holographic measurement techniques.

Demos

Content Creation of Immersive Audio for Loudspeaker and Headphone Reproduction

Thursday, 26th of September, 13:30 - 14:20   Room: WS
Friday, 27th of September, 14:30 - 15:20   Room: WS

Georg Meier, Erik Bjoerkman, Stanislaw Gorlow, Brett Crockett

Dolby Atmos is an immersive audio format that combines standard audio channels with audio objects that can be positioned anywhere in a 3D space, including overhead. On playback, the channels and objects are rendered for the speaker layout of the playback system in order to deliver a custom, immersive audio experience using any Dolby Atmos-enabled playback system from mobile devices and home theaters to cinema theaters and music clubs. During this demonstration the basics of the Dolby Atmos Production Suite and workflows for immersive audio productions will be presented. More specifically, this demonstration will focus on best practices for music productions and binaural headphone content creation when using the Dolby Atmos Production Suite, however, the same production tools and audio format applies to the creation and rendering for Dolby Atmos compatible speaker systems. Two immersive music mixes will be broken down in two hands-on demo sessions. On Thursday an electronic piece will be presented and on Friday a pop/rock piece will be presented. Some extra headphones will be available on-site but it is recommended to bring your own if you want to follow the demonstration in detail.

On the eve of World War I - Interactive Walk-in 3D Audio Soundscape

Thursday, 26th of September, 16:00 - 18:00   Room: TV Demo
Friday, 27th of September, 10:00 - 12:00   Room: TV Demo
Friday, 27th of September, 16:00 - 18:00   Room: TV Demo

Werner Bleisteiner, Helge Schwarz, Christian Schimmöller

This piece presents part 1 of a projected but unfinished radio documentary on World War I. Conceptualised as collaboration between BR, IRT, BBC R&D and Radio France in 2013, it was intended to become a prototype to explore the creative and technical possibilities of object-based audio for immersion, accessibility, interactivity and personalisation by providing new quality features like 3D audio, multiple languages, dialogue enhancement, changing perspectives/storylines and responsive narratives. The original production, comprised of some 200 audio tracks for the soundscape along with the soundbites, quotes and music elements, is now transferred into the TU-Ilmenau Media Lab's Auditory Augmented Reality system to create an interactive walk-in 3D audio application, allowing the user both to select any of the scenes and move within freely to explore them. The individual soundscape is triggered through the motion tracking system and rendered in real-time into binaural sound reproduced over the listener's headphones.

icsa2019@tu-ilmenau.de

Moving Sound Pictures

Thursday, 26th of September, 16:30 - 18:00   Room: AV1
Friday, 27th of September, 15:30 - 17:00   Room: AV1

Konstantina Orlandatou

Moving Sound Pictures is a project where users have the opportunity to interactively explore images of well-known artists in a VR environment through playful actions. On the one hand, this interaction generates music and, on the other hand, information about the painting, e.g. historical aspects of the painting, painting techniques, etc. is transmitted. The project is a work-on-progress whose goal is the transfer of knowledge to people who have no artistic background. In this phase of the project, the visual artists Wassily Kandinsky, El Lissitzky, Piet Mondrian and Kasimir Malevich are in the foreground. The paintings "Merry Structure" by Wassily Kandinsky, "R.V.N.2" by El Lissitzky, "Suprematism" by Kasimir Malevich and "The large composition with red, yellow, blue and black" by Piet Monrdian are reconstructed in a VR environment built in Unreal Engine. At the moment the audio is being spatialized binaurally using the HTC Vive Pro toolkit but the goal is to have the possibility to use an external 3d audio sound system for the reproduction of sound and simulate the 3d space position of the paintings' objects during interaction to a 3d audio space.

Audiovisual Zoom Tablet Demo

Thursday, 26th of September, 17:00 - 18:00   Room: AV2
Friday, 27th of September, 11:00 - 12:00   Room: AV2

Axel Plinge, Emanuel Habets

The VR Player Audiovisual Zoom on an Android tablet plays 360 degree MPEG-H videos with Ambisonics sound. The user can look around by moving the tablet in space or swiping, pinch to zoom in, and change rendering parameters. The acoustic zoom is realized by parametric processing allowing to focus on individual sound sources. The playback over loudspeakers utilizes crosstalk cancellation, so binaural 3D sound is possible without headphones.

Immersive and Interactive Audio Production with the Gewandhausorchester Leipzig using MPEG-H

Thursday, 26th of September, 12:00 - 13:30   Room: H1554
Friday, 27th of September, 15:30 - 17:00   Room: H1554

Jakob Mäsel, Andreas Silzle

In the course of our field test for immersive and interactive audio production in classical music at Gewandhaus zu Leipzig we recorded three different concerts, i.a. Beethoven's 9th symphony conducted by Gewandhauskapellmeister Andris Nelsons. We present four different immersive and interactive demo items from these recordings on a 3D soundbar. The demo sessions build on the correspondent paper we present in the course of the conference. Please consider that space is limited to 9 seats, so come early. The demo takes approximately 30 minutes.

Flat, baffled 3D arrays

Friday, 27th of September, 17:00 - 18:00   Room: LL

Svein Berge

Does a 3D microphone array need to have a three dimensional shape? In this workshop / demo, we will review some basic theory about 3D microphone arrays and the physical constraints that they are subject to. This will be used to explain how to design a flat, disc-shaped array to acquire high-quality recordings in three dimensions, either in higher-order ambisonic or binaural format. All with off-the shelf components and standard manufacturing techniques. A few prototypes of this kind have already been made, and will be used to demonstrate how the theory translates into practice.

Paper Authors

Orientation

Campus Plan

Helmholtz-Building – Overview
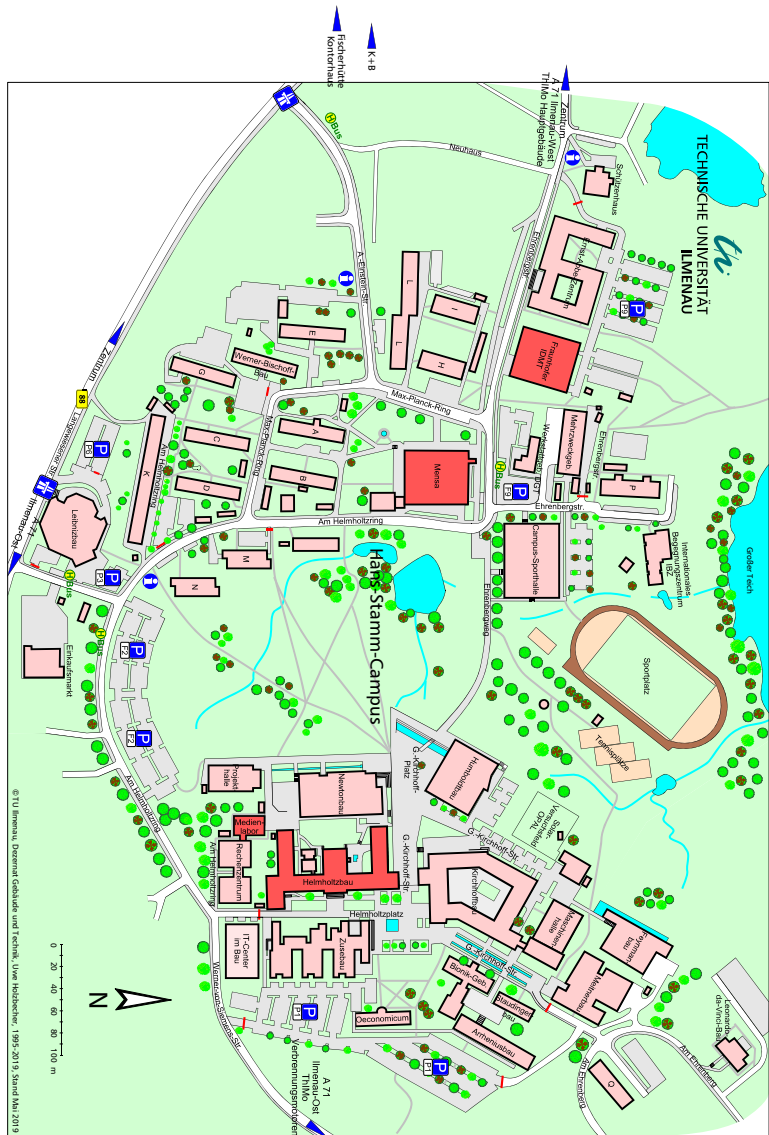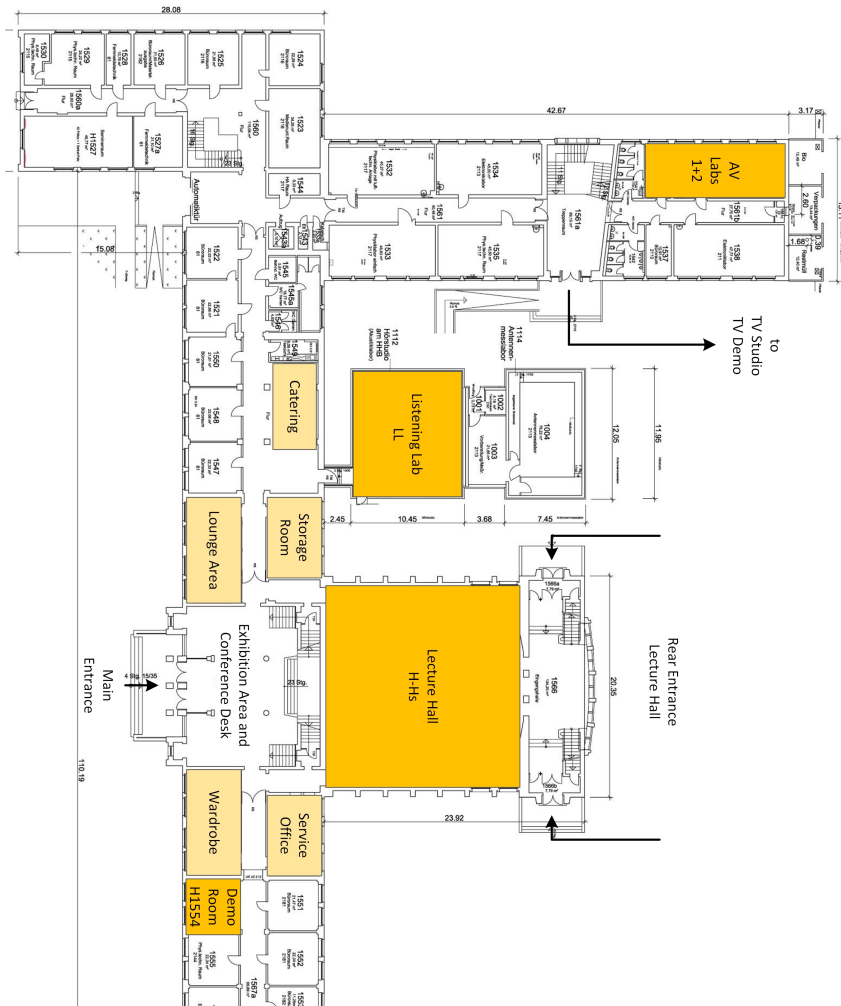
Helmholtz-Building – Ground level (Entrance hall)

lecture hall

lecture hall

conference desk

#3 Microtech Gefell

meeting area

main entrance
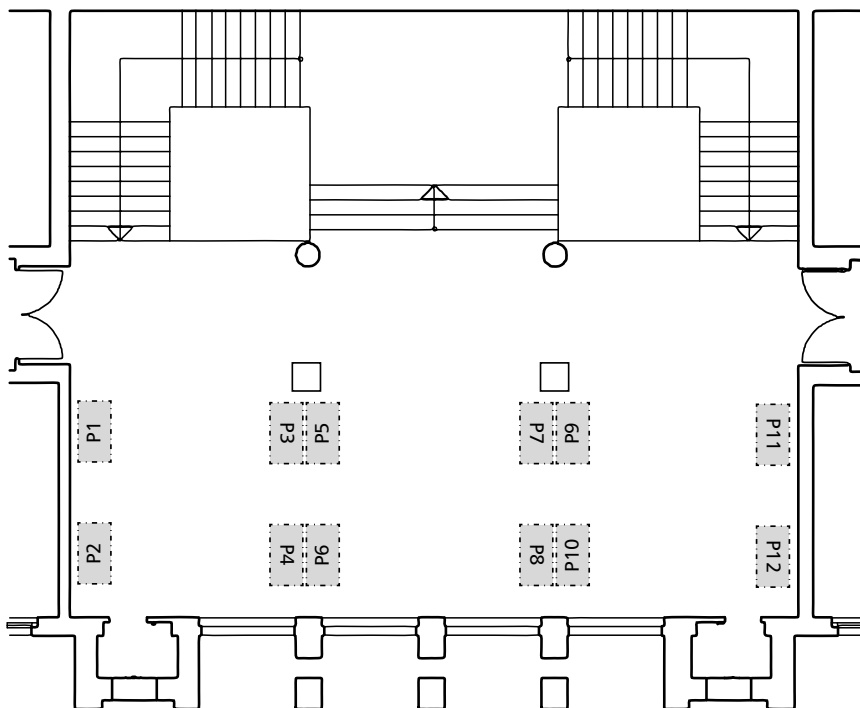
#2 ME Geithain
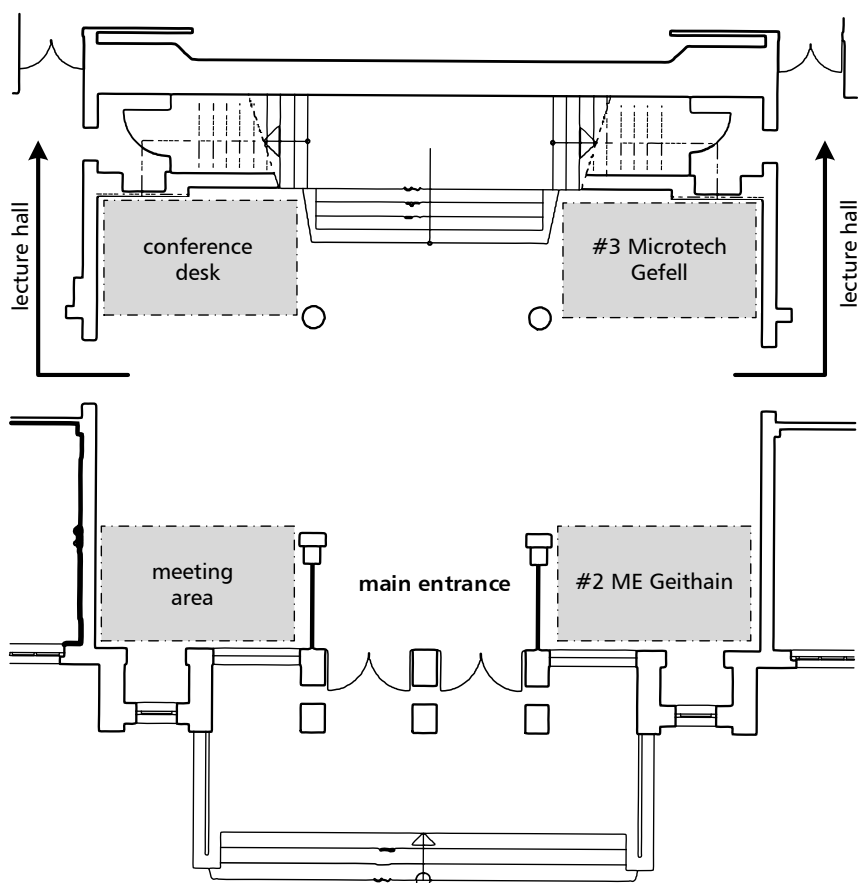
Helmholtz-Building – Level 1 (Exhibition)
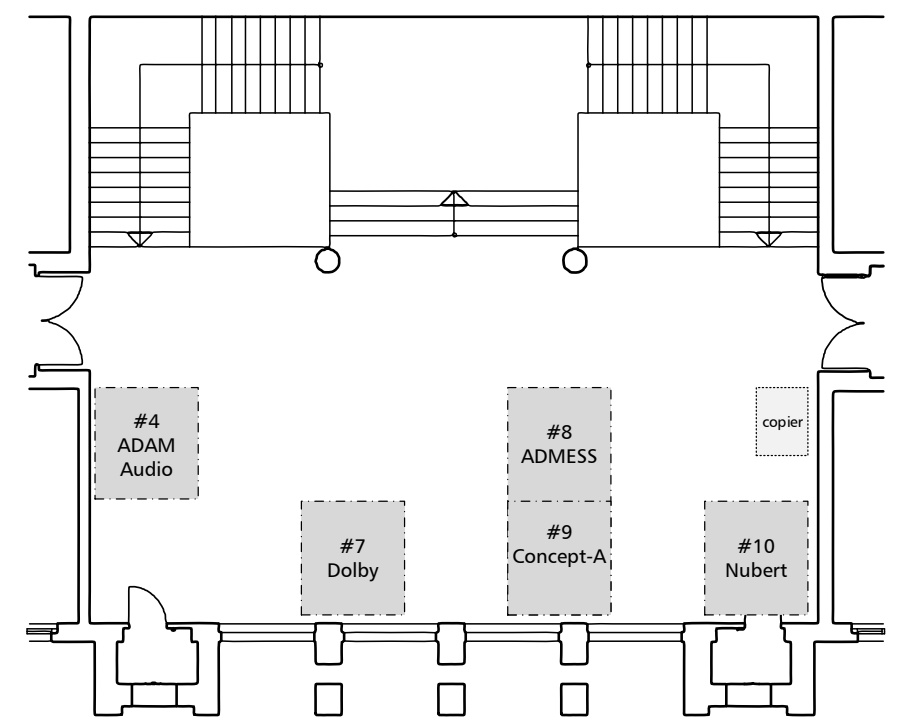
Helmholtz-Building – Level 2 (Posters)

Fraunhofer IDMT – Demos and Workshops

WaveStudio

3DP - 3D Prensentation room

Wardrobe

Main entrance

Seminar Room
SR-0.10/0.12

Entrance from
parking space

Bathroom
(Men)

Bathroom
(Women)

Air-Hockey
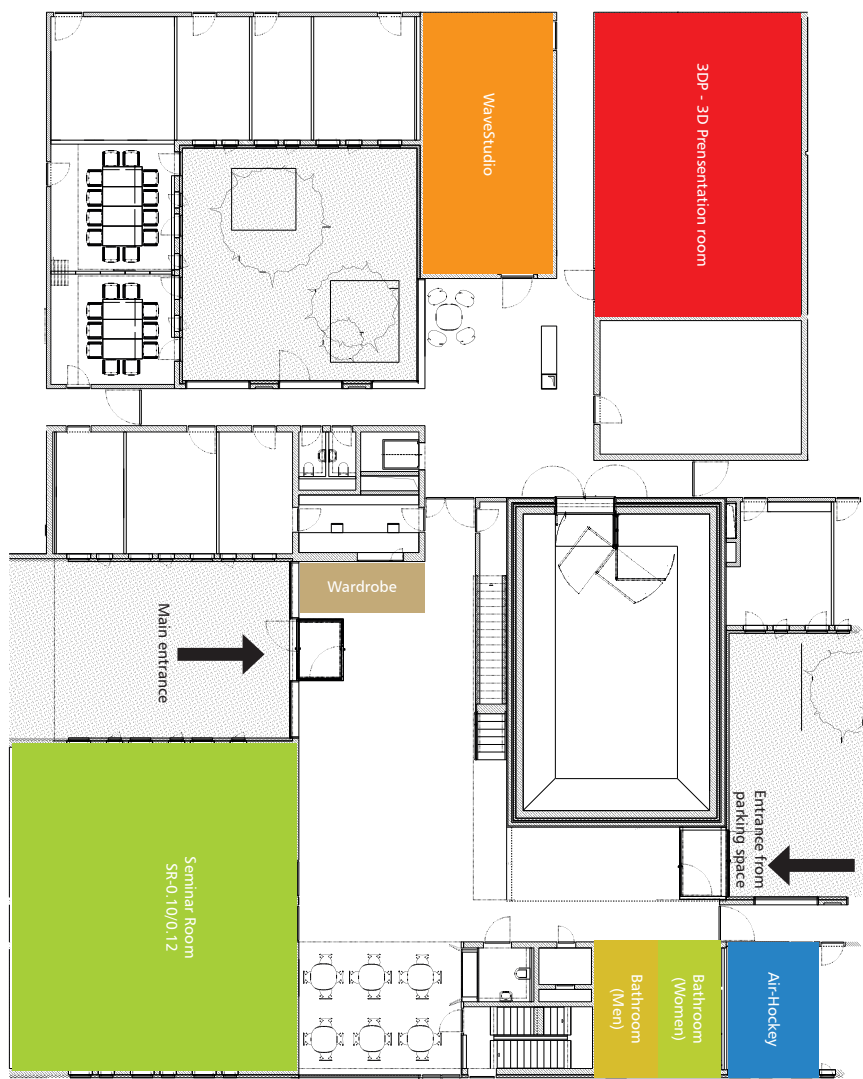
Imprint

Publisher

- ▷ Technische Universität Ilmenau
- ▷ Fakultät für Elektrotechnik und Informationstechnik
- ▷ Institut für Medientechnik
- ▷ PF 10 05 65
- ▷ 98684 Ilmenau

Conference Commitee

General Chairs

- ▷ Karlheinz Brandenburg
- ▷ Stephan Werner
- ▷ Monique Rodegast

Co-Chairs

- ▷ Stefani Renner
- ▷ Thomas Sporer

Technical Program Chairs

- ▷ Karlheinz Brandenburg
- ▷ Thomas Sporer
- ▷ Stephan Werner

Workshops and Demo Chairs

- ▷ Annika Neidhardt
- ▷ Florian Klein

Booklet

- ▷ Bernhard Fiedler
- ▷ Claudia Stirnat
- ▷ Stephan Werner
- ▷ Steve Göring

References

- ▷ Images (Frontcover)
  - ◇ Fraunhofer IDMT WaveStudio https://www.idmt.fraunhofer.de/de/business_units/acoustics/ equipment.html
  - ◇ TU Ilmenau TV Studio https://www.tu-ilmenau.de/de/mt/das-institut/laborerechnerpools/medienlabor-ii/

Sponsors and Supporters

facebook
Reality Labs

ADAM AUDIO
Adam Audio GmbH

ADMESS
Admess Elektronik GmbH

concept **A**
acoustics+architecture
concept-A GmbH

DOLBY
Dolby

Microtech Gefell GmbH

musikelectronic geithain GmbH
Musikelectronic Geithain GmbH

MCS
THÜRINGEN
Media & Communication Systems (MCS) GmbH Thüringen

mh
acoustics
mh acoustics

nubert
Nubert electronic GmbH

R E M I C
M I C R O P H O N E S
remic microphones ApS

Thüringer Aufbaubank
Die Förderbank.
Thüringer Aufbaubank

Hosts

Partners

German Acoustical Society

Audio Engineering Society