

IMMERSIVE AUDIO TECHNOLOGIES FOR VR APPLICATIONS

Adrian Murtaza

Senior Research Engineer

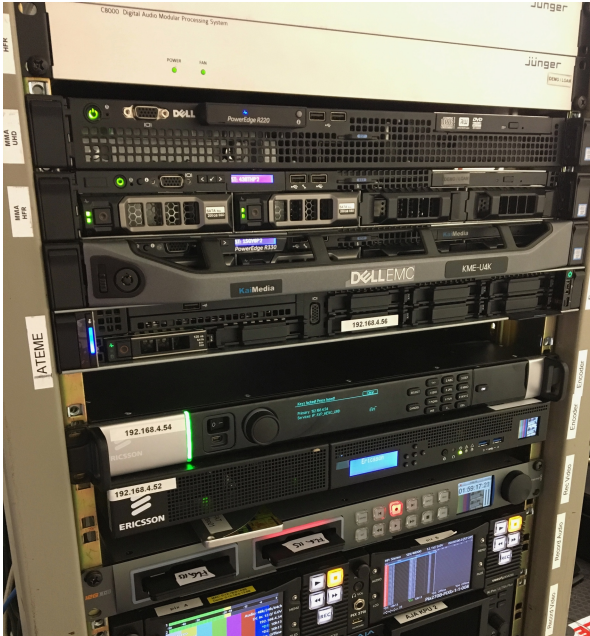
Technical Standards & Business Development

adrian.murtaza@iis.fraunhofer.de



MPEG-H Audio

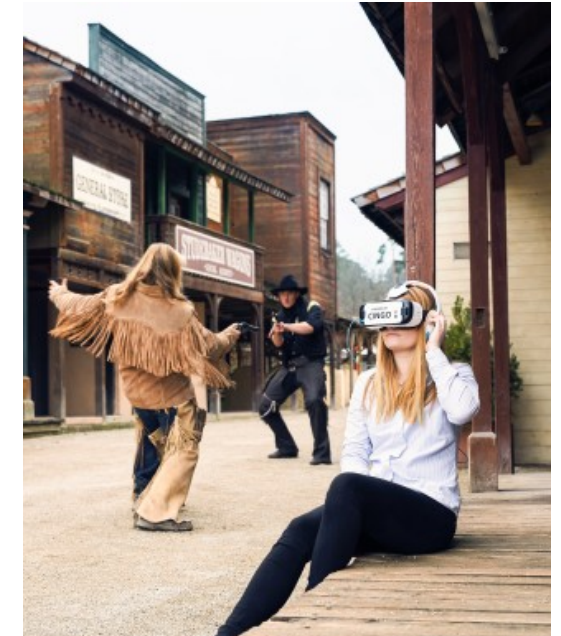
The Next Generation Audio Standard for Broadcast, OTT and VR/AR



Live MPEG-H Audio DVB-T2/S2
Transmission during French
Tennis Open



DVB DASH streaming on a set-top box running
Vewd Core and Vewd HbbTV Module with MPEG-H Audio

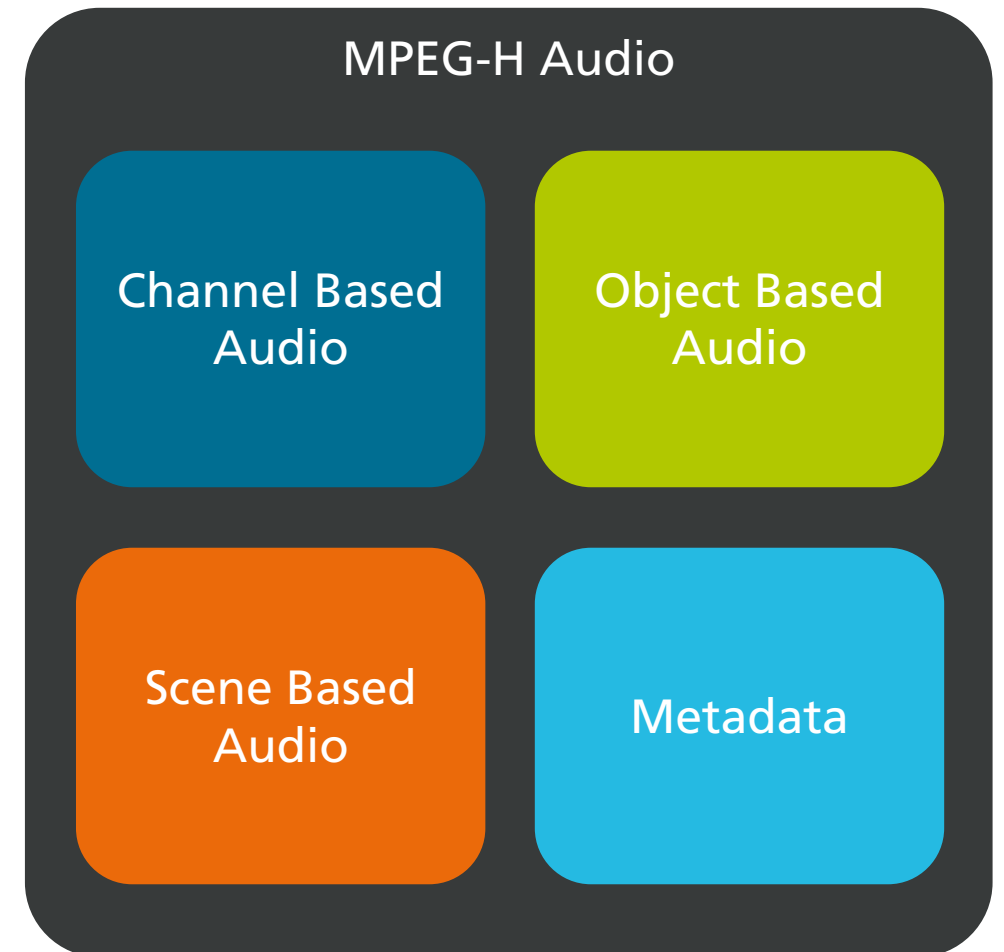


Fraunhofer Cingo® –
Binaural VR Audio Renderer
using MPEG-H Audio

MPEG-H Audio

Flexible Content Formats

- Content Formats:
 - Channel-based audio (e.g., 7.1+4H, 5.1+4H etc.)
 - Object-based audio (static and dynamic objects)
 - Higher Order Ambisonics (HOA including FOA support)
- Flexible combination:
 - From complete mixes
 - To any combination of channel signals, objects and HOA



MPEG-H Audio

Introduction and Features

Interactivity



Personalization of audio presentation

- Adaptation to the user's preferences/situation
- Turn the dialogue up or down
- Select your own version of the audio

Immersion



Compelling sound experience (sound comes from all directions)

- A viewer becomes part of the audience
- Conveniently delivered to consumer's homes with 3D soundbars

Universal Delivery



Play on any device with delivering best possible sound experience

- Home Theater
- Soundbars
- Mobile devices
- Headsets

MPEG-H Audio

Complete integrated audio solution

- MPEG-H Audio is more than a codec, it is a complete audio system
 - Integrated loudness and DRC solution
 - Integrated rendering and down-mix
- Flexibility on loudspeaker layouts
 - stereo, binaural, 5.1, 5.1+4H, 7.1+4H and more
- Several versions from one audio stream
 - Multi language, Dialogue Enhancement, various effects etc.
 - Total audio bitrate much lower compared to multiple full mixes

MPEG-H Audio

The Next Generation Audio Standard for Broadcast



The Digital Video Broadcasting (DVB) has included MPEG-H Audio in the DVB specification ETSI TS 101 154..



The Advanced Television Systems Committee has included MPEG-H Audio in its ATSC 3.0 suite of standards as ATSC Standard A/342-3.



The Telecommunications Technology Association (TTA) has selected MPEG-H Audio as the sole audio system for ATSC 3.0 in South Korea.



The Brazilian Association of Technical Standards (ABNT) has specified MPEG-H Audio into their data coding specification ABNT/CEE-085.



MPEG-H Audio regular service started in South Korea in May 2017.
First and currently only regular terrestrial UHDTV service worldwide using NGA.



The Consumer Technology Association has included MPEG-H Audio in its Recommended Practice for ATSC 3.0 Television Sets as CTA-CEB32.5.



SCTE/ISBE has included MPEG-H Audio in its coding and carriage suite of standards for cable television.

...

MPEG-H Audio in South Korea

Terrestrial UHDTV Service in South Korea

- **First and currently only regular** terrestrial UHDTV service worldwide using a **Next Generation Audio** Codec
- Regular service started in May 2017



Demo of terrestrial UHD broadcasting with MPEG-H 121st MPEG Meeting in Gwangju, Korea

- HEVC
- MPEG-H 3D Audio
- MMT
- DASH



SAMSUNG SAMSUNG RESEARCH AMERICA

• UHF Antenna

19

MPEG-H Audio

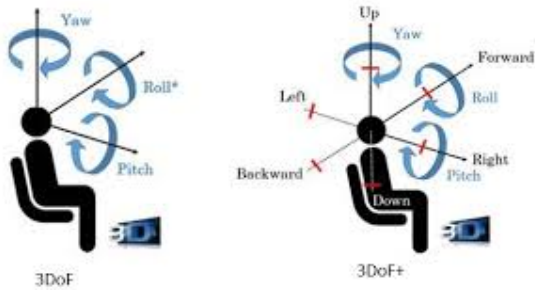
Next Generation Audio Standard for OTT and Hybrid delivery

- CTA WAVE has defined the CMAF Media Profile for MPEG-H Audio:
 - Based on ISO Base Media File Format
 - Fully interoperable with DVB DASH, HbbTV and DASH-IF
- MPEG-H Audio enables
 - seamless configuration changes (e.g., for add-insertion)
 - seamless adaptive bitrate switching
 - interactivity and personalization of content
 - Hybrid delivery, e.g., for a football game:
 - Main commentator on broadcast
 - Alternative commentators on broadband



MPEG-H Audio

Next Generation Audio Standard for VR/AR



MPEG OMAF

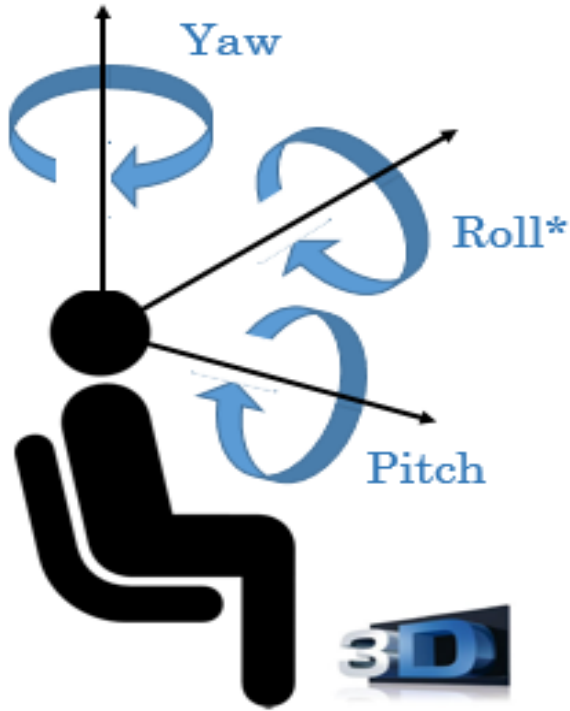


INDUSTRY FORUM



MPEG-H Audio

OMAF 3D Audio Baseline Profile



- MPEG has finalized the Omnidirectional Media Format (OMAF) standard for delivery of 3DoF media content including audio and video media profiles and presentation profiles.
- Two audio media profiles were specified:
 - **OMAF 3D audio baseline profile (MPEG-H 3D Audio)**
 - OMAF 2D audio legacy profile (AAC)
- The **OMAF 3D audio baseline profile** is the only audio profile included in the two presentation profiles:
 - OMAF viewport-independent baseline presentation profile
 - OMAF viewport-dependent baseline presentation profile

MPEG-H Audio - Next Generation Audio Standard for VR/AR

OMAF 3D Audio Baseline Profile included in VR-IF



- The VR Industry Forum (VRIF) has published its first set of VR Industry Guidelines, which includes topics relevant for VR content production, distribution, security and consumption
 - MPEG-H Audio has been included as the only Audio Media Profile.

Table 4: Overview of OMAF media profiles for audio

Media Profile	Codec	Profile	Level	Max. Sampling Rate	Brand	Section
3D Audio Baseline	MPEG-H 3D Audio	Low Complexity	1, 2 or 3	48 kHz	'oab1'	10.2.2

<https://www.vr-if.org/wp-content/uploads/vrif2018.018.09-clean.pdf>

MPEG-H Audio - Next Generation Audio Standard for VR/AR

OMAF 3D Audio Baseline Profile included in 3GPP



- 3GPP has approved a technical specification for 360° video streaming in its work item “**VRStream**” , it will be included in 3GPP release 15, the most recent 3GPP release and the first one to specify 5G.
- The recommended audio format for use in VRStream services is the MPEG OMAF 3D Audio Baseline Profile, based on MPEG-H Audio.
- The VRStream specification defines audio and video media profiles for mobile operators that want to offer fully immersive 360° video services cost-effectively on a broad range of mobile devices connected over the most advanced mobile networks, such as upcoming 5G networks.

VRStream Specification: <http://www.3gpp.org/DynaReport/WiSpec--770025.htm>

MPEG-H Audio - Next Generation Audio Standard for VR/AR

OMAF 3D Audio Baseline Profile included in 3GPP



- MPEG OMAF 3D Audio Baseline Profile (MPEG-H Audio) is recommended as the only audio format for VRStream services, as part of 3GPP's:
 - Packet-switched Streaming Service (PSS) and
 - Multimedia Broadcast Multicast Service (MBMS).

VRStream Specification: <http://www.3gpp.org/DynaReport/WiSpec--770025.htm>

MPEG-H Audio

Next Generation Audio Standard for VR/AR

- VR demos based on the MPEG OMAF, 3GPP and VRIF specifications using the OMAF 3D Audio Baseline Profile at:
 - **Fraunhofer booth 8.B80**
 - **VRIF booth 8.F19**

■ **Live360TV at booth 8.F27**

- a research collaborative project, led by ATEME together with several French high tech companies and research institutes
- end to end chain from capture to devices with an immersive video and audio top quality.
- The demonstration will rely on
 - an OTT distribution of a high quality 360 video TV with advanced audio and video consumption techniques for bandwidth reduction based on “tiling”.
 - a portable VR playback from Telecom ParisTech.

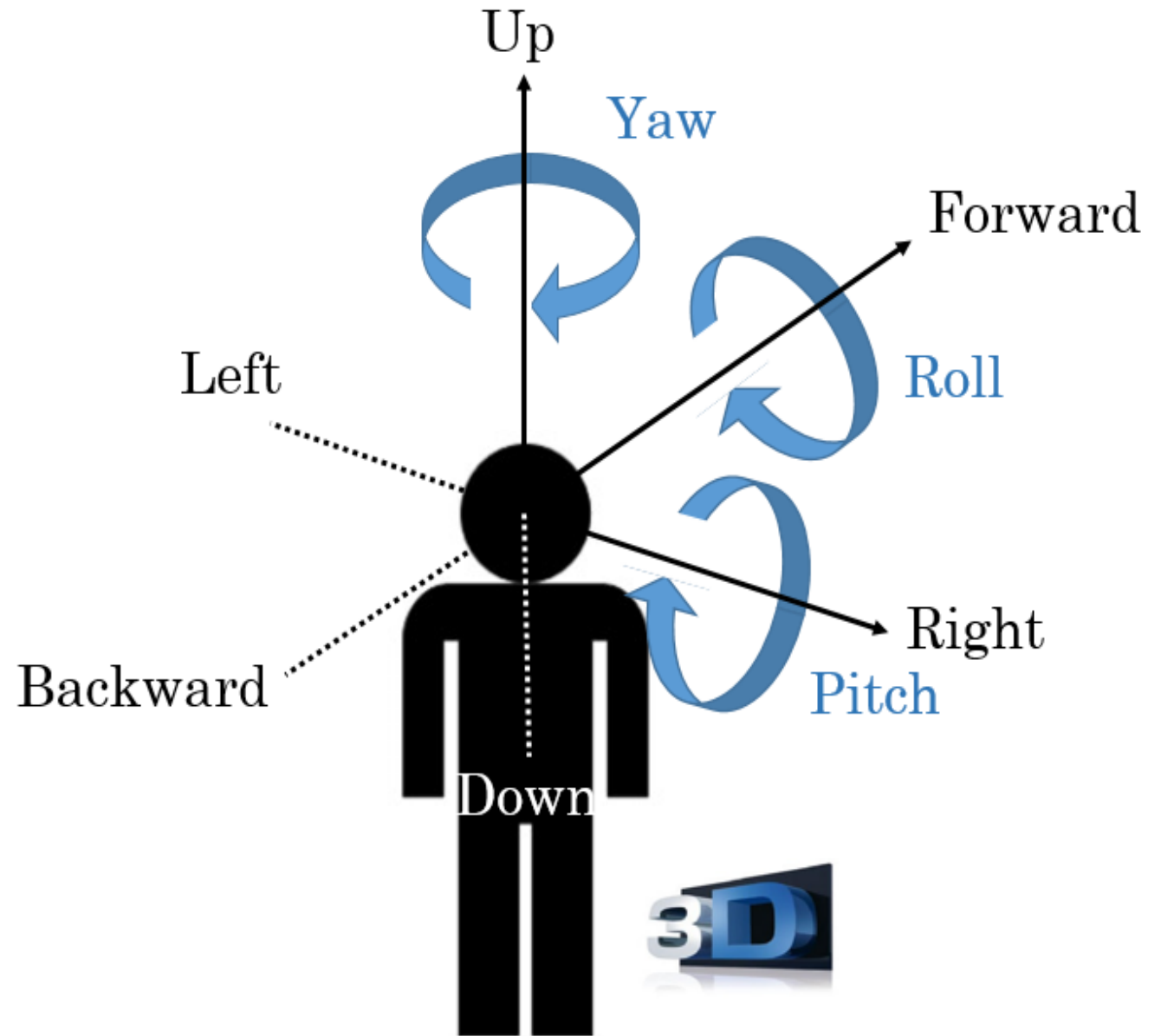
From MPEG-H to MPEG-I Audio

Immersive Audio for 6DoF

MPEG-I Audio

Immersive Audio for 6DoF

- MPEG Audio Subgroup has started the work on the next audio standard for 6DoF:
 - Requirements specification and architecture definition ongoing
 - New Test Methodology Evaluation under development
 - Call For Proposals probably in 2019 and final standard available by 2021!



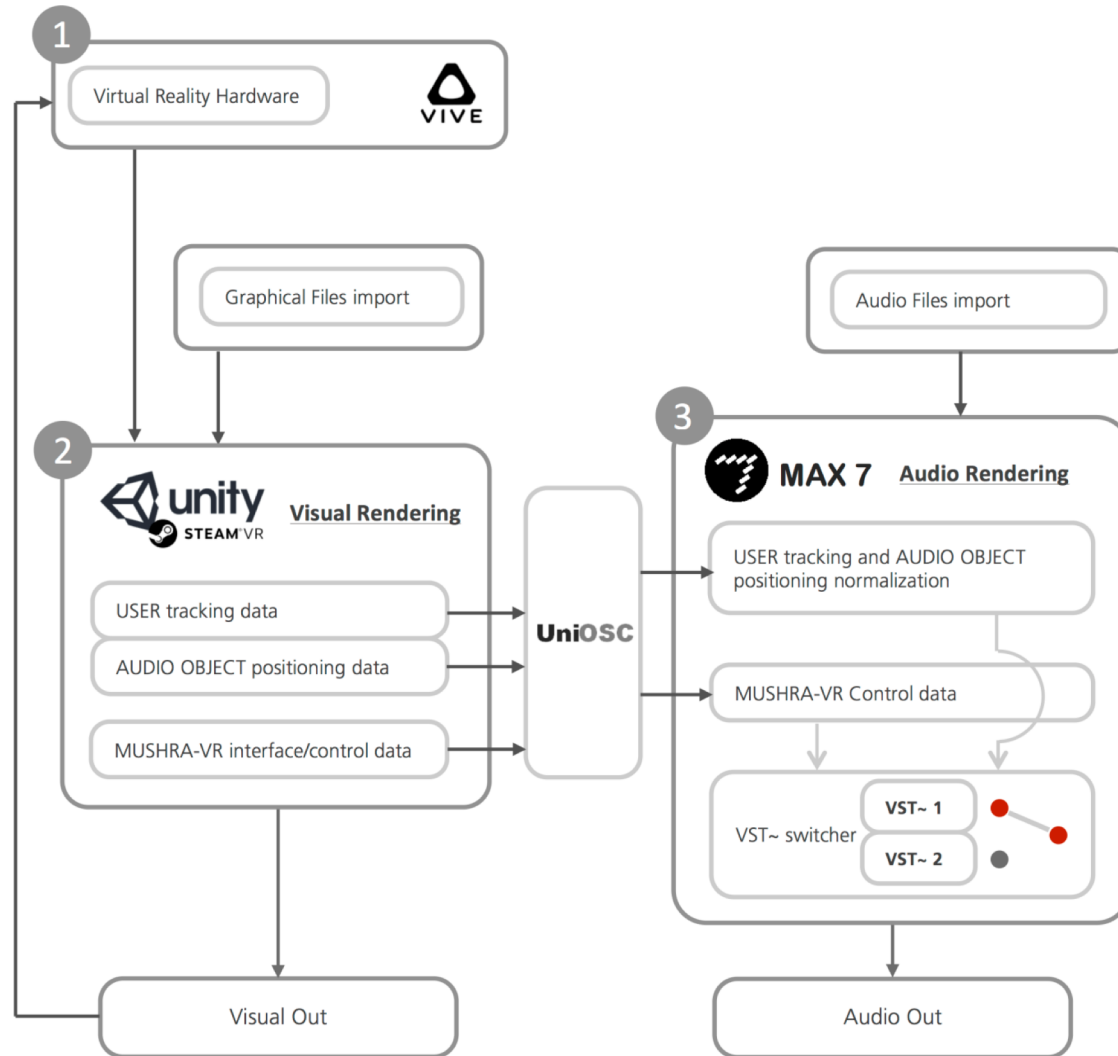
MPEG-I Audio

Immersive Audio for 6DoF

- **What is planned to be standardized in MPEG-I Audio**
 - **Audio Codec: MPEG-H Audio!** (as specified in OMAF/3GPP/VRIF)
 - MPEG Audio Subgroup will not work on improving the audio compression!
 - the MPEG-H Audio system is already fulfilling all requirements for MPEG-I Audio
 - **New Audio 6DoF Renderer**
 - 6DoF Rendering Technology
 - Metadata for MPEG-I 6 DoF rendering
 - Metadata for environmental acoustic effects, including scene graphs describing physical environment

MPEG-I Part 4: Immersive Audio

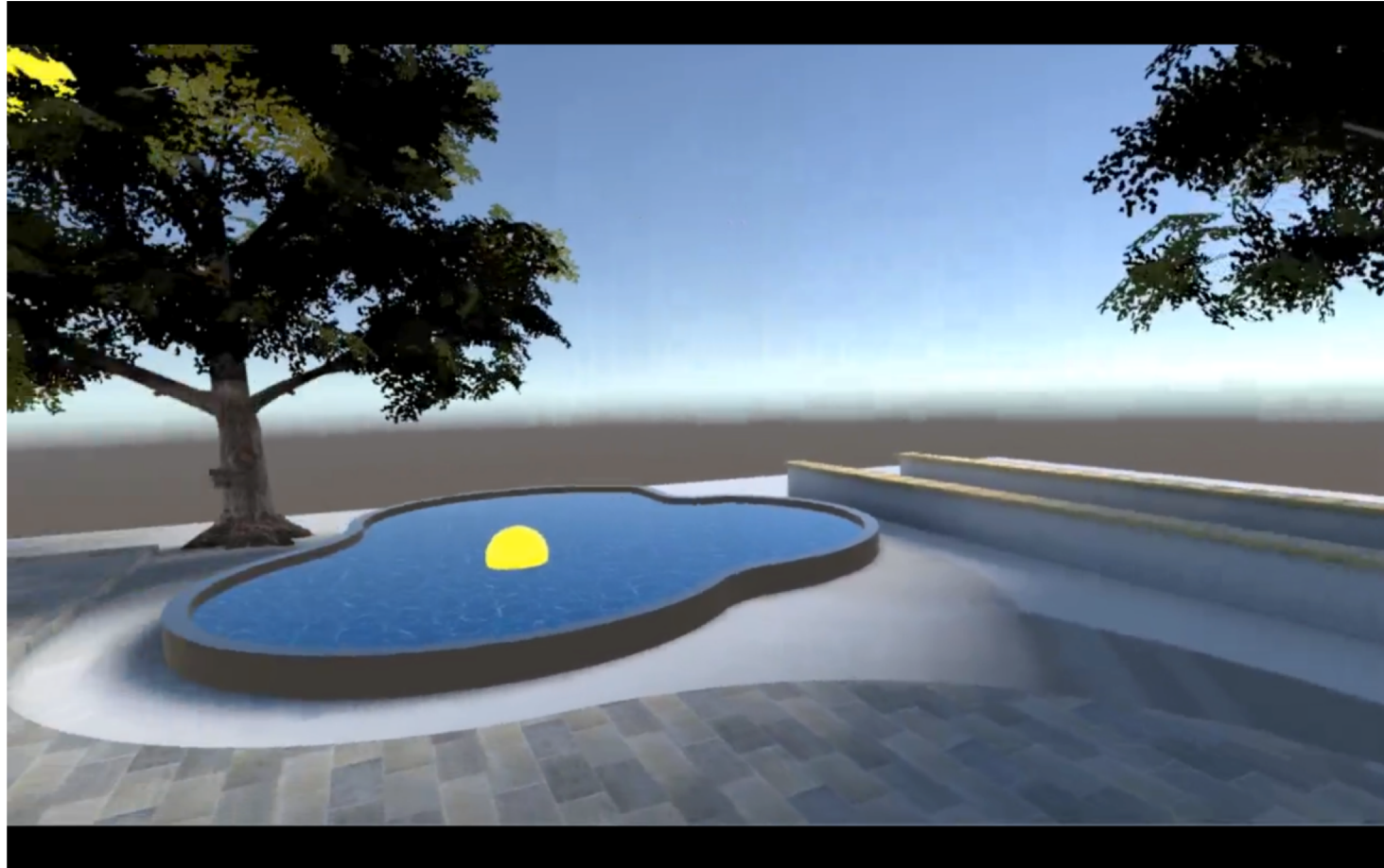
Common Evaluation Platform



- HTC VIVE
- Visual Rendering by Unity
- Audio Rendering by MAX7
- Renderers as VST plugins
- PC/Windows

MPEG-I Part 4: Immersive Audio

Example of a Scene in the Evaluation Platform



(Computer graphics; yellow sphere could be audio object)

MPEG-I Part 4: Immersive Audio

MUSHRA-Style Testing (“MUSHRA-VR”)



Virtual MUSHRA-style panel pops up

MPEG-I Part 4: Immersive Audio

Architecture and requirements

- General Requirements
- Requirements on Metadata
- Requirements on Rendering
- Flexible Interfaces and extensions (allow future/external tools and plugins)
- Support for Social VR use cases
- Interoperability between 3DoF and 6DoF platforms
-

THANK YOU!

Adrian Murtaza

Senior Research Engineer

Technical Standards & Business Development

adrian.murtaza@iis.fraunhofer.de

<http://www.iis.fraunhofer.de/amm/>