Overview of technologies for immersive visual experiences: capture, processing, compression, standardization and display

Marek Domański

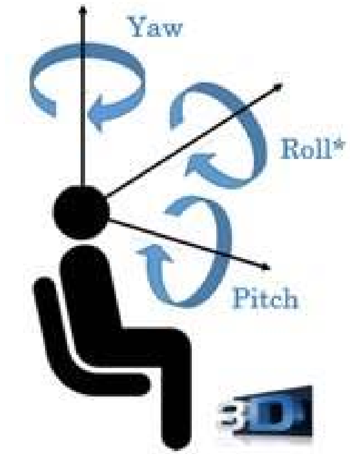
Poznań University of Technology

Chair of Multimedia Telecommunications and Microelectronics

Poznań, Poland

Immersive visual experiences

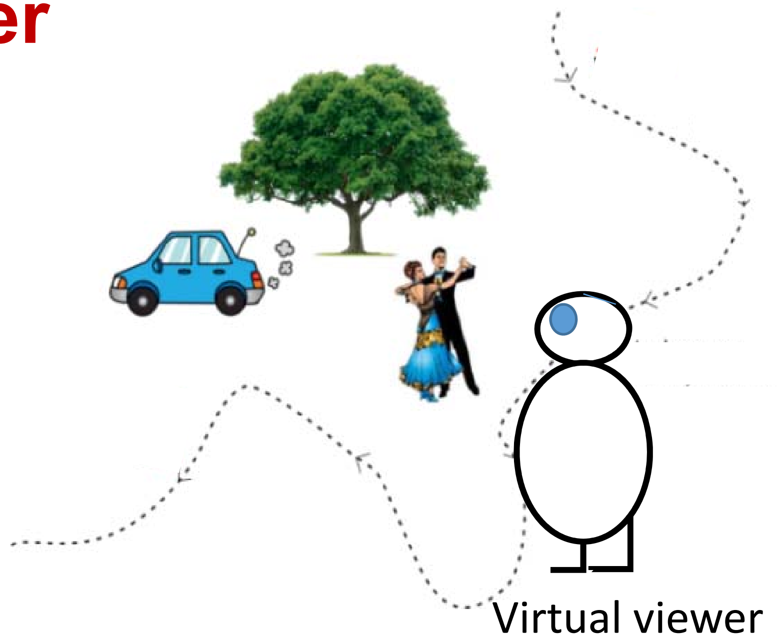
- **Arbitrary direction of viewing**



System : 3DoF

- **Arbitrary location of viewer**

- virtual navigation
- free-viewpoint television



- **Both**

System : 6DoF

Immersive video content

- **Computer-generated**
- **Natural content**
  - Multiple camera around a scene

    also depth cameras, light-field cameras
  - Camera(s) located in a center of a scene

    360-degree cameras
  - Mixed

    e.g.: wearable cameras

Video capture
Common technical problems

- **Synchronization of cameras**

  Camera hardware needs to enable synchronization

  Shutter release error $<$ Exposition interval

- **Frame rate**

  High frame rate needed – Head Mounted Devices

Common task for immersive video capture:

Calibration

Camera parameter estimation:

- **Intrinsic** – the parameters of individual cameras – remain unchanged by camera motion
- **Extrinsic** – related to camera locations in real word – do change by camera motion (even slight !)

Out of scope of standardization

- Improved methods developed

Depth estimation

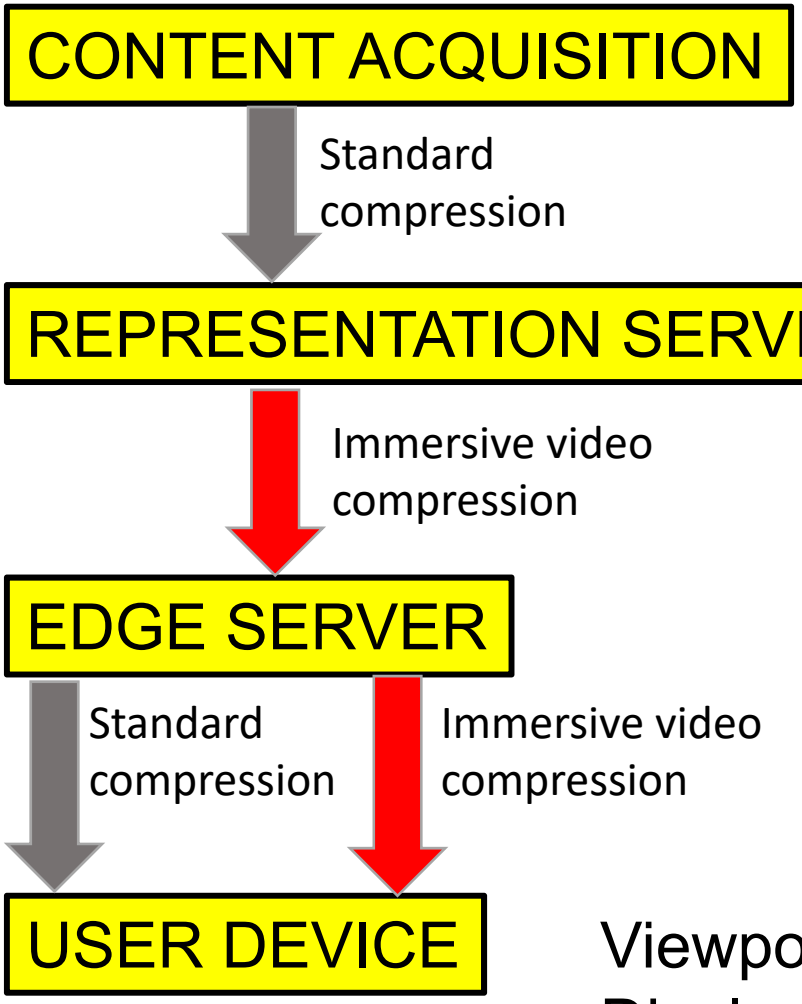
- **By video analysis** – from at least 2 video sequences
  - Computationally heavy
  - Huge progress recently
- **By depth cameras** – diverse products
  - Infrared illuminate of the scene
  - Limited resolution mostly

Out of scope of standardization

MPEG
Test video sequences

- Large collection of video sequences with depth information
- Multiview, omnidirectional, lightfield, volumetric, etc.
- Used for experiments in standardization and algorithm testing

Immersive video path



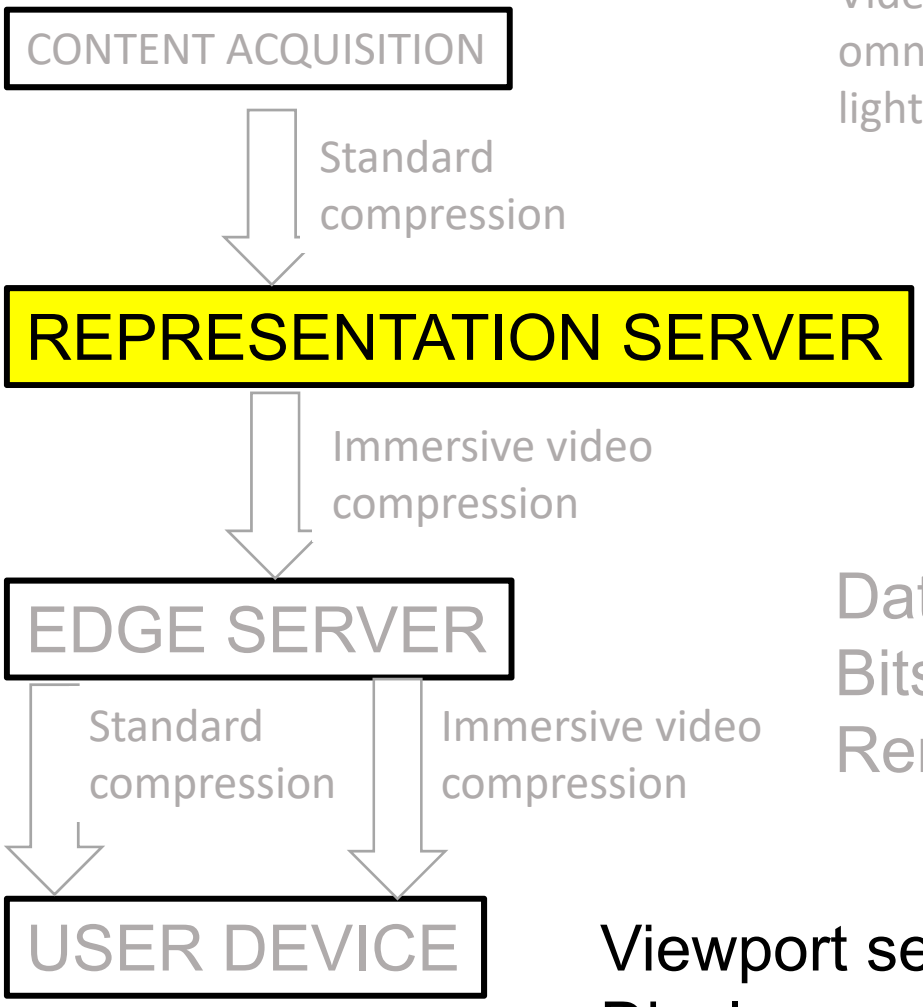
Video cameras, depth cameras, omnidirectional cameras, lightfield cameras

System calibration,
video correction,
depth estimation,
scene representation , …

Data buffering and restoration
Bitstream personalization
Rendering (partially?)

Viewport selection, Rendering (partially?), Display

Immersive video path



Video cameras, depth cameras, omnidirectional cameras, lightfield cameras

System calibration, video correction, depth estimation, scene representation , …

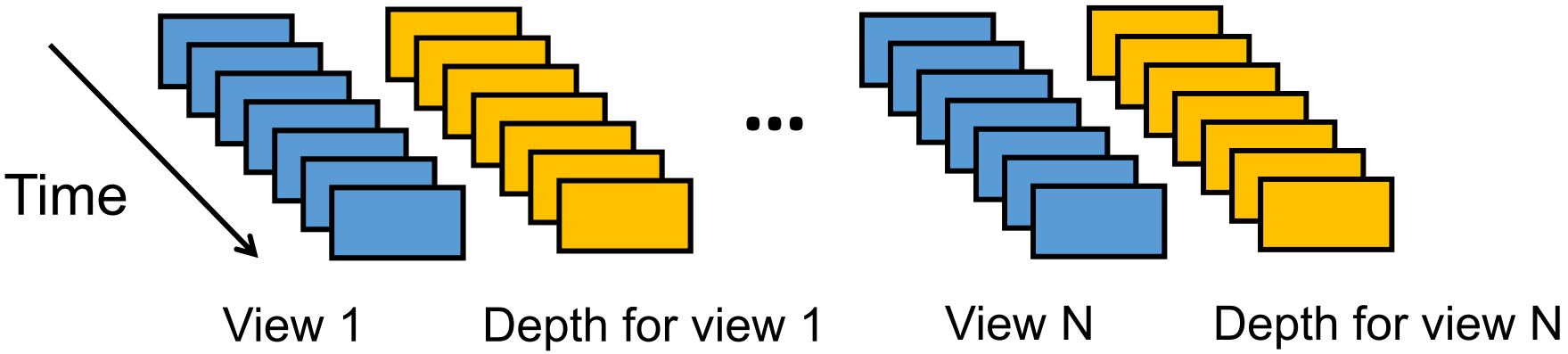
Data buffering and restoration
Bitstream personalization
Rendering (partially?)

Viewport selection, Rendering (partially?), Display

Scene representations

Basic representations used by MPEG

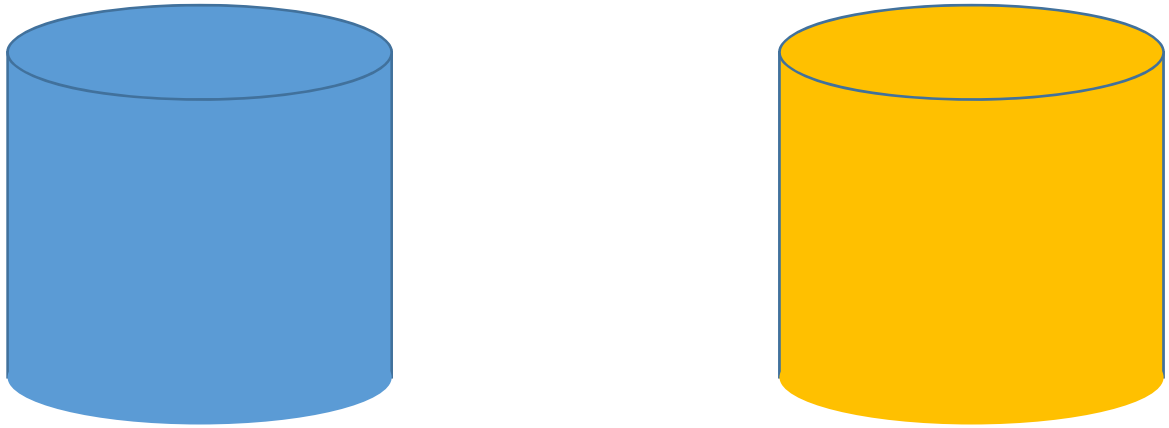
- **Multiview plus depth**



Scene representations

Basic representations used by MPEG

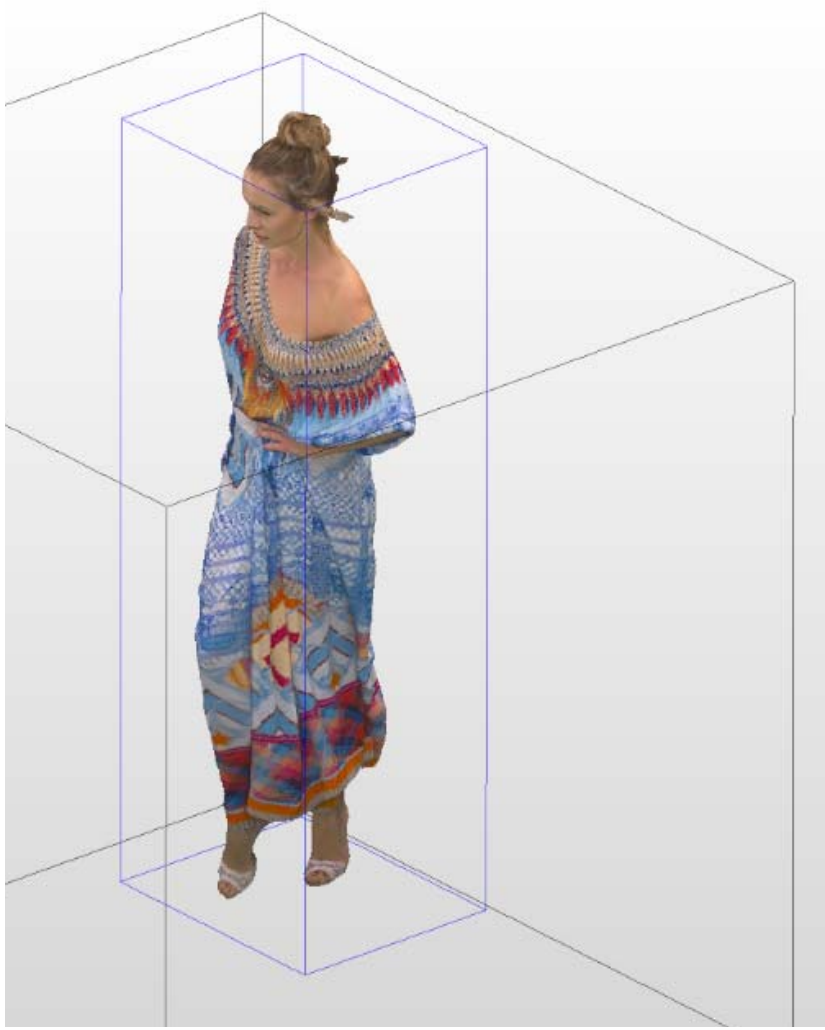
- **360-degree + 360-degree depth**



Scene representations

Basic representations used by MPEG

- **Point clouds**

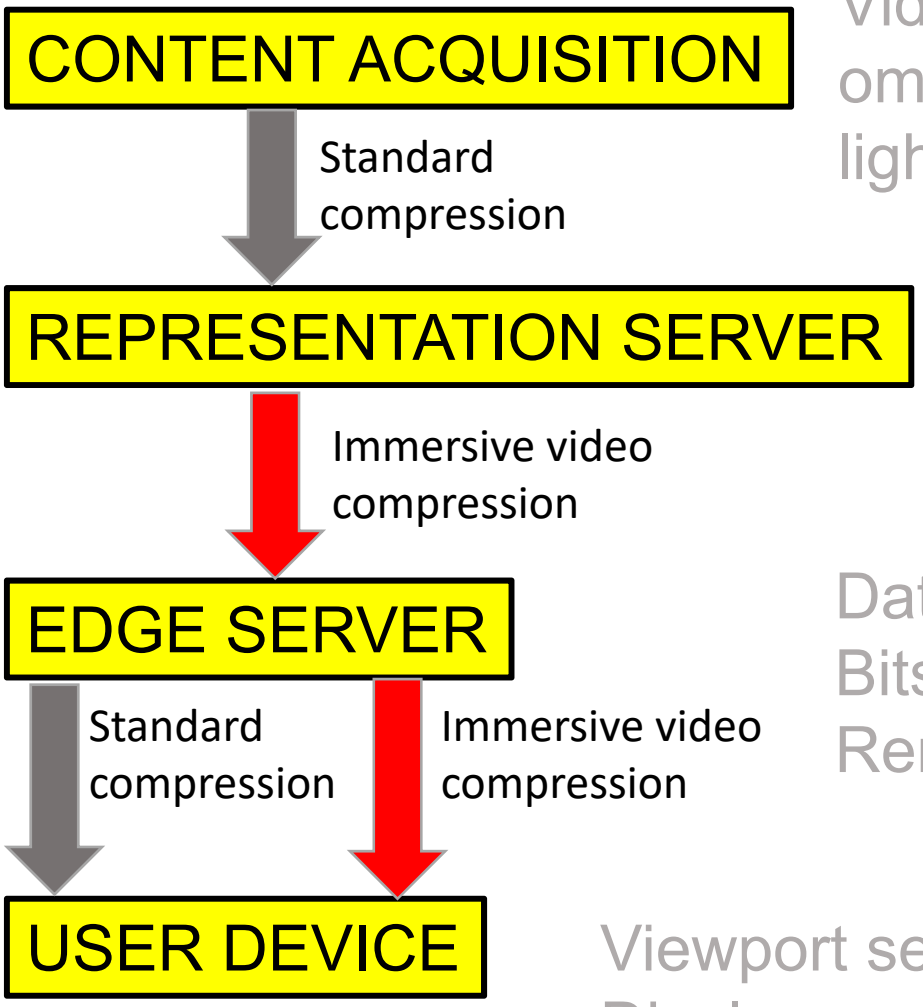


Immersive video compression

Available techniques
+ standards under development

mostly try to exploit
general compression techniques

Immersive video path



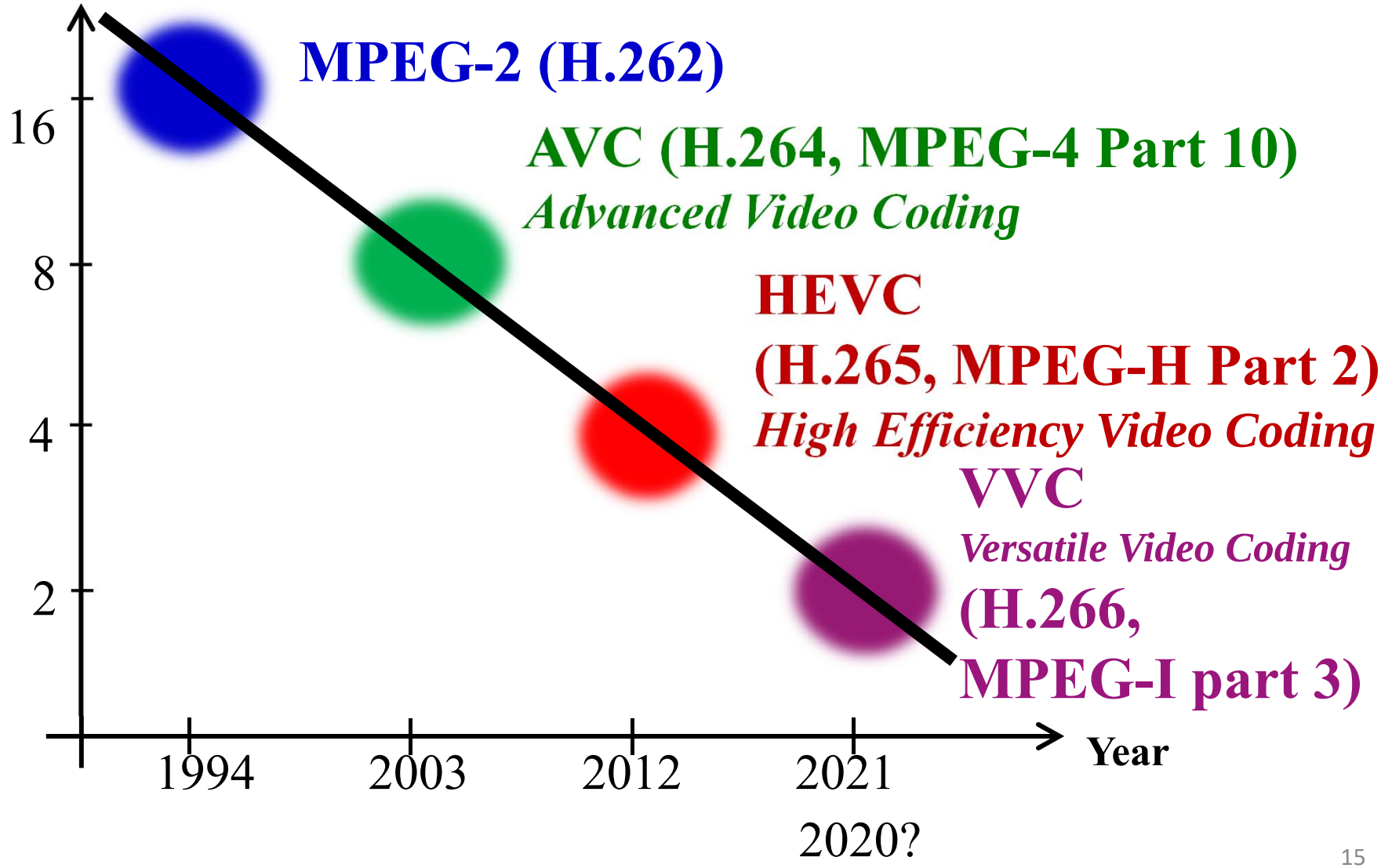
Video cameras, depth cameras, omnidirectional cameras, lightfield cameras

System calibration,
video correction,
depth estimation,
scene representation , …

Data buffering and restoration
Bitstream personalization
Rendering (partially?)

Viewport selection, Rendering (partially?), Display

Approximate bitrate for HD (1920×1080, 25 frame/s) [Mb/s]



Compressed video bitrates

$$B \approx A \cdot V \quad [\text{Mbps}]$$

A – technology coefficient,

$A = 4$ for MPEG-2, $A = 1$ for HEVC,

$A = 2$ for AVC, $A = 0.5$ for VVC (expect).

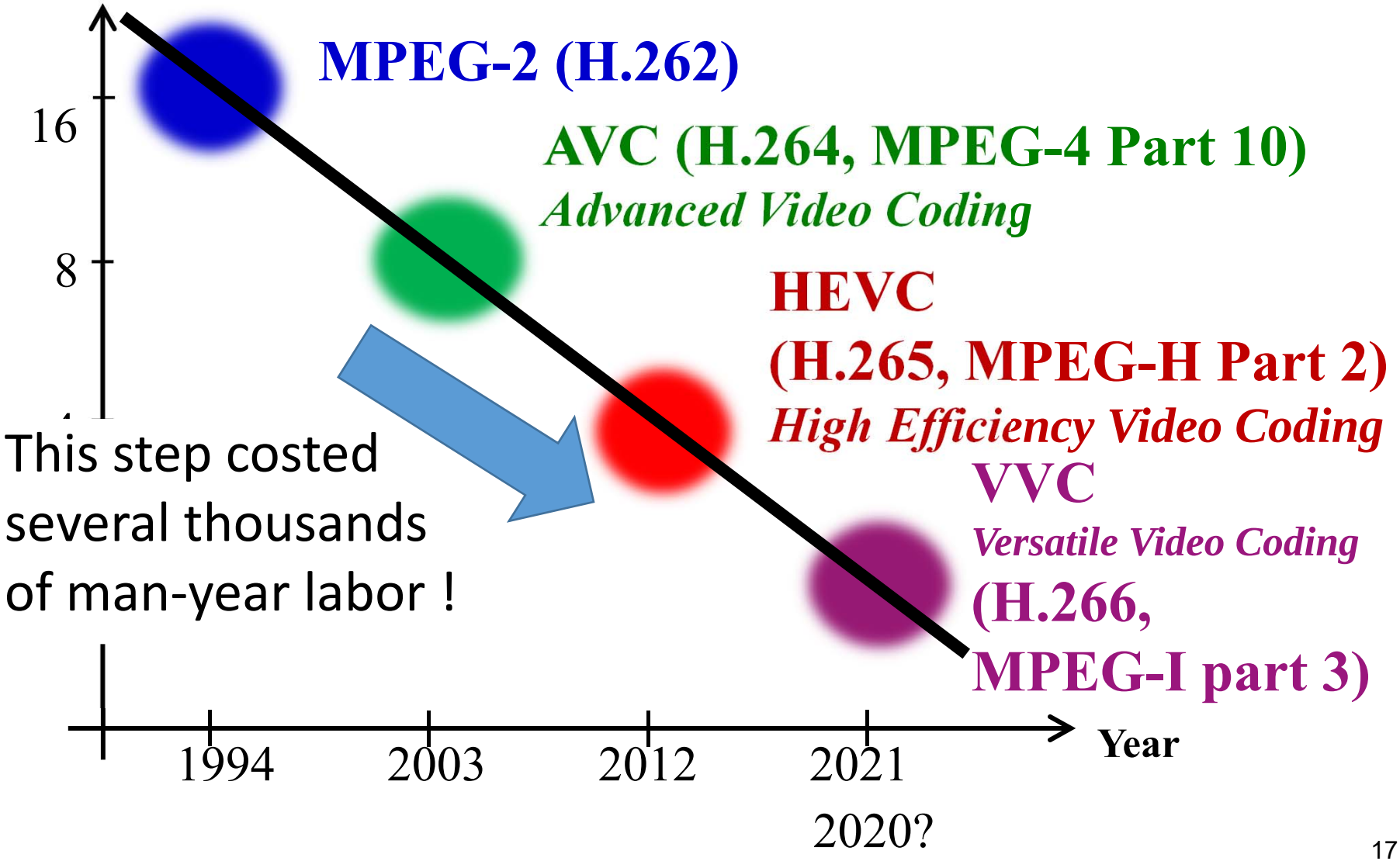
V – format coefficient,

V=1 for SDTV (720×576, 25*i*),

V=4 for HDTV (1920×1080, 25*i*),

V=16 for UHDTV (3840×2160, 50*p*),

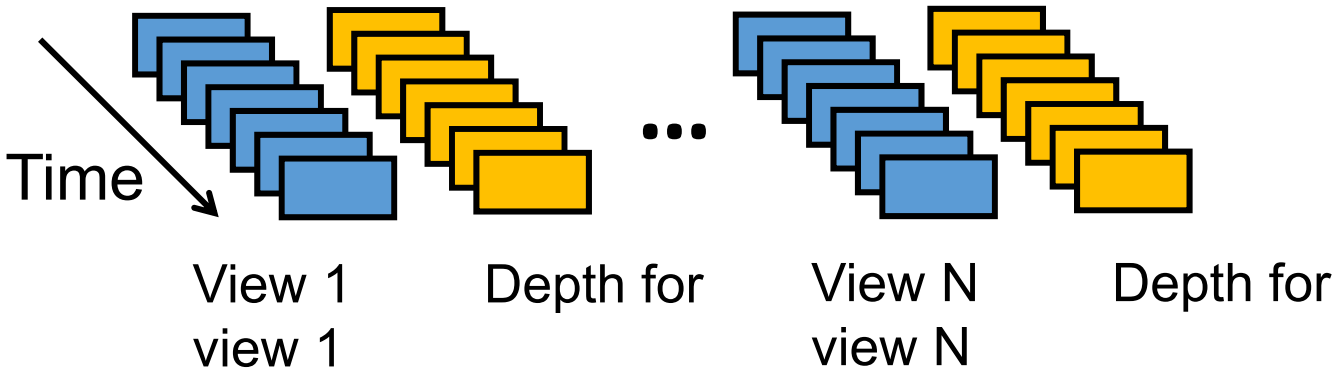
Approximate bitrate for HD (1920×1080, 25 frame/s) [Mb/s]



Immersive video compression

Enormous cost
of development
of new video technology
from scratch

Compression of multiview plus depth video



Coding exploits inter-view redundancy :

Multiview extensions of video coding standards

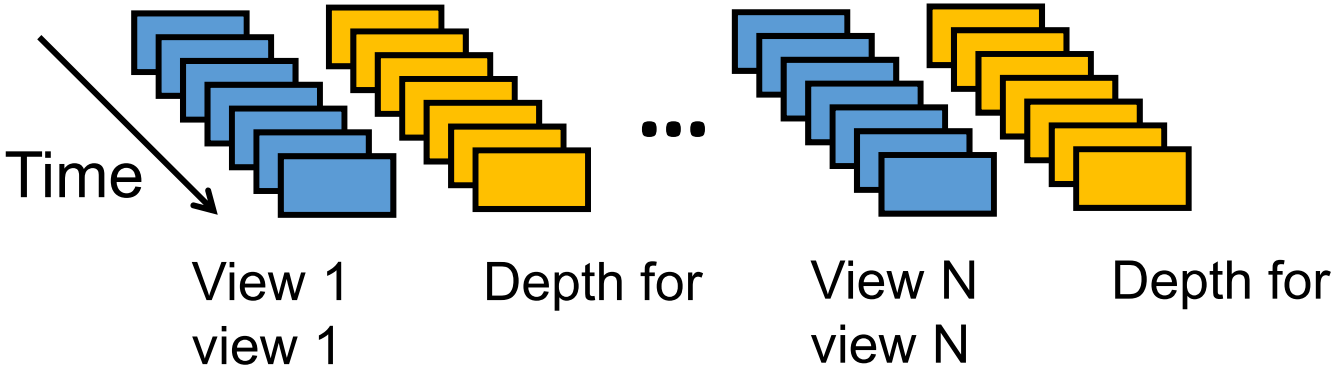
MPEG-2, AVC, HEVC (standardized)

Additionally exploits depth information:

3D extensions of video coding standards

AVC, HEVC (standardized)

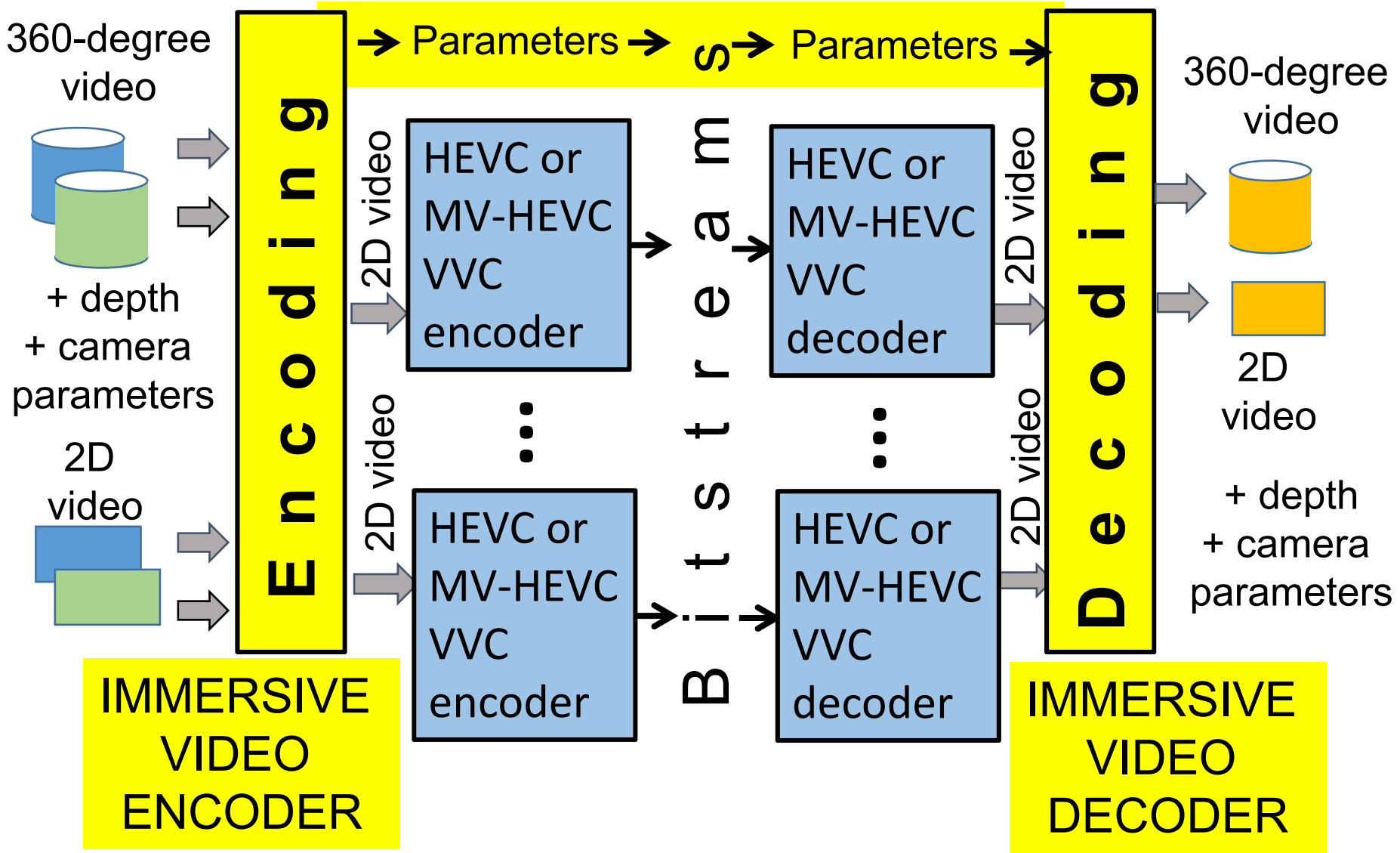
Compression of multiview plus depth video



Efficiency of Multiview/3D coding

- 10 – 50 % bitrate reduction vs. simulcast
- Efficient for dense camera locations on a line

Immersive video compression



Video examples

Inputs to HEVC encoders

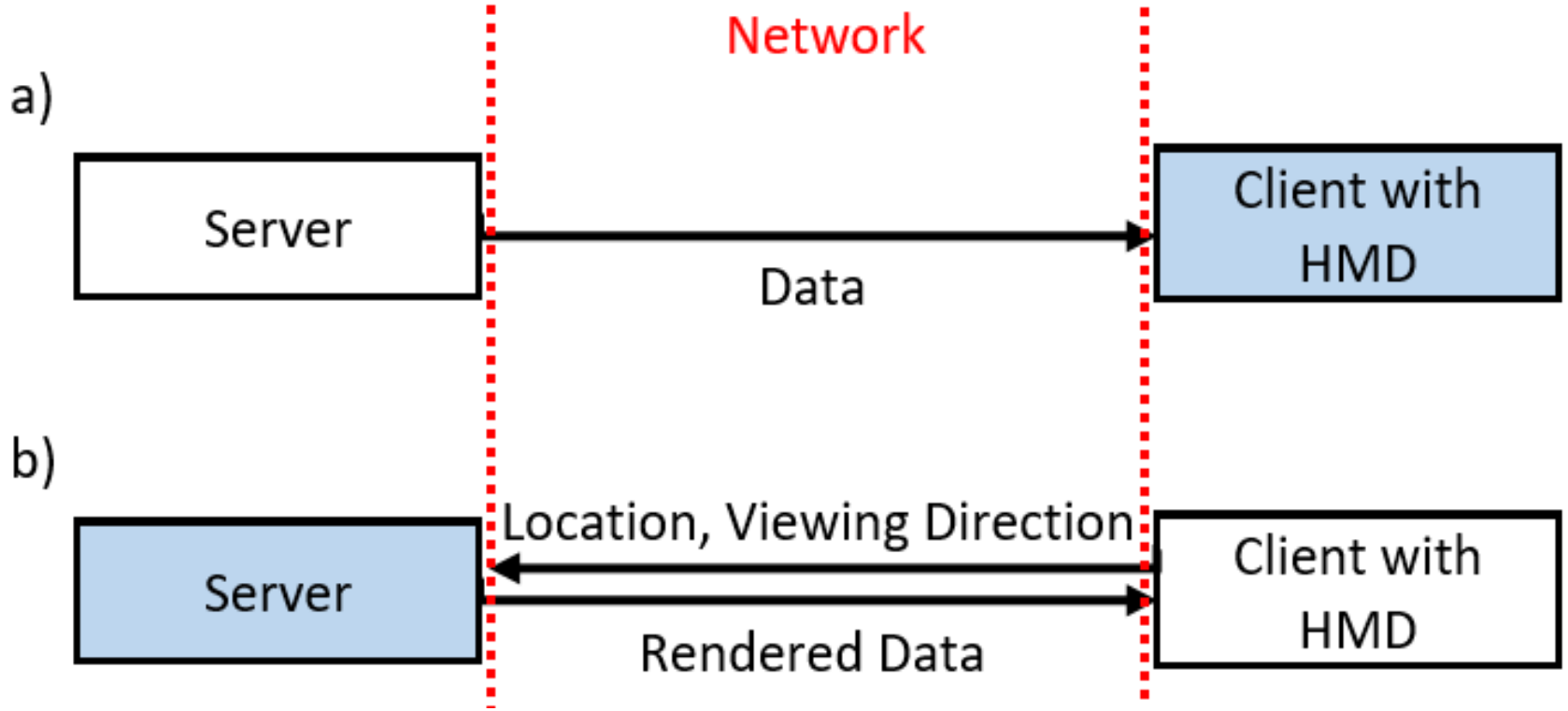
Similar approach

Video-based Point Cloud

Audio:

MPEG surround
MPEG-H 3D Audio

Edge server – user device link



Thank you