upf. Universitat Pompeu Fabra Barcelona

# T4 – Future in codecs

(First of all, new technical concept:

IPTV: Internet Protocol TeleVision

Wide concept to express all video delivered through the internet)

Guillem will explain better next class :-)

Universitat Pompeu Fabra Barcelona

# It’s the CODEC WARS!!!

streaming media
THE BUSINESS & TECHNOLOGY OF ONLINE VIDEO

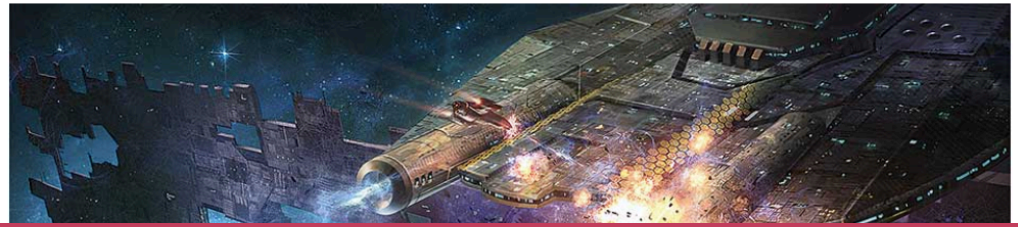
upf. Universitat Pompeu Fabra Barcelona





July 17, 2018
By Jan Ozer Contributing Editor
Featured Articles

# Return of the Codec Wars: A New Hope—a Streaming Summer Sequel



# How the ‘codec wars’ started? Let’s get back to 2004/5/6…

- Internet access was at higher speeds. Democratization of internet

- Video over IP started to work out

- Moore’s law: now computing is bigger, so we can do harder processing encoding

Universitat Pompeu Fabra Barcelona

# ...plus some curious companies started to success


NETFLIX

- Company used to rent movies on demand by ordinary mail

- Was a competitor of Blockbuster

- Started to use data-mining and big data collecting

- Started an emergency strategy of sharing online its content, before bankrupcy in 2008


YouTube

- Company created by 3 former Paypal employees in February 2005
- The idea was to use their service to upload content and insert it into user's website. Myspace users started to use this service in their comments and profiles.
- Nike launched a Ronaldinho ad in late 2005, other companies started to focus on it. On october 2006 Google buys it for 1650M$

Universitat Pompeu Fabra Barcelona

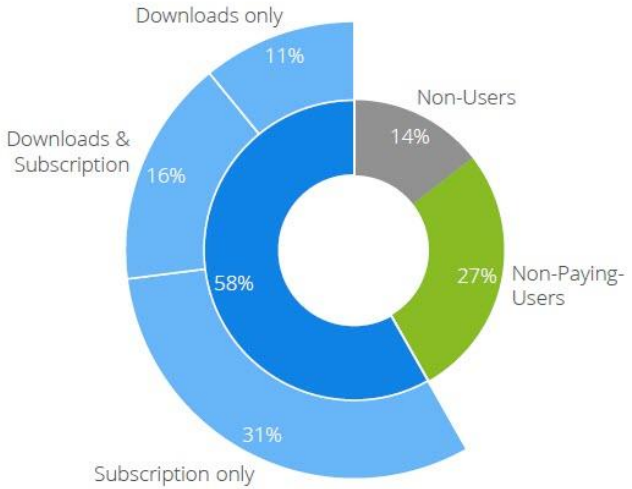
# NOWADAYS

Universitat Pompeu Fabra Barcelona

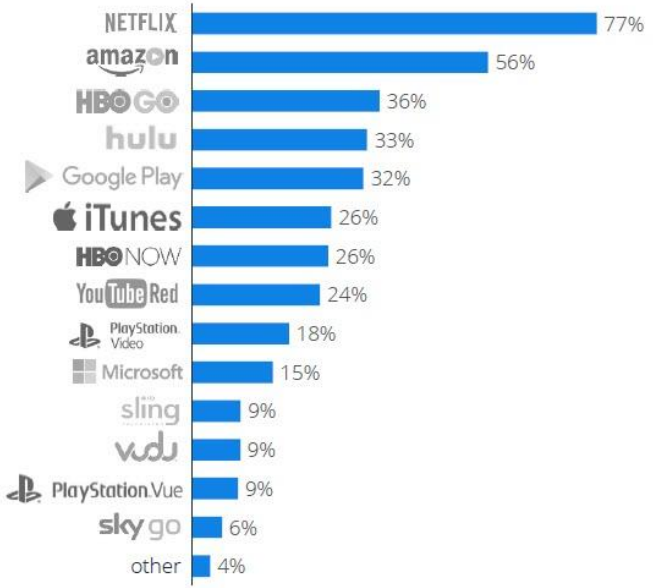
# More than half of U.S. onliners pay for VoD services, with Netflix having the biggest user base

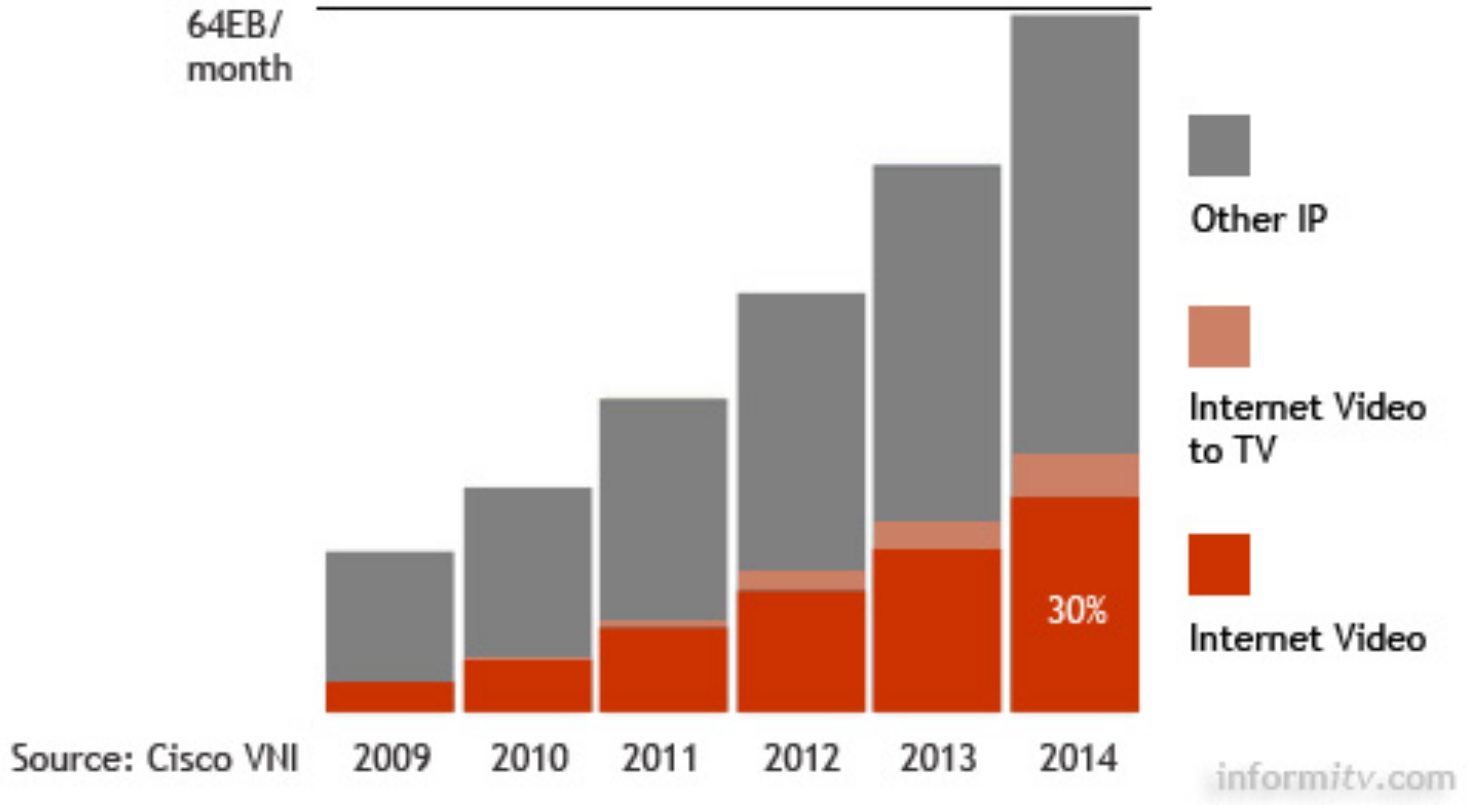
VoD users and brands

**Share of VoD users among U.S. onliners**




**Distribution of paying[1] U.S. VoD users among services**




64EB/ month
Other IP
Internet Video to TV
Internet Video
30%
Source: Cisco VNI
2009
2010
2011
2012
2013
2014
informitv.com

Universitat Pompeu Fabra Barcelona

# How much volume of worldwide internet traffic you think video represents?

**Now real estimations is that IPTV represents between**

# 70-80%

**Of internet traffic. That means 7 or 8 out of 10 bytes through internet are VIDEO**

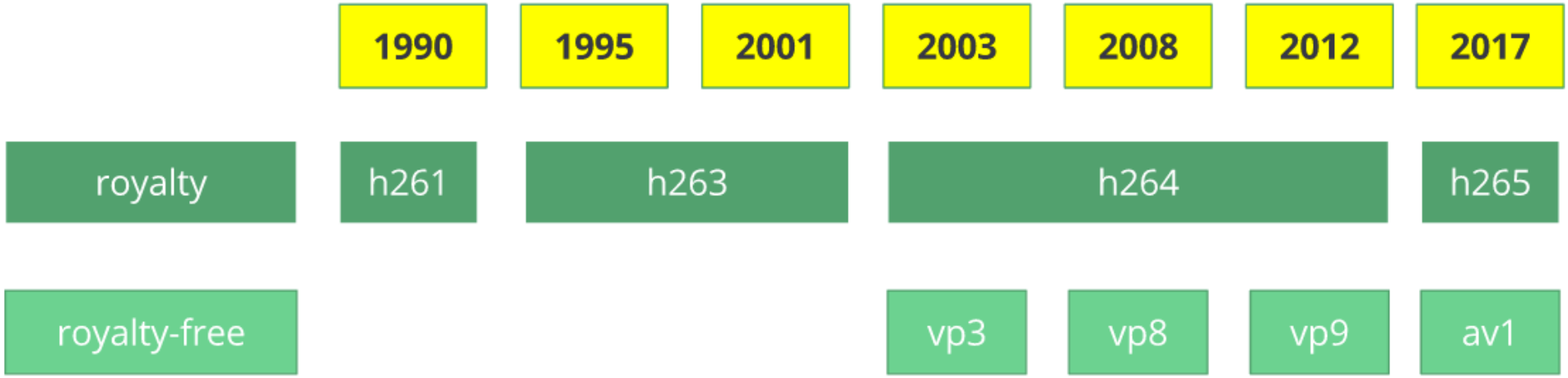
# So new challenges are all focused on IPTV:

- Reduce data (decrease storage costs)

- Reduce bitrate (decrease information - fastest delivery)

- Reduce complexity (decrease latency)

**ALL OF THAT WITHOUT LOSING ‘QUALITY’**


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

# … and all this take us to the CODEC WARS!

| | H.264 | HEVC | VP9 | AV1 |
|---|---|---|---|---|
| **Revenue Side** | | | | |
| **Reach** | | | | |
| Computers | 100% | Only Safari/Edge/ No Chrome or Firefox | 95% | 80% in 6 months/CPU not yet known |
| Mobile with hardware | 100% | 50% | 50% | 2 years |
| Smart TV/STB/OTT | 100% | 100% | under 100% | 2 years |
| **Features** | | | | |
| Live | 100% | Many options | Little support | Challenging |
| Live transcode | 100% | Many options | Wowza & Nimble | Challenging |
| Low latency | 100% | Some options | unknown | Challenging |
| HDR | Not optimal (reach of 10-bit AVC unknown) | 100% | HLG; no Dolby Vision or HDR10 | 2 years |
| Quality | 1 - lowest of the bunch | H.264 @ 60% | H.264 @ 60% | VP9 @ 70% |
| **Cost Side** | | | | |
| Encoding time | 1x | ~ 4x H.264 | ~2X H.264 | 1000+ x VP9 |
| Content royalty cost | PPV/Subscripion | Velos? | None | No current std. support |
| FUD Factor | Nokia/Motorola | Others not in pool | Feels low risk | Some risk |

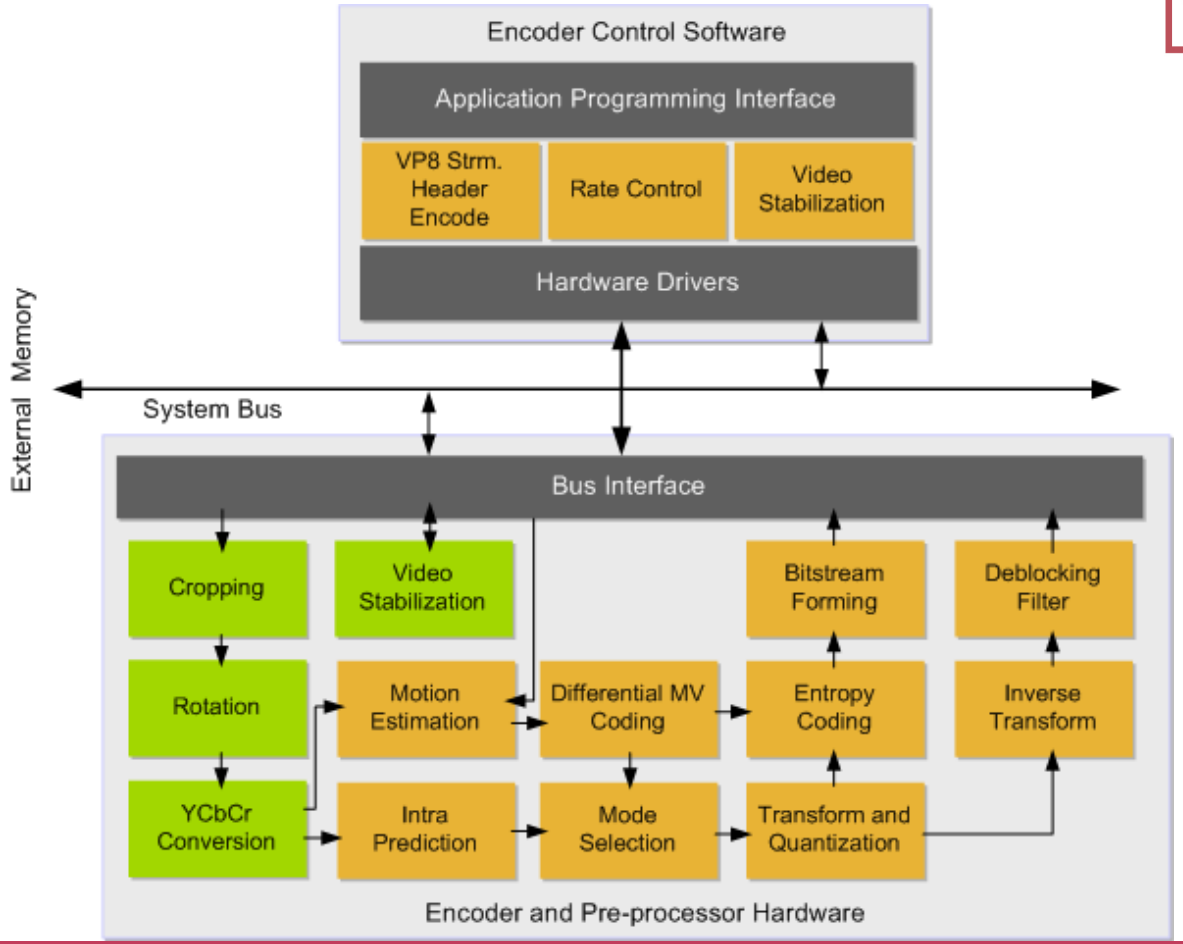
- If you get a patent, you get paid.
- If you develop better codecs, you can offer better services and better user experiences for your customers. You can push or block other companies' codecs in your hardware to avoid them to success (VHS vs BetaMax example).
- It's the first time in history where big companies are fighting with codecs to get a bigger share of the market (but not only codecs: i.e. Twitch vs Youtube)

upf. Universitat Pompeu Fabra Barcelona

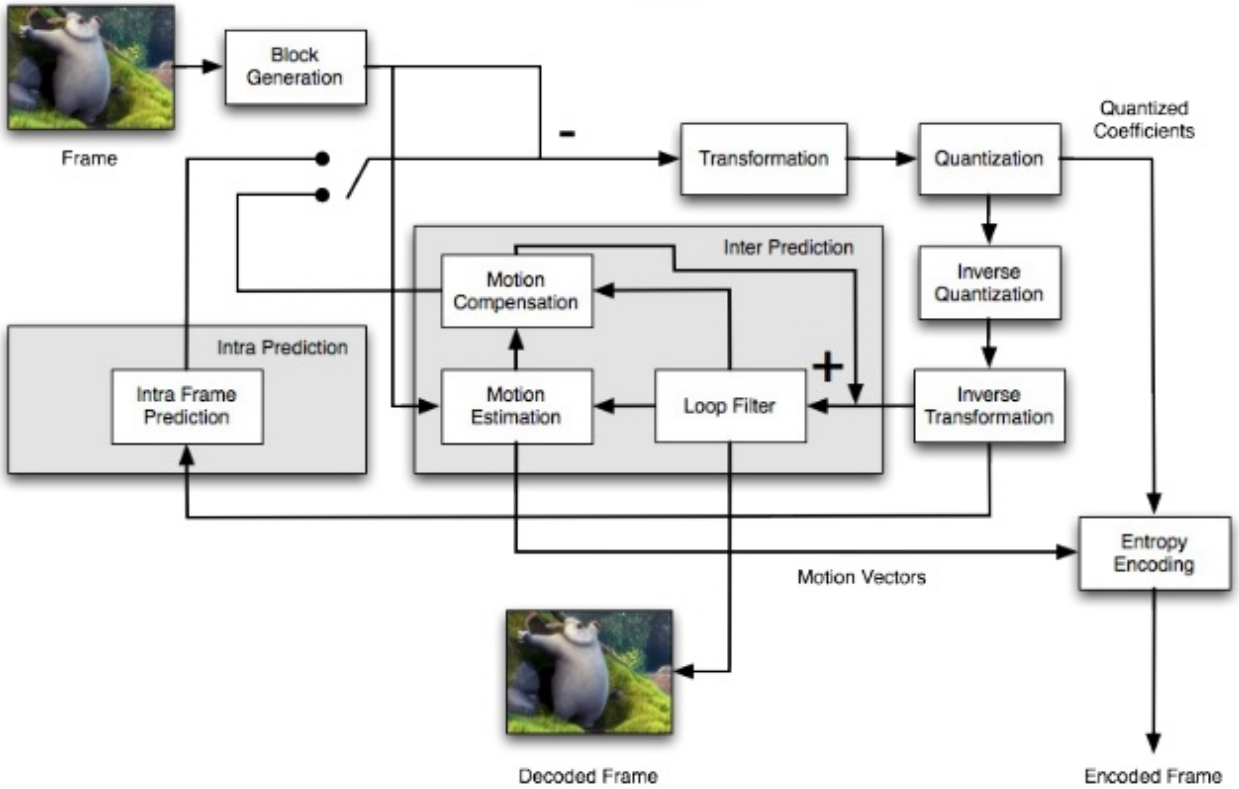
# VP8

**In May 2010, after the purchase of On2 Technologies, Google provided an irrevocable patent promise on its patents for implementing the VP8 format, and released a specification of the format under the Creative Commons Attribution 3.0 license. That same year, Google also released libvpx, the reference implementation of VP8, under the revised BSD license.**

**Opera, Firefox, Chrome, and Chromium support playing VP8 video in HTML5 video tag. Internet Explorer officially supports VP8 with a separate codec. According to Google VP8 is mainly used in connection with WebRTC and as a format for short looped animations, as a replacement for the Graphics Interchange Format (GIF)**


Encoder Control Software
Application Programming Interface
VP8 Strm. Header Encode
Rate Control
Video Stabilization
Hardware Drivers
External Memory
System Bus
Bus Interface
Cropping
Video Stabilization
Bitstream Forming
Deblocking Filter
Rotation
Motion Estimation
Differential MV Coding
Entropy Coding
Inverse Transform
YCbCr Conversion
Intra Prediction
Mode Selection
Transform and Quantization
Encoder and Pre-processor Hardware

# VP8 Encoding Overview



- Very similar to .h264

- Changes the prediction modes (motion compensation)

- It uses DCT with 4x4 macroblocks and Hadamard transform

- VP8 only supports progressive scan video signals with 4:2:0 chroma subsampling and 8 bits per sample

- Macro blocks can comprise 4×4, 8×8, or 16×16 samples

- Motion vectors have quarter-pixel precission

# Concept of *altframes* (alternative frames)

A maximum of three frames can be referenced for temporal prediction: the last Golden Frame (may be an intra frame), alternate reference frame, and the directly preceding frame. The so-called alternate reference frames (altref) can serve as reference-only frames for displaying them can be deactivated.

**In this case the encoder can fill them with arbitrary useful image data, even from future frames, and thereby serve the same purpose as the b-frames of the MPEG formats**

upf. Universitat Pompeu Fabra Barcelona

# Matroska files

# Matroska - .mkv files

**The Matroska Multimedia Container is a free, open-standard container format, a file format that can hold an unlimited number of video, audio, picture, or subtitle tracks in one file.**

**Matroska is similar in concept to other containers like AVI, MP4, or Advanced Systems Format (ASF), but is entirely open in specification, with implementations consisting mostly of open source software.**

# "Matroska" is derived from matryoshka (Russian: матрёшка [mɐˈtrʲɵʂkə])

**In 2010, it was announced that the WebM audio/video format would be based on a profile of the Matroska container format together with VP8 video and Vorbis audio.**

**It’s also used for illegal BlueRay Disc rips**

**…and finally added support to the next codec we’ll see**

upf. Universitat Pompeu Fabra Barcelona

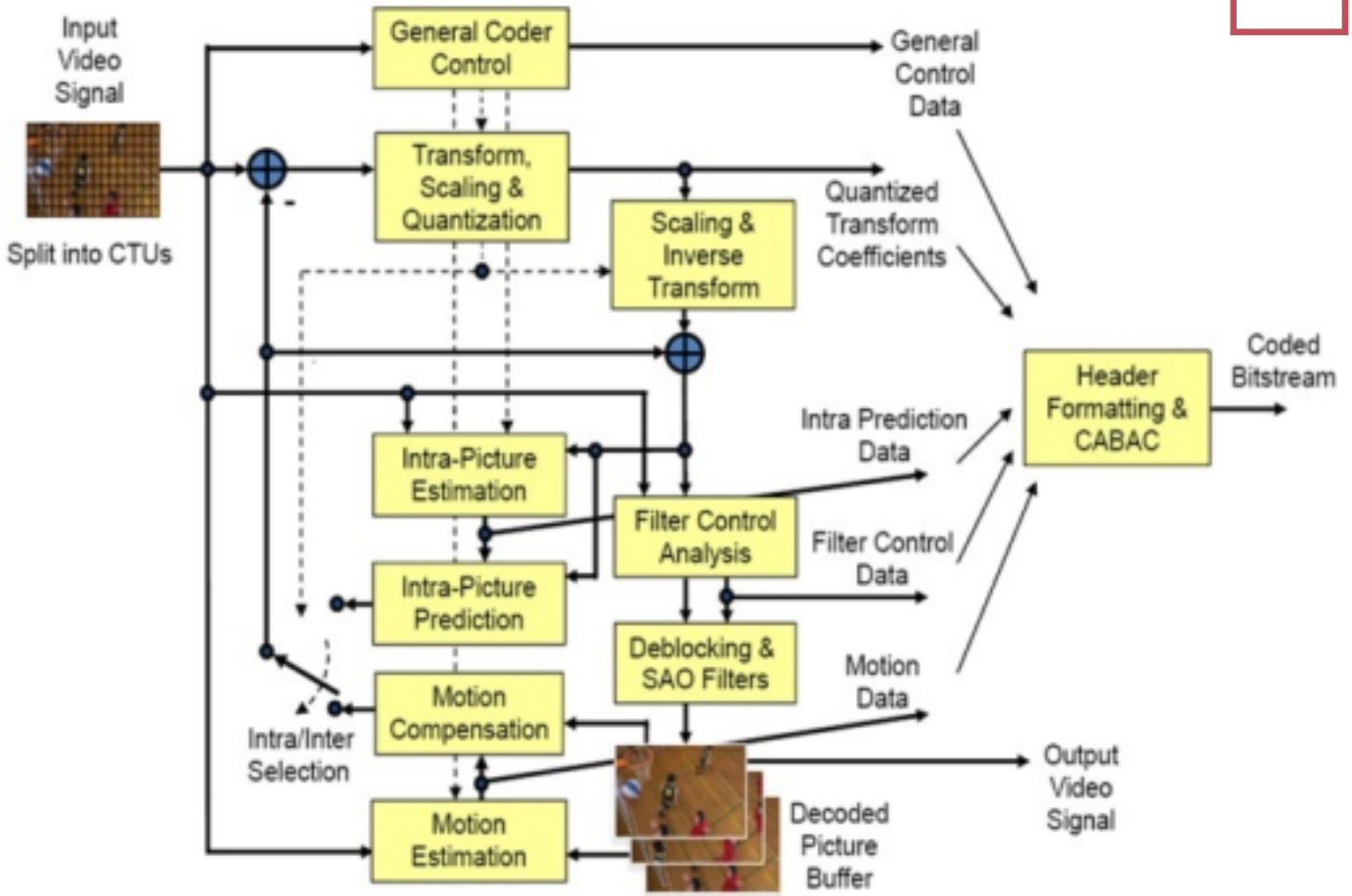
# .h265

- Also known as HEVC (High Efficiency Video Coding

- .h264 successor

- Offers from 25% to 50% better data compression at the same level of video quality, or substantially improved video quality at the same bit rate. It supports resolutions up to 8192×4320, including 8K UHD

- While .h264 uses the integer discrete cosine transform (DCT) with 4x4 and 8x8 block sizes, HEVC uses integer DCT and DST transforms with varied block sizes between 4x4 and 32x32. The High Efficiency Image Format (HEIF) is based on HEVC. As of 2019, HEVC is used by 43% of video developers, and is the second most widely used video coding format after .h264.


Input Video Signal
Split into CTUs
General Coder Control
General Control Data
Transform, Scaling & Quantization
Quantized Transform Coefficients
Scaling & Inverse Transform
Header Formatting & CABAC
Coded Bitstream
Intra Prediction Data
Intra-Picture Estimation
Filter Control Analysis
Filter Control Data
Intra-Picture Prediction
Deblocking & SAO Filters
Motion Data
Motion Compensation
Intra/Inter Selection
Output Video Signal
Motion Estimation
Decoded Picture Buffer

- It has a basic version which you don't pay; but it's patented and theorically you should pay if you use .mp4 files

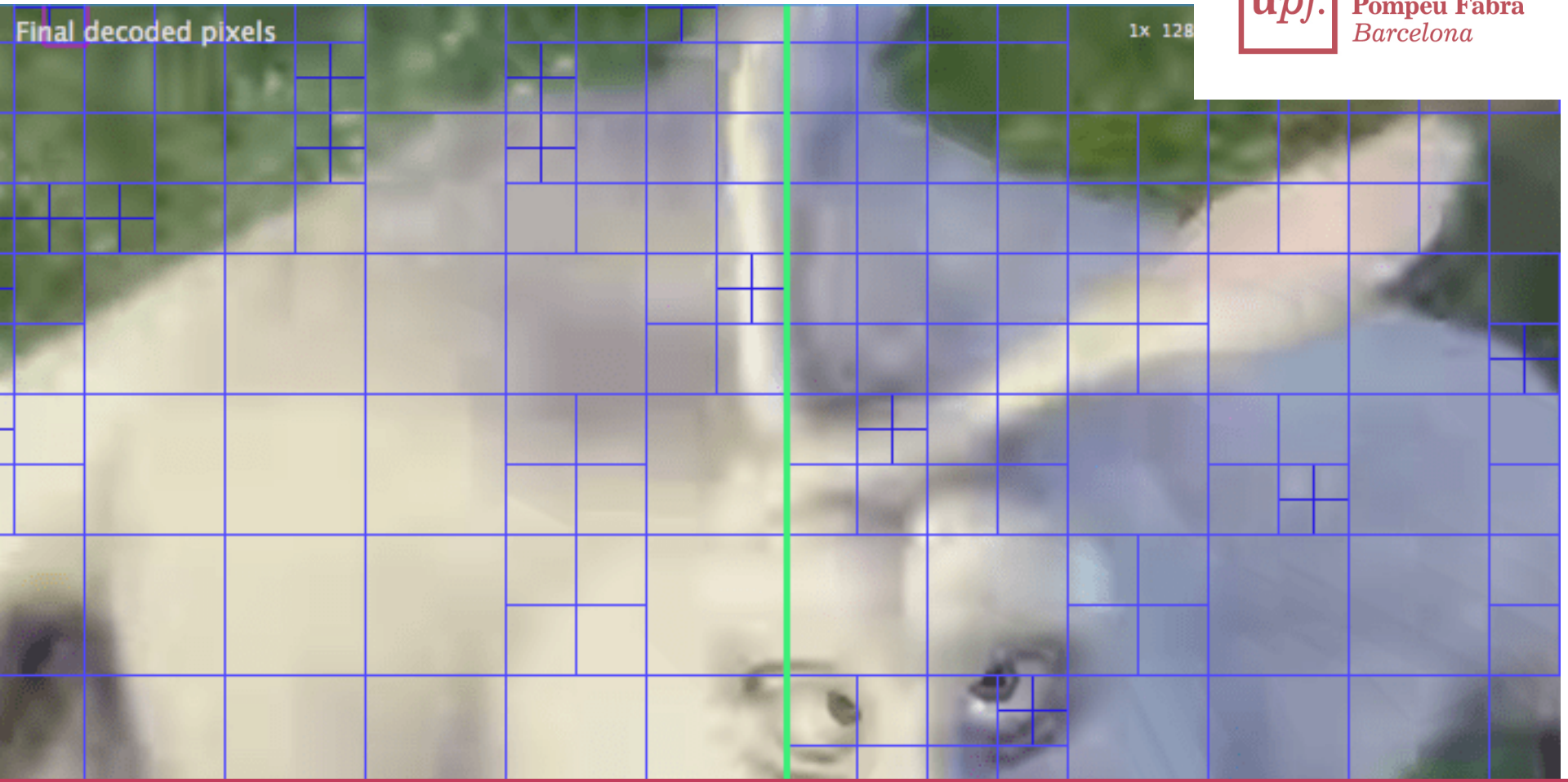
- The HEVC video coding layer uses the same "hybrid" approach used in all modern video standards: it uses inter-/intra-picture prediction and 2D transform coding.

- A HEVC encoder first proceeds by splitting a picture into block shaped regions for the first picture, or the first picture of a random access point, which uses intra-picture prediction.

- After the prediction methods are finished and the picture goes through the loop filters, the final picture representation is stored in the decoded picture buffer.Pictures stored in the decoded picture buffer can be used for the prediction of other pictures.

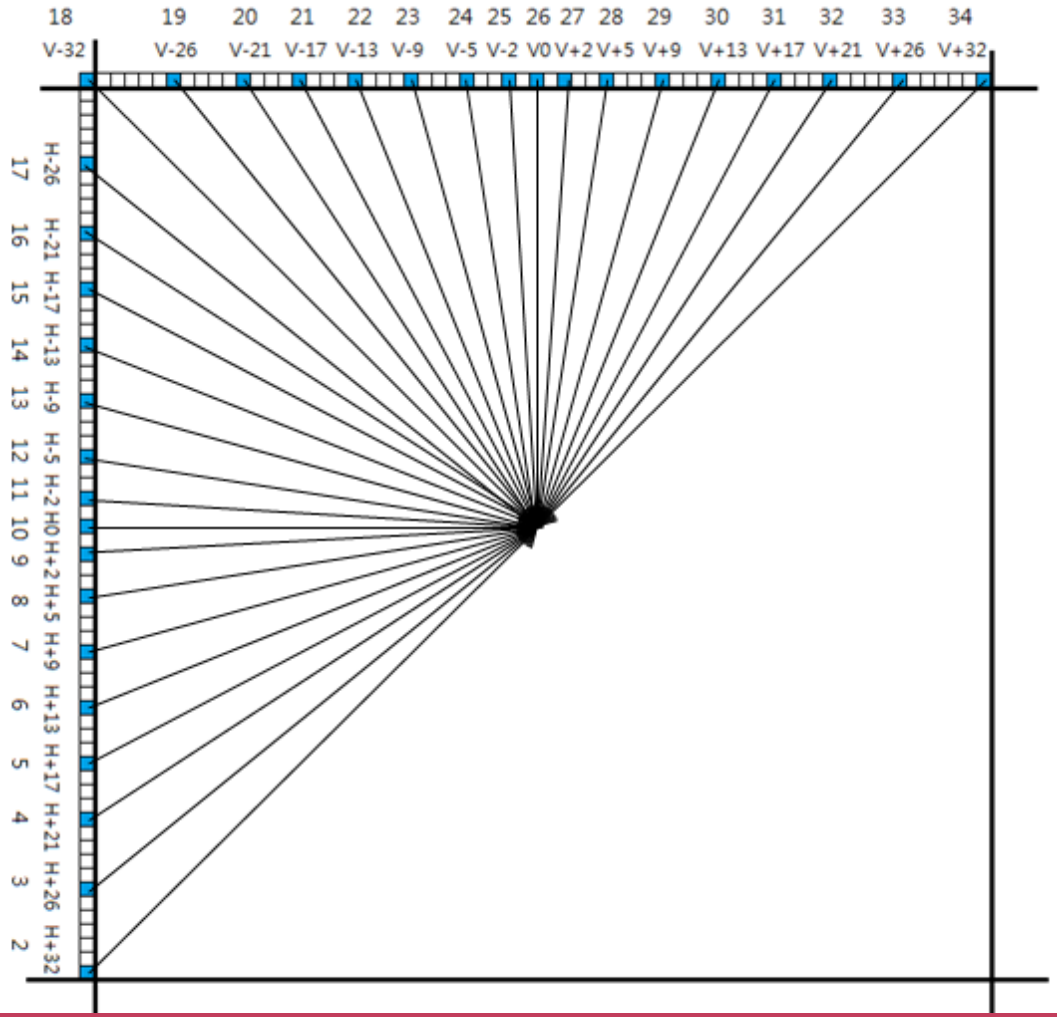
- Does not support interlaced video natively

- HEVC replaces 16×16 pixel macroblocks, which were used with previous standards, with coding tree units (CTUs) which can use larger block structures of up to 64x64 samples and can better sub-partition the picture into variable sized structures. HEVC initially divides the picture into CTUs which can be 64×64, 32×32, or 16×16 with a larger pixel block size usually increasing the coding efficiency.


Final decoded pixels
1x 128

- Uses a new revision of CABAC

- Introduces new modes for motion estimation: 33 vector possibilities


18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34
V-32 V-26 V-21 V-17 V-13 V-9 V-5 V-2 V0 V+2 V+5 V+9 V+13 V+17 V+21 V+26 V+32
17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2
H-26 H-21 H-17 H-13 H-9 H-5 H-2 H0 H+2 H+5 H+9 H+13 H+17 H+21 H+26 H+32

- Longer filter for chroma subsampling
- A lot of filters and small upgrades internally.
- Works with ‘profiles’ too. This is kind if we consider to encode with some presets
- Introduces new modes for motion estimation: 33 vector possibilities

## H.264/AVC

16x16 Macroblock

### Block coding structure

3 Intra partitioning

4 Inter partitioning

+4 sub-partitioning 8x8

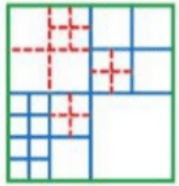
2 Transform sizes:
4x4, 8x8

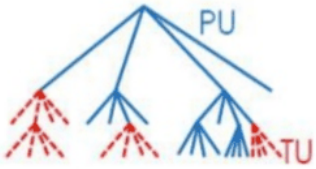
Up to 9 Intra prediction directions

## HEVC

Coding Unit
64x64 to 8x8

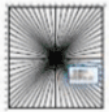
### Quadtree coding structure






Prediction Unit and Transform Unit partitioning

⇨ Multiples sizes/forms: 64x64 to 4x4



35 Intra prediction directions

Efficient spatio-temporal mv prediction

http://www.worldbroadcastingunions.org/wbuarea/library/docs/isog/presentations/2012B/2.4%20Vieron%20ATEME.pdf

# Some examples against .h264


H.265 - Tears of Steel 1080P at 400 kbps
H.264 - Tears of Steel 1080P at 400 kbps

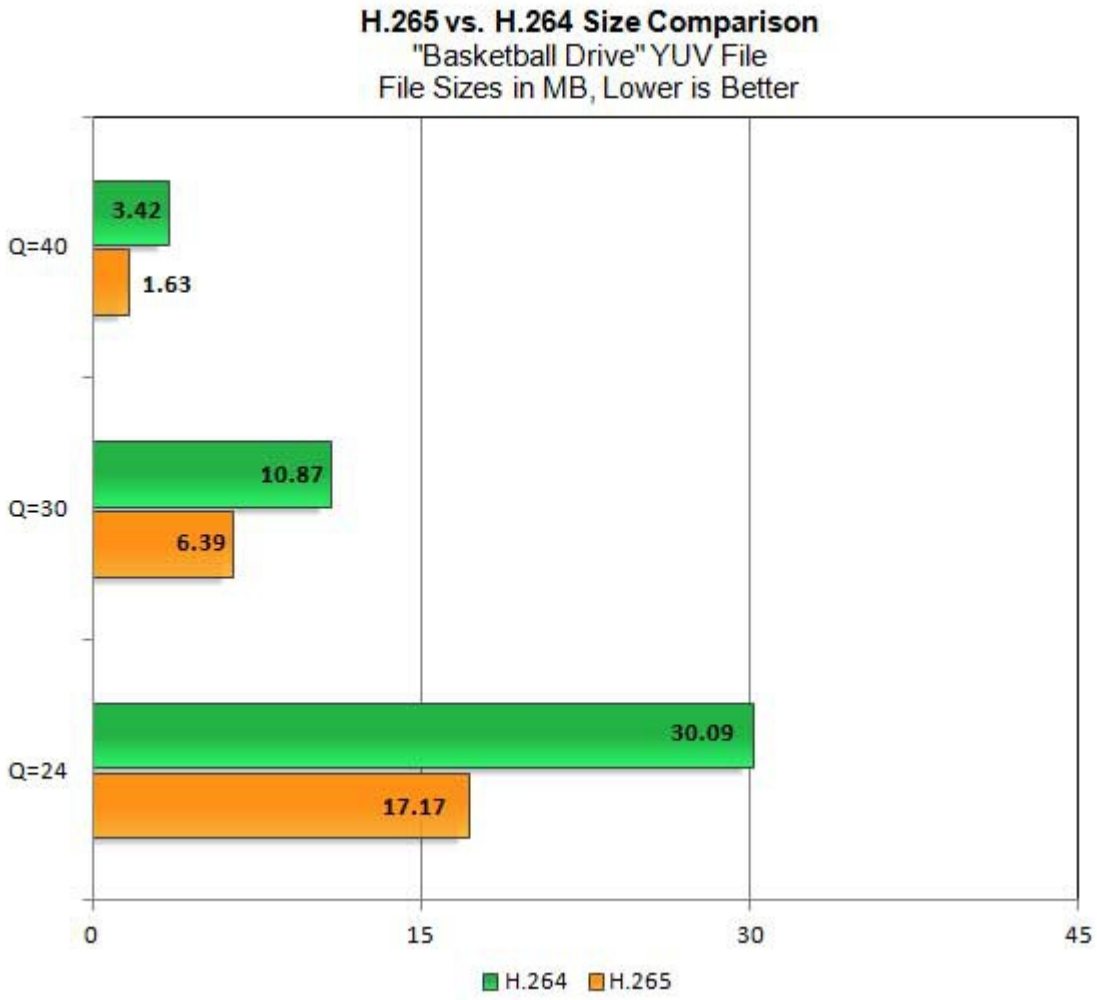
# Another example:

# https://youtu.be/qL22L0mRSDs

**Most important: it has a HUGE reduction of the size of the files, compared to .h264.**

**We can consider the compression is really better.**

**The estimated ratio is between -25% and -50% of the bitrate!!**


H.265 vs. H.264 Size Comparison
"Basketball Drive" YUV File
File Sizes in MB, Lower is Better
Q=40
3.42
1.63
Q=30
10.87
6.39
Q=24
30.09
17.17
0
15
30
45
H.264
H.265

# Cons:

## NOT fully supported on Chrome & Firefox

## It’s the codec wars!!!

upf. Universitat Pompeu Fabra Barcelona

# VP9

**VP9 is an open and royalty-free video coding format developed by Google.**

**VP9 is the successor to VP8 and competes mainly with MPEG's HEVC (H.265, as recently seen). At first, VP9 was mainly used on Google's video platform YouTube.**

**The emergence of the Alliance for Open Media, and its support for the ongoing development of the successor AV1, of which Google is a part, led to growing interest in the format.**
**In contrast to HEVC, VP9 support is common among web browsers. The combination of VP9 video and Opus audio in the WebM container, as served by YouTube, is supported by roughly 80% of the browser market (mobile included) as of June 2018.**

**The single holdout among major modern-day browsers is Apple's Safari (both desktop and mobile versions). Android has supported VP9 since version 4.4 KitKat.**

**Parts of the format are covered by patents held by Google. The company grants free usage of its own related patents based on reciprocity, i.e. as long as the user does not engage in patent litigations.**

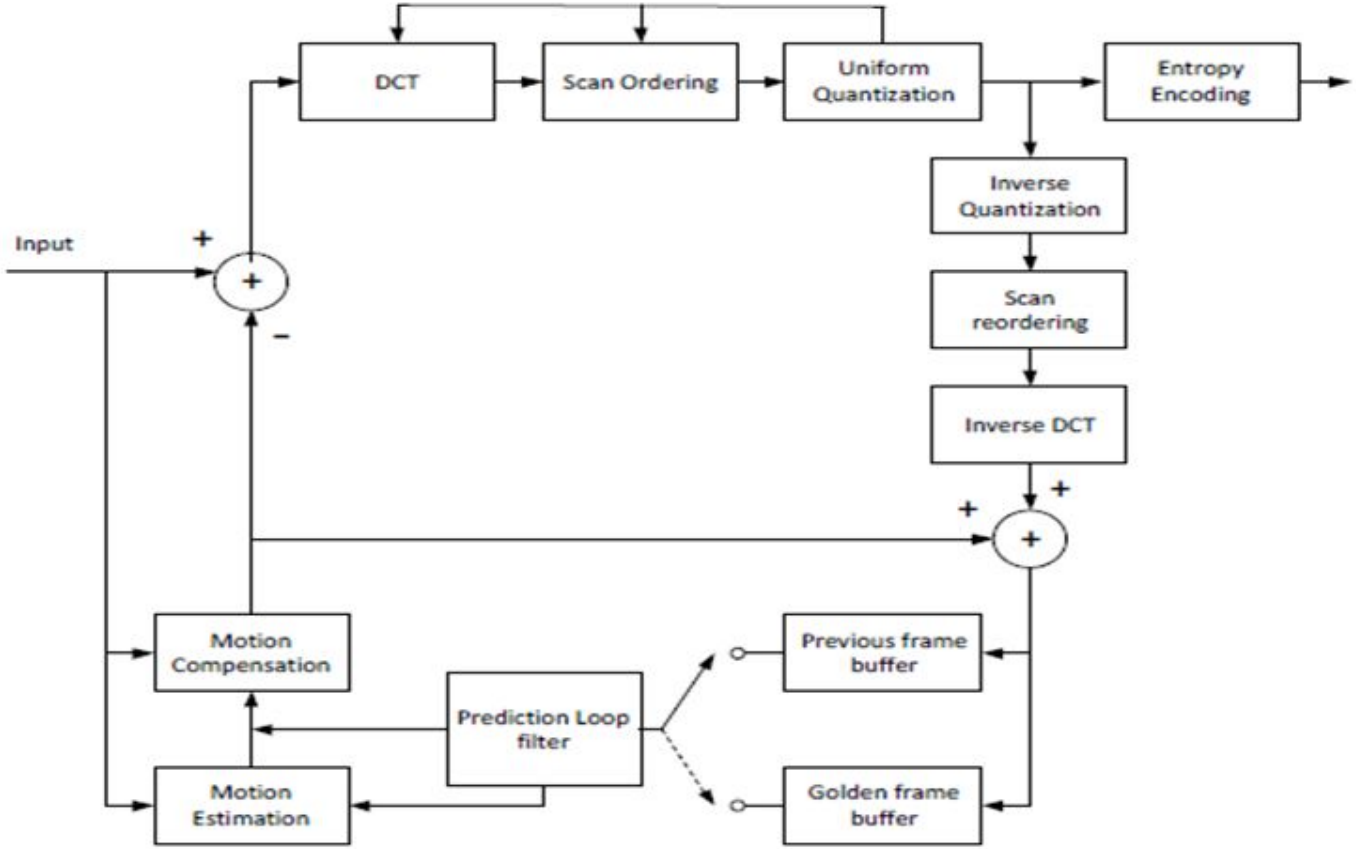



Fig 9: Encoder block diagram for VP9 [24]

- VP9 has many design improvements compared to VP8.

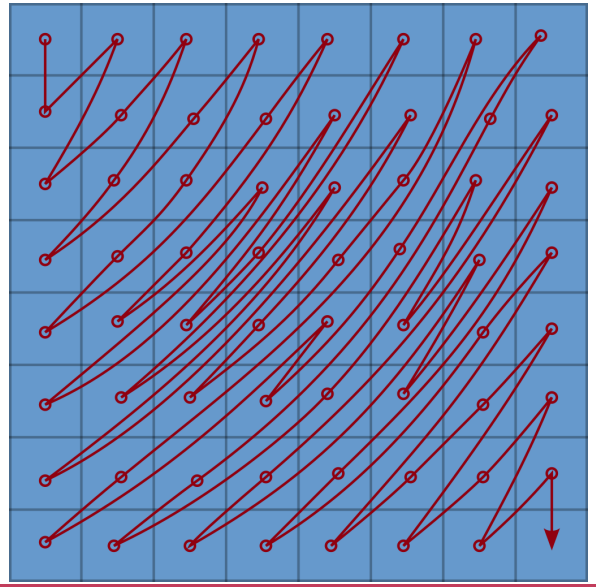
- Its biggest improvement is support for the use of coding units of 64×64 pixels. This is especially useful with high-resolution video.

- Also the prediction of motion vectors was improved. In addition to VP8's four modes (average/"DC", "true motion", horizontal, vertical), VP9 supports six oblique directions for linear extrapolation of pixels in intra-frame prediction.

# Small upgrades:

- eighth-pixel precision for motion vectors,
- three different switchable 8-tap subpixel interpolation filters,
- improved selection of reference motion vectors,
- improved coding of offsets of motion vectors to their reference,
- improved entropy coding,
- improved and adapted (to new block sizes) loop filtering,
- the asymmetric discrete sine transform (ADST),
- larger discrete cosine transforms (DCT, 16×16 and 32×32), and
- improved segmentation of frames into areas with specific similarities (e.g. fore-/background)

# Transform coefficients are scanned in a round pattern (increasing distance from the corner)



Universitat Pompeu Fabra Barcelona

# WebM

# Webm

- WebM is an audiovisual media file format, which is primarily intended to offer a royalty-free alternative to use in the HTML5 video and the HTML5 audio elements.

- It has a sister project WebP for images. The development of the format is sponsored by Google, and the corresponding software is distributed under a BSD license.

- The WebM container is based on a profile of Matroska. WebM initially supported VP8 video and Vorbis audio streams. In 2013, it was updated to accommodate VP9 video and Opus audio.

# (BONUS TRACK):

## Can you search about opus audio?

## You have 10 minutes. We launch a random number and that person jumps into stage

(let’s get into business!)

- Native WebM support by Mozilla Firefox, Opera, and Google Chrome was announced at the 2010 Google I/O conference.

- Didn’t had support on Microsoft Edge/Explorer until 2017

- Still not supported on Safari

upf. Universitat Pompeu Fabra Barcelona

# Some comparision

Universitat Pompeu Fabra Barcelona

# VP9 | VP8 | H.265 | x264 encoder test

# https://www.youtube.com/watch?v=Ctjm1kxw-BM

Universitat Pompeu Fabra Barcelona

# AV1 – real future?

# AV1

# AV1 ya supera en todo a HEVC: ya no hay motivos para no usarlo



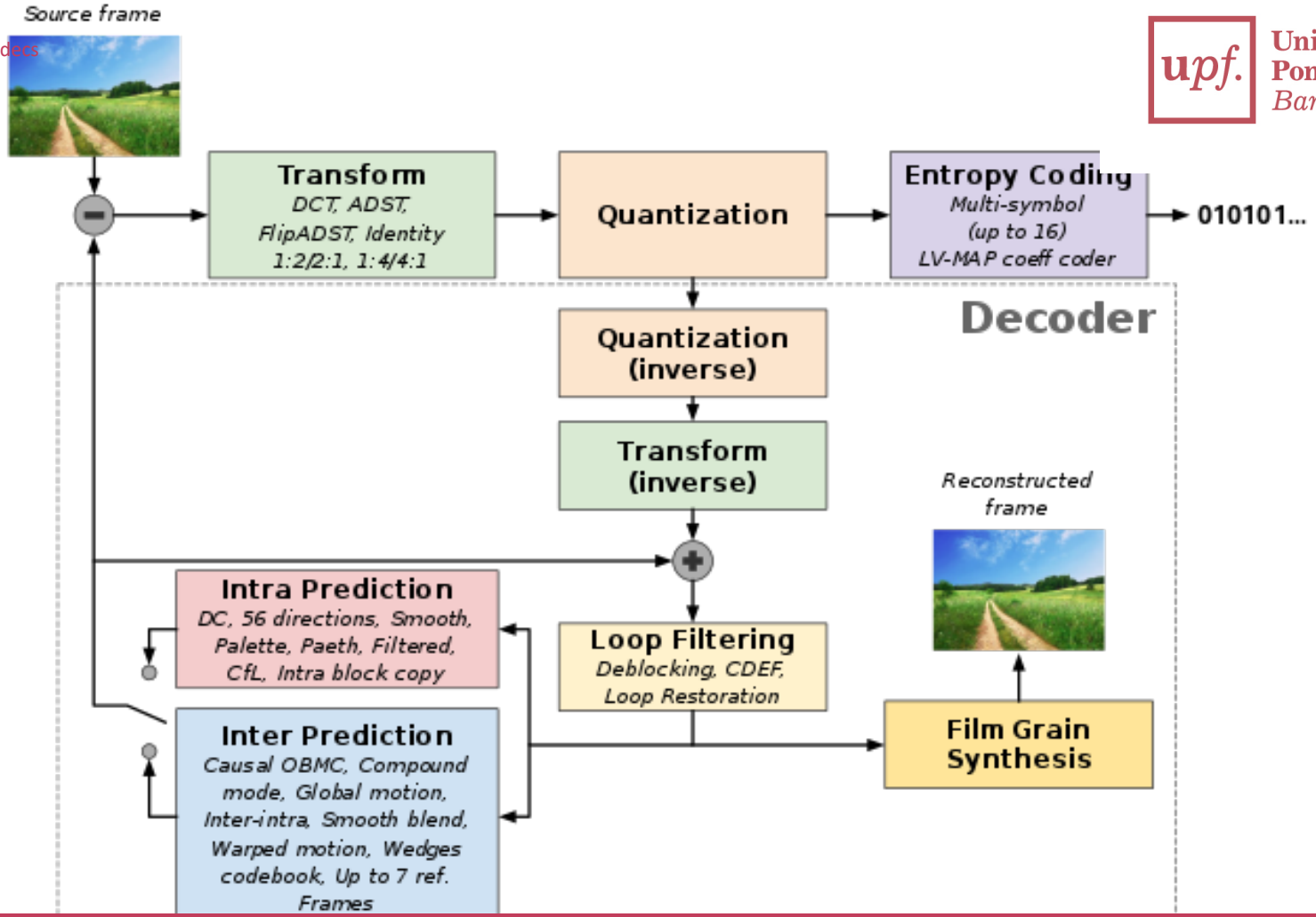
software

9

Los **códecs** que usamos a diario llevan años de perfeccionamiento detrás. Sus desarrolladores tienen que pulirlos antes de lanzarlos oficialmente, ya que no sólo es importante la calidad final que ofrecen, sino también su rendimiento, consumo de recursos, tiempo de codificación y tamaño final. Ahora, AV1 ha avanzado tanto que ya es mejor que **HEVC** en todo.

- AOMedia Video 1 (AV1) is an open, royalty-free video coding format designed for video transmissions over the Internet
- It was developed as a successor to VP9 by the Alliance for Open Media (AOMedia), a consortium founded in 2015 that includes semiconductor firms, video on demand providers, video content producers, software development companies and web browser vendors.

- **The AV1 bitstream specification includes a reference video codec. In Facebook testing that approximates real world conditions AV1 achieved 34%, 46.2% and 50.3% higher data compression than *libvpx-vp9*, *x264* high profile, and *x264* main profile respectively.**


Source frame
Transform
DCT, ADST,
FlipADST, Identity
1:2/2:1, 1:4/4:1
Quantization
Entropy Coding
Multi-symbol
(up to 16)
LV-MAP coeff coder
010101...
Decoder
Quantization
(inverse)
Transform
(inverse)
Reconstructed
frame
Intra Prediction
DC, 56 directions, Smooth,
Palette, Paeth, Filtered,
CfL, Intra block copy
Loop Filtering
Deblocking, CDEF,
Loop Restoration
Inter Prediction
Causal OBMC, Compound
mode, Global motion,
Inter-intra, Smooth blend,
Warped motion, Wedges
codebook, Up to 7 ref.
Frames
Film Grain
Synthesis

**(CODEC WARS)**

**AV1 was announced with the creation of the Alliance for Open Media on 1 September 2015. The Alliance's goal is to combine its members' technology and expertise to develop a video format that is royalty-free, has higher compression efficiency and is more suitable for use in browsers and on the web.**

**(CODEC WARS)**

**Several of the Alliance's founding members, including Google (VP10), Mozilla (Daala) and Cisco (Thor), had ongoing research projects into royalty-free video and incorporated aspects of each into AV1.**

Universitat Pompeu Fabra Barcelona

# (CODEC WARS)
# All this was from the wiki. Sounds great, isn't it?

# But… patent owner is Google!

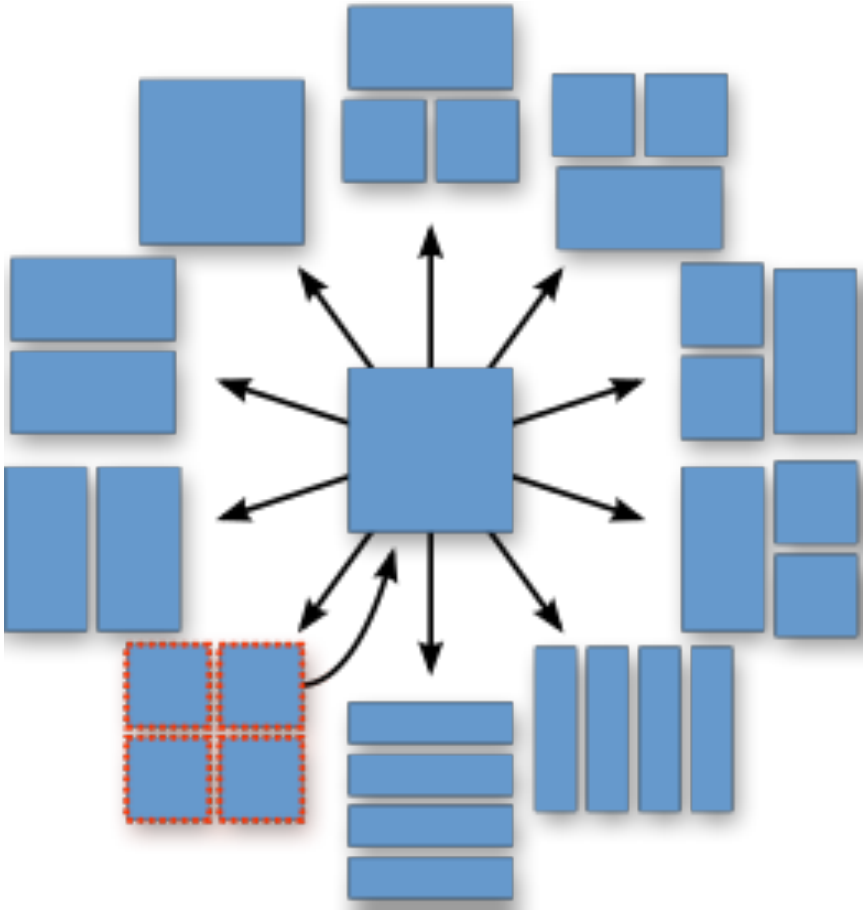
Universitat Pompeu Fabra Barcelona

# … let’s recap

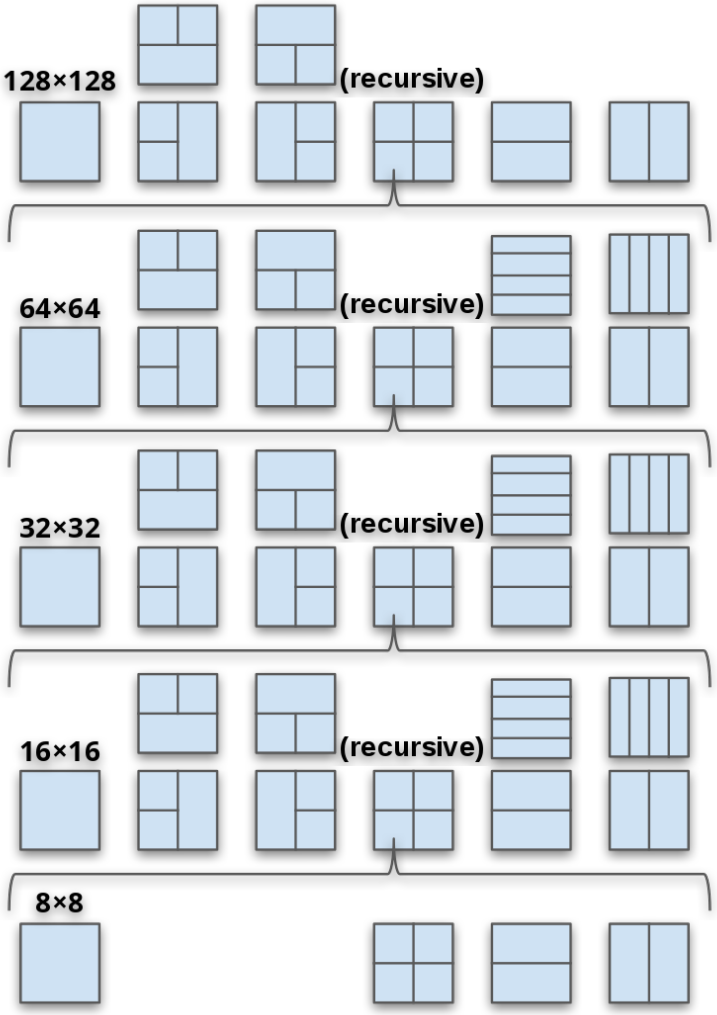
Universitat Pompeu Fabra Barcelona

# Technology

Frame content is separated into adjacent same-sized blocks referred to as superblocks. Similar to the concept of a macroblock, superblocks are square-shaped and can either be of size 128×128 or 64×64 pixels. Superblocks can be divided in smaller blocks according to different partitioning patterns. The four-way split pattern is the only pattern whose partitions can be recursively subdivided.



"T-shaped" partitioning patterns are introduced, a feature developed for VP10, as well as horizontal or vertical splits into four stripes of 4:1 and 1:4 aspect ratio. The available partitioning patterns vary according to the block size, both 128×128 and 8×8 blocks can't use 4:1 and 1:4 splits. Moreover, 8×8 blocks can't use "T" shaped splits.


128×128
(recursive)
64×64
(recursive)
32×32
(recursive)
16×16
(recursive)
8×8

This allows superblocks to be divided into partitions as small as 4×4 pixels.
Diagram of the AV1 superblock partitioning shows how 128×128 superblocks can be split all the way down to 4×4 blocks.

As special cases, 128×128 and 8×8 blocks can't use 1:4 and 4:1 splits, and 8×8 blocks can't use "T"-shaped splits.

# Motion estimation & compensation

- AV1 performs internal processing in higher precision (10 or 12 bits per sample), which leads to compression improvement due to smaller rounding errors in reference imagery.

**Predictions can be combined in more advanced ways (than a uniform average) in a block (compound prediction), including smooth and sharp transition gradients in different directions (wedge-partitioned prediction) as well as implicit masks that are based on the difference between the two predictors. This allows combination of either two inter predictions or an inter and an intra prediction to be used in the same block.**

**A frame can reference 6 instead of 3 of the 8 available frame buffers for temporal (inter) prediction while providing more flexibility on bi-prediction (ext_refs).**

**The Warped Motion (warped_motion) and Global Motion (global_motion) tools in AV1 aim to reduce redundant information in motion vectors by recognizing patterns arising from camera motion.**

They implement ideas that were tried to be exploited in preceding formats like e.g. MPEG-4 ASP, albeit with a novel approach that works in three dimensions. There can be a set of warping parameters for a whole frame offered in the bitstream, or blocks can use a set of implicit local parameters that get computed based on surrounding blocks.

**Switch frames (S-frame) are a new inter-frame type that can be predicted using already decoded reference frames from a higher-resolution version of the same video to allow switching to a lower resolution without the need for a full keyframe at the beginning of a video segment in the adaptive bitrate streaming use case.**

**Intra-prediction:**
**The "TrueMotion" predictor got replaced with a Paeth predictor which looks at the difference from the known pixel in the above left corner to the pixel directly above and directly left of the new one and then chooses the one that lies in direction of the smaller gradient as predictor.**

A palette predictor is available for blocks with very few (up to 8, dominant) colors like in some computer screen content. Correlations between the luminosity and the color information can now be exploited with a predictor for chroma blocks that is based on samples from the luma plane (cfl). In order to reduce discontinuities along borders of inter-predicted blocks, predictors can be overlapped and blended with those of neighbouring blocks (overlapped block motion compensation).

# Quantization

AV1 has new optimized quantization matrices (aom_qm). The eight sets of quantization parameters that can be selected and signaled for each frame now have individual parameters for the two chroma planes and can use spatial prediction. On every new superblock, the quantization parameters can be adjusted by signaling an offset.

# Quantization

AV1 has new optimized quantization matrices (aom_qm). The eight sets of quantization parameters that can be selected and signaled for each frame now have individual parameters for the two chroma planes and can use spatial prediction. On every new superblock, the quantization parameters can be adjusted by signaling an offset.

# Filters

**AV1 uses new filters that combined increase compression. It focuses on perceptual analysis, instead of spatial analysis: it's like applying psycho-acustics to an audio codec; but in this case into video.**

# Entropy coding

It uses Daala's entropy coder (daala_ec), a non-binary arithmetic coder. The use of non-binary arithmetic coding helps evade patents, but also adds bit-level parallelism to an otherwise serial process, reducing clock rate demands on hardware implementations.

# Entropy coding

This is to say that the effectiveness of modern binary arithmetic coding like CABAC is being approached using a greater alphabet than binary, hence greater speed, as in Huffman code (but not as simple and fast as Huffman code). AV1 also gained the ability to adapt the symbol probabilities in the arithmetic coder per coded symbol instead of per frame (ec_adapt).

# Companies starting to use it



**Just to mention a few.**

**A lot of companies joined the AOM Media, but still not implemented them (codec wars!)**

**https://www.ibc.org/manage/av1-codec-wars-erupt/3737.article**

Universitat Pompeu Fabra Barcelona

# You can check

# https://www.youtube.com/results?search_query=av1

# Thanks

franciscojavier.brines@upf.edu