3D Video Coding and Broadcasting summary

A. Tsanakas, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
Version 2.4.1

Artificial Intelligence & Information Analysis Lab

3D Video Coding and Broadcasting

- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

Introduction

- The past decade in the field of ***Digital Video Broadcasting*** (DVB) and multimedia video systems can be described a significant revolution.

- One of the important achievements in this decade is related to display technology and new television (TV) services, including Three-Dimensional TV display technology (3DTV).

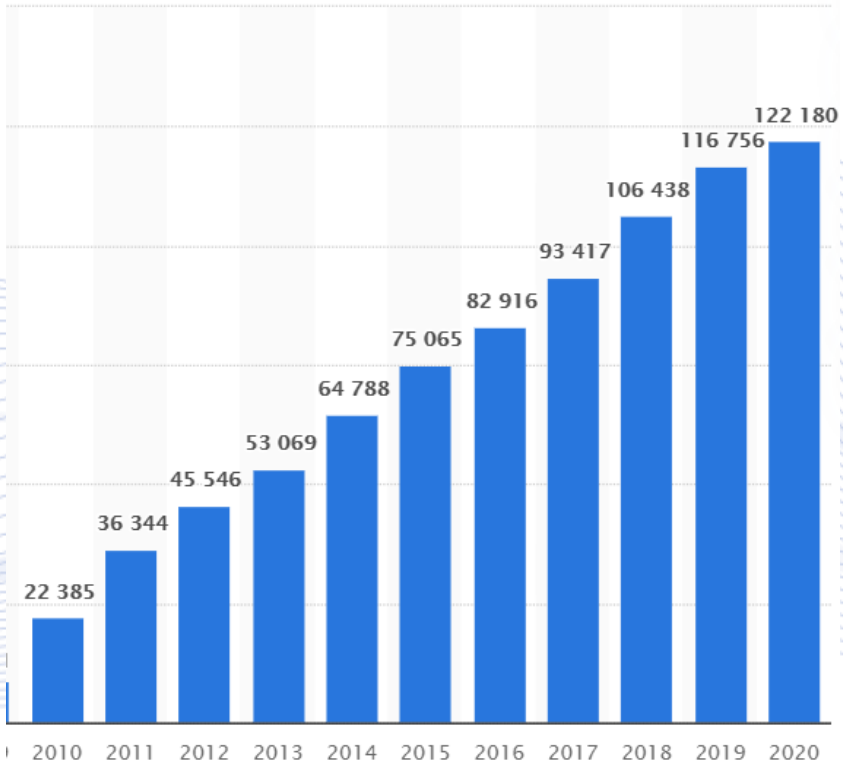
Introduction

The main ***categories*** of 3D video may be defined:

- ***Monoscopic 360^o video:*** a common video from many cameras is stitched to a panorama.
- ***Stereoscopic and binocular 360^o video:*** viewer can watch in an arbitrary position with many levels of spatial sensations
- ***6^o of freedom 360^o video***: a viewer has the ability to change his location for free.

Introduction

- Diagram with the number of **Digital 3D Cinema Screens** worldwide from 2010 to 2020 according to *www.statista.com.*



3D Video Coding and Broadcasting

- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

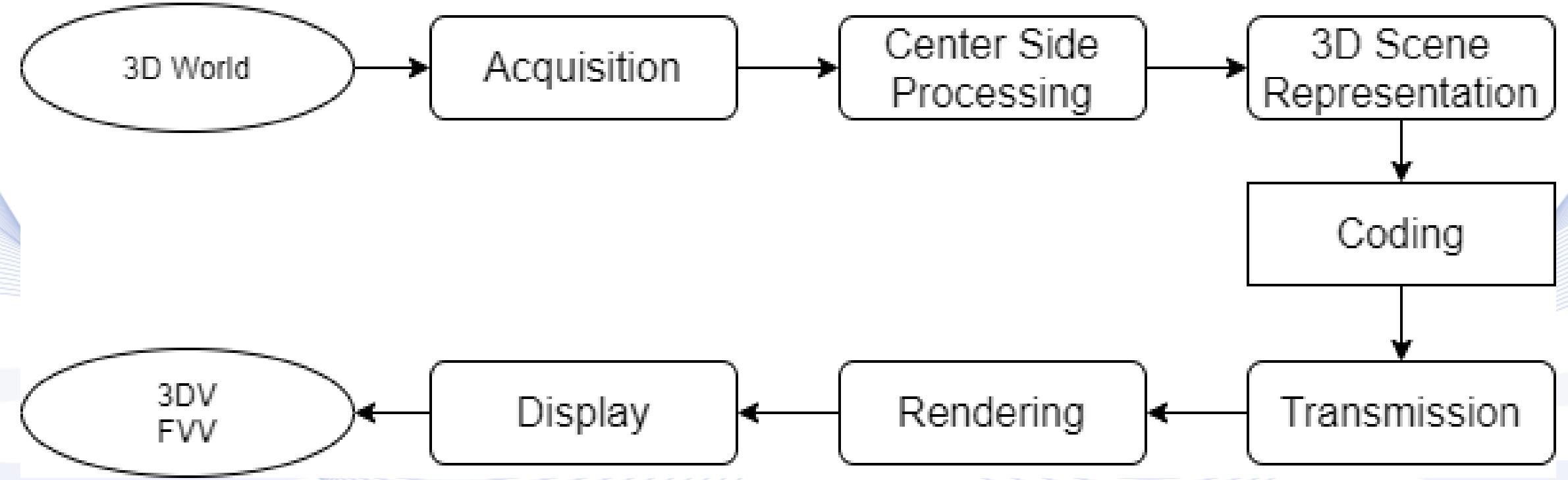
3D Video Processing Chain

The end-to-end 3D video system for short applications is divided into four main blocks:

- **3D Content Creation (acquisition, center side processing)**
- **3D Representation**
- **Delivery (Coding, Transmission, Decoding)**
- **Visualization (Rendering, Display)**

3D Video Processing Chain

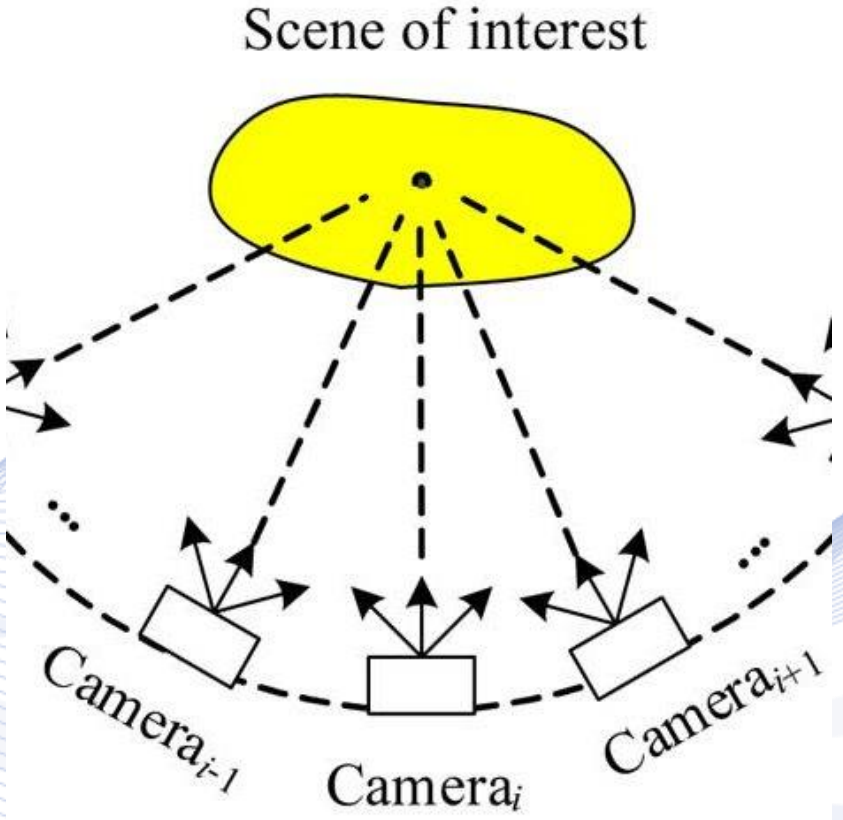
- The steps of ***3D Video Processing***



3D Video Processing Chain

❖ ***3D Video Acquisition Tools***

- **Stereoscopic** Camera Systems
- **Multi-view** Camera Systems
- **3D Depth Sensor** Camera



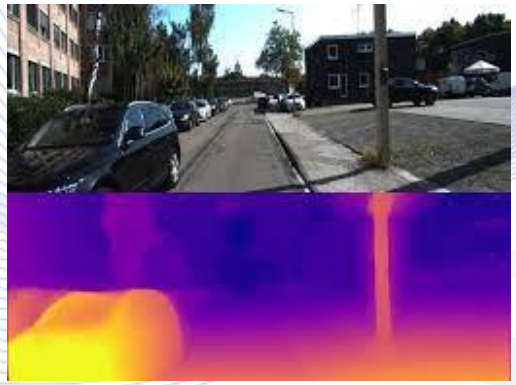
Multi-view Camera System [SX07]

3D Video Processing Chain

❖ ***Center Side Processing***

At this stage, we recover the geometry of the scene and reconstruct the 3d scene. Important algorithms used in this case are:

- ***Calibration***
- ***Rectification***
- ***Correction***
- ***Filtering***
- ***Depth Estimation***

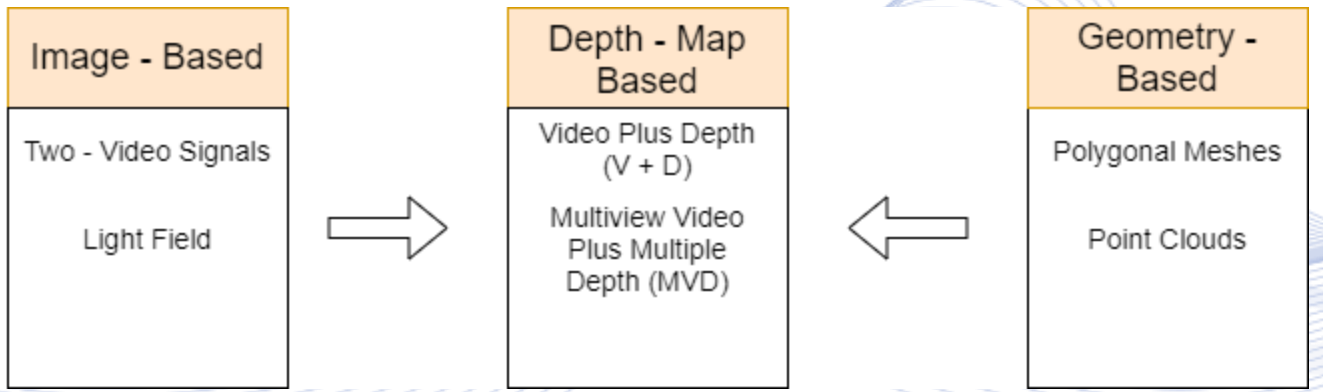


Depth map of an image [PSRA19]

3D Video Processing Chain

❖ ***3D Scene Representation***

- There are ***three main categories of 3D representation***: geometry-based, image-based and representation based on depth maps.



Three main classes of 3D Video Representations

3D Video Processing Chain

❖ ***3D Video Coding***

- Many different techniques have been proposed for 3D video coding in recent decades. the most useful tools for this purpose are:

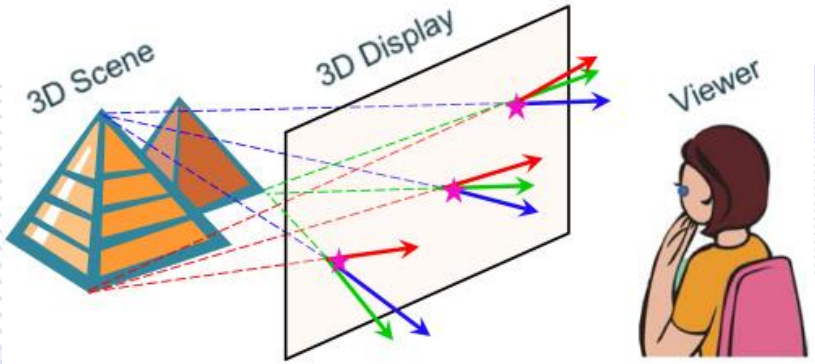
- **Multiview Video Coding (MVC)**
- **Multiview High Efficiency Video Coding (MV-HEVC)**
- **3D - High Efficiency Video Coding (3D-HEVC)**

3D Video Processing Chain

❖ ***Rendering stage - 3D Displays***

- The final piece (**Rendering stage** and **3D Displays**) of a 3D video system is vital to the end user, because it projects the 3D content.

- The **Rendering Stage** utilizes algorithms to render the data stored in the representation format.



3D Displays [G13]

3D Video Coding and Broadcasting

- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

Video Coding Standards

- A ***video coding standard*** is a decoding method for ***video compression*** and a descriptive document for a bitstream structure

- ***Video coding standards*** is the main toolkit utilized for the **compression process**

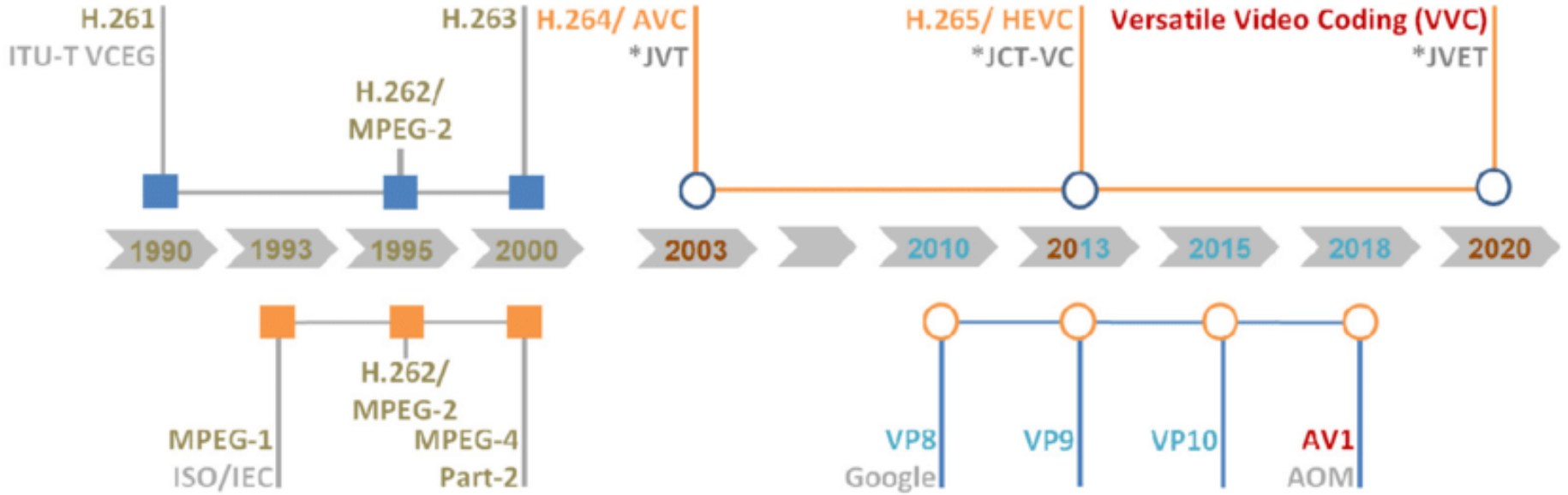
Video Coding Standards

❖ The most important ***video encoding standards*** are:

- ***MPEG – 2***
- ***MPEG - 4***
- ***AVC / H.264***
- ***HEVC / H.265***
- ***VVC***

Video Coding Standards

- ***Video coding standards development timeline ([PA20])***



Advanced Video Coding (AVC)

Advanced Video Coding (AVC), also known as H.264 or MPEG-4 Part 10, Advanced Video Coding (MPEG-4 AVC), is a video compression standard that uses block-oriented, motion-compensated integer-DCT coding.

High Efficiency Video Coding

- ***High-Efficiency Video Coding* (*HEVC*)**, (***H.265*** or **MPEG-H Part 2**), is the following coding standard to the popular AVC designed in the process of the MPEG-H project.

- ***HEVC***, in comparison to the ***AVC***, provides up to 50% better data compression at the same or even better level video quality at the same bit rate (bit/s).

High Efficiency Video Coding

- Some of the **new features of HEVC** are:

1. **Improved basic coding structure** called ***Coding Tree Unit (CTU)*** with size up to 64 × 64 pixels

2. **New directional modes** for ***intra-picture prediction***

3. ***Content-Adaptive Binary Arithmetic Coding (CABAC)*** as the only entropy coding

4. ***Sample Adaptive Offset (SAO)*** filter

Versatile Video Coding (VVC)

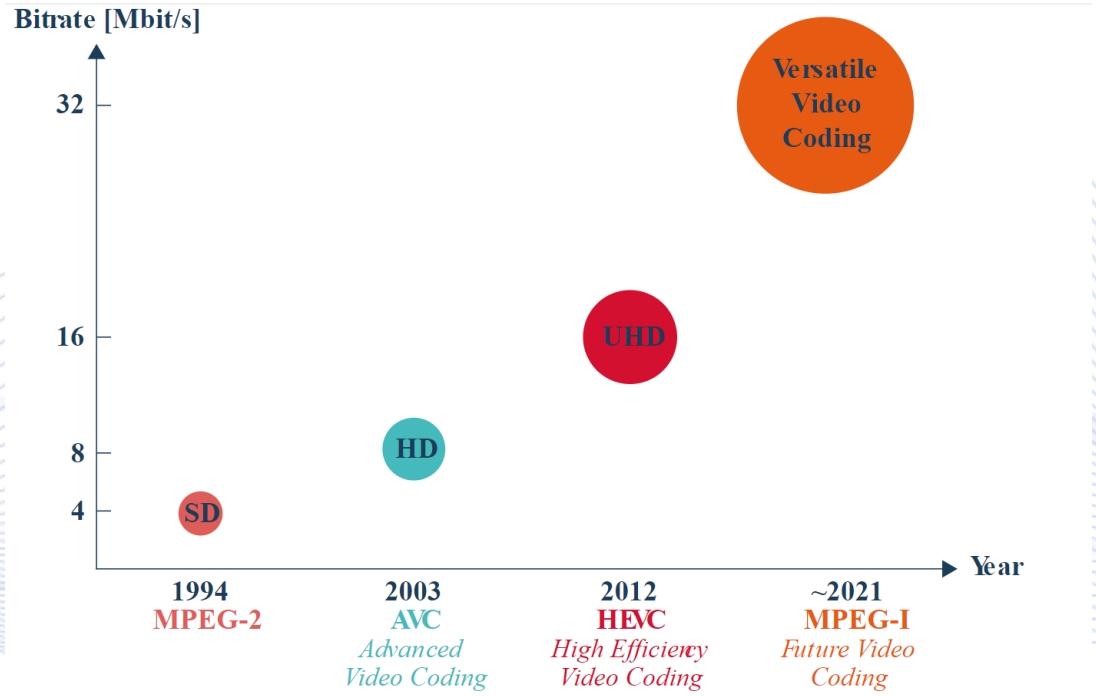
1. ***Versatile Video Coding (VVC),*** also known as ***H.266***, ***ISO/IEC 23090-3***, ***MPEG-I Part 3*** and ***Future Video Coding (FVC)*** developed by the **Joint Video Experts Team (JVET)** in July 6th 2020.

2. The goal is to commercialize and sell **4K streaming and broadcast.**

3. We expected the **decoding complexity** of VVC to be about **twice** that of HEVC.

Video Coding Standards

- ***Video quality (Bitrate)*** in different ***encoding techniques*** ([CIP-HSE])***.***



3D Video Coding and Broadcasting

- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

3D Video Coding

- ***Coding, transmission and decoding*** are significant aspects which influence the overall quality of the experience of the 3D video.

- We defined the 3D coding based on the different formats and the tools required to encode the 3D domain.

3D Video Coding and Broadcasting

- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

3D Coding by Format

Description of the ***coding process*** in the next 3 formats:

- ***Stereoscopic Frame-Compatible Format***
- ***Frame Sequential Stereoscopic Format***
- ***Video Plus Depth (V + D)***

Stereoscopic Frame-Compatible Format

- ***Frame-compatible formats*** depict the arrangement of the left and right images in a spatial multiplex, resulting in an image that can be treated like a normal **High Definition TV** (HDTV) image by the demodulator and compression decoder in the Receiver.

Stereoscopic Frame-Compatible Format

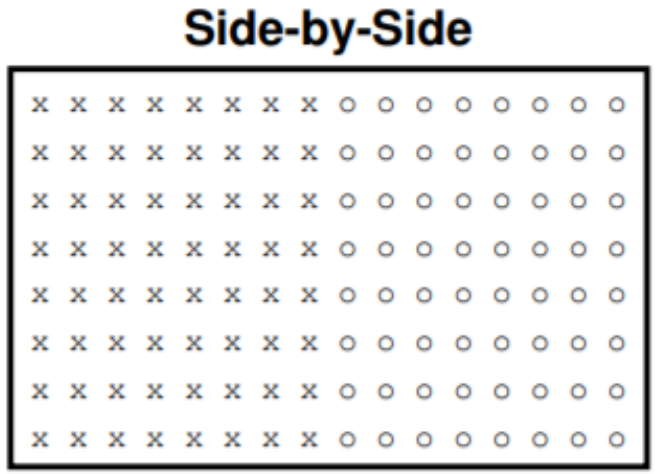
These ***two views*** can be transmitted at a lower resolution in the same space positioned, either:

- **Side-by-Side**
- **Top-and-Bottom**

Stereoscopic Frame-Compatible Format

❖ ***Side-By-Side (SBS)***

The ***SbS Format*** is the arrangement of the frame-compatible spatial multiplex such that the horizontally anamorphic left-eye picture is placed in spatial multiplex to occupy the first half of each line, and the right-eye picture is placed in the spatial multiplex to occupy the second half of each line.



Side-By-Side format ([VA10])

Stereoscopic Frame-Compatible Format

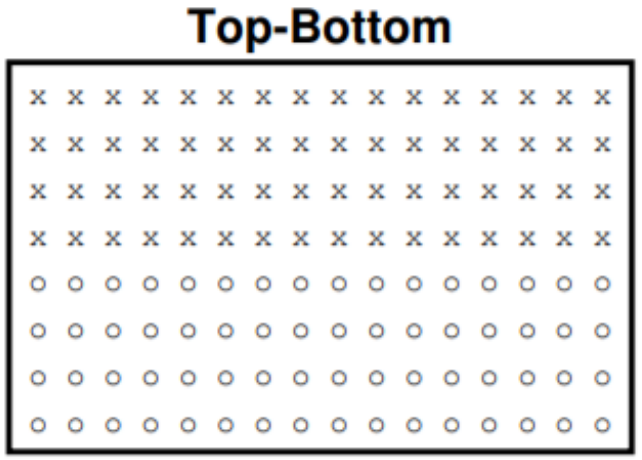
Example: **1280 × 800 pixels** in a full frame the left view and the right **view multiplexed** in the frame consisting of **640 × 800 pixels each**.

Side-By-Side format ([AESE-HVCF])

Stereoscopic Frame-Compatible Format

❖ ***Top-and-Bottom***

The ***TaB*** format is the arrangement of the frame-compatible spatial multiplex such that the vertically anamorphic left-eye picture is placed in the spatial multiplex to occupy the first (top) half of a single HDTV video frame, and the right-eye picture is placed in the spatial multiplex to occupy the second (bottom) half of a single HDTV video frame ([E12])



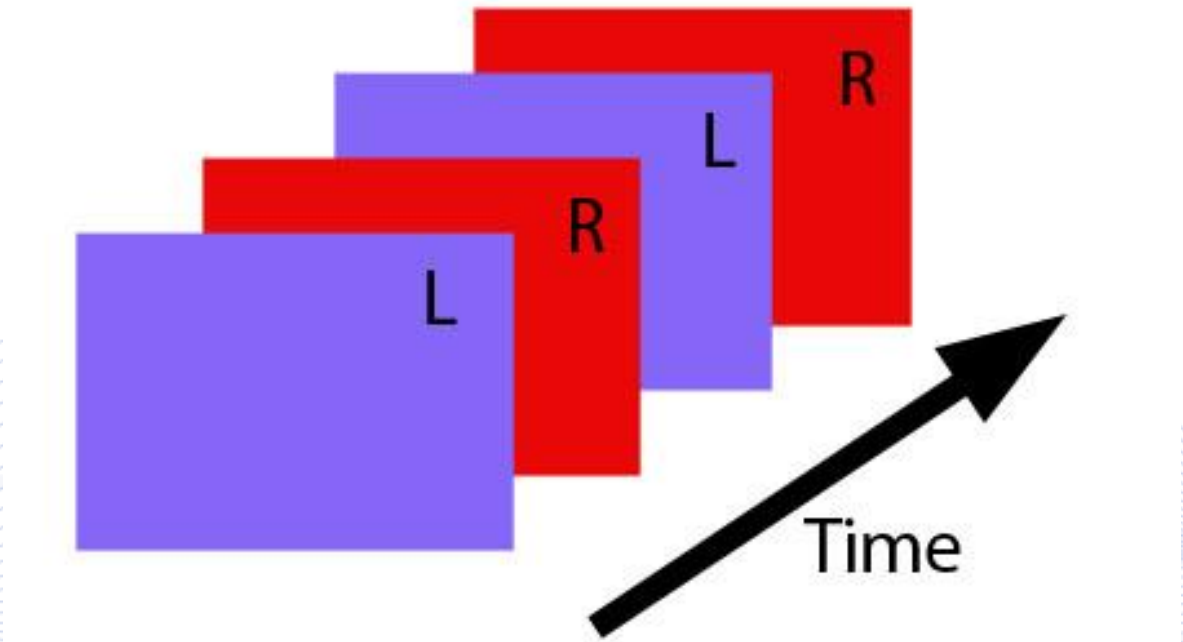
Top-and-Bottom Format([VA10])

Frame Sequential Stereoscopic Format

- A ***frame sequential signal*** in full spatial resolution picture is carried at typically *120 frames per second*, corresponding to a 60 Hz progressive scanning format for each view(2).

- The frames are formed ***into a sequence*** and received as left/right/left/… frames.

Frame Sequential Stereoscopic Format



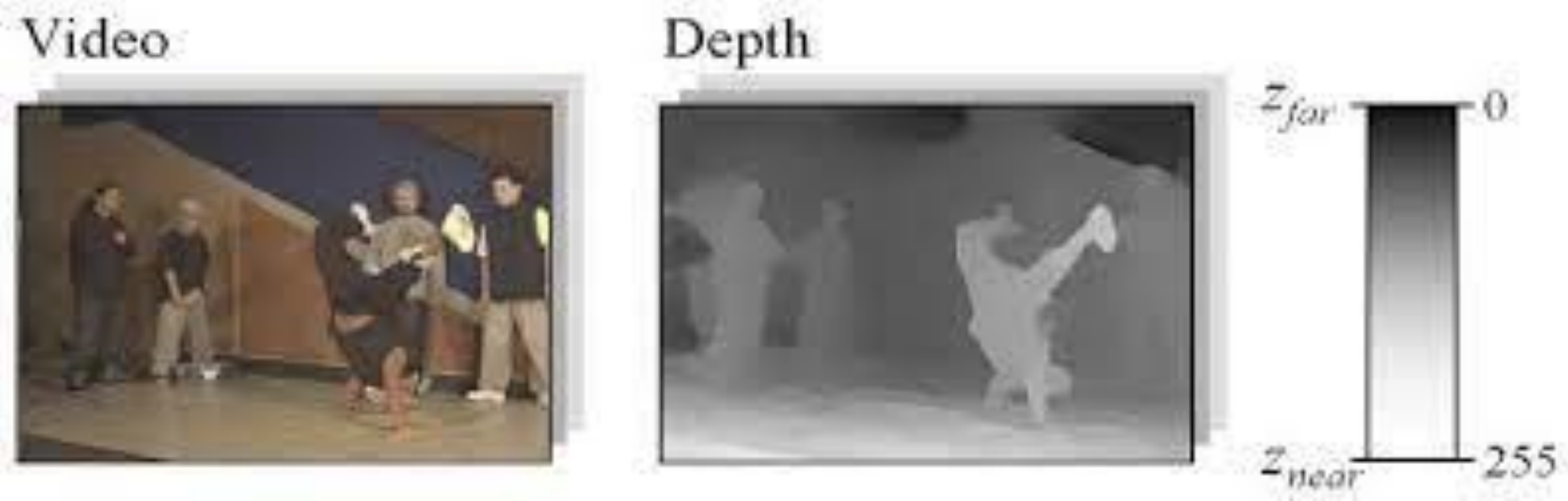
Frame Sequential Stereoscopic Format ([WIKI-FS])

Video Plus Depth (V + D)

- Video plus depth (one stream with associated depth map, marked as “V + D”) is another type of 3D formats and it involves a video signal and a per-pixel depth map.

- Additional depth data (either monochromatic or luminance video signal) for each pixel can be generated from either calibrated stereo or multi-view video according to its depth estimation.

Video Plus Depth (V + D)



Video Plus Depth ([MP07])

3D Video Coding and Broadcasting

- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

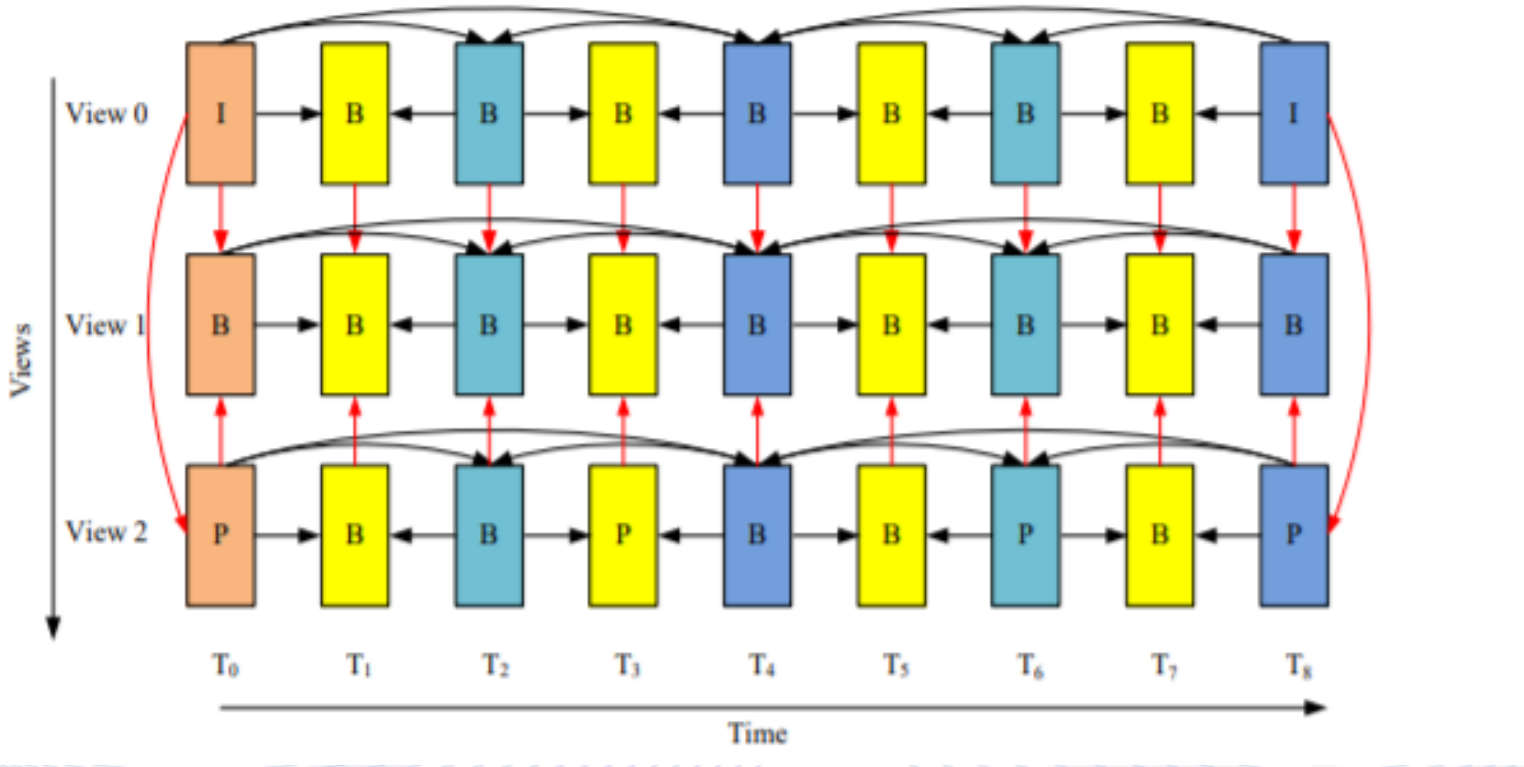
Multiview Video Coding (MVC)

- ***Multiview Video Coding*** (MVC or MVC 3D) is a stereoscopic video coding standard developed cooperatively by MPEG and VCEG and is an modification to the H.264 (MPEG-4 AVC) video compression standard.

- This method is particularly efficient for encoding video sequences, captured at the same time from multiple camera angles in a single video stream.

- It uses the 2D plus Delta method.

Multiview Video Coding (MVC)



Typical MVC prediction structure
[KKA12]

Multiview High Efficiency Video Coding (MV-HEVC)

- **MVC**, **Multiview High Efficiency Video Coding (MV-HEVC)** is based on the single-view variant of the codec.

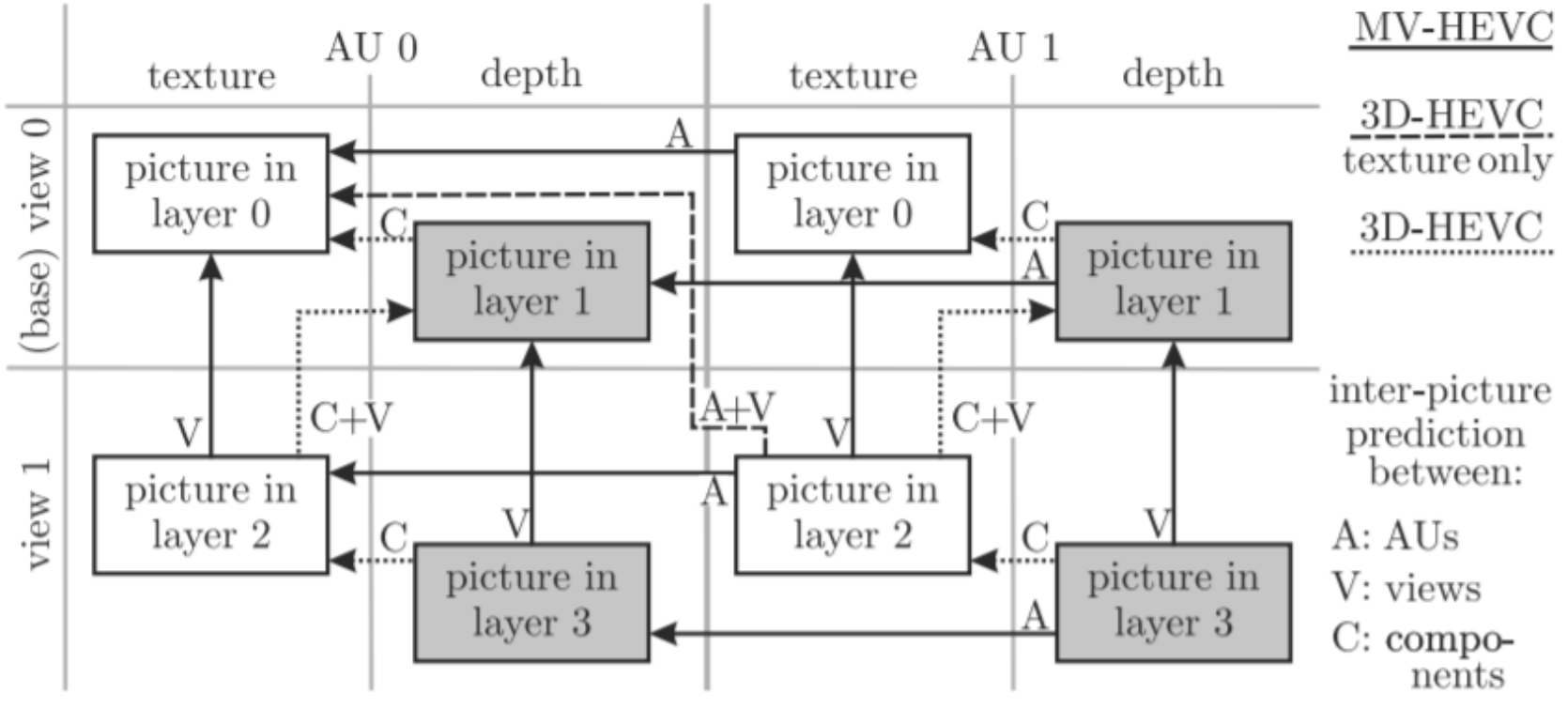
- The coding toolkit of **MV-HEVC** are the same as in **HEVC**, and the main adjustments are made in syntax only, in order to provide the support of encoding to more than one view.

3D-HEVC

- ***3D-HEVC*** computes the state-of-the-art technology for the compression of 3D video with accompanying depth.

- This technology not only ***compresses*** the depth but also ***exploits*** the depth in order to improve coding performance of the Multiview video.

3D-HEVC



MV-HEVC, 3D-HEVC texture-only and 3D-HEVC texture-and-depth – The typical coding structure and picture dependencies ([TG15])

3D Video Coding and Broadcasting

- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

3D Video Transmission

- With the quick development of different 3D video content, its transmission and delivery to the end user become a challenging issue to the existing network systems due to the ever-increasing capacity needs, different errors that may occur in the transmission chain, used format of 3D content, etc.

- This chapter describes several **video transport technologies** that are compatible with most of the existing **3D video formats**.

3D Video Coding and Broadcasting

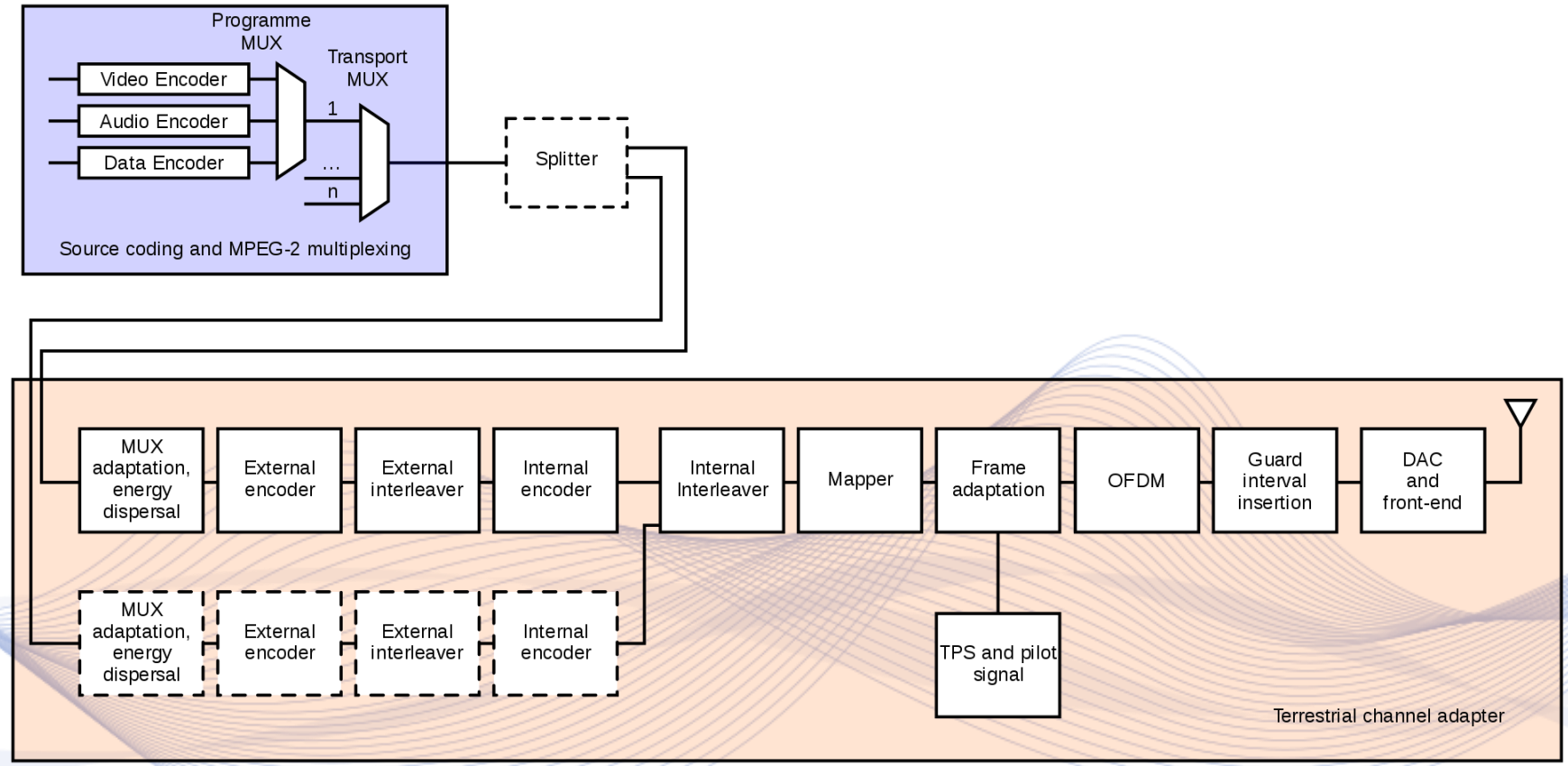
- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

3D Video in DVB Systems

- ***Digital Video Broadcasting (DVB)*** is a set of international open standards for digital television.

- ***DVB standards*** are maintained by the DVB Project, an international industry consortium, and are published by a ***Joint Technical Committee (JTC)*** of the ***European Telecommunications Standards Institute (ETSI), European Committee for Electrotechnical Standardization (CENELEC)*** and ***European Broadcasting Union (EBU).***

3D Video in DVB Systems



The chain of a ***DVB-T transmission system*** ([WIKI-DVB])

3D Video in DVB Systems

- Originally, ***MPEG-2 Systems*** were developed in order to be used in the transmission of 2D video but it has since undergone amendments to support the transmission of 3D video in stereo, multi-view video, and in ***texture plus depth formats***.

- Encoded 3D stereo video can be transported in several different ways over ***MPEG-2 TS systems***, some choices depending on how the left and right views are encoded.

3D Video Coding and Broadcasting

- Introduction
- 3D Video Processing Chain
- Video Coding Standards
- 3D Video Coding
  - 3D Coding by Format
  - 3D Coding Tools
- 3D Video Transmission
  - 3D Video in DVB Systems
  - 3D Video Delivery Over IP

3D Video Delivery Over IP

- ***Streaming over IP*** is a flexible transport system for multi-view video since **the transmission bitrate can be configured** according to requirements of video format and user equipment

- ***Streaming systems*** can be classified as:

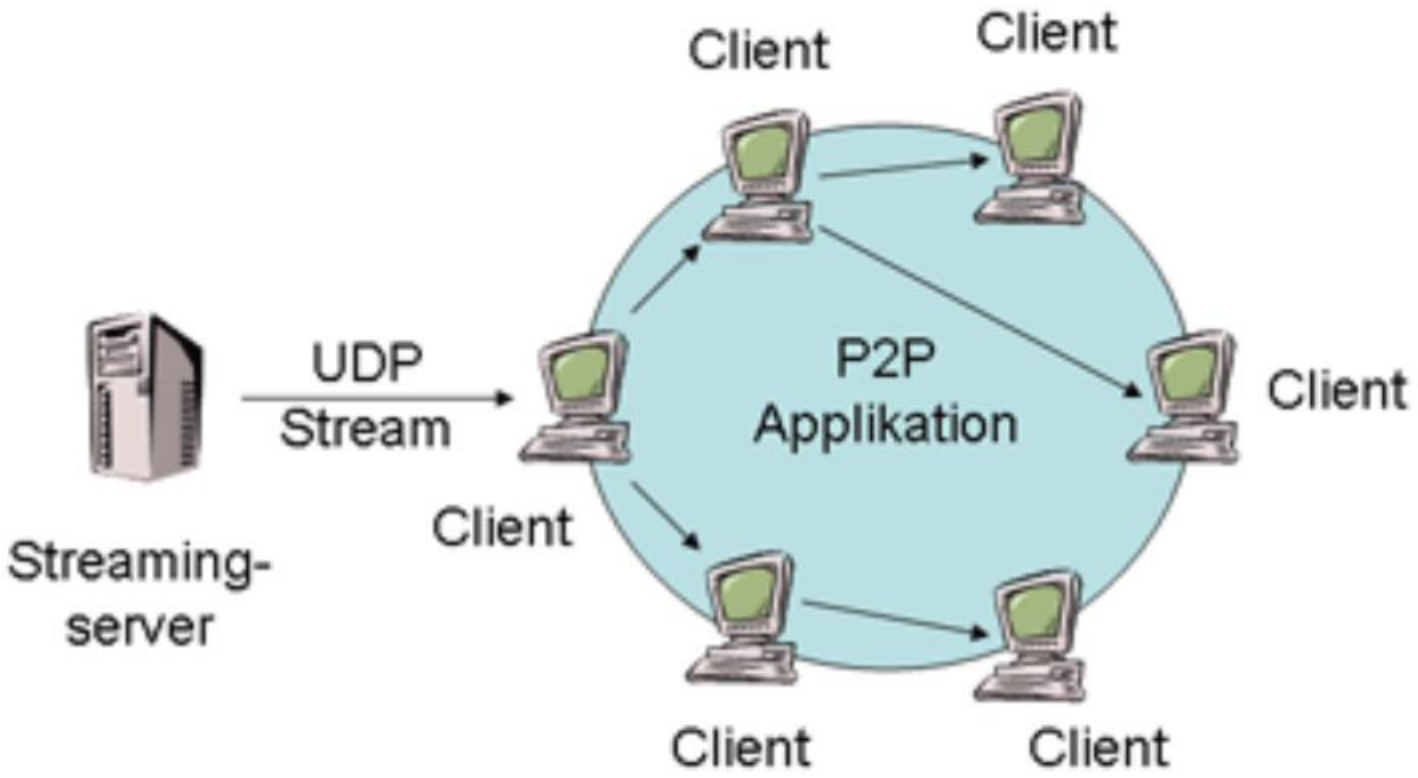
I. Server-Client systems
II. Peer-to-Peer (P2P) systems.

3D Video Delivery Over IP

❖ ***Server-Client Streaming System***

- That system consists of a **streaming server** and a **client** that communicate using a set of standard protocols.
- In the server-client model, typically a different stream is sent to each client.

3D Video Delivery Over IP



P2P streaming system ([CSG-P2P])

3D Video Delivery Over IP

❖ ***HTTP streaming***

- **HTTP streaming** includes support for adaptive streaming (bitrate switching) to allow clients to dynamically switch between streams of varying quality and chunk size during playback, in order to adapt to changing network conditions and available **central processing unit (CPU)** resources.

3D Video Delivery Over IP

❖ ***Dynamic Adaptive Streaming over HTTP***

- **MPEG-DASH (Dynamic Adaptive Streaming over HTTP)** is an international standard, published in April 2012, for adaptive bitrate **HTTP-based** streaming that is audio/video codec agnostic in order to address this fragmentation problem.

Bibliography

[PIT2021] I. Pitas, “Computer vision”, Createspace/Amazon, in press.
[PIT2017] I. Pitas, “Digital video processing and analysis ” , China Machine Press, 2017 (in Chinese).
[PIT2013] I. Pitas, “Digital Video and Television ” , Createspace/Amazon, 2013.
[NIK2000] N. Nikolaidis and I. Pitas, 3D Image Processing Algorithms, J. Wiley, 2000.
[PIT2000] I. Pitas, “Digital Image Processing Algorithms and Applications”, J. Wiley, 2000.

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr