# **ACCESSIBILITY FOR NEXT GENERATION AUDIO**

## ADLAB PRO 2019

Christian Simon Fraunhofer Institute for Integrated Circuits (IIS) March 20, 2019, Barcelona



1

## Agenda

- Introduction
- What is Object Based Audio?
- The MPEG-H Audio Codec
- Accessibility features
- MPEG-H Demo
- Q & A



## Introduction



## Fraunhofer-Gesellschaft

# The largest organization for applied research in Europe

- 69 institutes and research units
- Offices worldwide
- 24.500 staff
- \$2.1 billion annual research budget totaling





## Joseph von Fraunhofer (1787 – 1826)



#### Researcher

 Discovery of the "Fraunhofer lines" in the solar spectrum

#### Inventor

 New methods for processing lenses

#### Entrepreneur

 Director and partner in a glassworks





## Fraunhofer Institute for Integrated Circuits (IIS)

- Audio and Media Technologies Division
- Inventor of mp3
- In billions of devices with AAC and its successors
- Close partnership with Friedrich Alexander University Erlangen-Nürnberg



## About me

- Tonmeister from Film University Babelsberg, Germany
- Strong interest in accessibility for AV media:
  - Speech intelligibility optimization
  - Audio description production

## Since 2016 at the Fraunhofer IIS

- Soundlab group
- Focus on Object Based Audio



## **Motivation**

#### To learn:

- Actual requirements for audio description productions?
- Needs for future production workflows ?

#### To present:

- What is Object Based Audio?
- What can it do for accessibility in AV media?
- How does it work ?



What is Object Based Audio?





## **Current production**

- A new mix is done for every audio format, for example stereo and 5.1
- Additional deliveries are needed for audio description and clean audio
- All mixes need to be broadcasted or streamed in parallel



## **Object Based Audio**

- Audio components are delivered separately to the playback device
- Audio is rendered at playback device
- It adds metadata to the audio, for example
  - Presets, e.g. for audio description
  - Loudness of components
  - Gain and position metadata
  - Labels for different languages



The MPEG-H Audio Codec



## MPEG-H Audio Flexible Content Formats

- Content Formats:
  - Channel signals
  - Objects
  - Scene-based (Ambisonics)
- Flexible combination:
  - From complete mixes
  - to any combination of channel signals, objects and HOA





## **MPEG-H** features

- Interactivity
- Immersion
- Universal delivery



## **Features: Interactivity**

#### Personalization of audio presentation

- Adaption for the user's preference or situation
- Change the mix, for example:
  - select different sound presets
  - turn the dialog up and down
  - change positions of sound components



## **Features: Immersion**

#### Sound from all directions

- Listener becomes part of the audience
- Conveniently delivered to consumer's homes with 3D soundbars
- Binaural over headphones



## **Features: Universal delivery**

- Play on any device with delivering the best possible sound experience
  - Home theater
  - Soundbars
  - Mobile devices
  - Headsets



## MPEG-H Audio Standardization

#### **MPEG-H Audio – A single solution**

- Broadcast
- Streaming
- VR/AR Applications





## 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
  - Currently available in the Seoul metropolitan area and other areas
  - 2020 nationwide service





# MPEG-H Trials in Europe 2018

- Eurovision Song Contest
  - MPEG-H Immersive Sound Live Production
- French Tennis Open
  - Broadcast via Satellite and DVB-T
- European Athletics Championships
  - EBU Ultra High Definition Trial





Accessibility features



## Audio description production with Object Based Audio

- Production workflow stays mostly the same
- Instead of a new channel based mix for the AD, the mix is saved by metadata
- Film mix and AD are not mixed together
- The voice over stays separately as audio object
- The volume automation is stored as metadata



## Audio data and metadata





## **Advantages of MPEG-H AD features 1**

- For the user:
  - Audio description is part of the regular broadcast or stream
  - Audio description is automatically played if available
  - Interactivity for level and position of the audio description speaker
  - Audio description users can also enjoy multichannel mixes



## **Advantages of MPEG-H AD features 2**

- For the broadcaster:
  - Lower bandwidth, as the AD is only a mono track + metadata
  - One stream includes all audio signals, no second file needed
  - All channel-formats accessible



## Speech intelligibility optimization with Object Based Audio

- Two different use cases:
  - Speech is available separately
    - New production
    - Archive material with available separate voice over
  - Speech is not available separately
    - Most archive material



## If speech is available separately

- Speech can directly be authored a object
- A Dialog+ preset can be created with larger SNR
- Gain interactivity can be provided for more personalization



## If speech is not available separately

- Speech can be extracted from the mix by dialog separation
- The algorithms for this task are improving constantly
- The extracted speech can be authored as object like described on last page



## **Production tools**

## There are tools for postproduction and live operation

- Scene creation, monitoring, authoring
- Metadata as xml or "Control Track"

### Post production:

- Standalone: MPEG-H Authoring Tool, adds metadata to existing content
- Plugin: MPEG-H Plugin, adds metadata in a Digital Audio Workstation

https://www.iis.fraunhofer.de/en/ff/amm/dl/software/mhapi.html



## Production Tool Example:

## **MPEG-H Plugin**

Track Preset Auto   AUTHORING e <factory default=""> E BYPASS   MPEG-H Authoring Plugin - + E COMPARE SAFE Native</factory>						
Fraunhofer IIS Components Prese	ts	Monitoring	Loudness	Export	Settings About	
+ Ch + Obj + SwG Remove	🔲 Component prop					
Mix Complete Main   Unknown   2.0	Description AD english	Indicato	or color	1		
AD german Audio Description   German   Dynamic Object	Position Fix	ked • Switchg	roup			
AD english Audio Description   English   Dynamic Object	Content Kind Audio Description	Content	: Language			
					🞾 Input routing	
	Gain Min Max OdB OdB	Azimuth Min Max O° O°	Elevation Min Max 0° 0°		Subtrack inputs (DAW) 4-4 (1 channel)	



## **MPEG-H** Demo



**Questions & Answers** 



## **THANK YOU!**

Christian Simon christian.simon@iis.fraunhofer.de

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

