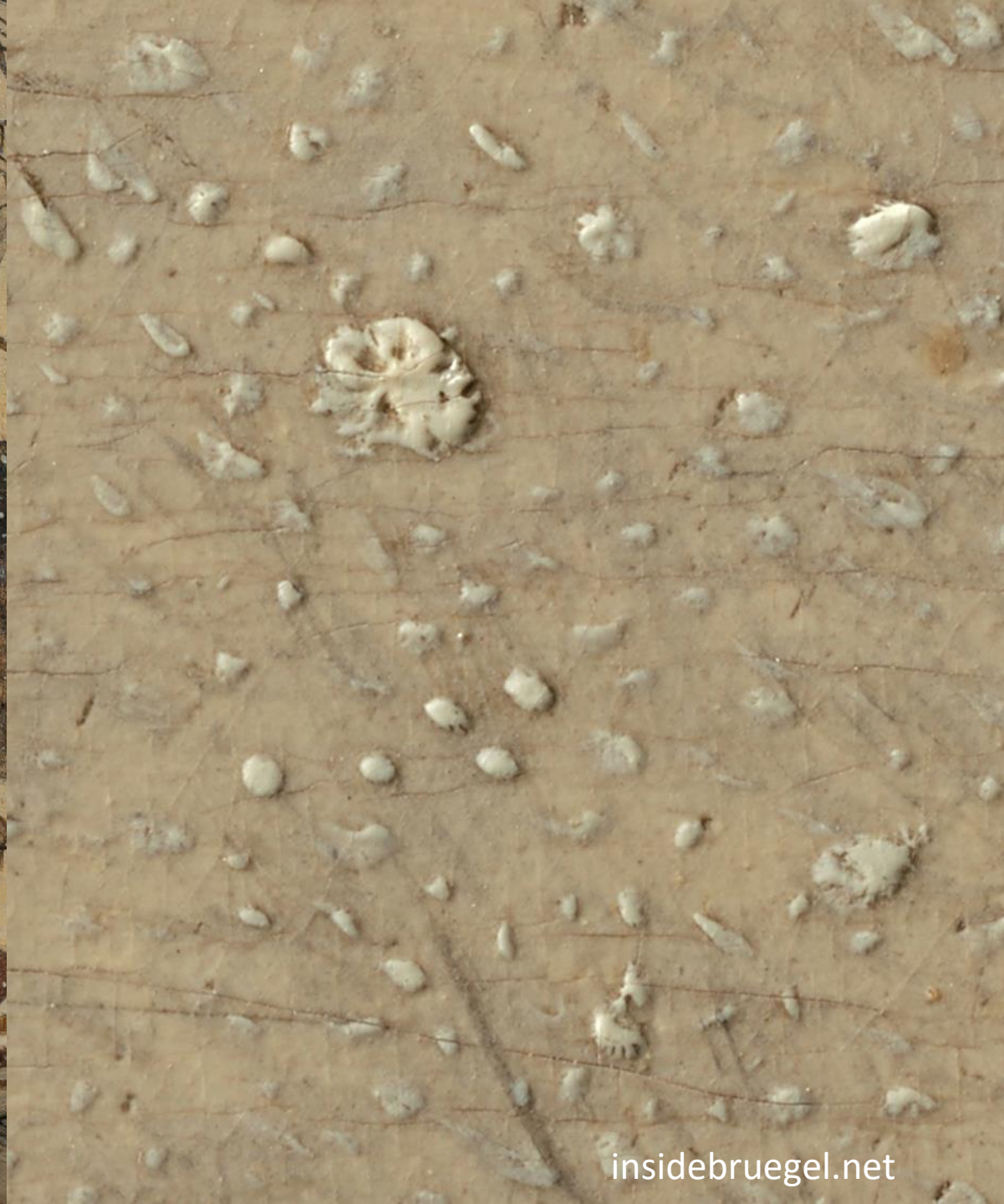
Standardisation activities for 3D modalities

3D in Cultural Heritage, Rome June 6th 2023

Frederik Temmermans

imec / Vrije Universiteit Brussel

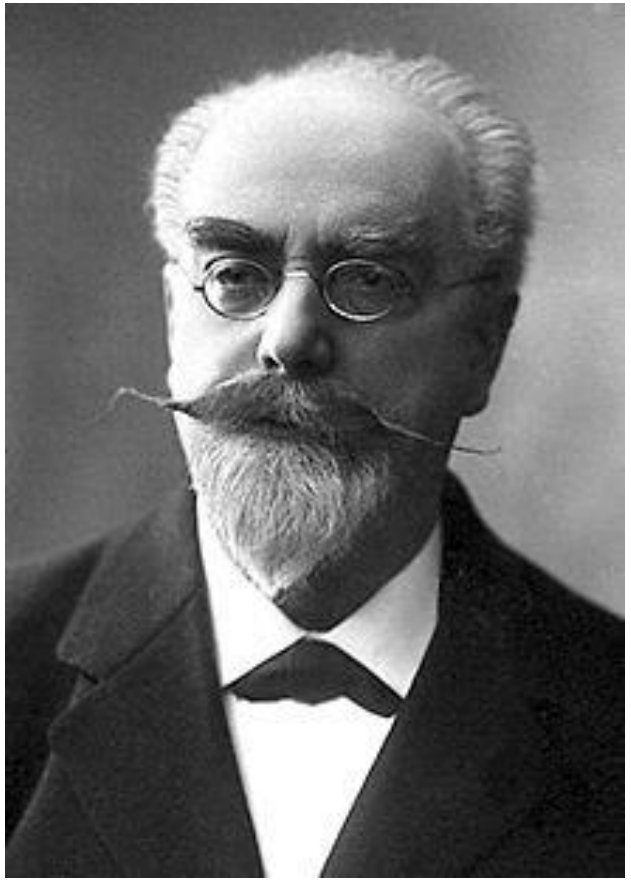


The most perfect photograph currently shows only one aspect of reality; it reduces to a unique image fixed on a plane, as a drawing or a painting would be traced by hand.

Can we ask photography to render all the richness that the direct view of an object offers?

”

Gabriel Lippmann



Gabriel Lippmann (1845 – 1921)
Lippmann is remembered as the inventor of a method for reproducing colors by photography, based on the interference phenomenon.
1908 Nobel Prize in Physics

Content quality

Quality of experience

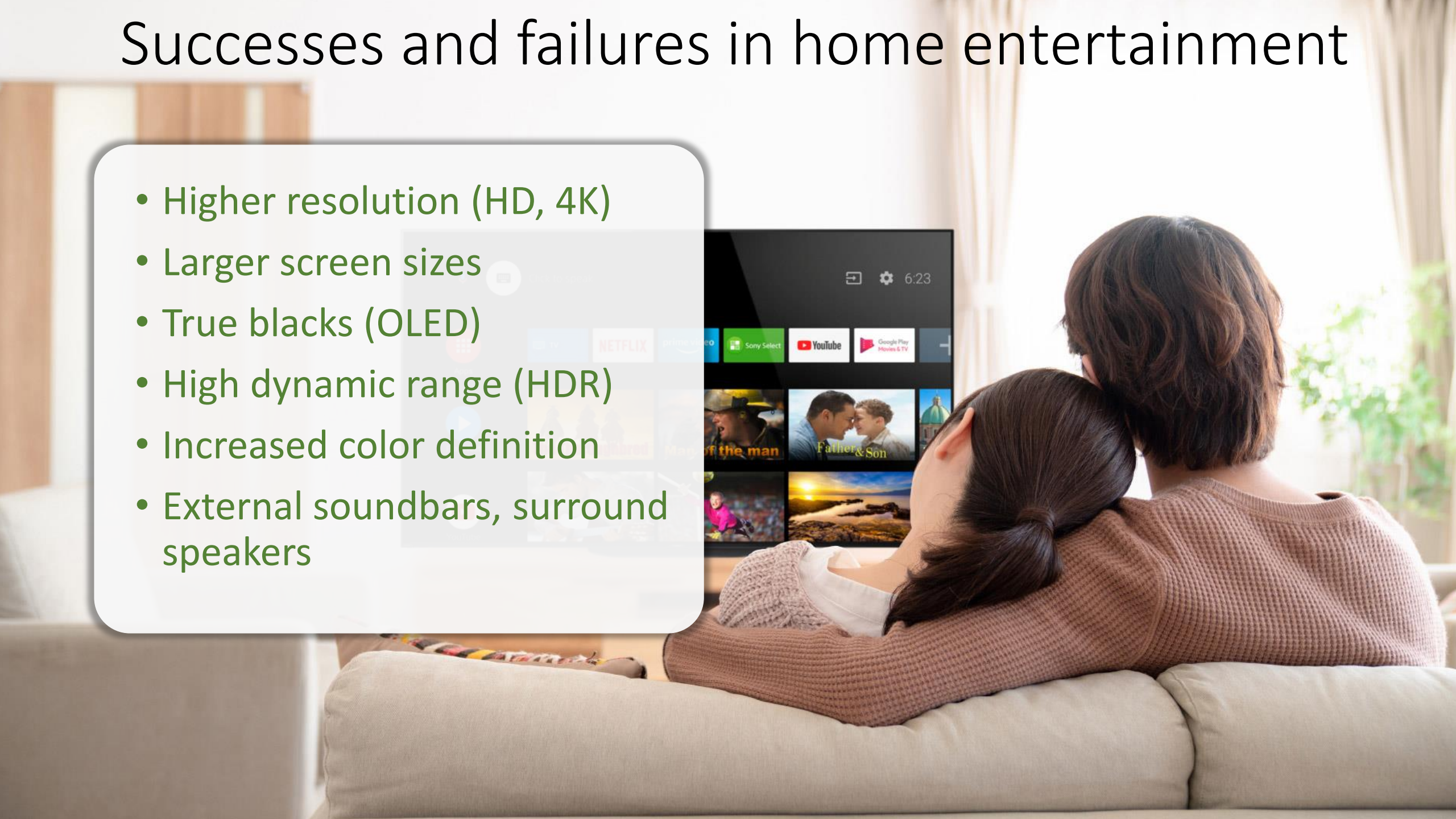
Interoperability

Content quality

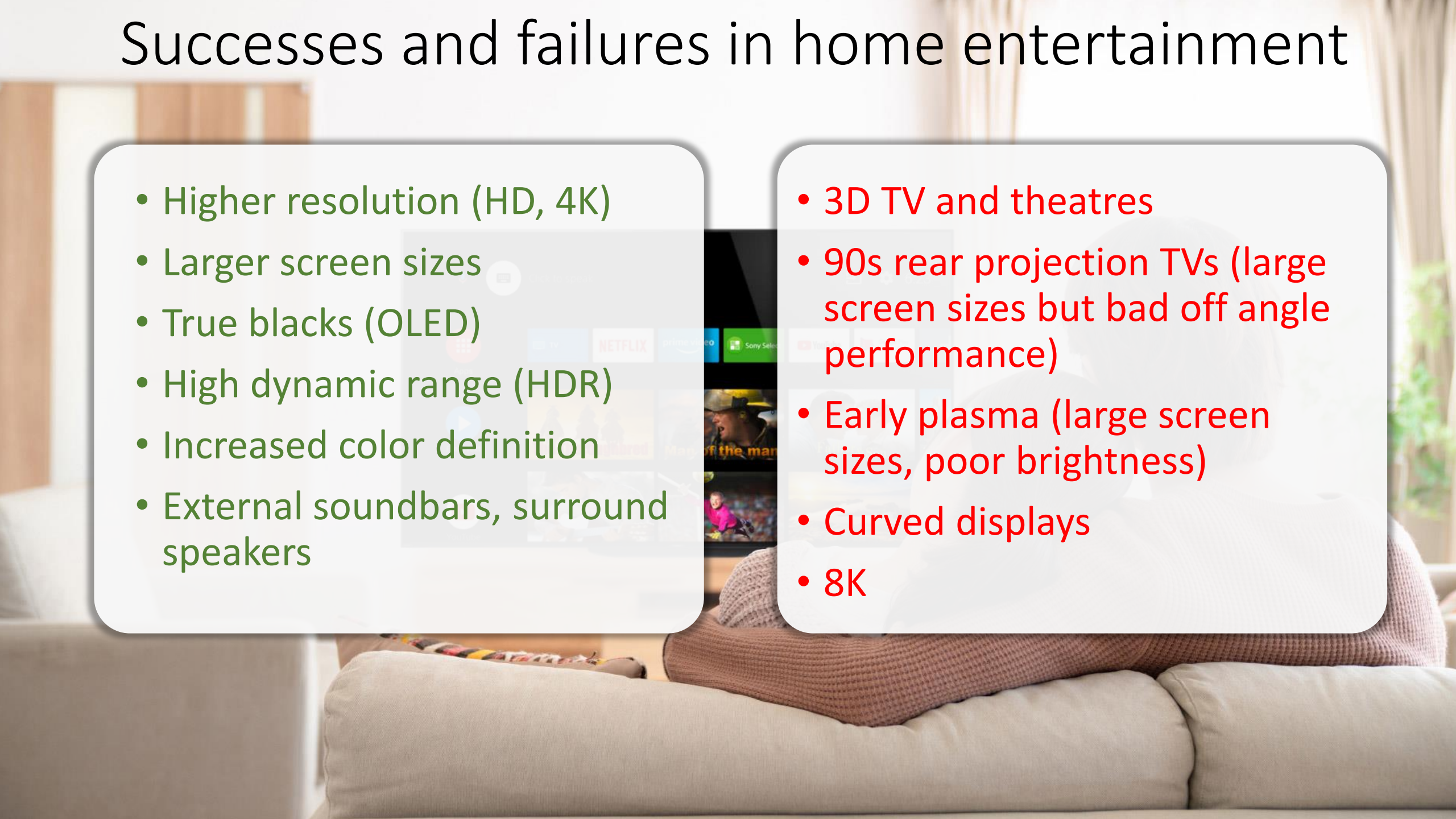
Successes and failures in home entertainment

Successes and failures in home entertainment

- Higher resolution (HD, 4K)
- Larger screen sizes
- True blacks (OLED)
- High dynamic range (HDR)
- Increased color definition
- External soundbars, surround speakers



Successes and failures in home entertainment



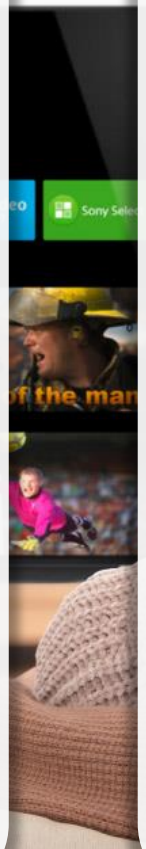
- Higher resolution (HD, 4K)
- Larger screen sizes
- True blacks (OLED)
- High dynamic range (HDR)
- Increased color definition
- External soundbars, surround speakers

- 3D TV and theatres
- 90s rear projection TVs (large screen sizes but bad off angle performance)
- Early plasma (large screen sizes, poor brightness)
- Curved displays
- 8K

Successes and failures in home entertainment

- Higher resolution (HD, 4K)
- Larger screen sizes
- True blacks (OLED)
- High dynamic range (HDR)
- Increased color definition
- External soundbars, surround speakers

Improved quality

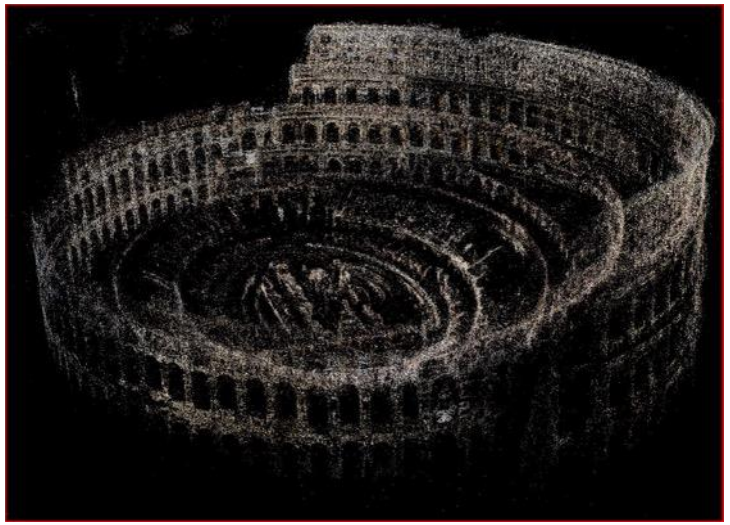


- 3D TV and theatres
- 90s rear projection TVs (large screen sizes but bad off angle performance)
- Early plasma (large screen sizes, poor brightness)
- Curved displays
- 8K

New functionality
at the cost of quality

Maintaining visual quality in 3D is hard

- Transition from 2D to 2D often entails a step back in visual quality:
  - Stereo vision reduces spacial resolution by a factor 2
  - 360 object captures have a significantly lower resolution in the 3rd dimension
  - Point clouds and meshes are sparse compared to 2d photography
- Restrictions apply to capturing, representation and rendering

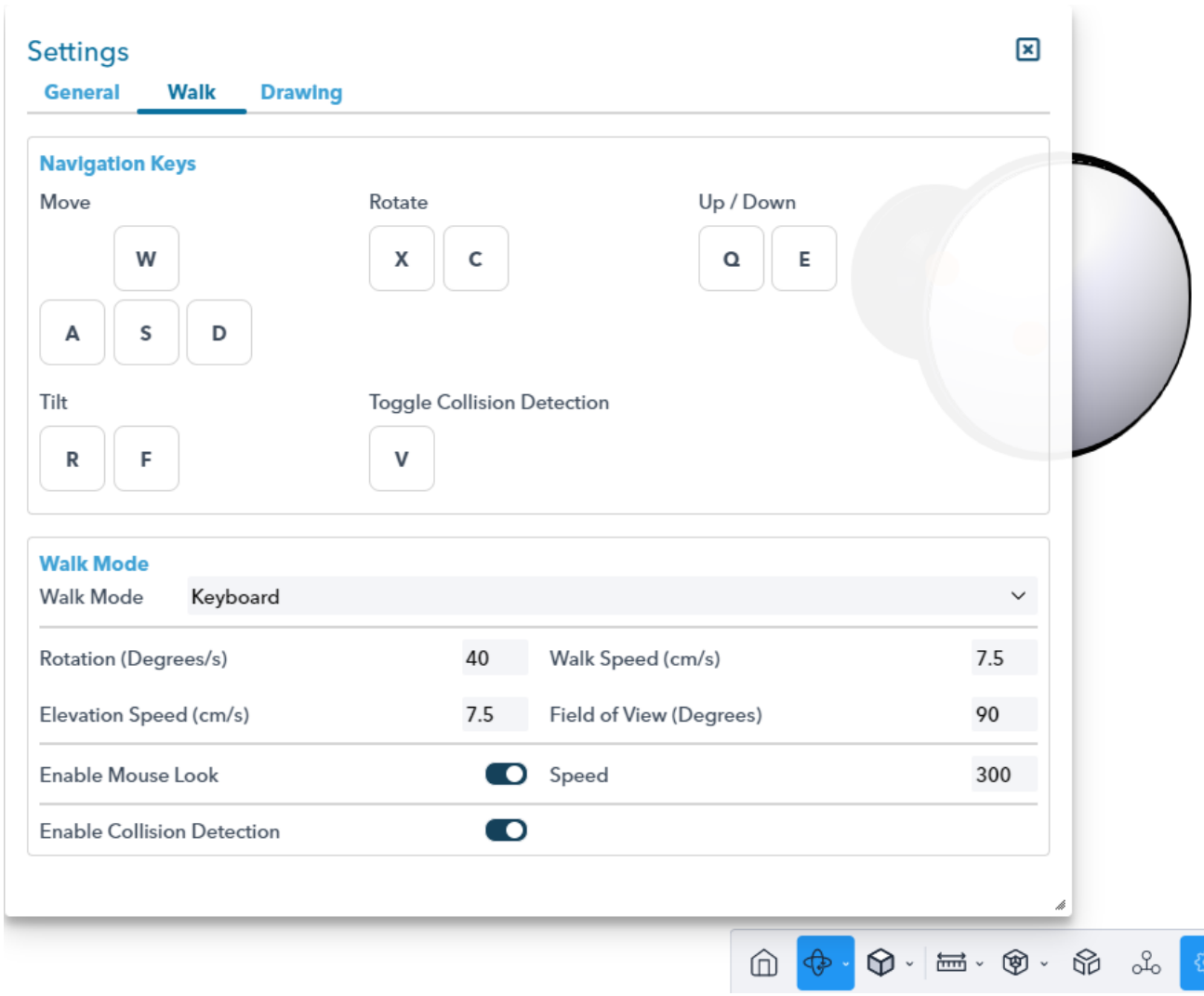


Maintaining visual quality in 3D is hard

- High standards for capturing
- Representation formats that can efficiently cope with increased data
- Big steps to be made in rendering

Quality of experience

Complicated graphical user interfaces



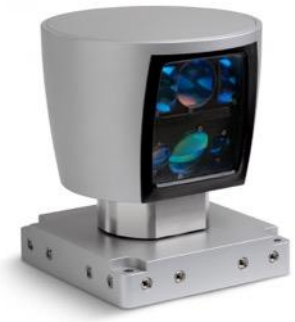
Aditional hardware requirements

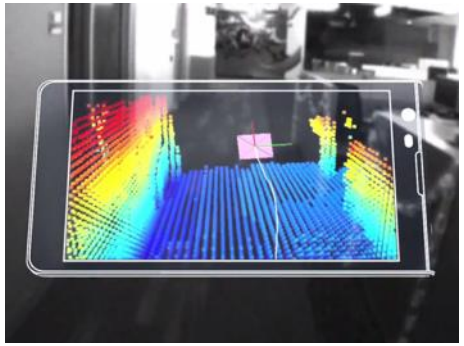
Reference field limitations

Interoperatbility

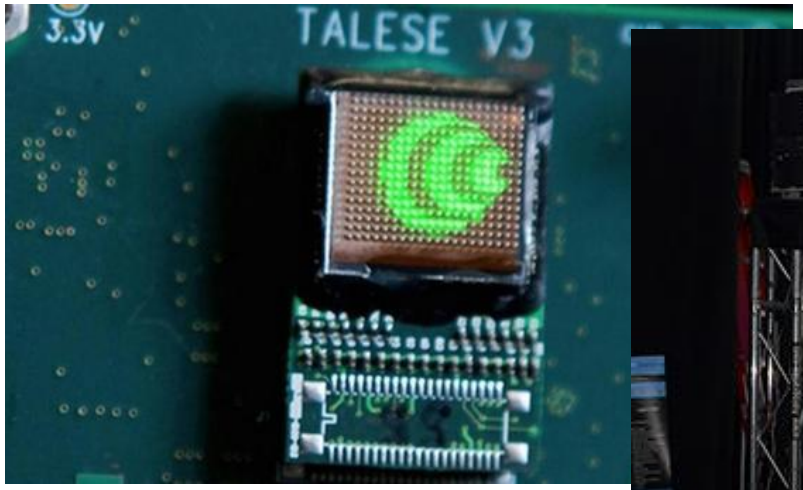

Baltimore Street, Hanover, Pa.
ITU
ISO
IEC
W3C®
IEEE

Capturing

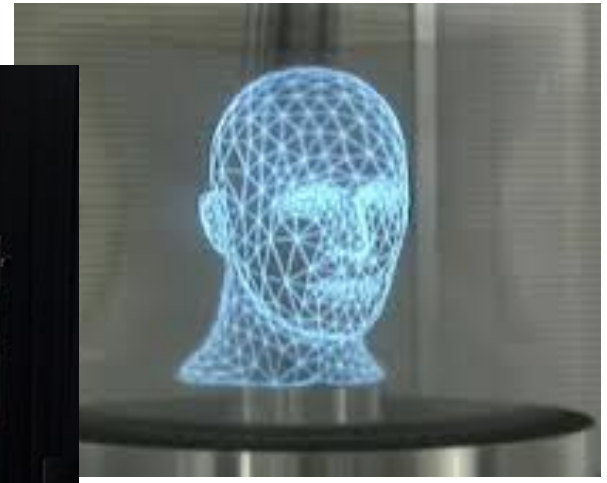


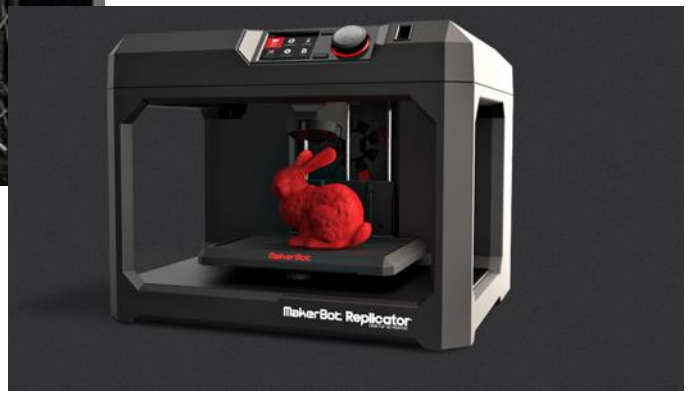


Rendering


3.3v
TALESE V3

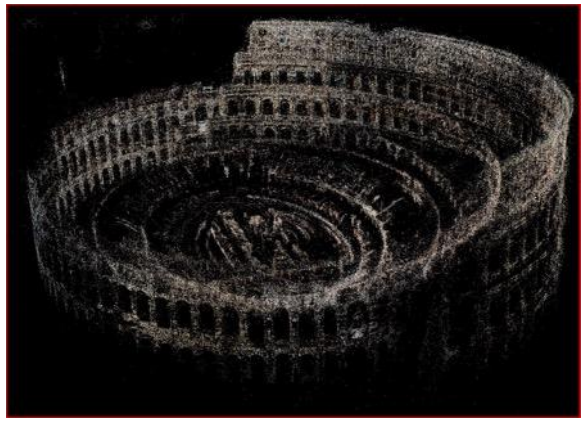

HOLOVIZIO™ C80 Glasses-free 3D Cinema System by HOLOGRAFIKA



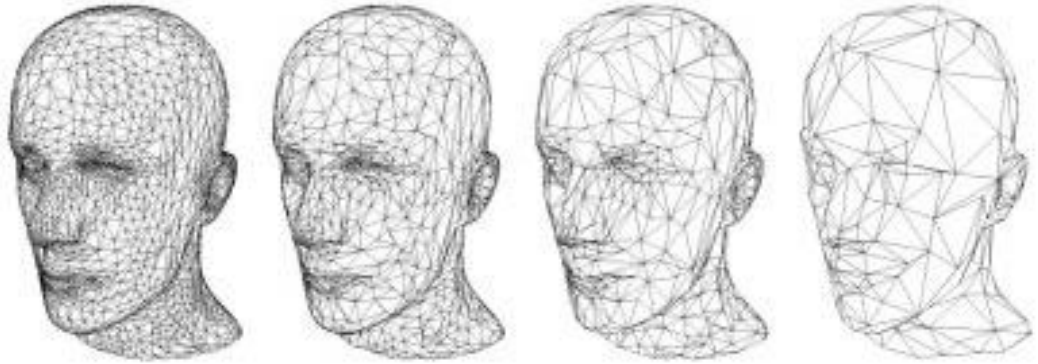

MakerBot Replicator

3D representations

Point clouds



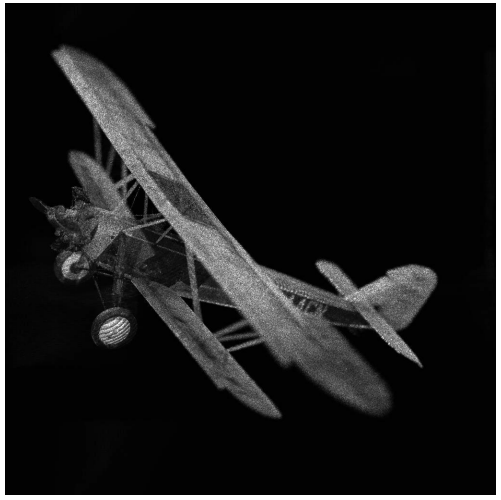
Meshes



Lightfield / 360

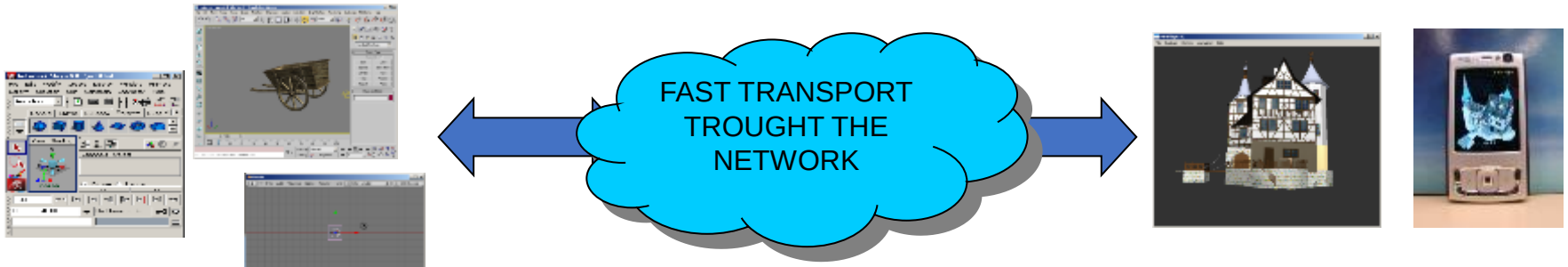
Stereoscopic

Holographic



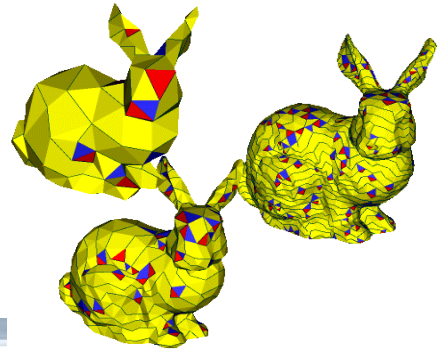
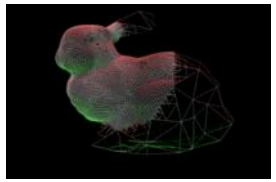
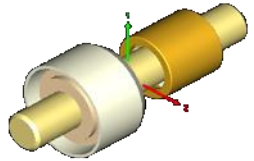
MPEG 3D Graphics Coding

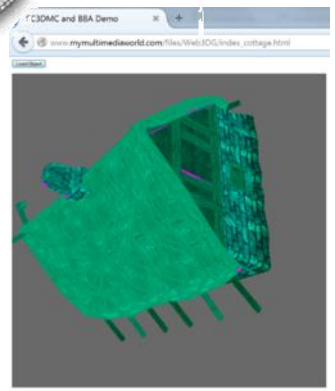
- MPEG 3D Graphics Coding : an ISO Working Group (SC 29 WG 7) aiming at developing standards for graphics compression
- Application agnostic
- For a variety of raw 3D data representation models
- Applying traditional and AI-based signal processing (prediction, frequency transformations, quantisation, entropy encoding)
- Explorations on new approaches for 3D graphics coding

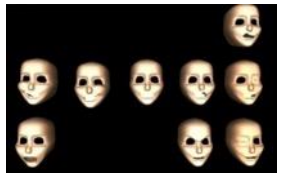


Traditional 3D Graphics models

- Traditionally, the graphics representations rely on (mathematical) models
- Set of parameters to represent the geometry, the appearance and the animations
- Obtained by using authoring tools





MPEG 3D Graphics and Haptics Coding history: 1st pillar

1992/4 — MPEG-1/2 (AV content)

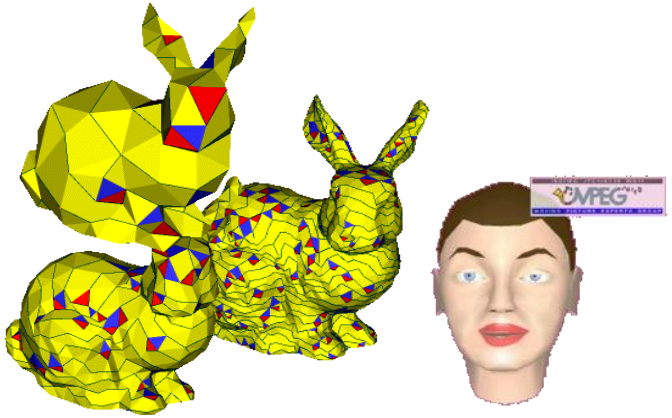
1997 — VRML

1998 — MPEG-4 v.1

- Part 11 - BIFS:
  - -Binarisation of VRML
  - -Extensions for streaming
  - -Extensions for server command
  - -Extensions for 2D graphics
  - - *Real time augmentation with audio & video*
- Part 2 - Visual:
  - - 3D Mesh compression
  - - Face animation

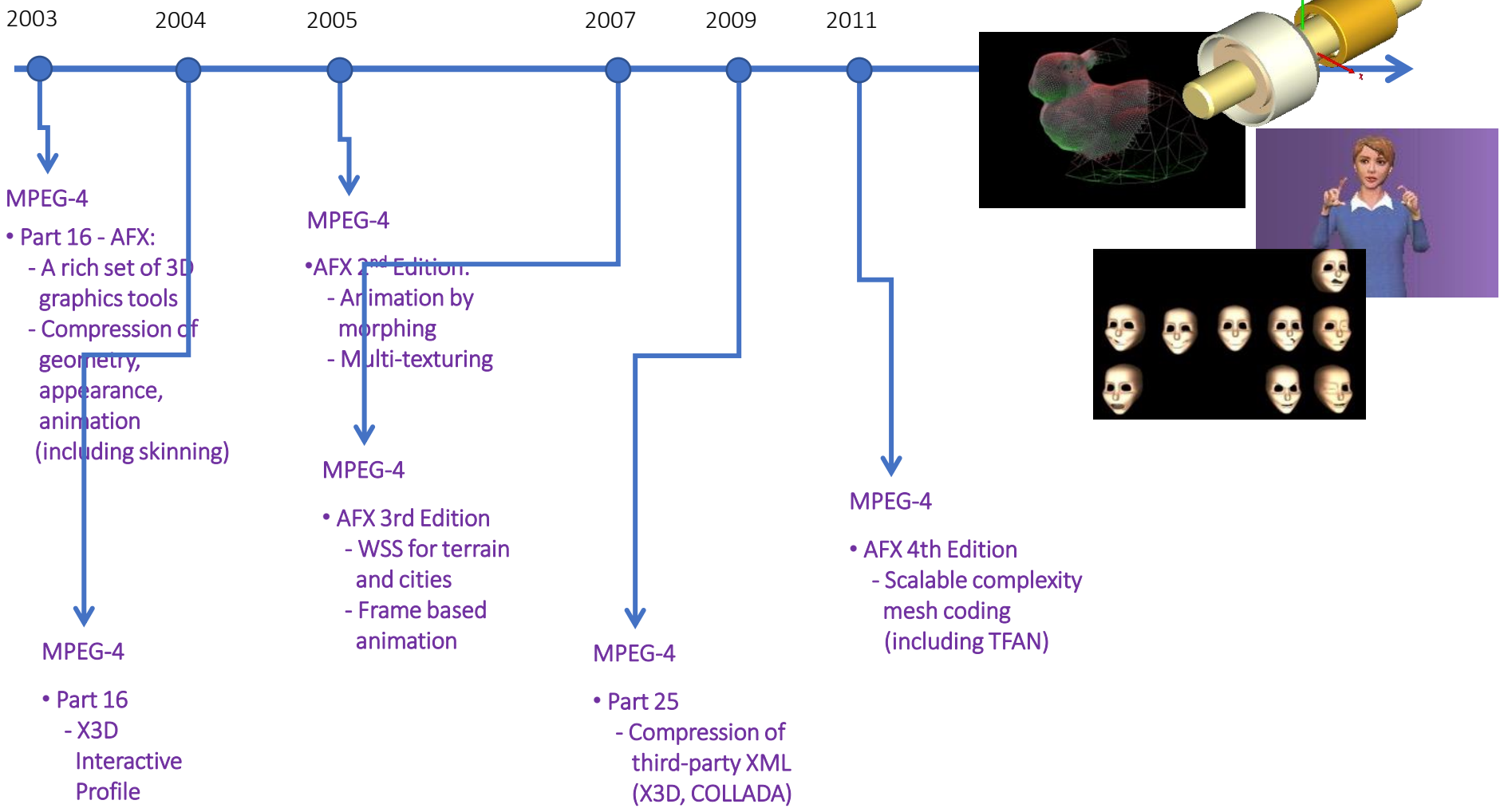
1999 — MPEG-4 v.2

- Part 2 – Visual
  - - Body animation



First form of broadcast signal augmentation

MPEG 3D Graphics and Haptics Coding history: 1st pillar

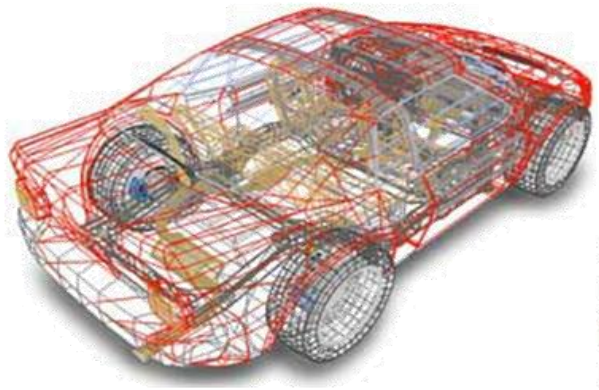


MPEG 3D Graphics and Haptics Coding history: 1st pillar

2011

MPEG-4

- Part 16 - AFX:
  - AMD 1
  - Efficient representation of 3D meshes with multiple attributes



2014

MPEG-4

- AFX AMD2:
  - Multi-resolution mesh coding

2014

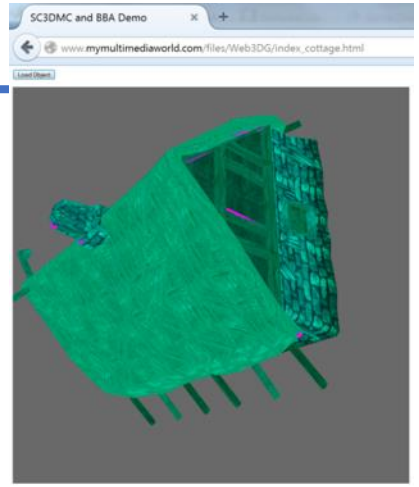
ISO/IEC 23002-4

- Graphics Tool Library

2016

MPEG-4

- AFX AMD3
  - Printing material and 3D graphics coding for browsers



Traditional MPEG 3D Graphics Coding tools

MPEG-4 Part 16 - AFX (*Animation Framework eXtension*) – traditional graphics

Shapes

- ❑ IFS surfaces
- ❑ Patches
- ❑ Subdivision surf.
- ❑ Wavelet SS
- ❑ Mesh Grid
- ❑ Solids
- ❑ Triangle FAN

Textures

- ❑ Visual Texture Coding
- ❑ Synthesized texture
- ❑ Procedural texture
- ❑ Depth Image-based Rep.
- ❑ Point Texture

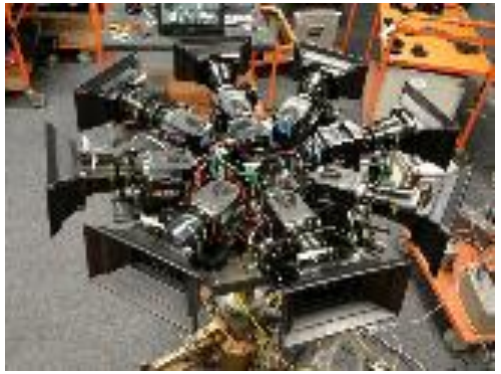
Animation

- ❑ Interpolators
- ❑ Bone-based anim.
- ❑ Morphing
- ❑ Frame-based animation coding

MPEG 3D Graphics: Encoded binary format for each item
- highly efficient representation
- transmission through various networks and terminal devices
- streaming capabilities

Paradigm change in 3D Graphics

- Major advancements in capturing systems allow to “reconstruct” 3D models of the reality (multi cameras, LIDARs, …)
- 3D graphics technologies are now used as a medium for representing the real data
- The content is not anymore a (complex) mathematical model but simple and straightforward representations: point clouds and meshes
  - Two MPEG standards for point clouds coding
  - Currently working on mesh coding



Dynamic Point Clouds

800,000 points -> 1 000 Mbps (uncompressed)

Compression is required in order to make PC usable

Point Clouds Compression in MPEG

V-PCC 04/2020

G-PCC 1/2021

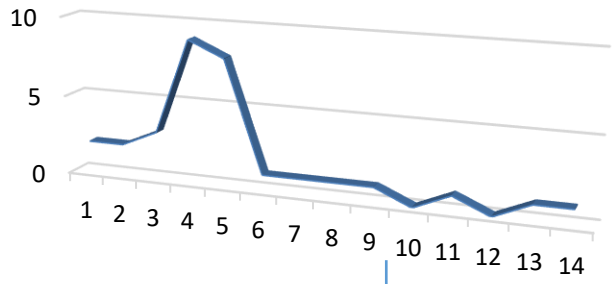
2014 2015 2016 2017 2018 2019 2020

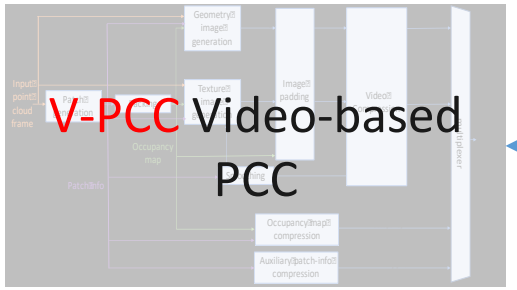
MPEG initiated the work on PCC

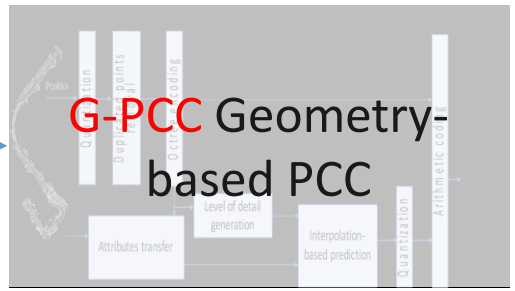
In April 2017 MPEG issued a Call for Proposals

First Committee Draft issued in October 2018

9 technology leading companies responded and MPEG evaluated them in October 2017

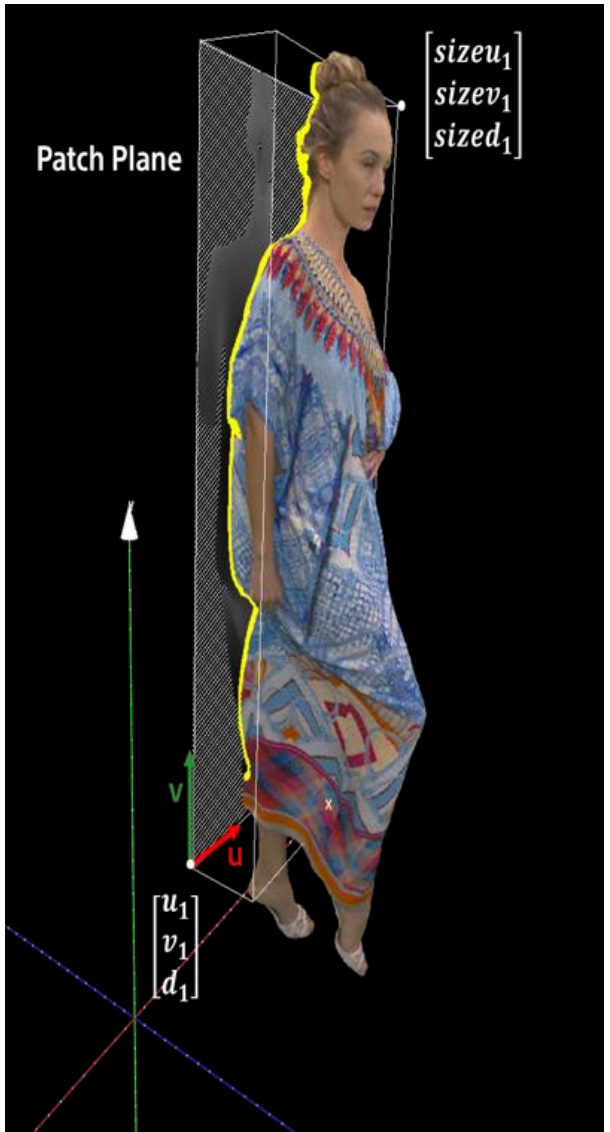






Video-based Point Clouds Compression

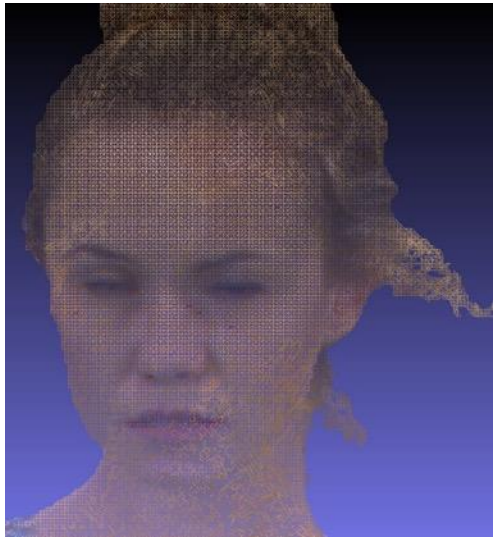
Principle: projecting the points on a plane and encode the projection



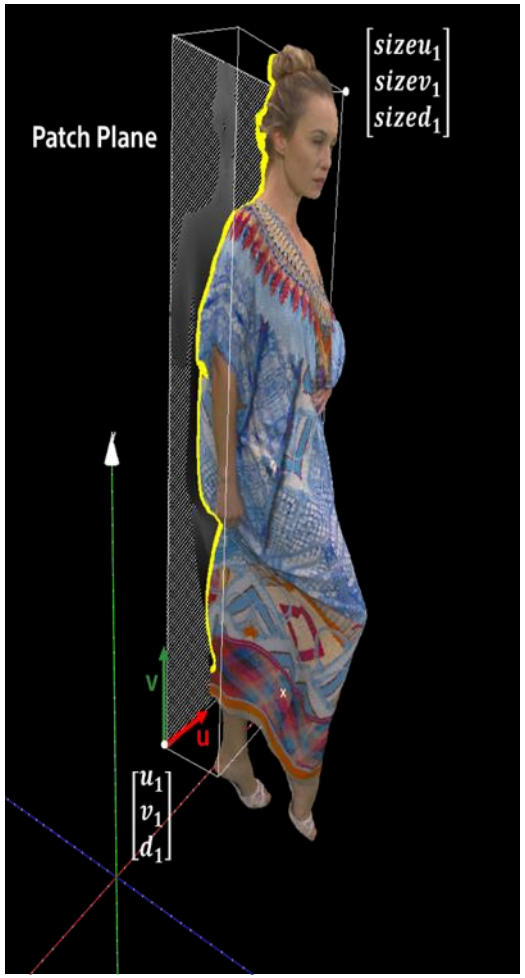
Compression of meshes

- The mesh-based representation is more compact than point clouds
- The graphics hardware is optimized for meshes, rendering point clouds is not direct
- Operations on meshes are easier to perform (because of the connection between vertices)

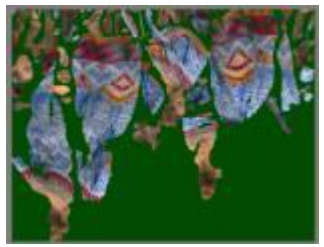
(graphics cards are optimized for triangles)

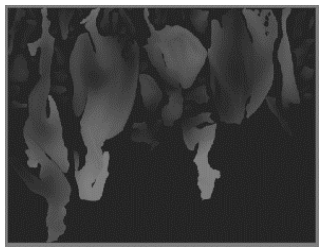


What is the V3C framework?

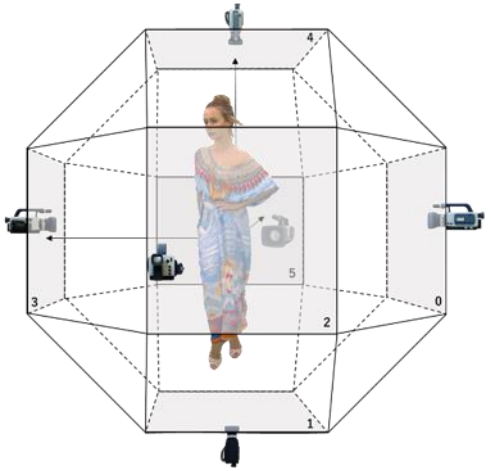


Projecting the 3D object on 2D planes and encode the 2D data by using video coding



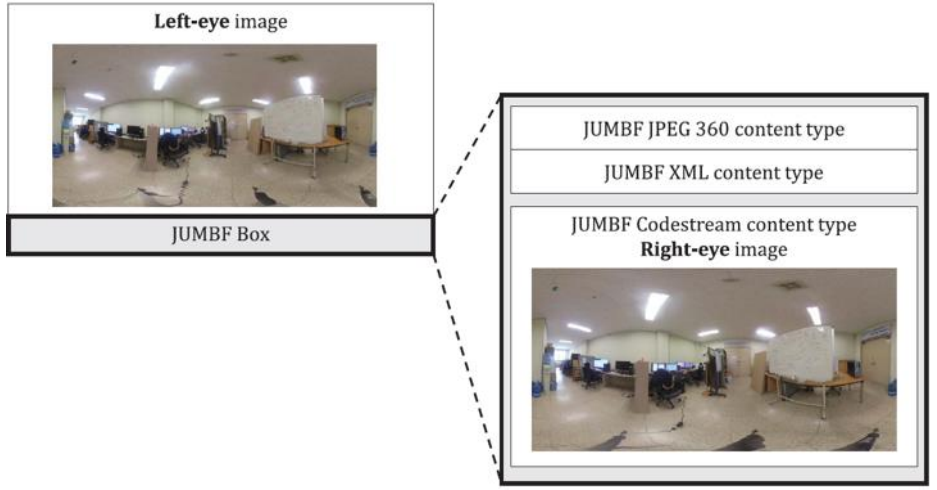


Optimize the projection by choosing appropriate planes per patch

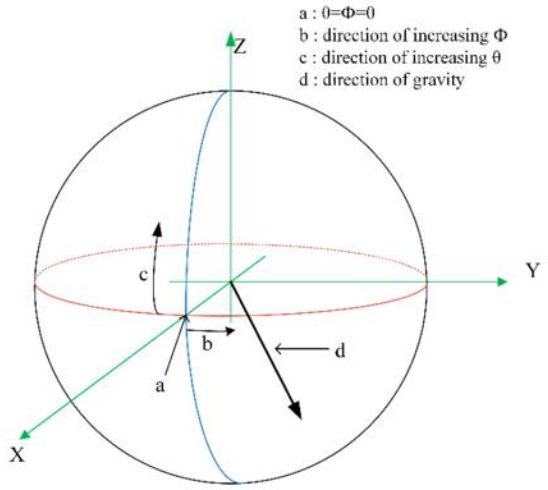


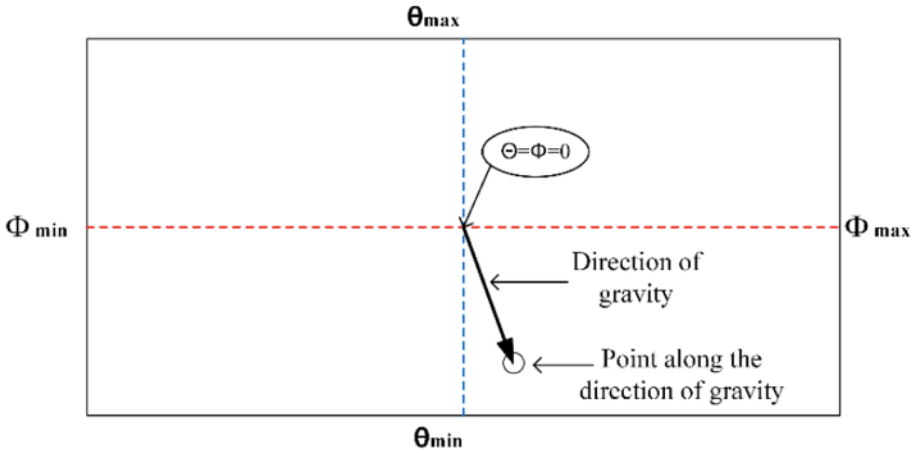
6 + 12 projection planes

JPEG 360 (ISO/IEC 19566-6)



- Part of **JPEG Systems framework**
- **Multi image** format
- Supports mapping between **spherical surface** and **equirectangular** representations
- Focus on **backwards compatibility** with common JPEG formats

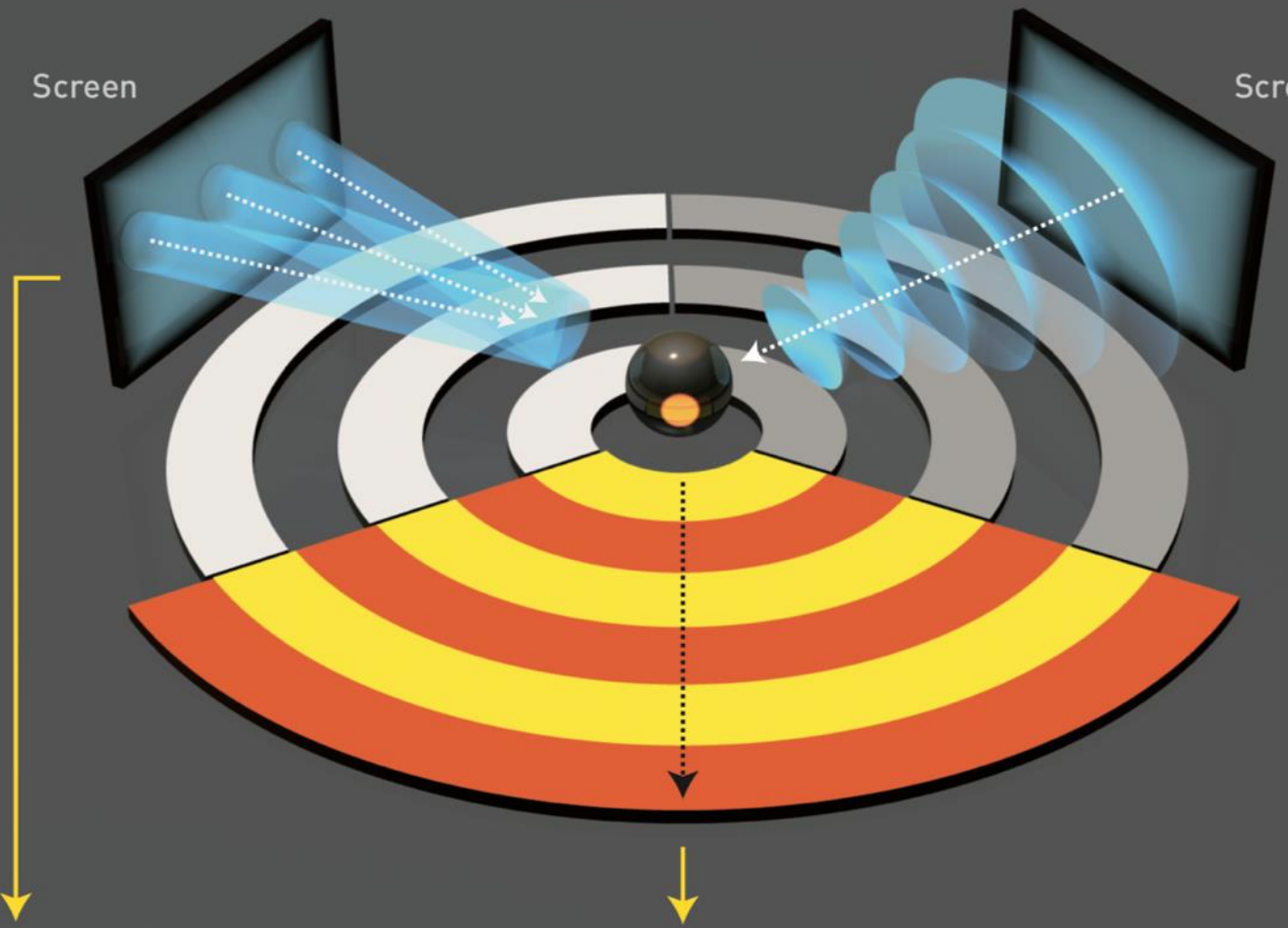




JPEG Pleno (ISO/IEC 21794)

© 2018 Smalley et al., “Volumetric Displays, Turning 3-D Inside-Out”, Optics & Photonic News, June 2018.

3-D display families



Ray (lightfield)

Parallax : easy
Accommodation : hard
Occlusion : easy
View angle : moderate
Virtual image formation : easy

Point (volumetric)

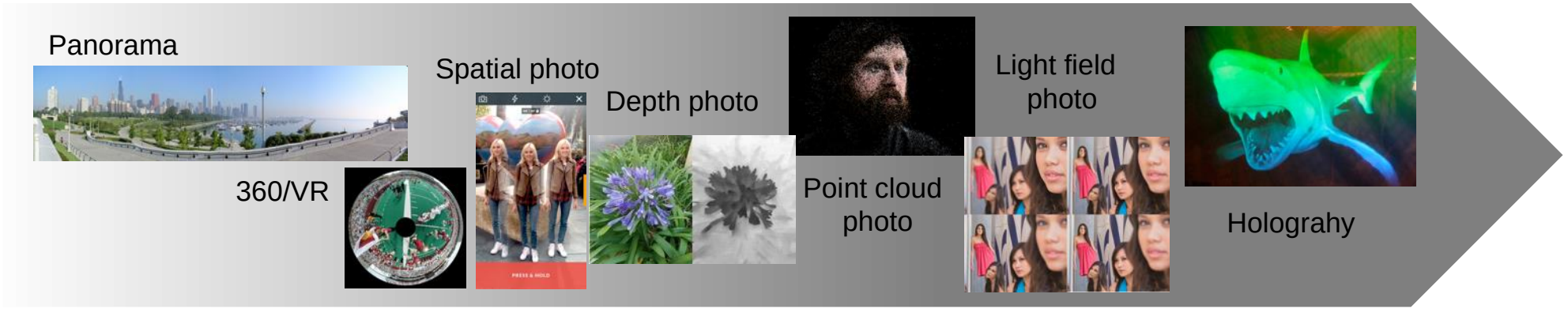
Parallax : easy
Accommodation : trivial
Occlusion : difficult
View angle : easy
Virtual image formation : impossible?

Wave (holographic)

Parallax : easy
Accommodation : easy
Occlusion : easy
View angle : hard
Virtual image formation : easy

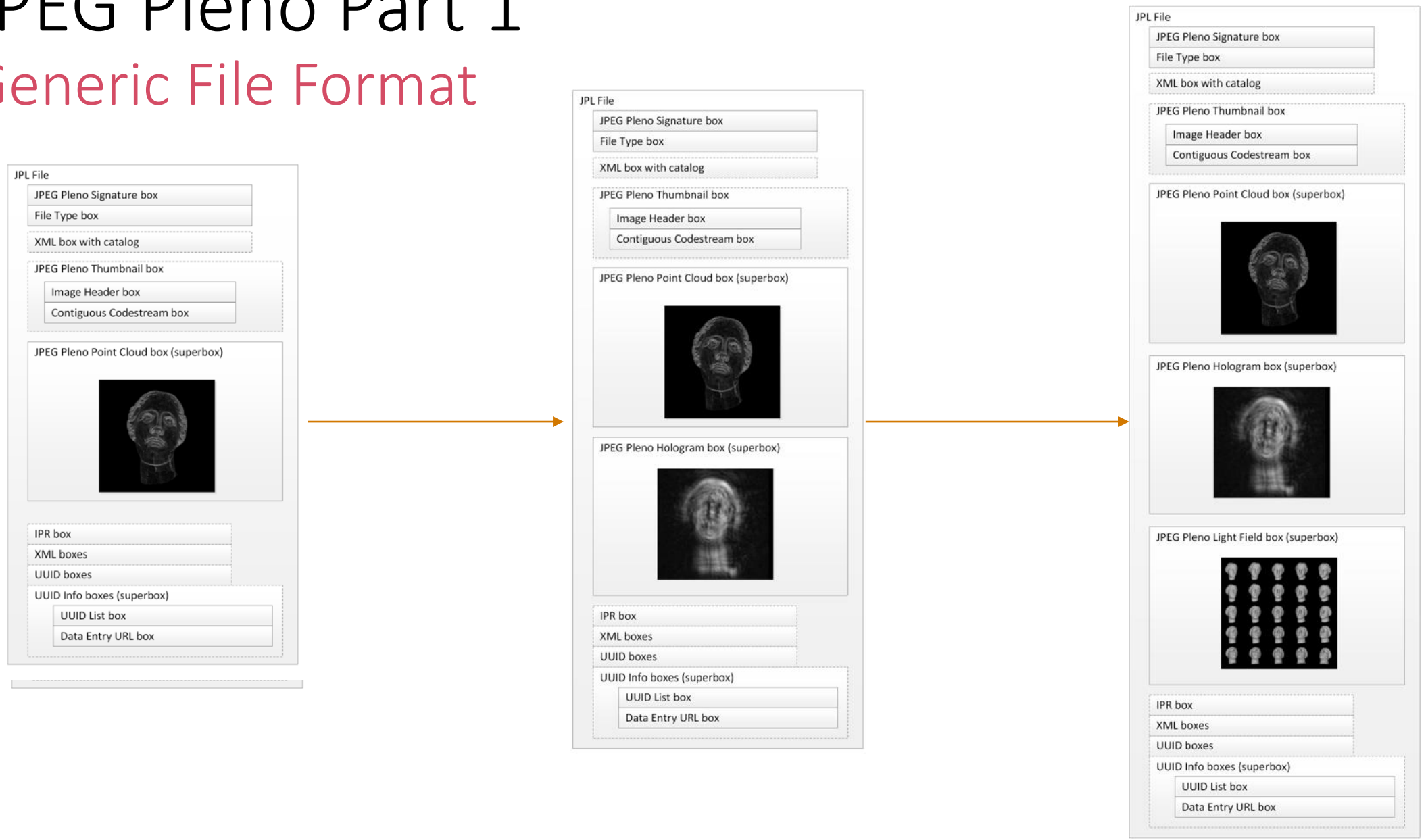
JPEG Pleno Design Principles

- A framework but not a super format
- One or limited number of representation models
- Support representation of point cloud, light field and holography
- Where needed, backward compatible with legacy JPEG standards
- Investigate quality evaluation methodologies



JPEG Pleno Part 1

Generic File Format



JPEG Pleno Light Field

- **Depth-based view prediction**.
- **Scene depth** provides information on 2D point **correspondences** between views.
- Algorithm:
  - **Encode** texture and depth at a **reference view** location.
  - **Predict** at **side view** locations using texture and depth of the reference.
  - **Encode residue**.

JPEG Pleno Light Field

JPEG Pleno Quality Assessment standardization effort

- JPEG has launched a new standardization effort known as JPEG Pleno Quality Assessment.

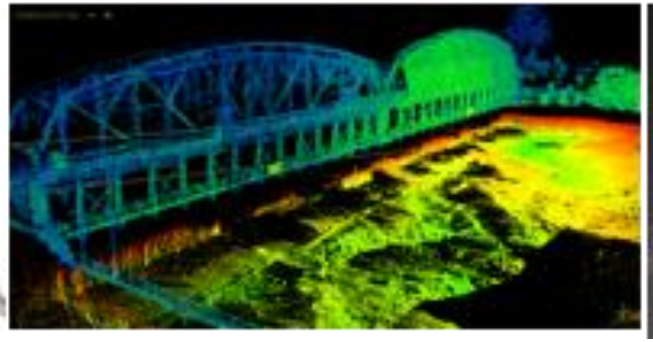
- The scope of JPEG Pleno Quality Assessment is the creation of a quality assessment standard defining a framework including subjective and objective quality assessment methodologies for lossy decoded data of plenoptic modalities in the context of multiple use cases.

- It aims at providing a quality assessment standard for plenoptic modalities, addressing the light field modality on its first phase.

JPEG Pleno Point Cloud

- Several **acquisition solutions** produce light field data (time of flight, lidar, …)
- Need for standards that support **human and machine** consumption
- The scope of the JPEG Pleno Point Cloud is the creation of a **learning-based coding** standard for point clouds

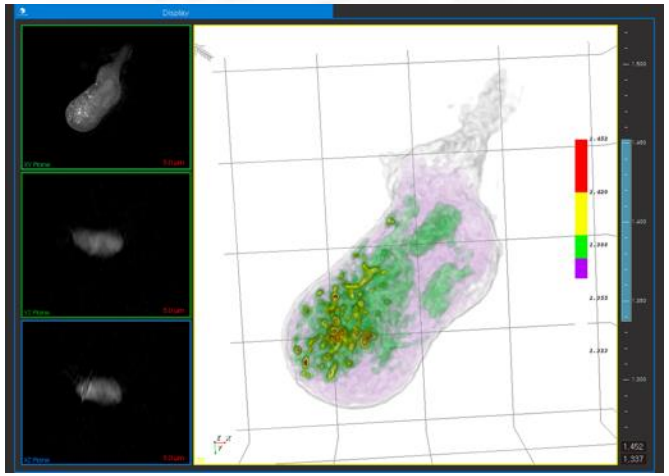
- The standard targets **interactive human visualization** and **3D processing computer vision tasks**

JPEG Pleno Holography

Nano- to Macroscale Applications

© 2018 Ovizio



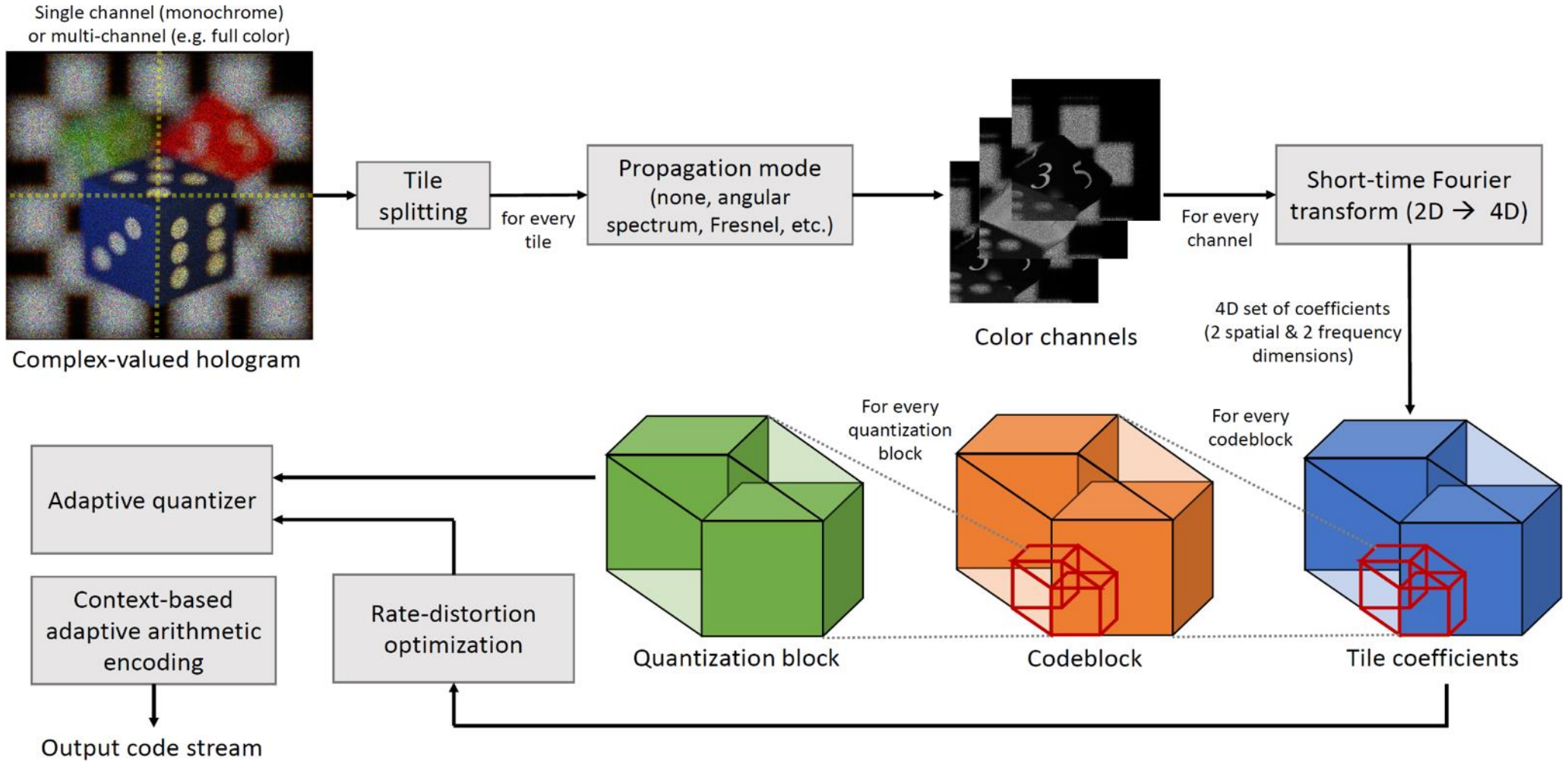
© 2019 Tomocube

© 2019 www.ultimate-holography.com

© 2018 Brigham Young University
www.youtube.com/watch?v=qUSiw87mQck

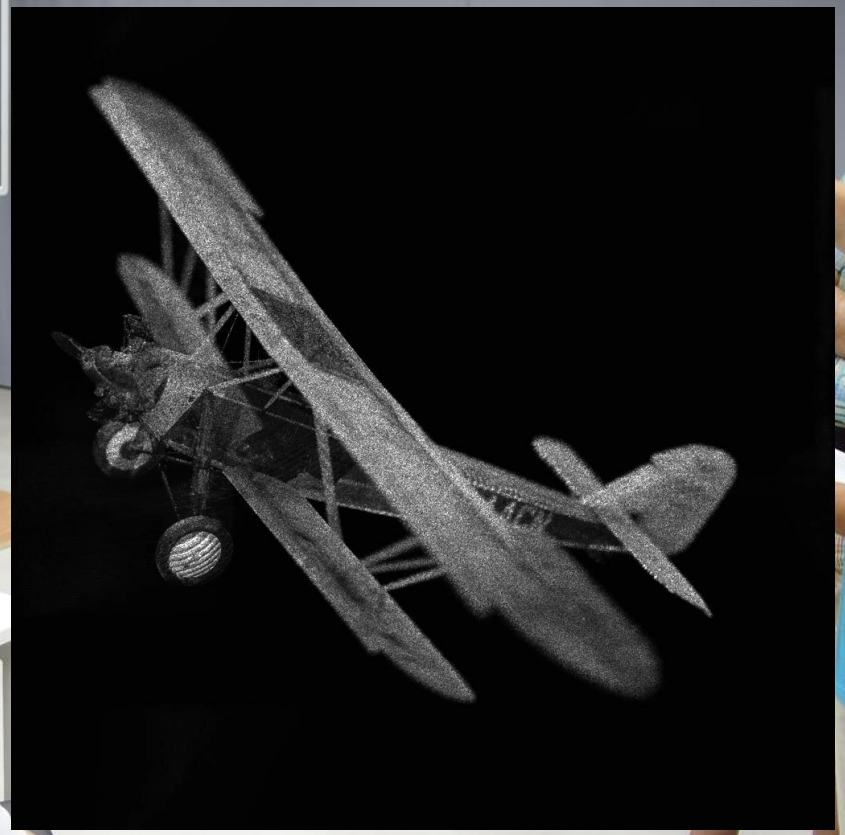
JPEG Pleno Holography

INTERFERE codec



JPEG Pleno Holography

Objective & Subjective Quality Assessment



A. Ahar et al. “Suitability Analysis of Holographic vs Light Field and 2D Displays for Subjective Quality Assessment of Fourier Holograms”, Optics Express, 2020.
A. Ahar et al., "Comprehensive performance analysis of objective quality metrics for digital holography", Signal Processing: Image Communication, 2021

Conclusions

- Representing data in 3D is complex and diverse
- No single 3D modality for all applications
- Capturing, data representation and rendering need to be assessed for each specific use case
- Succesful adoption will require high quality content, a good user experience and interoperable interfaces
- Standardization activities focusing on quality, efficiency and adressing functional needs

Thank you!

- More information:
  - https://www.mpeg.org/structure/coding-of-3d-graphics/
  - https://jpeg.org