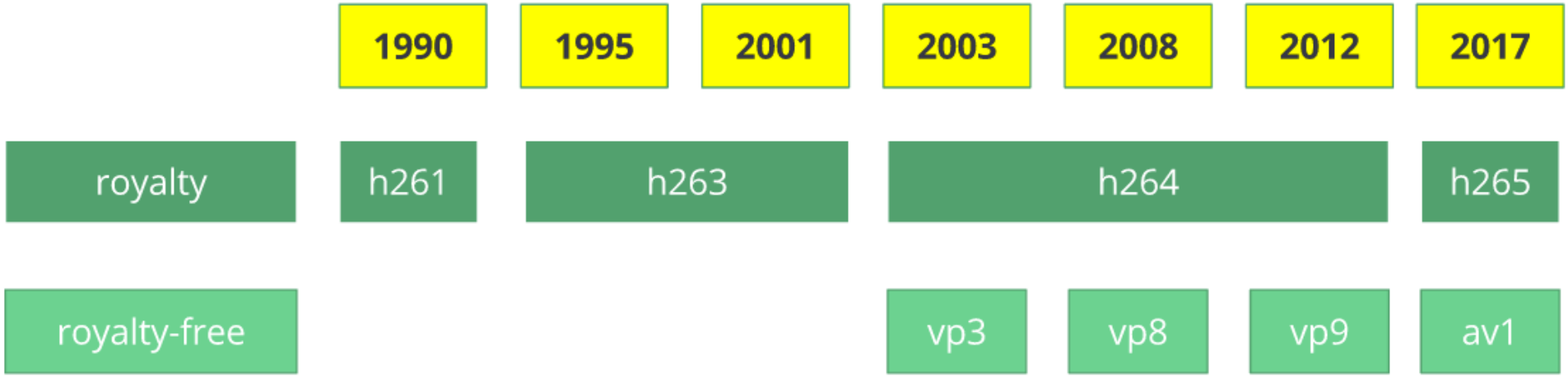
upf. Universitat Pompeu Fabra Barcelona

T3 – MPEG4, h.264 & Broadcasting standards

upf. Universitat Pompeu Fabra Barcelona

MPEG4

Moving Picture Experts Group


1990
1995
2001
2003
2008
2012
2017
royalty
h261
h263
h264
h265
royalty-free
vp3
vp8
vp9
av1

- **Introduced in late 1998**

- **Designated a standard for a group of audio and video coding formats and related technology agreed upon by the ISO/IEC Moving Picture Experts Group (MPEG) (ISO/IEC JTC1/SC29/WG11) under the formal standard ISO/IEC 14496 – Coding of audio-visual objects**

- **Uses of MPEG-4 include compression of AV data for web (streaming media) and CD distribution, voice (telephone, videophone) and broadcast television applications.**

·HDTV appears (720p and 1080p)

720p

- 1280×720 pixels
- Also called *HD ready* or *standard HD*
- It's a progressive HDTV signal format with 720 horizontal lines and an aspect ratio (AR) of 16:9, normally known as widescreen HDTV (1.78:1).

- **The frame rate is standards-dependent, and for conventional broadcasting appears in 50 progressive frames per second in former PAL/SECAM countries (Europe, Australia, others), and 59.94 frames per second in former NTSC countries (North America, Japan, Brazil, others).**

It can change or vary depending on the standard (i.e. broadcasting standards, end of this lesson)

1080p

- 1920×1080 pixels
- Also known as Full HD or FHD
- It's a set of HDTV. The p stands for progressive scan, i.e. non-interlaced.
- The term usually assumes a widescreen aspect ratio of 16:9, implying a resolution of 2.1 megapixels.

- Applications of the 1080p standard include television broadcasts, Blu-ray Discs, smartphones, Internet content such as YouTube videos and Netflix TV shows and movies, consumer-grade televisions and projectors, computer monitors and video game consoles. Small camcorders, smartphones and digital cameras can capture still and moving images in 1080p resolution.

(not to confuse with 720i and 1080i:

both have 720p and 1080p as resolution but they work with interlaced video)

·It was inspired in Apple QuickTime format

·It was inspired in Apple QuickTime format

- MPEG-4: A graphics and video compression algorithm standard that is based on MPEG-1 and MPEG-2 and Apple QuickTime technology. Wavelet-based MPEG-4 files are smaller than JPEG or QuickTime files, so they are designed to transmit video and images over a narrower bandwidth and can mix video with text, graphics and 2-D and 3-D animation layers. MPEG-4 was standardized in October 1998 in the ISO/IEC document 14496.

Introduced Video Object (VO) and Video Object Plane (VOP)

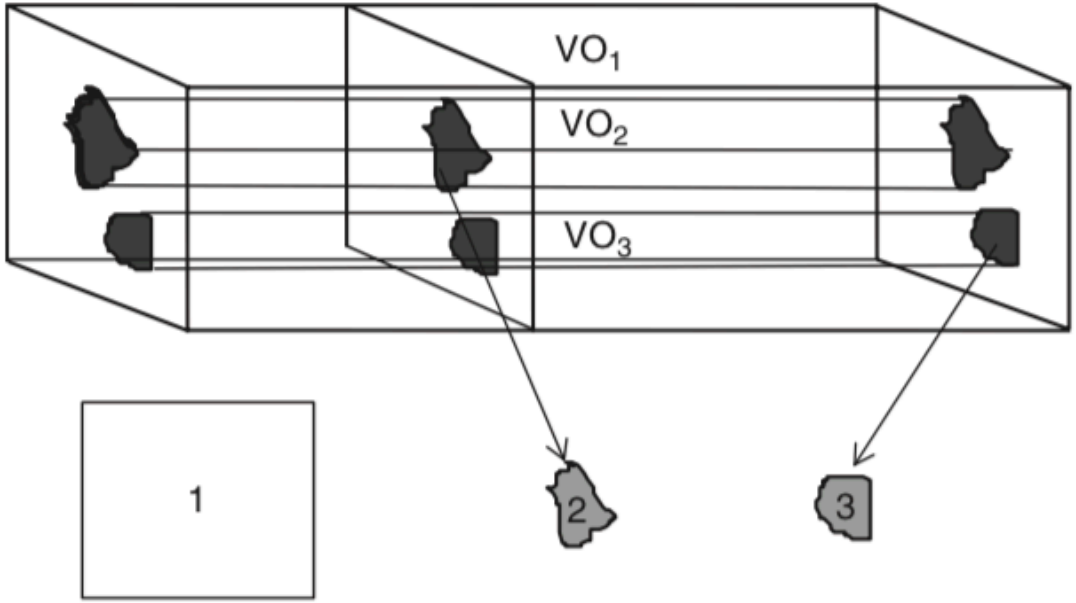


FIGURE 18.1
Video object definition and format.
(a) Video object, (b) VOPs.

Thanks to this, new enhancements:

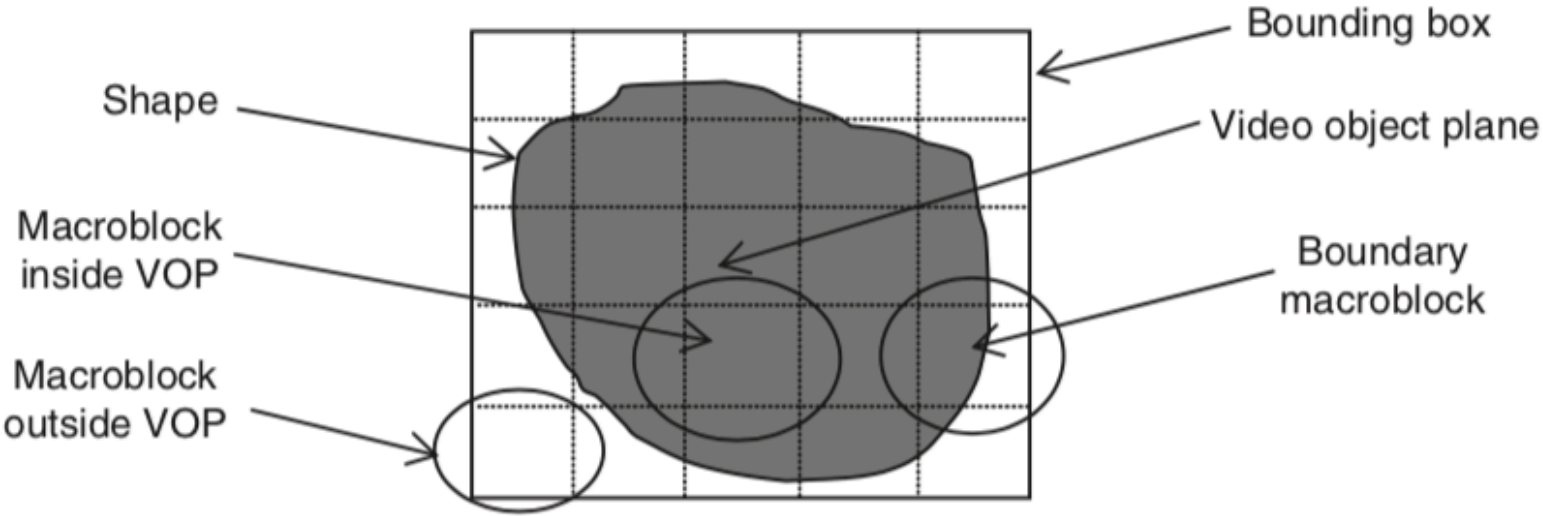


FIGURE 18.3
A VOP is represented by a bounding rectangular box.

- **Shape coding: MPEG-4 video coding is the first to make effort at providing a standardized approach to compress the shape information of objects and contain the compressed results within a video bitstream. Video data can be coded on an object basis…**

...The information in the video signal is decomposed to shape, texture, and motion. This information is then coded and transmitted within the bitstream. The shape information is provided in binary format or gray-scale format. The binary format of shape information consists of a pixel map

- **Sprite coding: A sprite is an especially composed VO that is visible throughout an entire piece of video sequence. For example, the sprite generated from a panning sequence contains all the visible pixels of the background throughout the video sequence. Portion of the background may not be seen in certain frames due to the occlusion**

- **Interlaced video coding**

- **Wavelet based texture coding**

·Generalized Spatial and Temporal Scalability

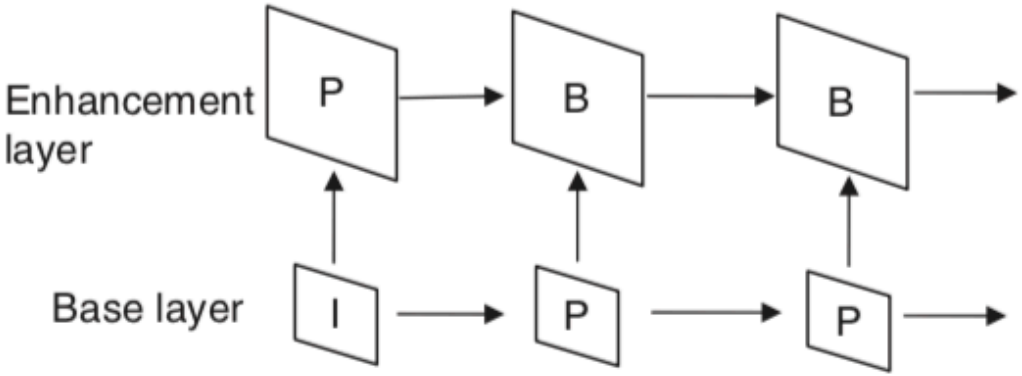


FIGURE 18.11
Illustration of spatial scalability.

·Generalized Spatial and Temporal Scalability

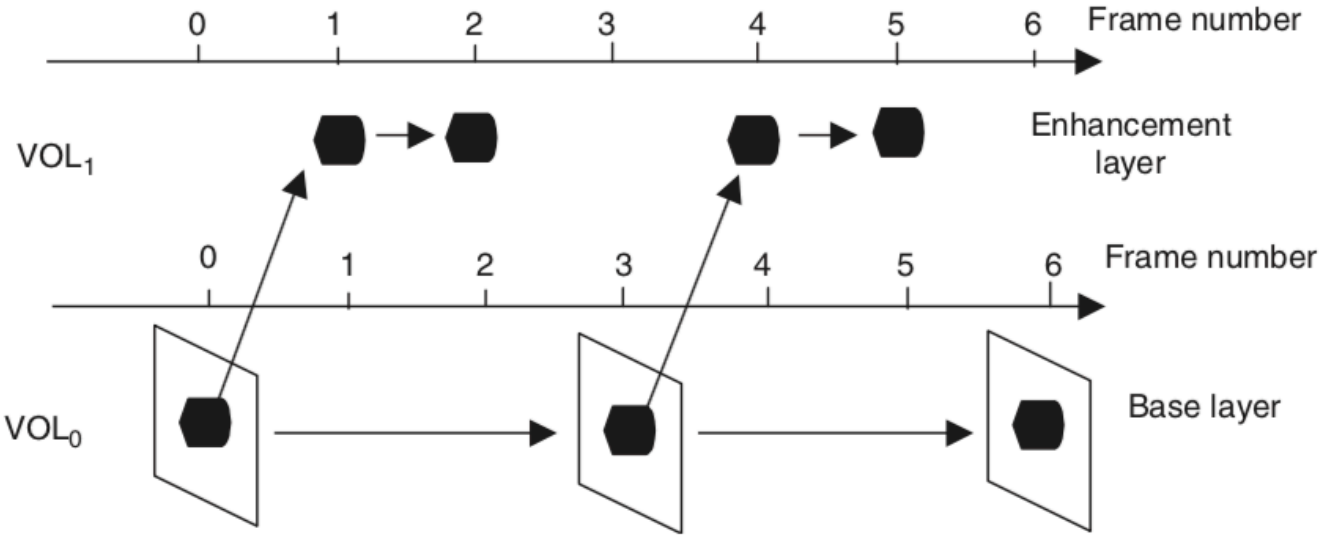
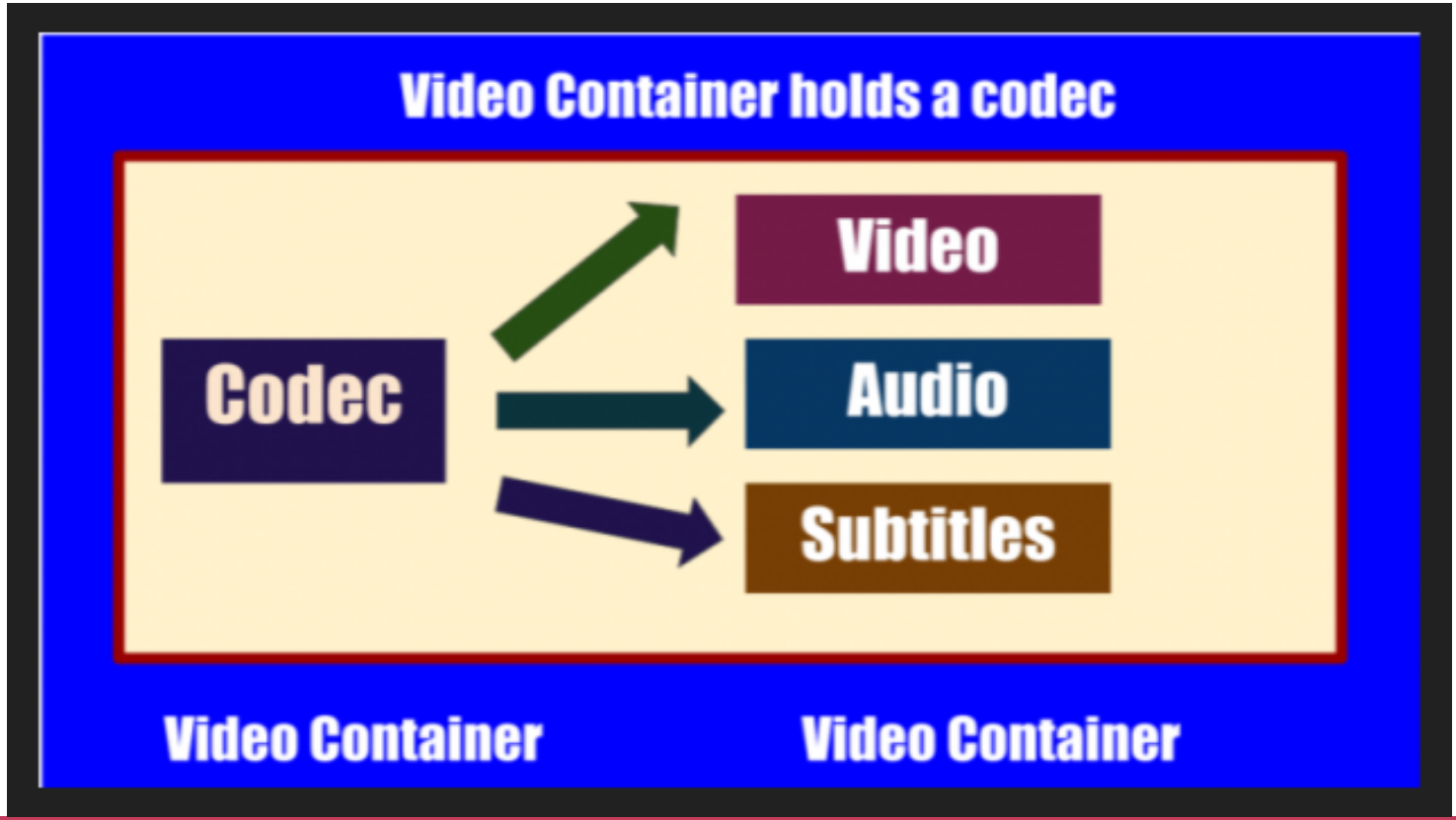


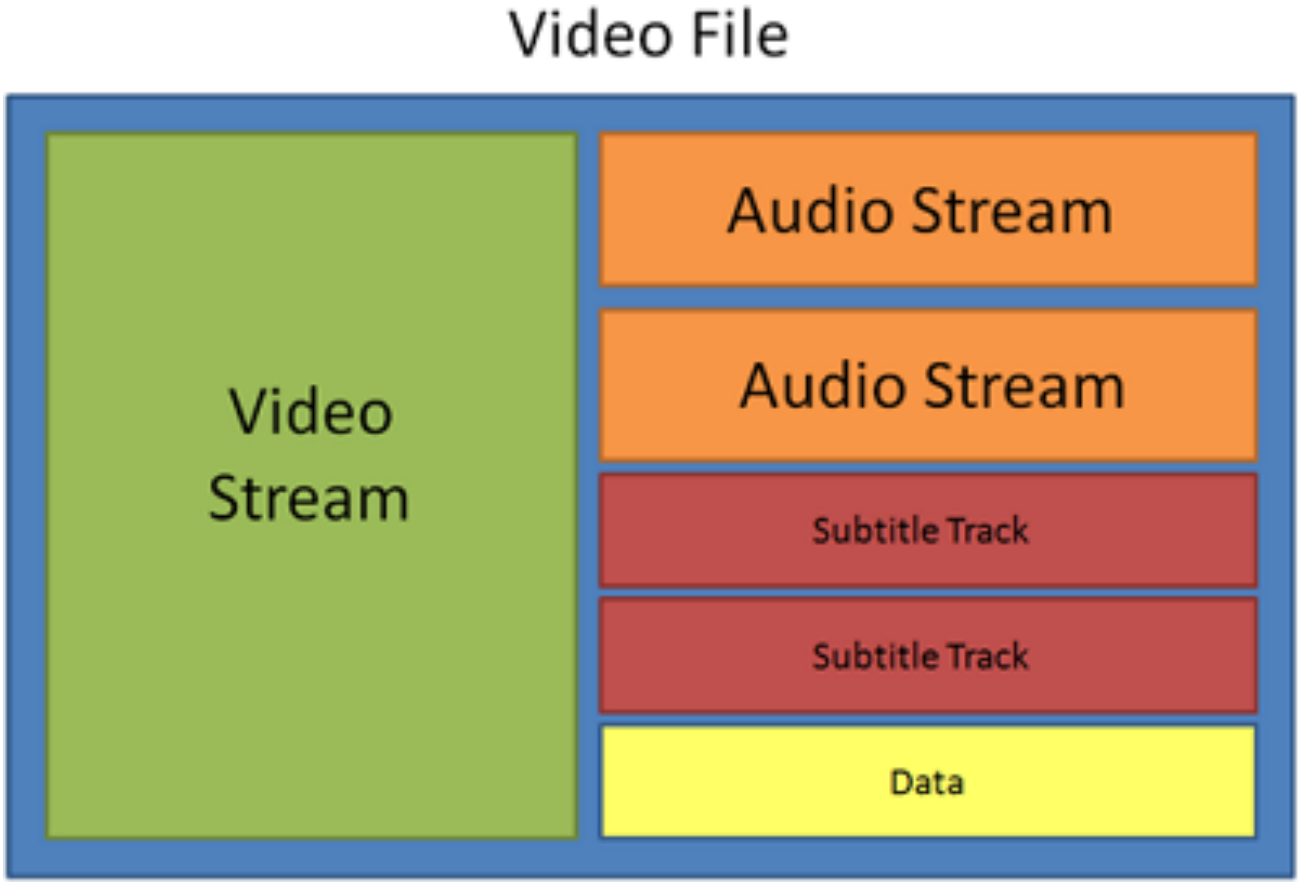
FIGURE 18.12
An example of temporal scalability.

- **Error resilience: MPEG4 provides error robustness and resilience to allow access of image and video data over a wide range of storage and transmission media**

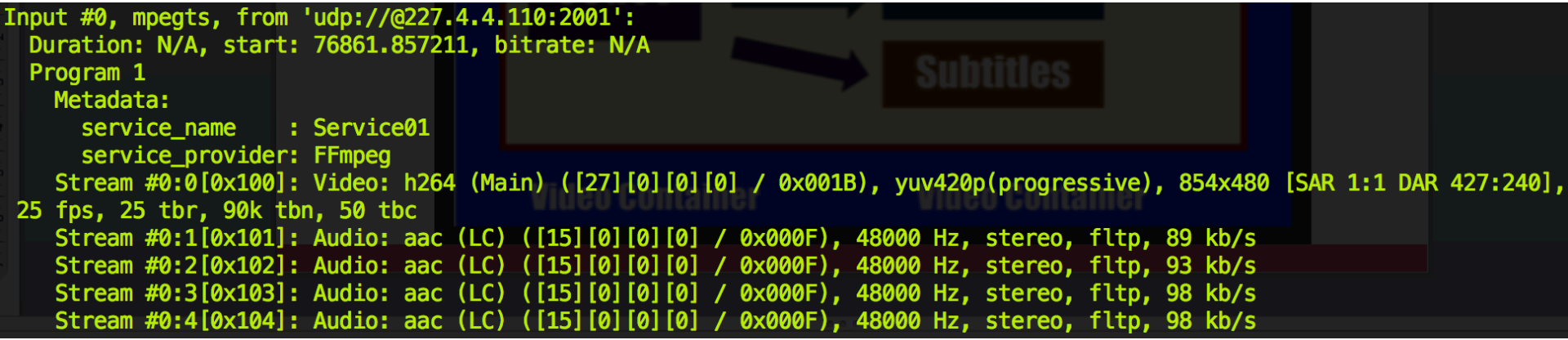
.mp4 File Format: introducing the container concept

One file can have multiple tracks as much as we want


Video Container holds a codec
Codec
Video
Audio
Subtitles
Video Container
Video Container


Video File
Video Stream
Audio Stream
Audio Stream
Subtitle Track
Subtitle Track
Data

Universitat Pompeu Fabra Barcelona



When we used to see .mpg or .mp2 files, those where unique videos with 1 track encoded with mpeg and mpeg2 codec respectively.

Now the container lets us to mix codecs.

I.e.: 1 MPEG video track, 2 MP3 audio tracks, 1 AAC audio track, 1 close caption/subtitles track

Also, MPEG-4 introduces the possibility of adding DRM

What's DRM?

Digital rights management (DRM) tools or technological protection measures (TPM) are a set of access control technologies for restricting the use of proprietary hardware and copyrighted works. DRM technologies try to control the use, modification, and distribution of copyrighted works (such as software and multimedia content), as well as systems within devices that enforce these policies

It’s a completely universe. You can find DRM by software and DRM by hardware.

It’s all about encryption and computing, if you’re more interested in the topic:

https://en.wikipedia.org/wiki/Digital_rights_management

...and the MPEG4 introduced for the first time a new codec for video

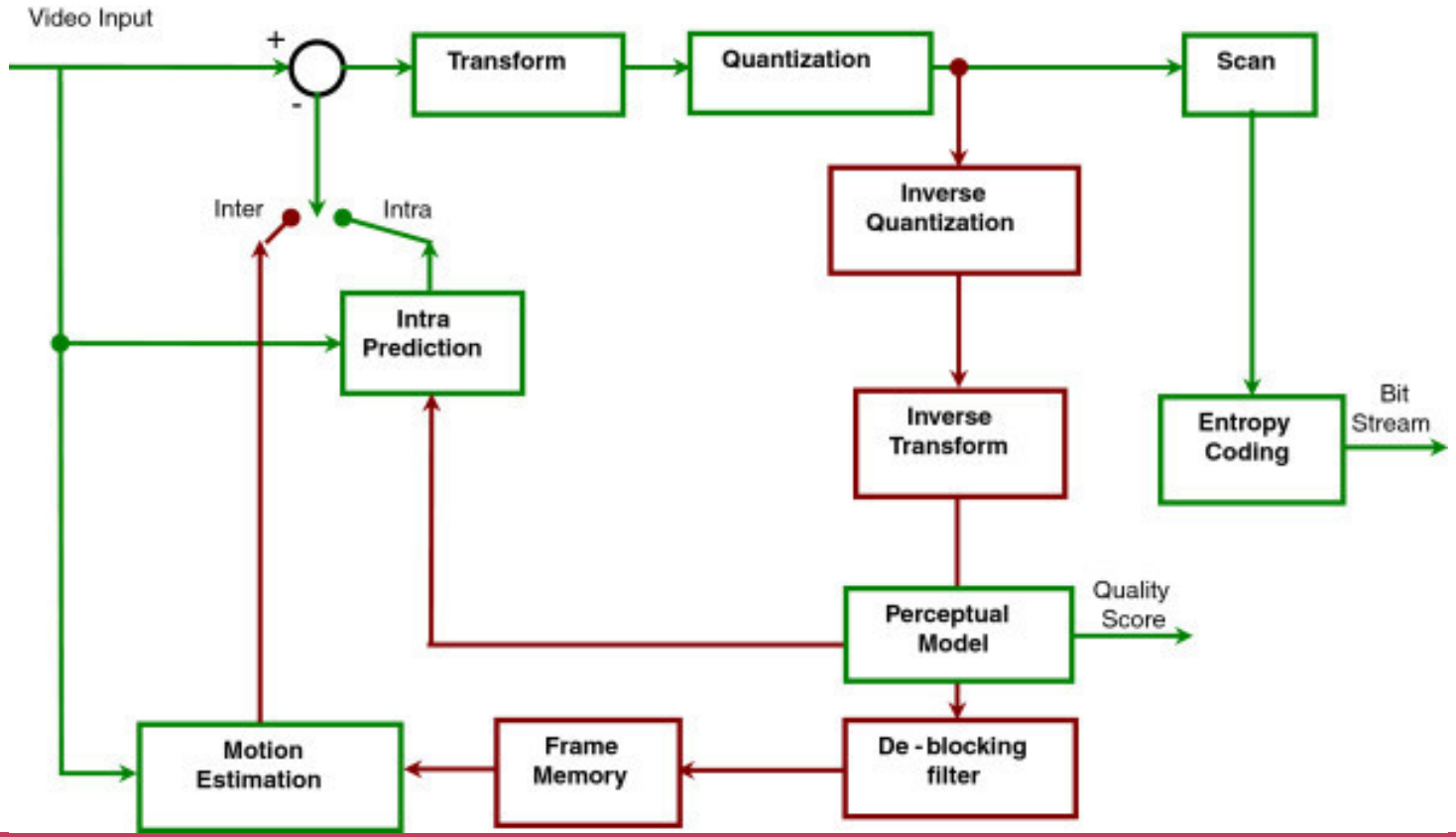
Universitat Pompeu Fabra Barcelona

h.264

Also known as AVC (Advanced Video Coding) or MPEG-4 Part 10, Advanced Video Coding (MPEG-4 AVC), is a video compression standard based on block-oriented, motion-compensated integer-DCT coding. It is by far the most commonly used format for the recording, compression, and distribution of video content, used by 91% of video industry developers as of September 2019. It supports resolutions up to and including 8K UHD (*only for eagles*)

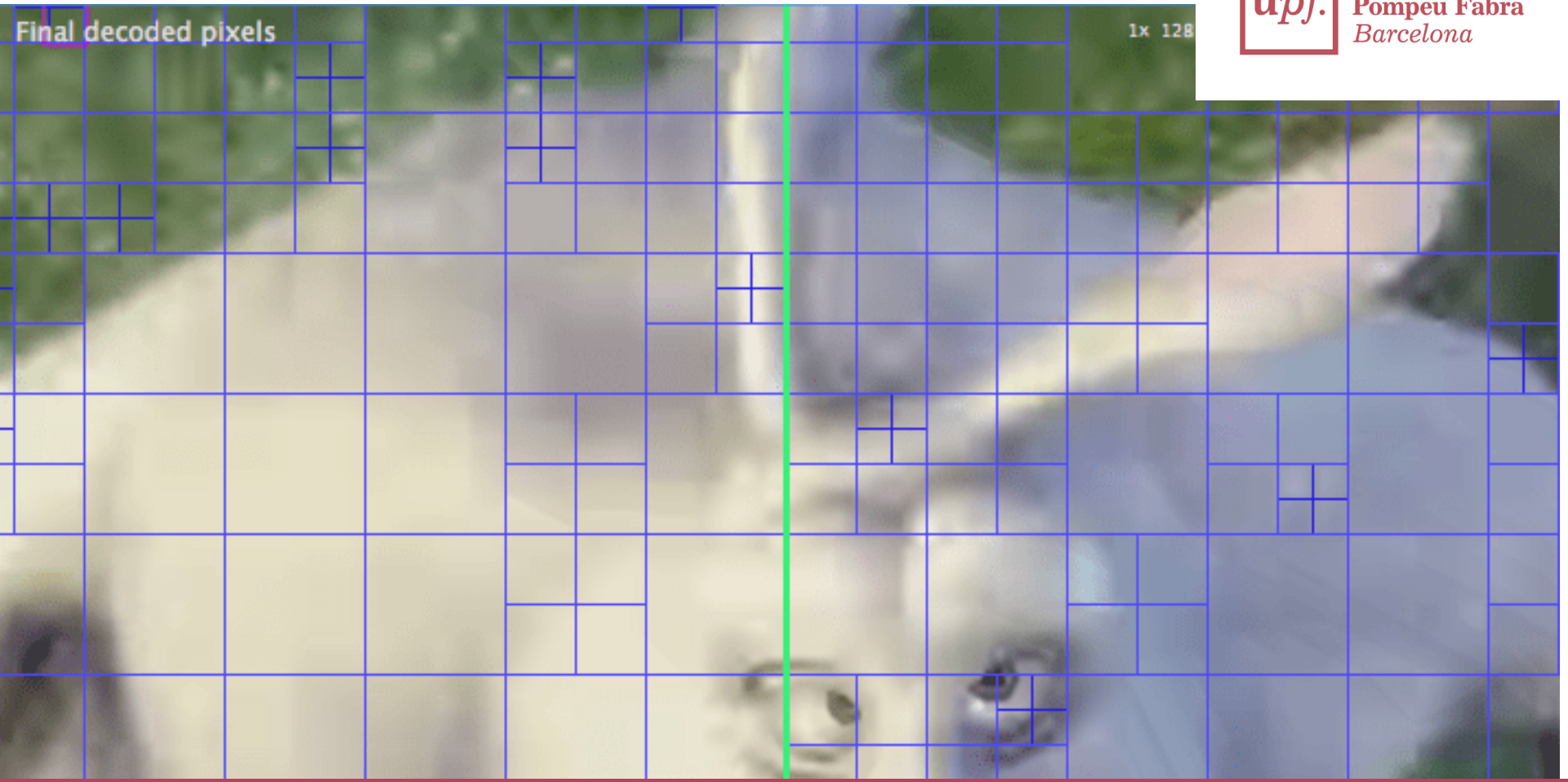
As MPEG2 was born due to achieve a broadcasting tech upgrade: compress interlaced broadcasting, and get digital video into satellite, cable, etc.;

MPEG4 was born just to get better compression bitrate


Video Input
+
-
Transform
Quantization
Scan
Inverse Quantization
Inter
Intra
Intra Prediction
Inverse Transform
Entropy Coding
Bit Stream
Perceptual Model
Quality Score
Motion Estimation
Frame Memory
De - blocking filter

2 main advantages:

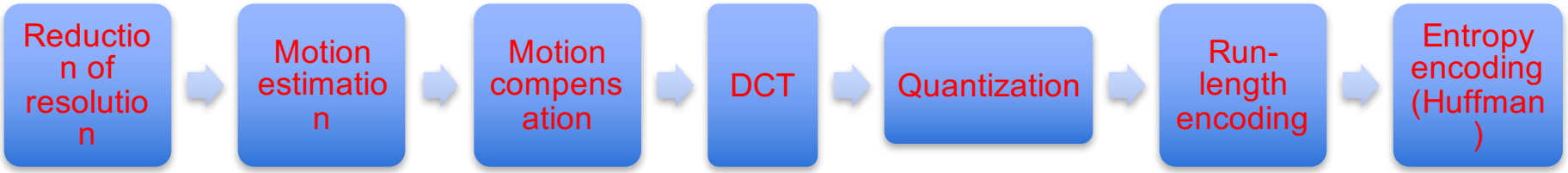
1) Motion Compensation: Instead of 16x16 macroblocks, it has now variable block size


Final decoded pixels
1x 128

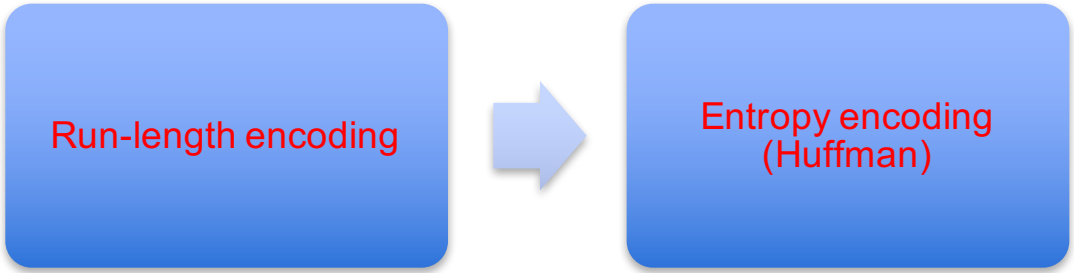
2 main advantages:

2) New Entropy Coding at the end of the workflow

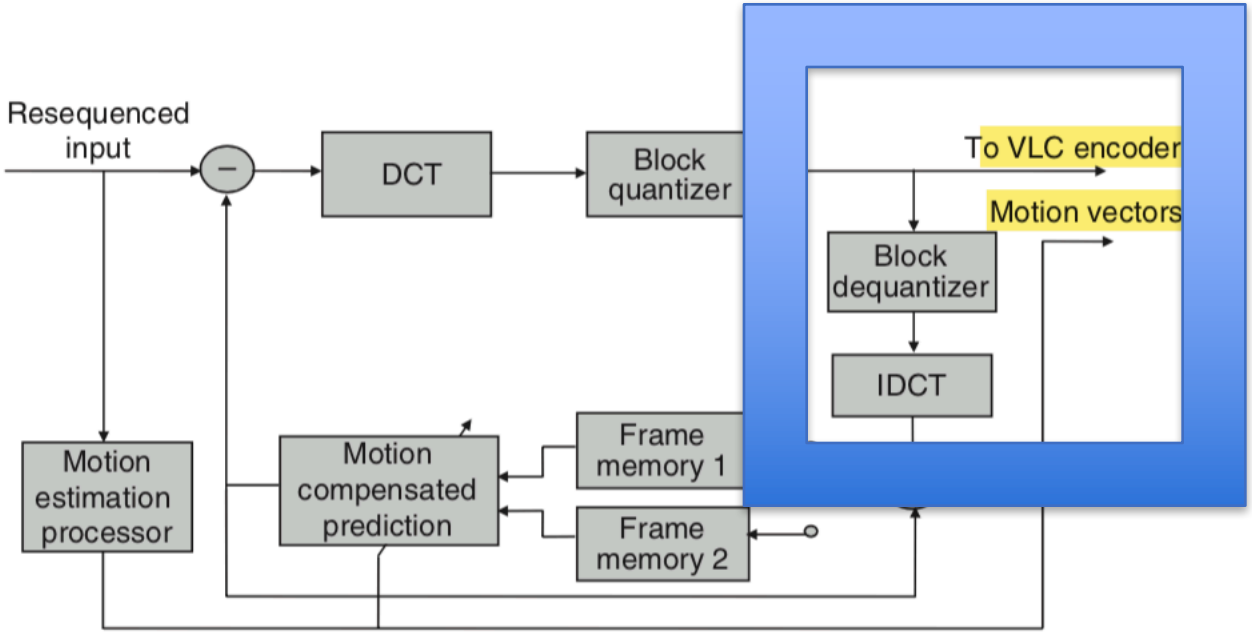
Remember MPEG2 Workflow?



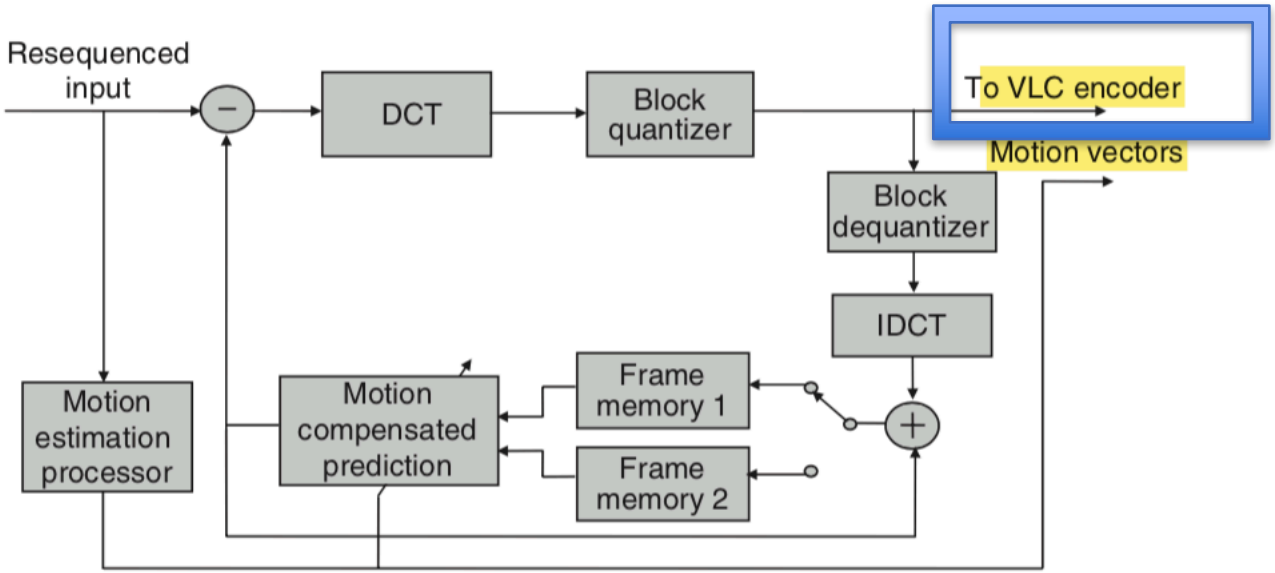
If we focus on the last steps



If we focus on the last steps



If we focus on the last steps



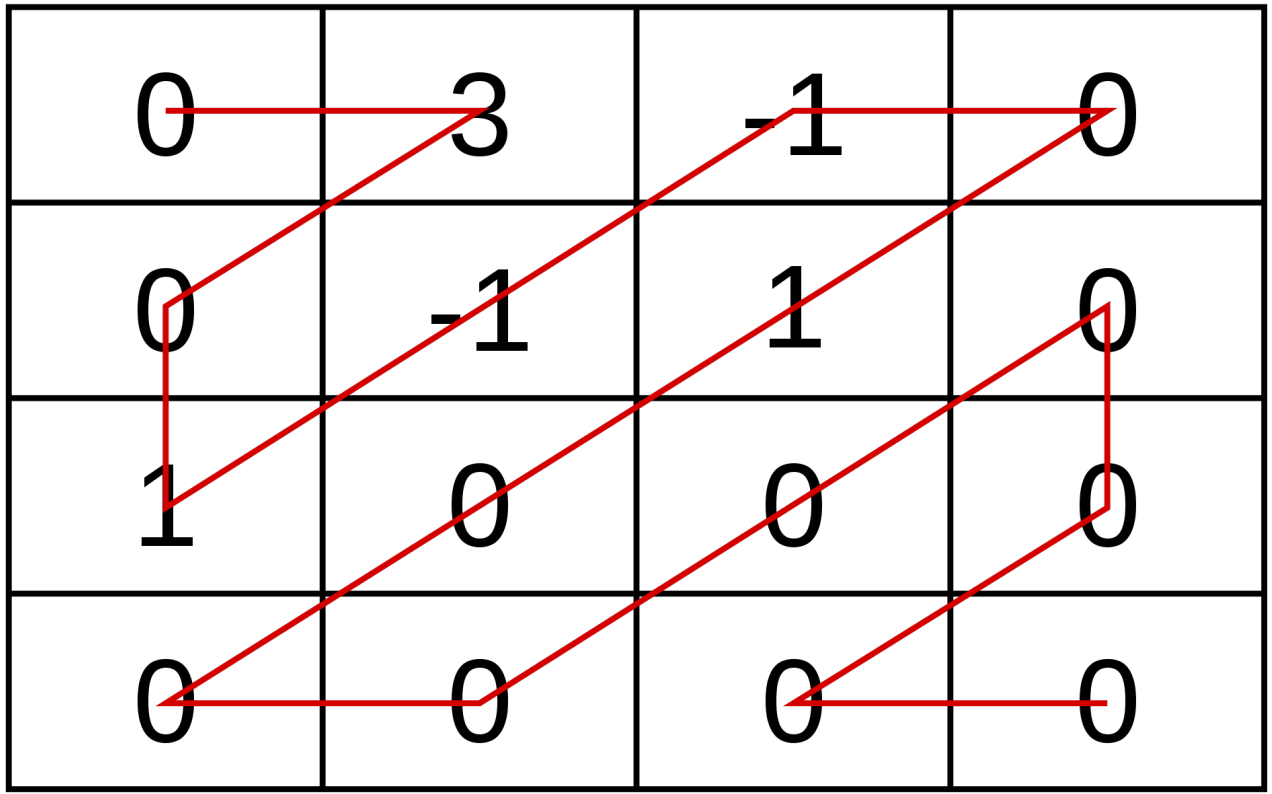
MPEG1 & MPEG2 used a variable length coding based on Hufman

MPEG4 introduces 2 algorithms: CAVLC and CABAC

They will substitute the RLE, and both are lossless too

4x4 block:

CAVLC



- **Context-adaptive variable-length coding (CAVLC) is a form of entropy coding used in H.264/MPEG-4 AVC video encoding.**

- **It is an inherently lossless compression technique, like almost all entropy-coders.**

- **In H.264/MPEG-4 AVC, it is used to encode residual, zig-zag order, blocks of transform coefficients**

- **CAVLC is supported in all H.264 profiles, unlike CABAC which is not supported in Baseline and Extended profiles.**

- **CAVLC is used to encode residual, zig-zag ordered 4x4 (and 2x2) blocks of transform coefficients. CAVLC is designed to take advantage of several characteristics of quantized 4x4 blocks (*not sure about this*)**

0,3,0,1,-1,-1,0,1,0… TotalCoeffs = 5 (indexed from highest frequency [4] to lowest frequency [0])

TotalZeros = 3

T1s = 3 (in fact there are 4 trailing ones but only 3 can be encoded as a "special case")

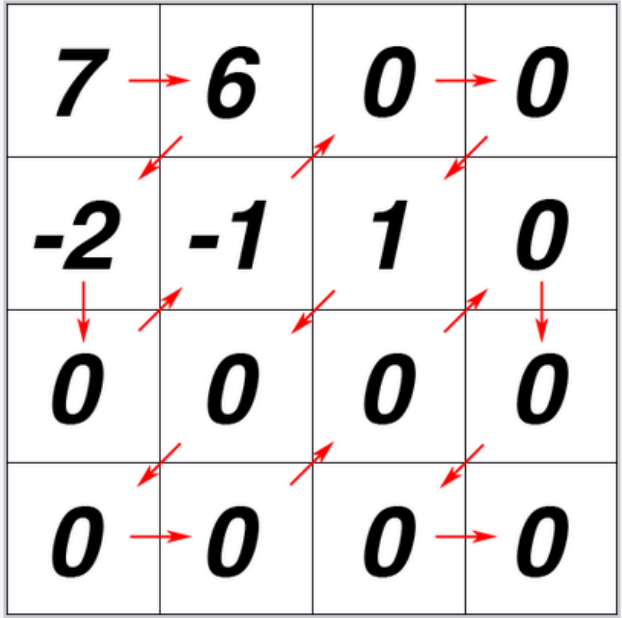
Encoding:

The transmitted bitstream for this block is 000010001110010111101101.

Decoding: The output array is "built up" from the decoded values as shown below. Values added to the output array at each stage are underlined.

The decoder has inserted two zeros; however, TotalZeros is equal to 3 and so another 1 zero is inserted before the lowest coefficient, making the final output array: 0, 3, 0, 1, -1, -1, 0, 1

Element	Value	Code
coeff_token	TotalCoeffs=5, T1s=3	0000100
T1 sign (4)	+	0
T1 sign (3)	-	1
T1 sign (2)	-	1
Level (1)	+1 (use Level_VLC0)	1
Level (0)	+3 (use Level_VLC1)	0010
TotalZeros	3	111
run_before(4)	ZerosLeft=3; run_before=1	10
run_before(3)	ZerosLeft=2; run_before=0	1
run_before(2)	ZerosLeft=2; run_before=0	1
run_before(1)	ZerosLeft=2; run_before=1	01
run_before(0)	ZerosLeft=1; run_before=1	No code required; last coefficient.



Escaneo de coeficientes del bloque 4x4.

Secuencia resultante: 7,6,-2,0,-1,0,0,1,0,0,0,0,0,0,0,0

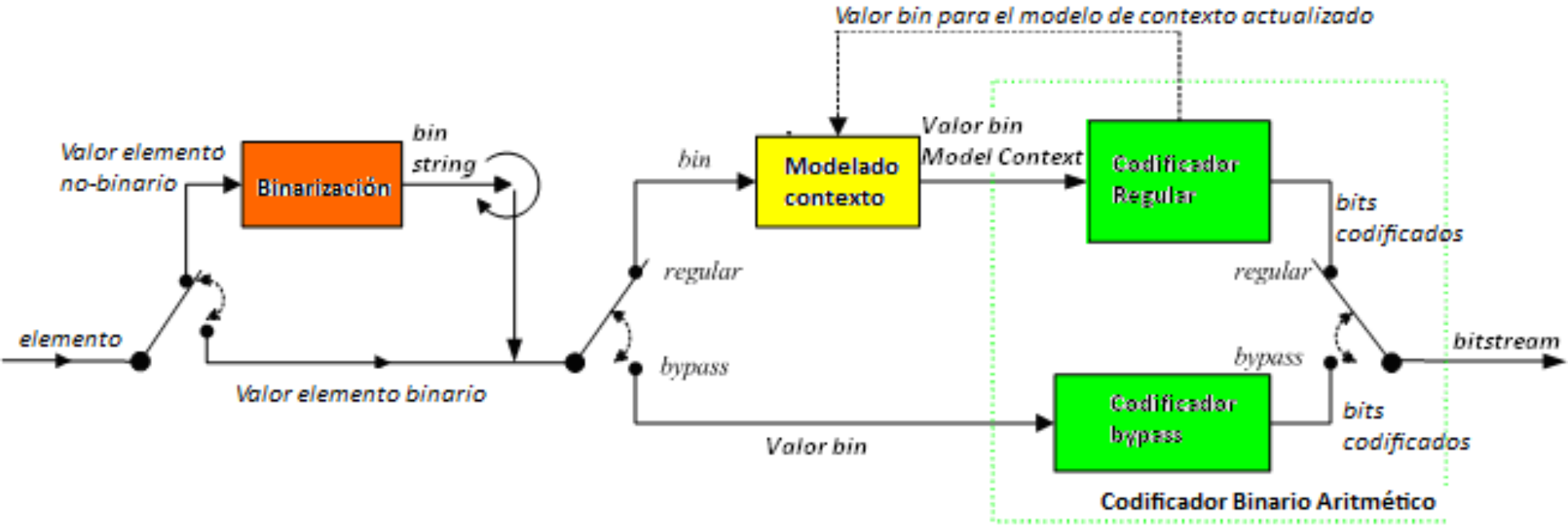
Secuencia resultante de la transformación	Secuencia que realmente considerará el codificador CAVLC
7,6,-2,0,-1,0,0,1,0,0,0,0,0,0,0,0	7,6,-2,0,-1,0,0,1

CABAC

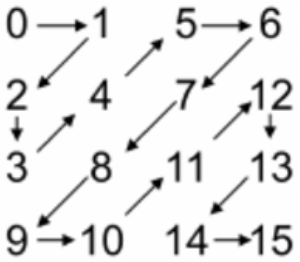
Context-adaptive binary arithmetic coding:

- Is a form of entropy encoding
- It is a lossless compression technique, although the video coding standards in which it is used are typically for lossy compression applications.
- CABAC is notable for providing much better compression than most other entropy encoding algorithms used in video encoding.

- **CAVLC requires considerably less processing to decode than CABAC, although it does not compress the data quite as effectively.**


Valor bin para el modelo de contexto actualizado
Valor elemento no-binario
Binarización
bin string
bin
Modelado contexto
Valor bin Model Context
Codificador Regular
bits codificados
regular
regular
elemento
bypass
bypass
bitstream
Valor elemento binario
Valor bin
Codificador bypass
bits codificados
Codificador Binario Aritmético

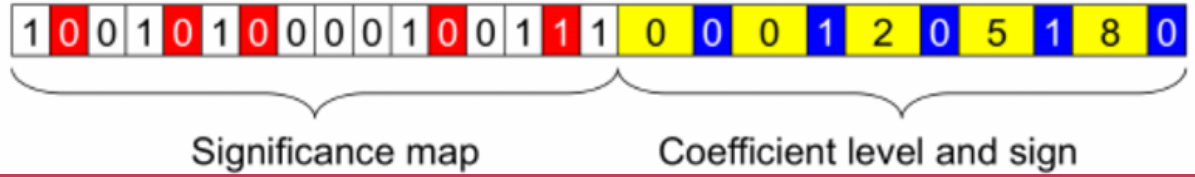
9	0	0	-1
-6	0	0	0
3	1	0	0
0	0	0	0



Zig-zag Scan
(forward for significance map, reverse for coefficient level & sign)

significant_coeff_flag	1	0	1	1	0	0	1	0	1
last_significant_coeff_flag	0		0	0			0		1
coeff_abs_level_minus1	8		5	2			0		0
coeff_sign_flag	0		1	0			1		0

Signaling order



It's estimated the bitrate savings can be as much as 50% or more compared to MPEG-2

As h.264 it’s newer, it’s easier to find examples and explanations about it on the internet, so…

Universitat Pompeu Fabra Barcelona

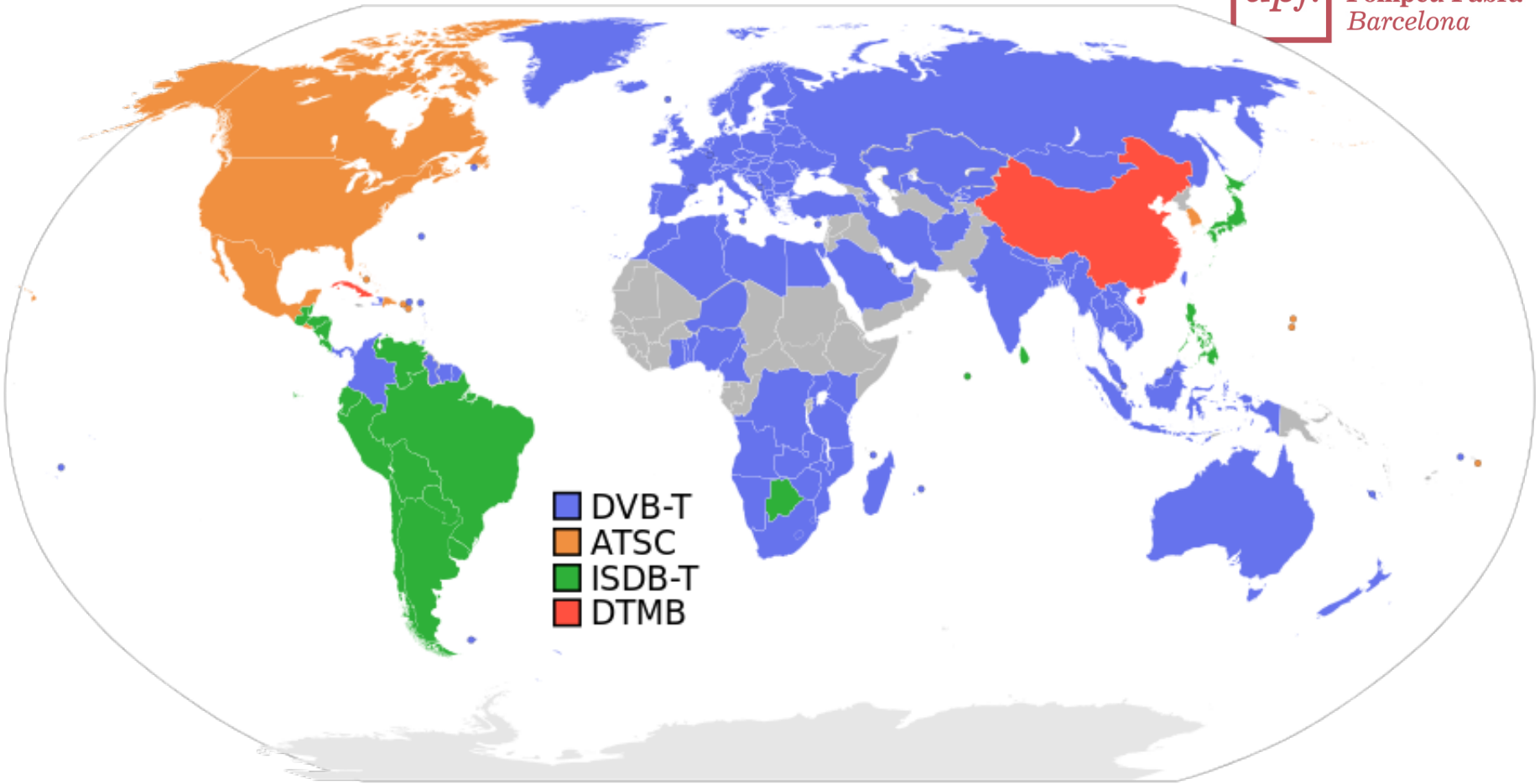
https://www.youtube.com/watch?v=PmoEsPWEdOA

upf. Universitat Pompeu Fabra Barcelona

Broadcasting standards

NOTE: I avoided to talk about MPEG-TS, because it's more a telecom. engineering class

You just need to know that Transport Stream exists. Think about that as an IPTV MPEG container


DVB-T
ATSC
ISDB-T
DTMB

DVB standard

The signals can be

Aerial -> From air

Satellite -> from satellite outside our planet

Cable -> physical cable

The letter that follows the standard marks the transmission medium

DVB Digital Video Broadcasting

Example with DVB standard:

DVB-T -> Terrestrial

DVB-T2 -> Terrestrial upgraded

DVB-C -> Cable

DVB-S -> Satellite

(This can be applied to all broadcasting standards: ISDB-T, ATSC-C, etc.)

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).

Video: MPEG2, h.264

Most common is MPEG2 for SD channels and .h264 for HD channels

Audio: AAC, Dolby Digital (AC-3) and MP3

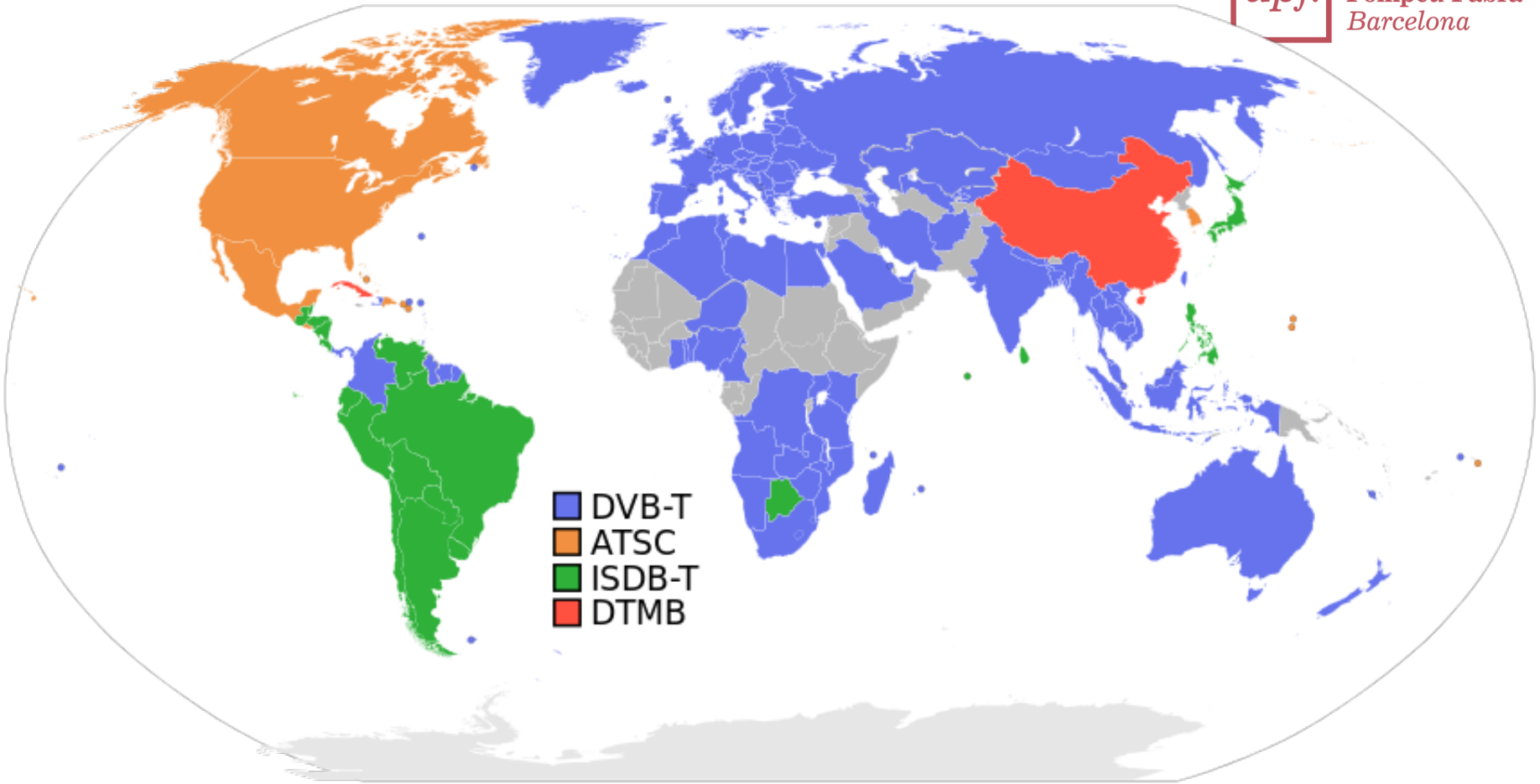
DVB

```
  Duration: N/A, start: 65915.868267, bitrate: 15640 kb/s
  Program 1
    Stream #0:0[0x44]: Audio: mp1 ([6][0][0][0] / 0x0006), 0 channels, s16
    Stream #0:1[0x45](spa): Audio: mp2 ([3][0][0][0] / 0x0003), 48000 Hz, stereo, s16, 256 kb/s
    Stream #0:2[0x46](spa): Audio: mp2 ([3][0][0][0] / 0x0003), 48000 Hz, stereo, s16, 128 kb/s
    Stream #0:3[0x47]: Audio: mp2 ([3][0][0][0] / 0x0003), 48000 Hz, stereo, s16, 256 kb/s
    Stream #0:4[0x48]: Video: mpeg2video (Main) ([2][0][0][0] / 0x0002), yuv420p, 720x576 [SAR 64:45 DAR 16:9], 15000 kb/s,
25.58 fps, 25 tbr, 90k tbn, 50 tbc
At least one output file must be specified
root@rec-es-albacete-01:~#
```

DVB

```
  Duration: N/A, start: 43177.641811, bitrate: N/A
  Program 103
    Metadata:
      service_name    : ?ZDFinfo
      service_provider: ?ZDFvision
    Stream #0:0[0x68]: Video: mpeg2video (Main) ([2][0][0][0] / 0x0002), yuv420p(tv, top first), 720x576 [SAR 64:45 DAR 16:9
], 25 fps, 25 tbr, 90k tbn, 50 tbc
    Stream #0:1[0x69](deu): Audio: mp2 ([3][0][0][0] / 0x0003), 48000 Hz, stereo, fltp, 256 kb/s
    Stream #0:2[0x6a](deu): Audio: mp2 ([3][0][0][0] / 0x0003), 48000 Hz, stereo, fltp, 192 kb/s (visual impaired) (descript
ions)
    Stream #0:3[0x6b](deu): Audio: ac3 ([6][0][0][0] / 0x0006), 48000 Hz, stereo, fltp, 448 kb/s
    Stream #0:4[0x6c](mul): Audio: mp2 ([3][0][0][0] / 0x0003), 48000 Hz, stereo, fltp, 192 kb/s
    Stream #0:5[0x55]: Unknown: none ([5][0][0][0] / 0x0005)
    Stream #0:6[0x6d](deu): Subtitle: dvb_teletext ([6][0][0][0] / 0x0006)
    Stream #0:7[0x6e](deu): Subtitle: dvb_subtitle ([6][0][0][0] / 0x0006) (hearing impaired)
At least one output file must be specified
root@rec-de-datacenter-01:~#
```

ISDB


DVB-T
ATSC
ISDB-T
DTMB

The Integrated Services Digital Broadcasting (ISDB; Japanese: 統合デジタル放送サービス, Tōgō dejitaru hōsō sābisu) is a Japanese standard for digital television (DTV) and digital radio used by the country's radio and television networks. ISDB replaced NTSC-J analog television system and the previously used MUSE Hi-vision analogue HDTV system in Japan, and will be replacing NTSC, PAL-M and PAL-N in South America and the Philippines.

Digital Terrestrial Television Broadcasting (DTTB) services using ISDB-T started in Japan in December 2003 and in Brazil in December 2007 as a trial. Since then, many countries have adopted ISDB over other digital broadcasting standards.

Video: MPEG2, h.264

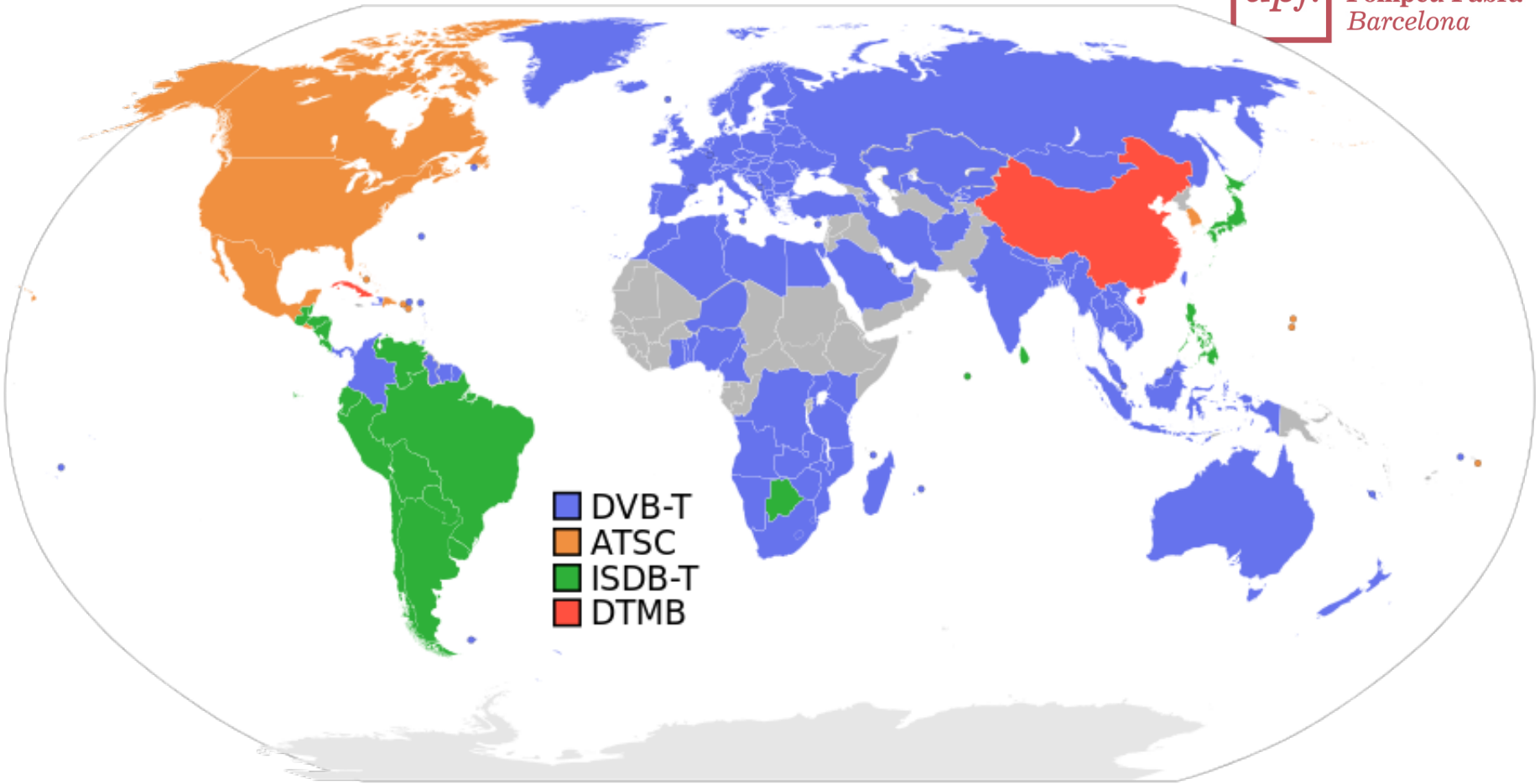
- Most common is MPEG2 for SD channels and .h264 for HD channels
- NOTE: ISDB-Tb (Brazilian & Latam) use only .h264

Audio: AAC

ISDB

```
      service_provider: GAMA TV
    Stream #0:12[0x32]: Audio: aac_latm (HE-AAC) ([17][0][0][0] / 0x0011), 48000 Hz, stereo, fltp
    Stream #0:13[0x33]: Unknown: none ([17][0][0][0] / 0x0011)
    Stream #0:14[0x34]: Video: h264 (High) ([27][0][0][0] / 0x001B), yuv420p(tv, smpte170m, bottom first), 720x480 [SAR 40:3
3 DAR 20:11], 29.97 fps, 29.97 tbr, 90k tbn, 59.94 tbc
  Program 1026
    Metadata:
      service_name    : ?TELEAMAZONAS SD
      service_provider:
    Stream #0:6[0x37]: Video: h264 (High) ([27][0][0][0] / 0x001B), yuv420p(tv, fcc, top first), 720x480 [SAR 10:11 DAR 15:1
1], 29.97 fps, 59.94 tbr, 90k tbn, 59.94 tbc
    Stream #0:15[0x38]: Audio: aac_latm (HE-AAC) ([17][0][0][0] / 0x0011), 48000 Hz, stereo, fltp
  Program 1281
    Metadata:
      service_name    : TVC-HD
      service_provider: TVC
    Stream #0:16[0x3b]: Video: h264 (High) ([27][0][0][0] / 0x001B), yuv420p(tv, bt709, bottom first), 1920x1080 [SAR 1:1 DA
R 16:9], 29.97 fps, 29.97 tbr, 90k tbn, 59.94 tbc
    Stream #0:17[0x3c]: Audio: aac_latm (HE-AAC) ([17][0][0][0] / 0x0011), 48000 Hz, stereo, fltp
At least one output file must be specified
root@rec-ec-quito-03:~#
```

ATSC


DVB-T
ATSC
ISDB-T
DTMB

ATSC, Advanced Television Systems Committee (ATSC) standards are a set of standards for digital television transmission over terrestrial, cable, and satellite networks. It is largely a replacement for the analog NTSC standard, and like that standard, used mostly in the United States, Mexico and Canada. Other former users of NTSC, like Japan, have not used ATSC during their digital television transition because they adopted their own system called ISDB.

ATSC includes two primary high definition video formats, 1080i and 720p. It also includes standard-definition formats, although initially only HDTV services were launched in the digital format. ATSC can carry multiple channels of information on a single stream

Video: MPEG2, h.264

Most common is MPEG2 for SD channels and .h264 for HD channels

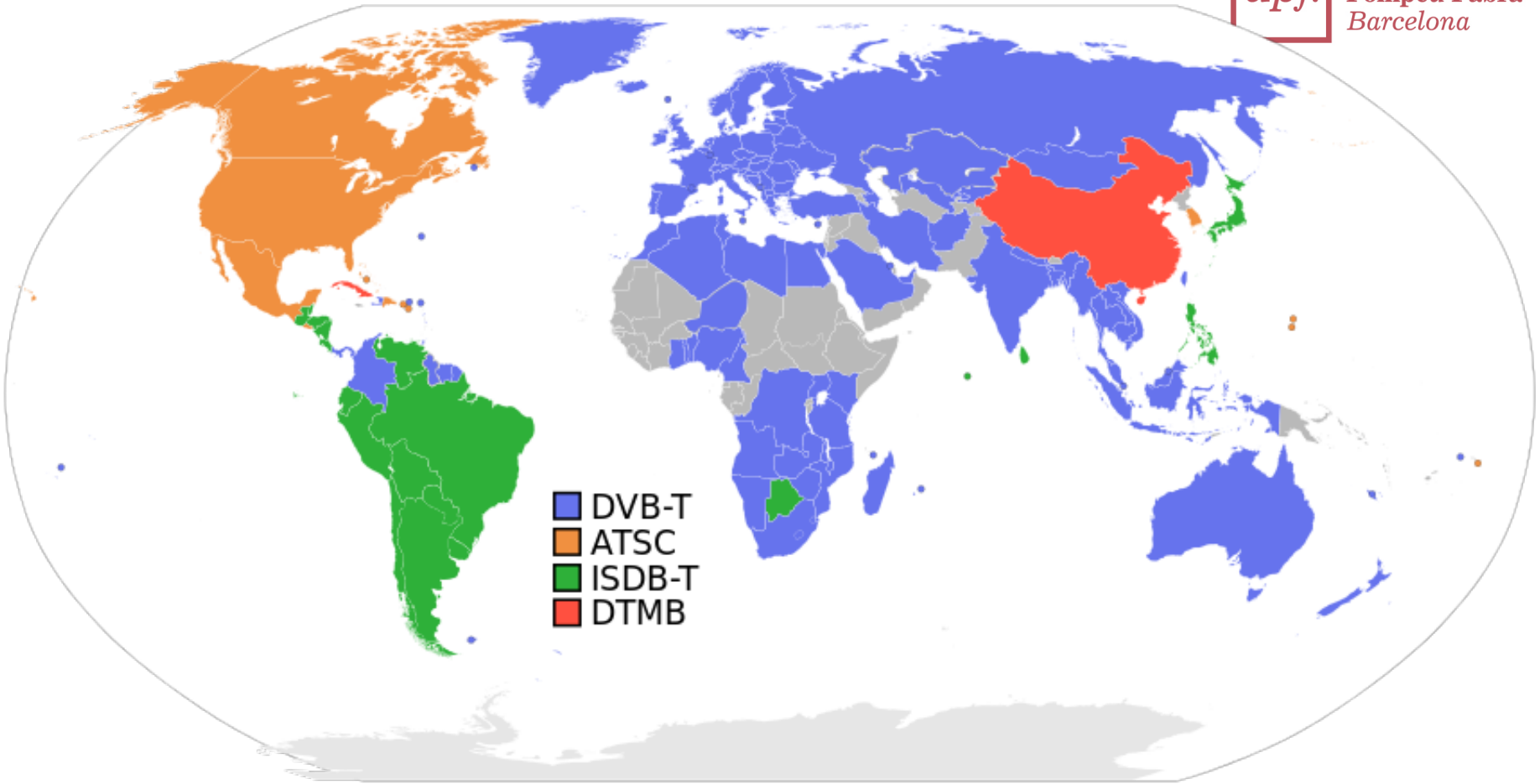
Audio: AC-3 (story behind: https://en.wikipedia.org/wiki/Grand_Alliance_(HDTV))

ATSC

```
Input #0, mpegts, from 'udp://@239.199.13.118:10118':
  Duration: N/A, start: 80343.747422, bitrate: N/A
  Program 301
    Stream #0:0[0x6e]: Video: h264 (High) ([27][0][0][0] / 0x001B), yuv420p(top first), 1920x1080 [SAR 1:1 DAR 16:9], Closed
 Captions, 29.97 fps, 29.97 tbr, 90k tbn, 59.94 tbc
    Stream #0:1[0x64]: Audio: ac3 ([129][0][0][0] / 0x0081), 48000 Hz, 5.1(side), fltp, 384 kb/s
    Stream #0:2[0x66]: Audio: ac3 ([129][0][0][0] / 0x0081), 48000 Hz, stereo, fltp, 192 kb/s
    Stream #0:3[0x8c]: Data: scte_35
At least one output file must be specified
root@rec-us-[redacted]-01:~#
```

```
  Program 1
    Stream #0:0[0x44](): Video: mpeg2video (Main) ([2][0][0][0] / 0x0002), yuv420p, 1920x1080 [SAR 1:1 DAR 16:9], 18053 kb/s
, 32.83 fps, 29.97 tbr, 90k tbn, 59.94 tbc
    Stream #0:1[0x45](spa): Audio: ac3 (AC-3 / 0x332D4341), 48000 Hz, 5.1(side), s16, 384 kb/s
At least one output file must be specified
root@rec-mx-mexicodf-01:~#
```

DTMB


DVB-T
ATSC
ISDB-T
DTMB

Previously known as DMB-T/H (Digital Multimedia Broadcast-Terrestrial/Handheld), the DTMB is a merger of the standards ADTB-T (developed by the Shanghai Jiao Tong University), DMB-T (developed by Tsinghua University) and TiMi (Terrestrial Interactive Multiservice Infrastructure); this last one is the standard proposed by the Academy of Broadcasting Science in 2002.

At first, neither Shanghai Jiao Tong University nor Tsinghua had enough political strength to make their own technology become the unique standard, so the final decision was to opt for a double standard, merged with the TIMI 3 standard, responding to a need for backward compatibility.

The DTMB was created in 2004 and finally became an official DTT standard in 2006.

Video: AVS, AVS+, MPEG2, h.264

- AVS is a chinese codec. Very similar to .h264 but they claim is better.

Audio: DRA, AAC, AC-3, MP2, MP3.

- DRA, Dynamic Resolution Adaptation is a chinese own audio codec. Same story

Thanks

franciscojavier.brines@upf.edu