KVS Media Interface

Aaron Xu
Solution Architecture, IoT Lab
亚马逊云科技

Basic Terminology

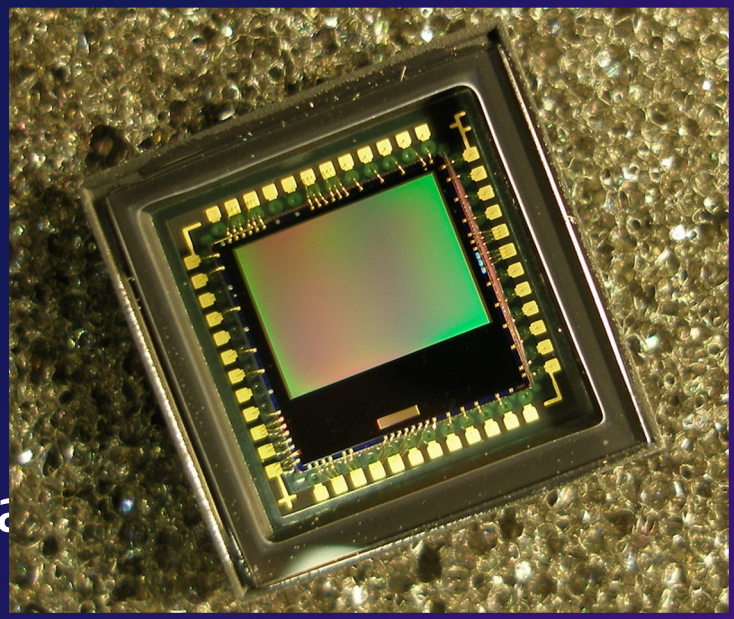
How images are captured? CMOS Sensor (APS)

Use less power

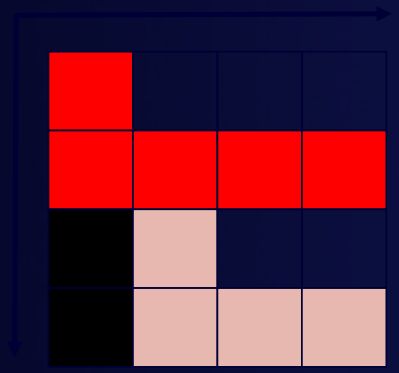
Transmit data faster than CCD

Cheaper

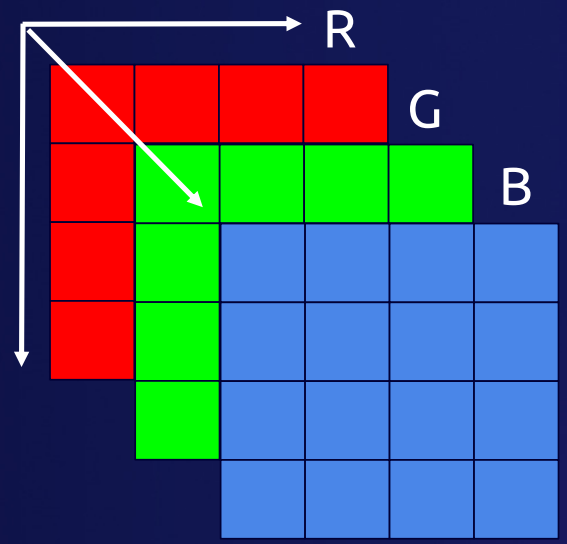
Most commonly in cell phone cameras, web ca



What is an image?



2D

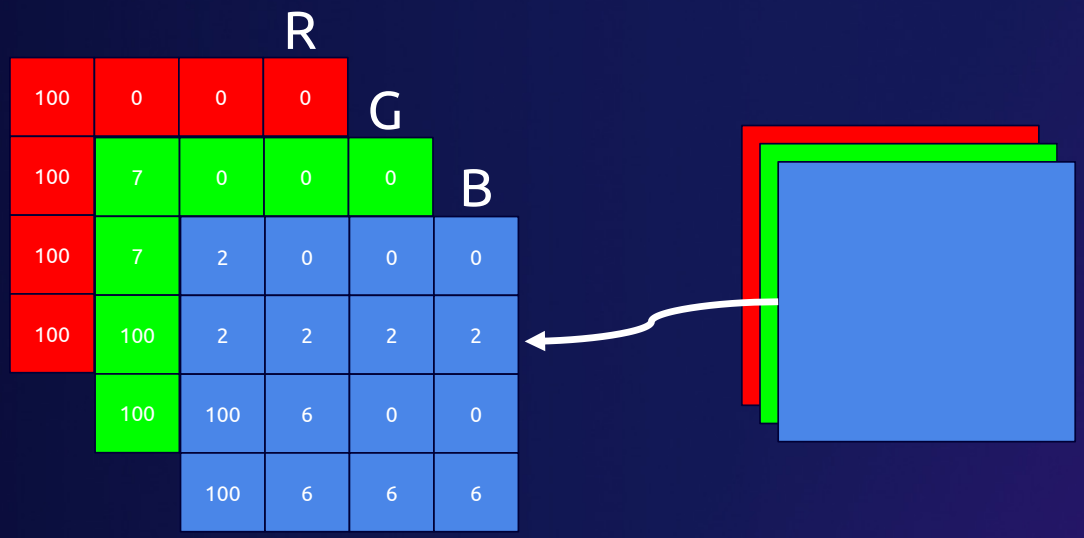


3D

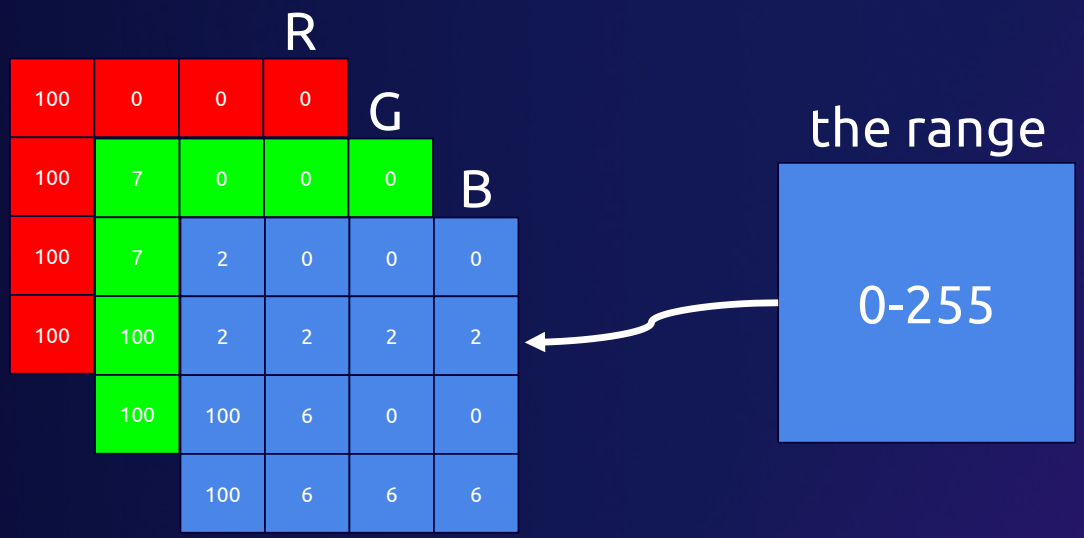
100	0	0	0	
100	7	0	0	0
100	7	2	0	0	0
100	100	2	2	2	2
	100	100	6	0	0
		100	6	6	6

color intensity

What is picture element (pixel) ?



What is bit (color) depth?

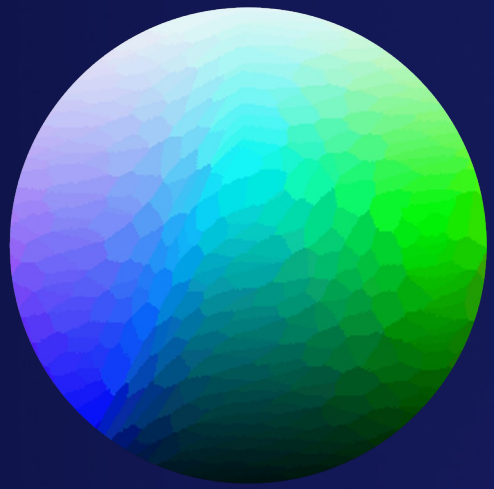


8R+8G+8B = 24 bits

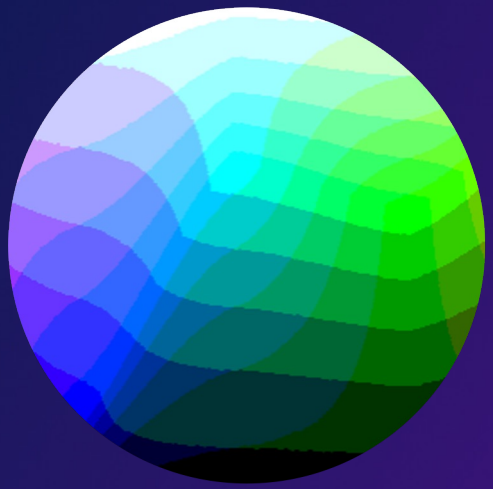
*it gives you 2^24 different colors

Color depth

24 bpp



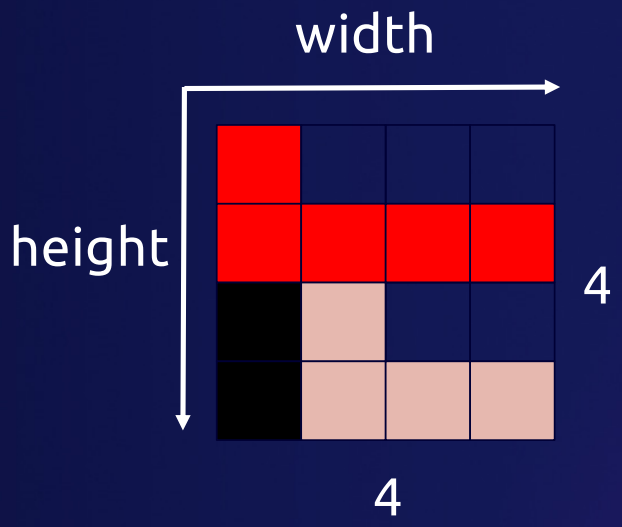
10 bpp



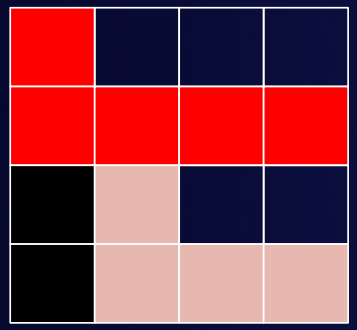
8 bpp

Color depth

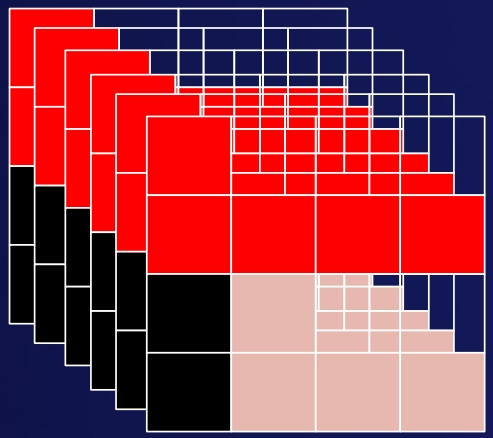
What is resolution?



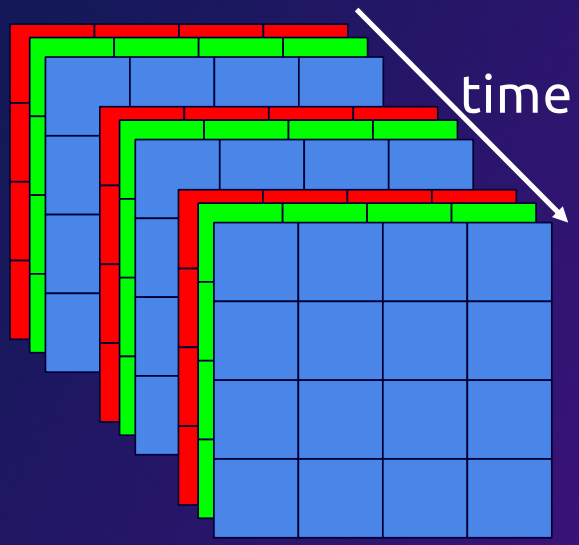
What is a video?



a single frame

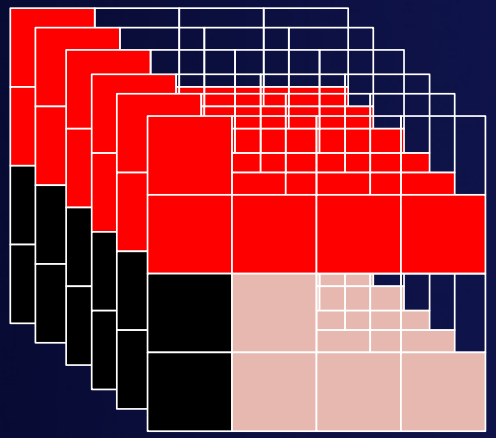


30 frames per sec (FPS)
framerate



4D

What is bitrate?



30 FPS

WIDTH * HEIGHT * BITS_PER_PIXEL * FPS

4 * 4 * 24 * 30

11,520 bits per second

Space needed to store 1h of video at 720p 30fps

WIDTH * HEIGHT * BITS_PER_PIXEL * FPS
1280 * 720 * 24 * 30
663,552,000 (663.552Mb) bits per second

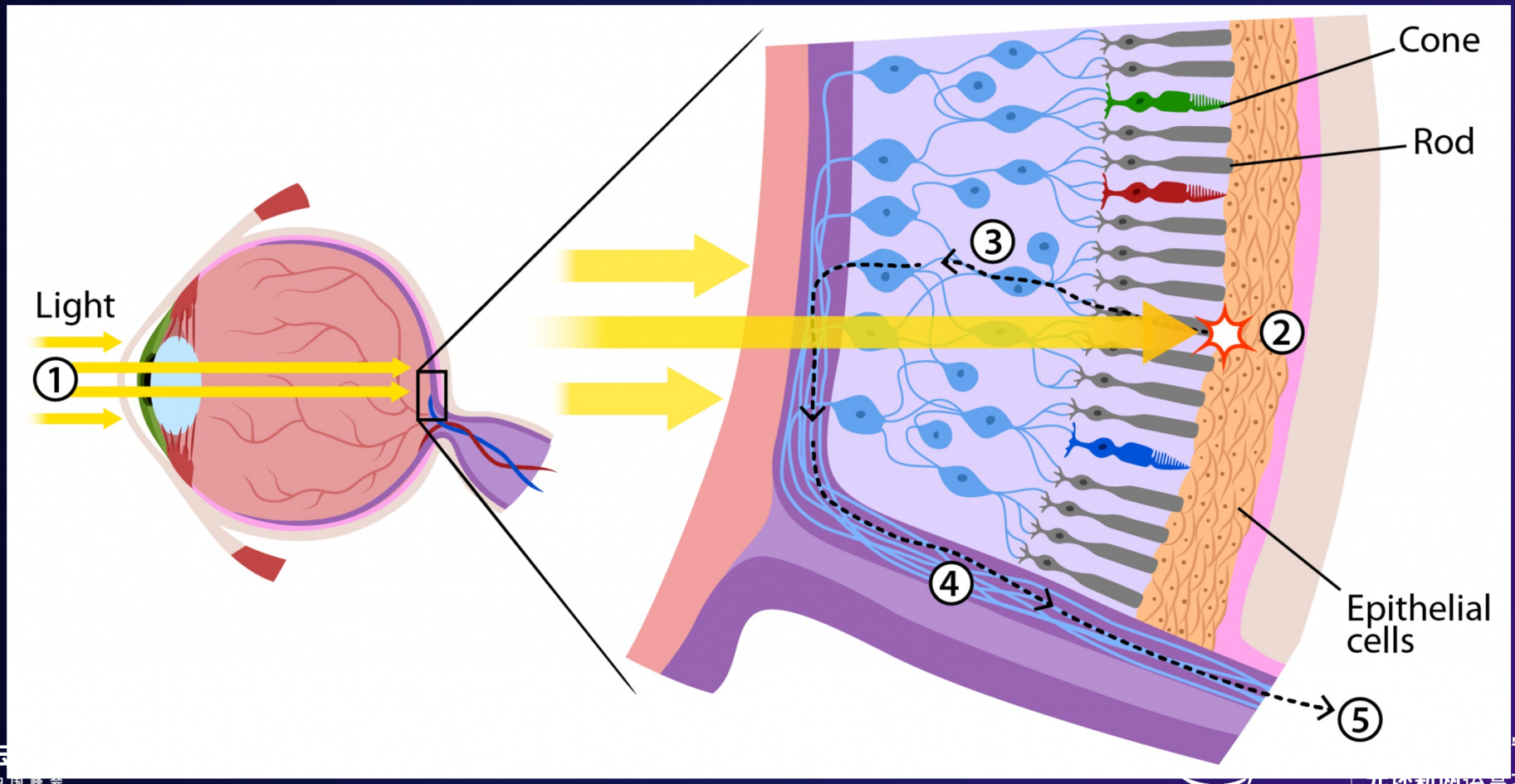
** without any compression technique at all.*

Review

image, pixel, bit depth, resolution, display aspect ratio, pixel aspect ratio, video, frame rate, interlaced, progressive, bitrate, ~~CBR, VBR, ABR~~

Exploiting our vision

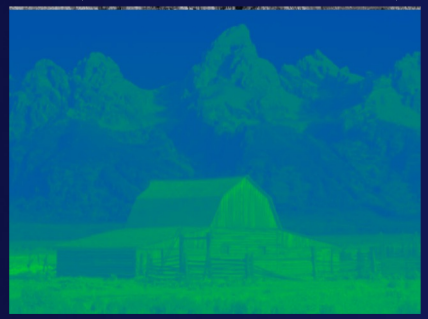
Our eyes - an oversimplification



Color space YUV (YCbCr, YPbPr)

Y (luma)

U (chroma blue)



V (chroma red)

From RGB to YCbCr

$$Y = 0.299R + 0.587G + 0.114B$$

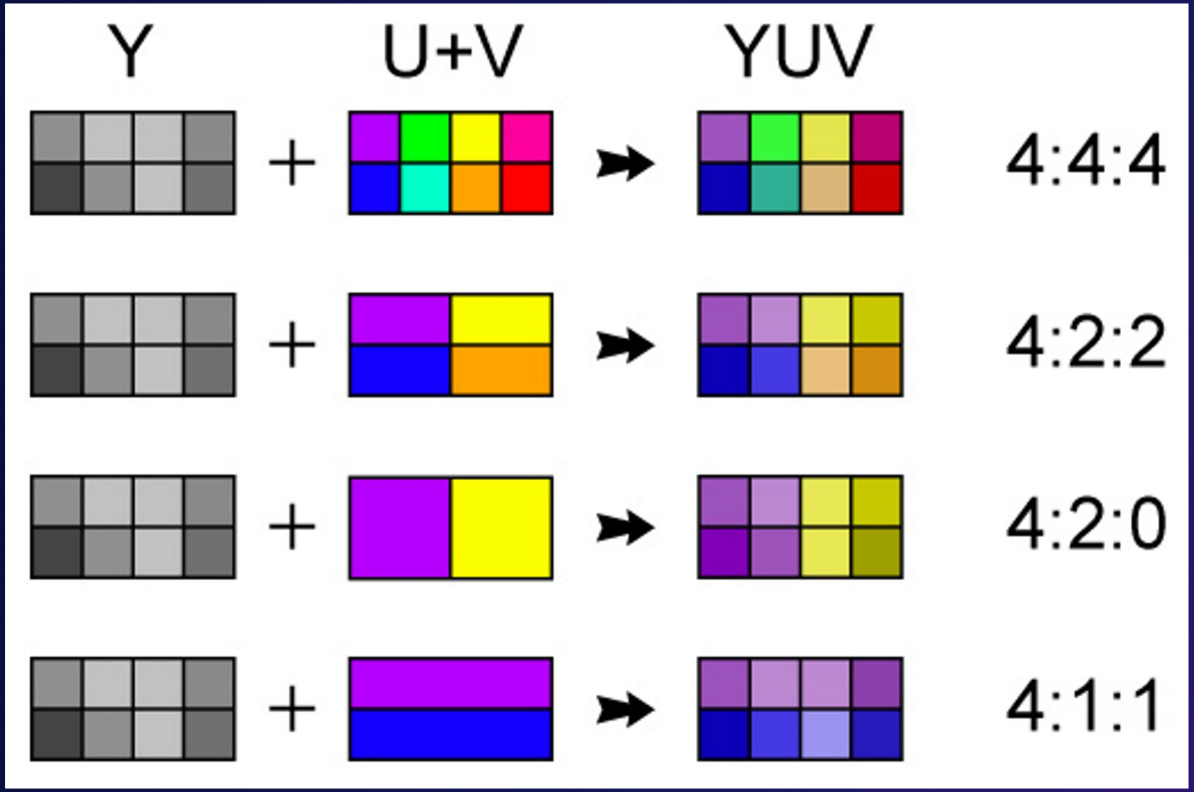
$$C_b = 0.564(B - Y) \quad | \quad C_r = 0.713(R - Y)$$

From YCbCr to RGB

$$R = Y + 1.402C_r \quad | \quad B = Y + 1.772C_b \quad | \quad G = Y - 0.344C_b - 0.714C_r$$

**ITU-R BT.601-7*

Chroma subsampling YUV 4:4:4 4:2:2 4:2:0



Space needed to store 1h of video at 720p 30fps

WIDTH * HEIGHT * BITS_PER_PIXEL * FPS
1280 * 720 * 24 * 30
663,552,000 (663.552Mb) bits per second

2,388,787,200,000 (278GB)

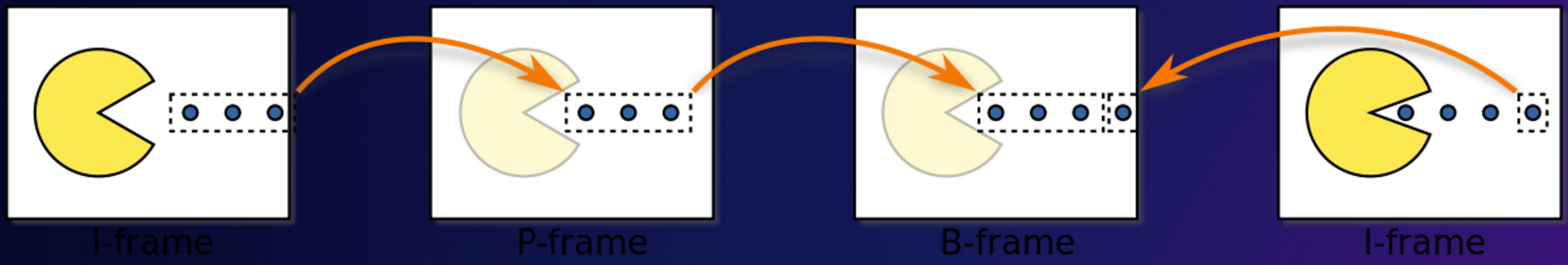
---

with chroma subsampling YUV420

WIDTH * HEIGHT * BITS_PER_PIXEL * FPS
1280 * 720 * 12 * 30
331,776,000 (331.776Mb) bits per second

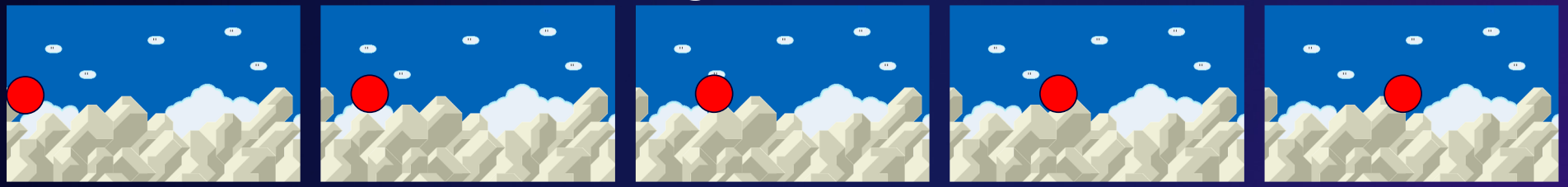
Correlations in time

Frame types

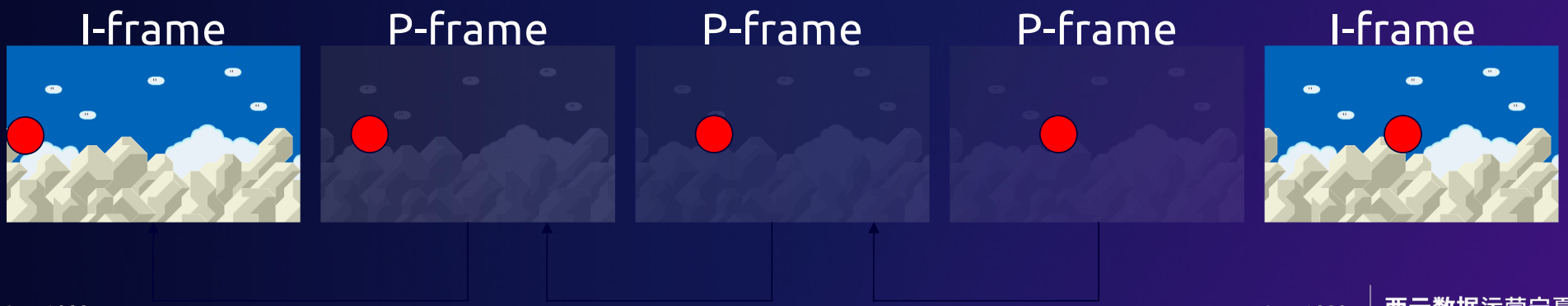


Temporal redundancy (inter prediction)

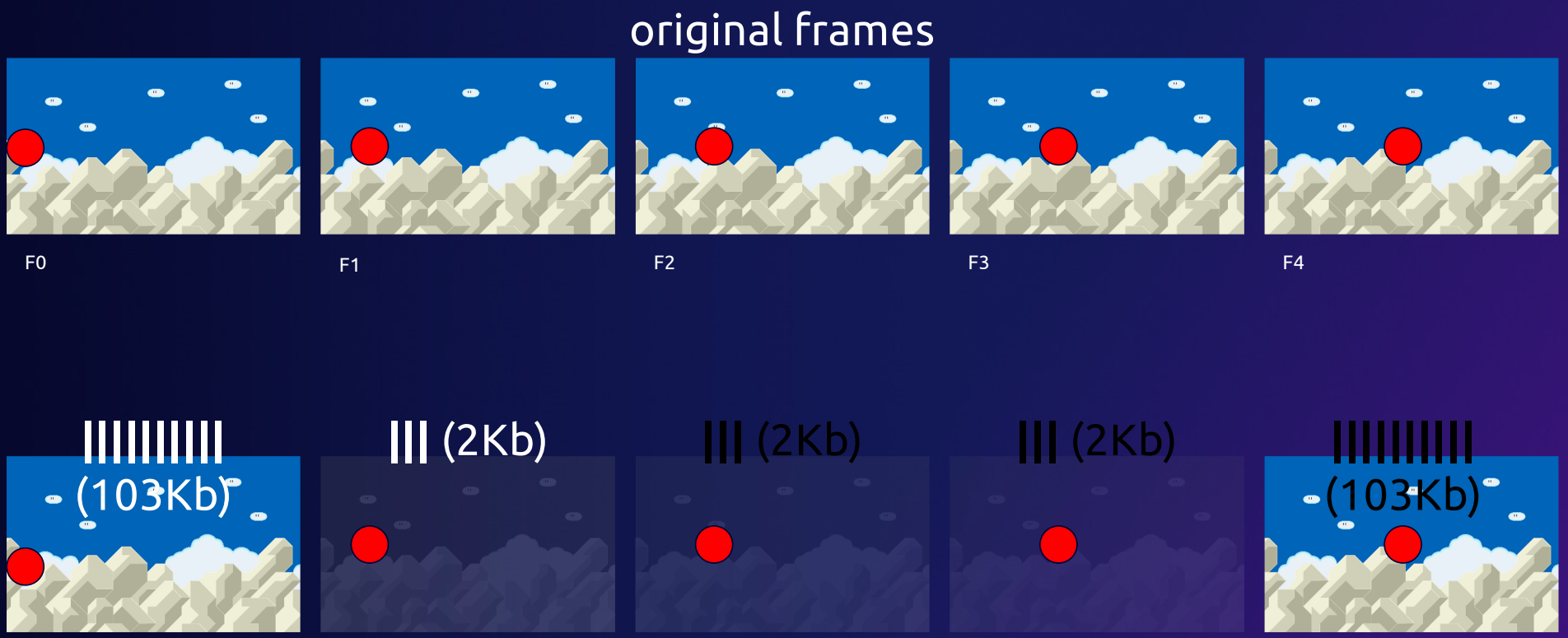
original frames

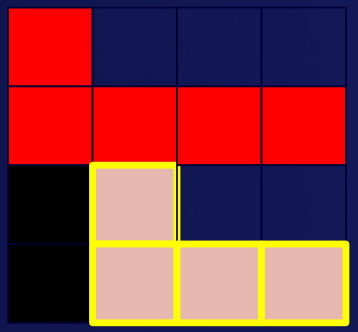


encoded frames



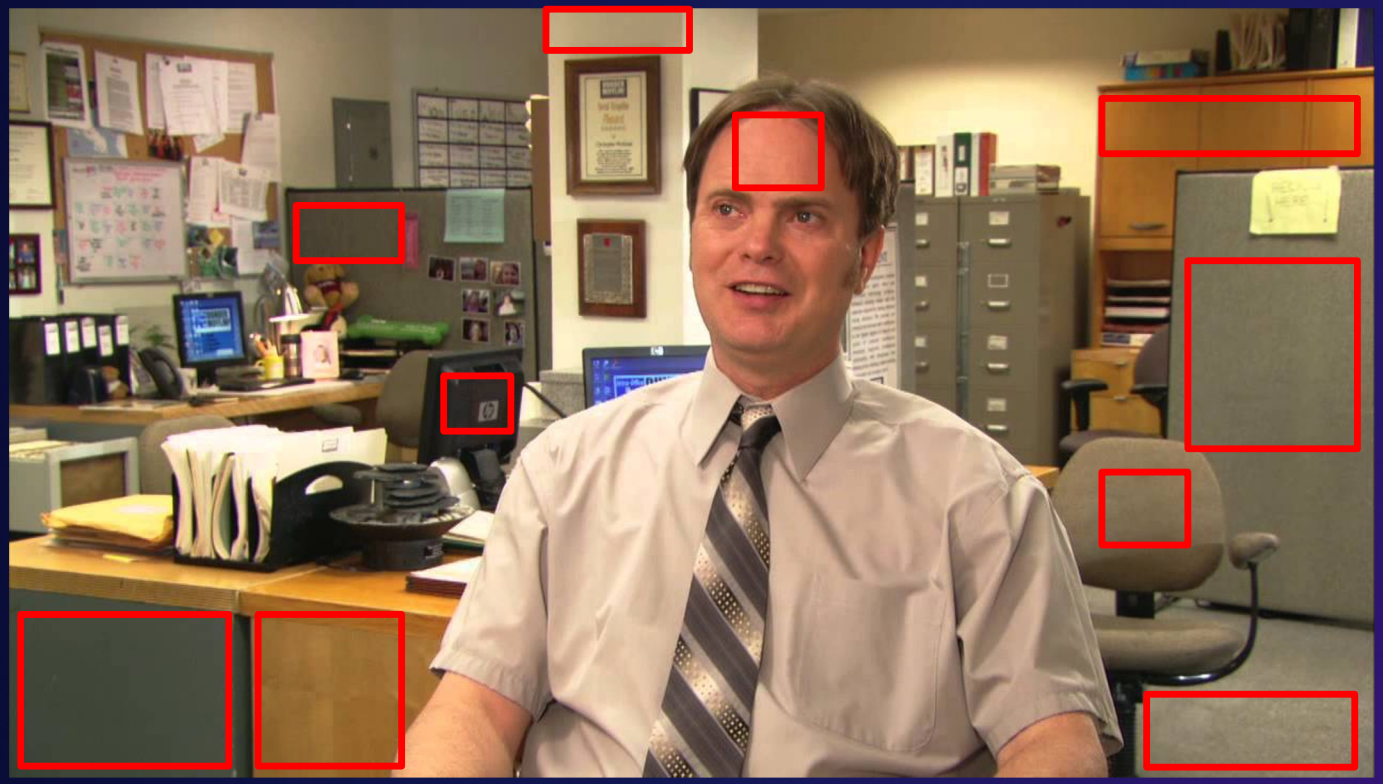
Temporal redundancy



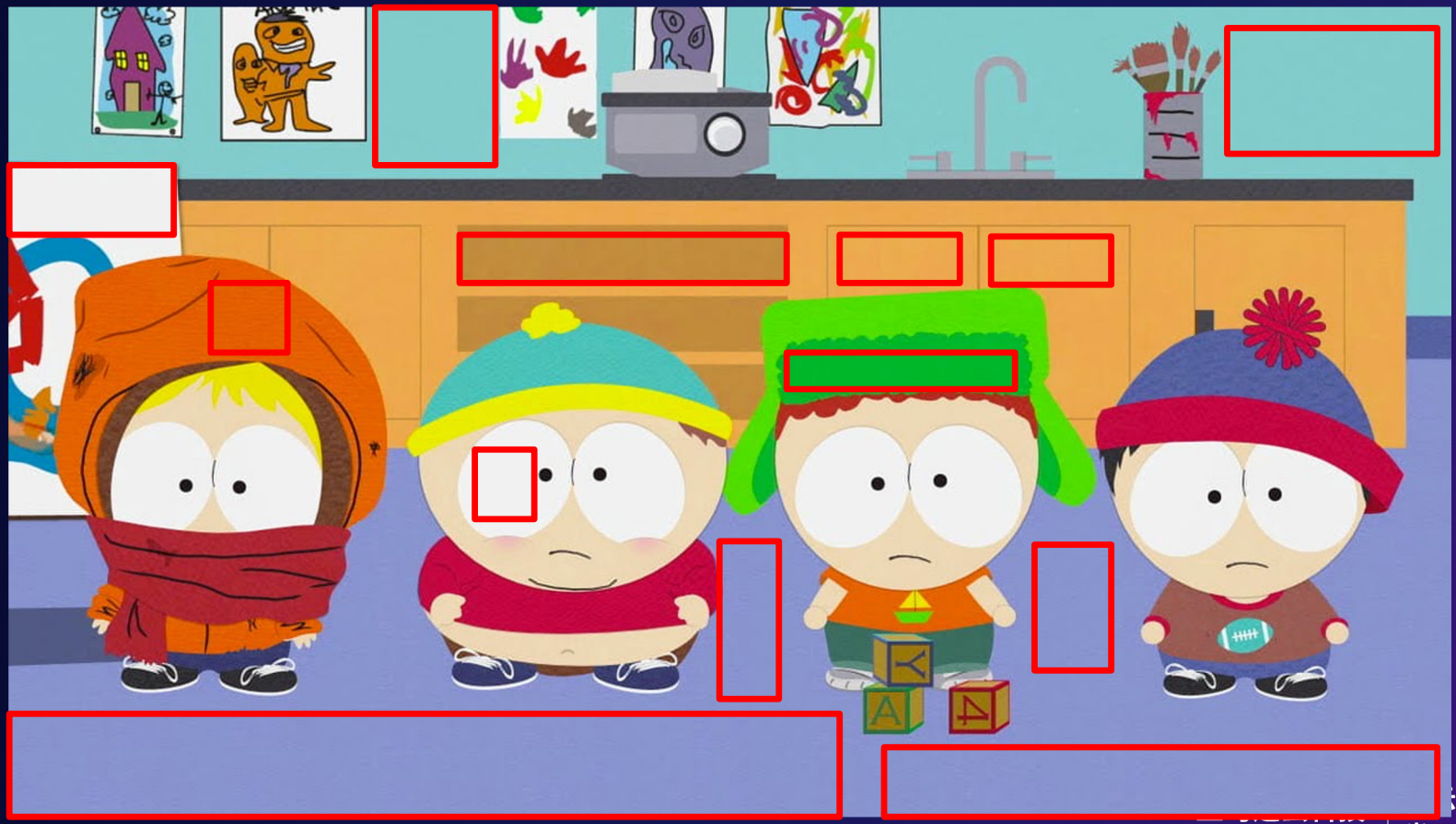


Correlations in space

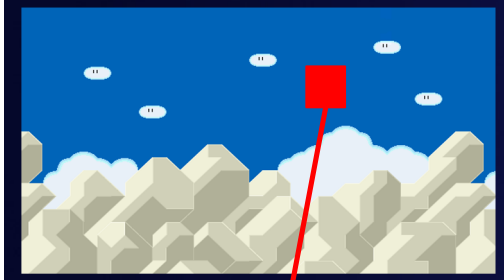
Lots of similarities



Lots of similarities



Spatial redundancy (intra prediction)



100	100	100	200
100	???	???	???
100	???	???	???
100	???	???	???

unknown values

100	100	100	200
100	100	100	200
100	100	100	200
100	100	100	200

direction of the prediction

—

100	100	100	200
100	100	100	200
100	100	100	200
100	100	120	210

real values

=

100	100	100	200
100	0	0	0
100	0	0	0
100	0	20	10

difference

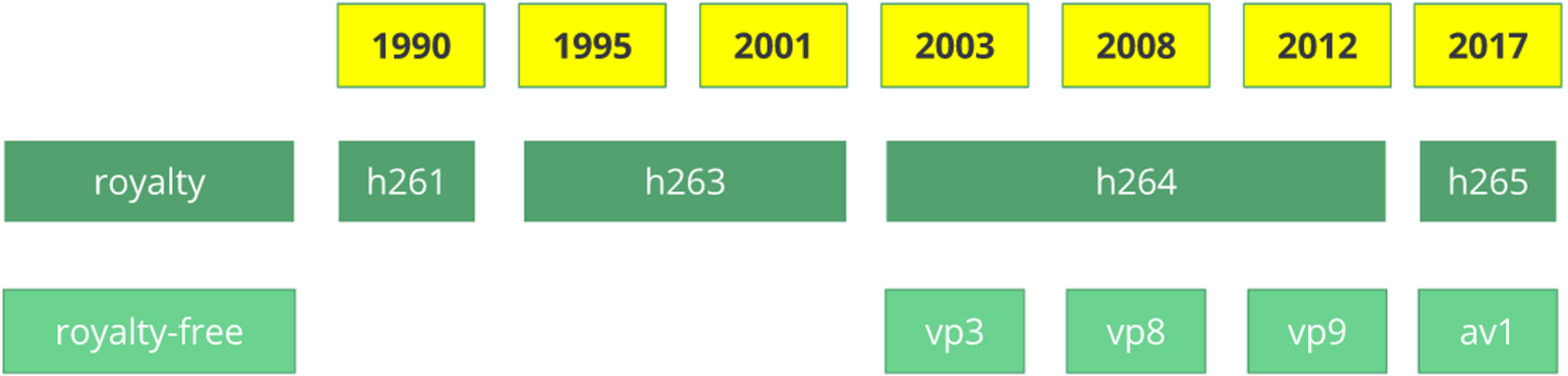
highly compressible

CODEC - enCOder / DECoder

CODEC

“A codec is a device or computer program for encoding or decoding a digital data stream or signal.”

History


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

Space needed to store 1h of video at 720p 30fps

with chroma subsampling YUV420

WIDTH * HEIGHT * BITS_PER_PIXEL * FPS

1280 * 720 * 12 * 30

331,776,000 (331.776Mb) bits per second

1,194,393,600,000 (139GB)

---

with H264 (chroma subsampling, motion estimation, intra prediction, CABAC…)

WIDTH * HEIGHT * BITS_PER_PIXEL * FPS

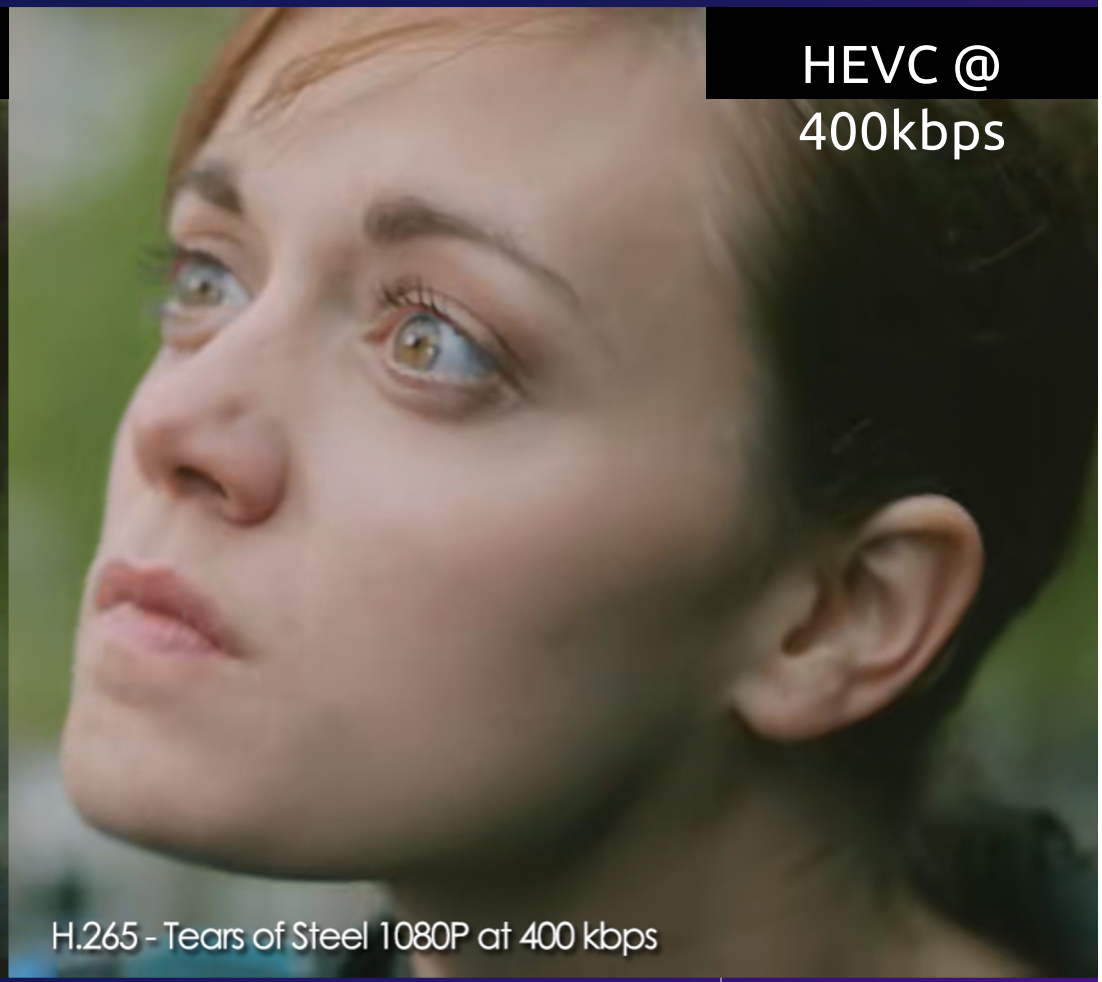
1280 * 720 * 0.031 * 30

857,088 (837Kb) bits per second

H264 vs H265


AVC @
4Mbps
00;02;36;13
HEVC @
2Mbps
00;02;36;13


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


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

H.264/AVC

16x16 Macroblock

Block coding structure

3 Intra partitioning

4 Inter partitioning

+4 sub-partitioning 8x8

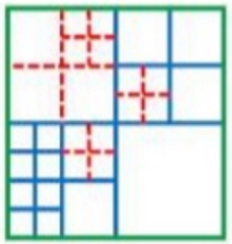
2 Transform sizes:
4x4, 8x8

Up to 9 Intra prediction directions

HEVC

Coding Unit
64x64 to 8x8

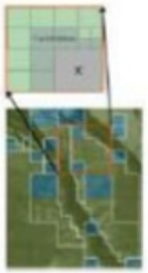
Quadtree coding structure



Prediction Unit and Transform Unit partitioning

⇨ Multiples sizes/forms: 64x64 to 4x4

35 Intra prediction directions

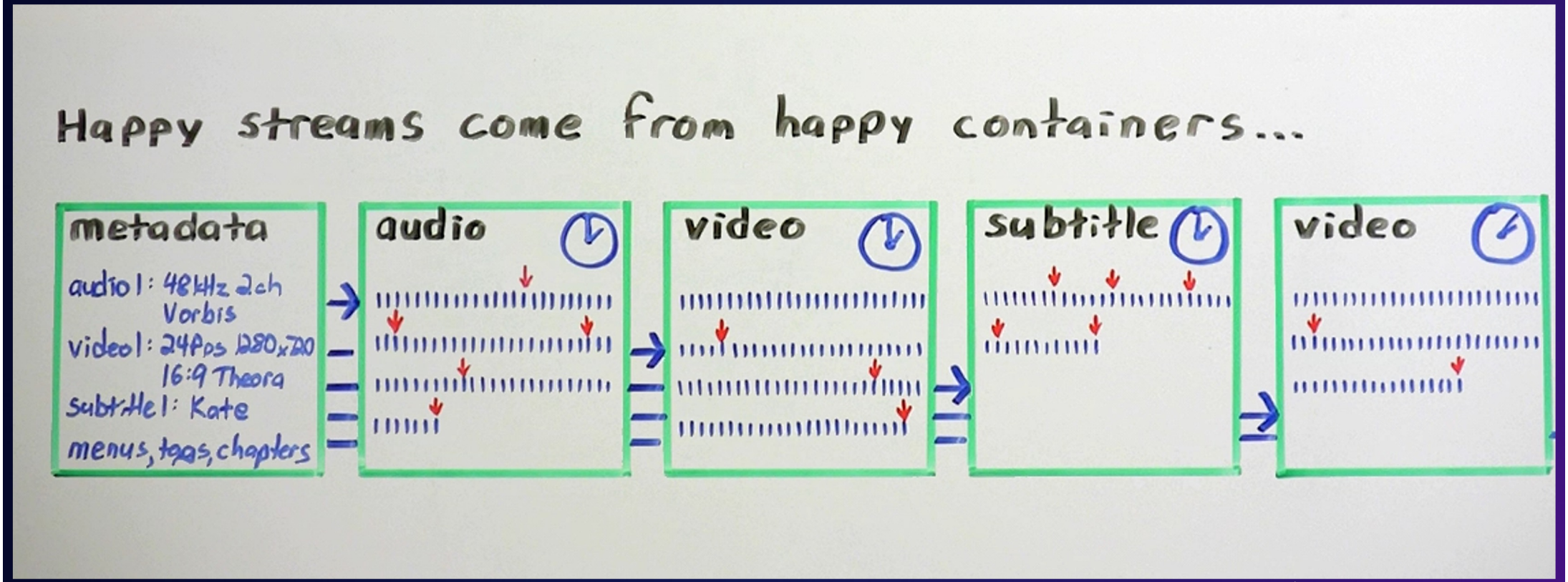


Efficient spatio-temporal mv prediction

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

Container

CODEC (VP9, H265) vs Container (.WEBM,.MP4)



Source https://xiph.org/video/vid1.shtml

Container vs CODEC

Containers

- OGG
- MP4
- WMA
- AVI
- MKV, WebM
- TS
- MOV

CODEC

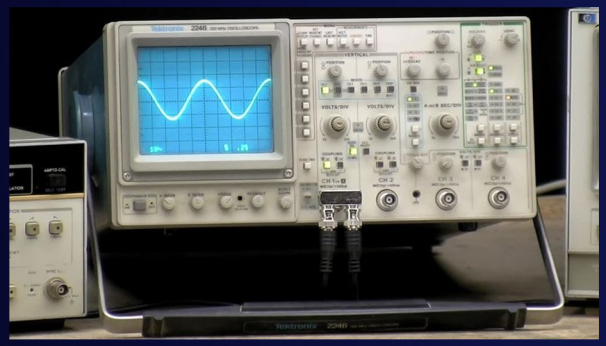
- H264 / AVC
- H265 / HEVC
- MPEG-4
- VP9
- AV1
- Theora
- Daala

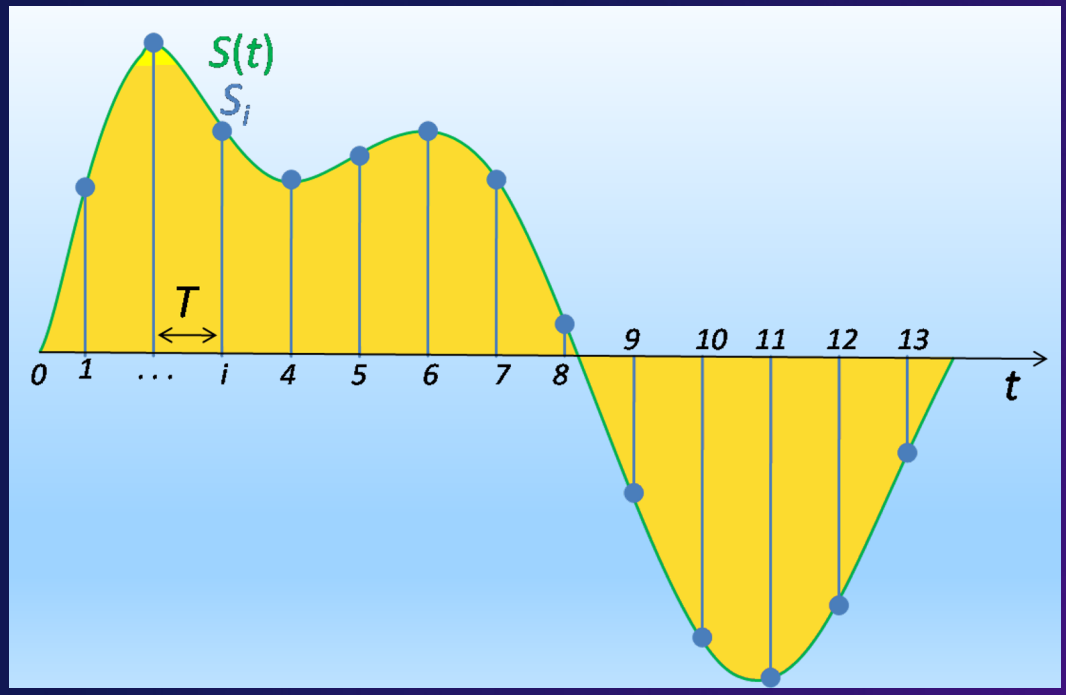
Video formats supported [edit]

	MPEG-1	MPEG-2	MPEG-4 (A)SP	H.264/ MPEG-4 AVC	H.265/ HEVC	VC-1/ WMV	Real Video	Theora	Microsoft MPEG4 V2	VP8	VP9	MVC
AVI	Yes	Yes	Yes	Yes[41]	No	Yes	Yes	Yes	Yes	Yes	No	?
Matroska	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes[42]
MP4	Yes	Yes	Yes	Yes	Yes	Yes[43]	Yes	Yes	Yes	Yes	Yes	Yes
MXF	Yes	Yes	Yes	Yes	?	Yes	No	?	?	?	?	?
Ogg	Yes	Yes	Yes	Yes	?	Yes	No	Yes	?	?	?	?
QuickTime	Yes	Yes	Yes	Yes	No	Yes	No	Yes	?	?	?	?

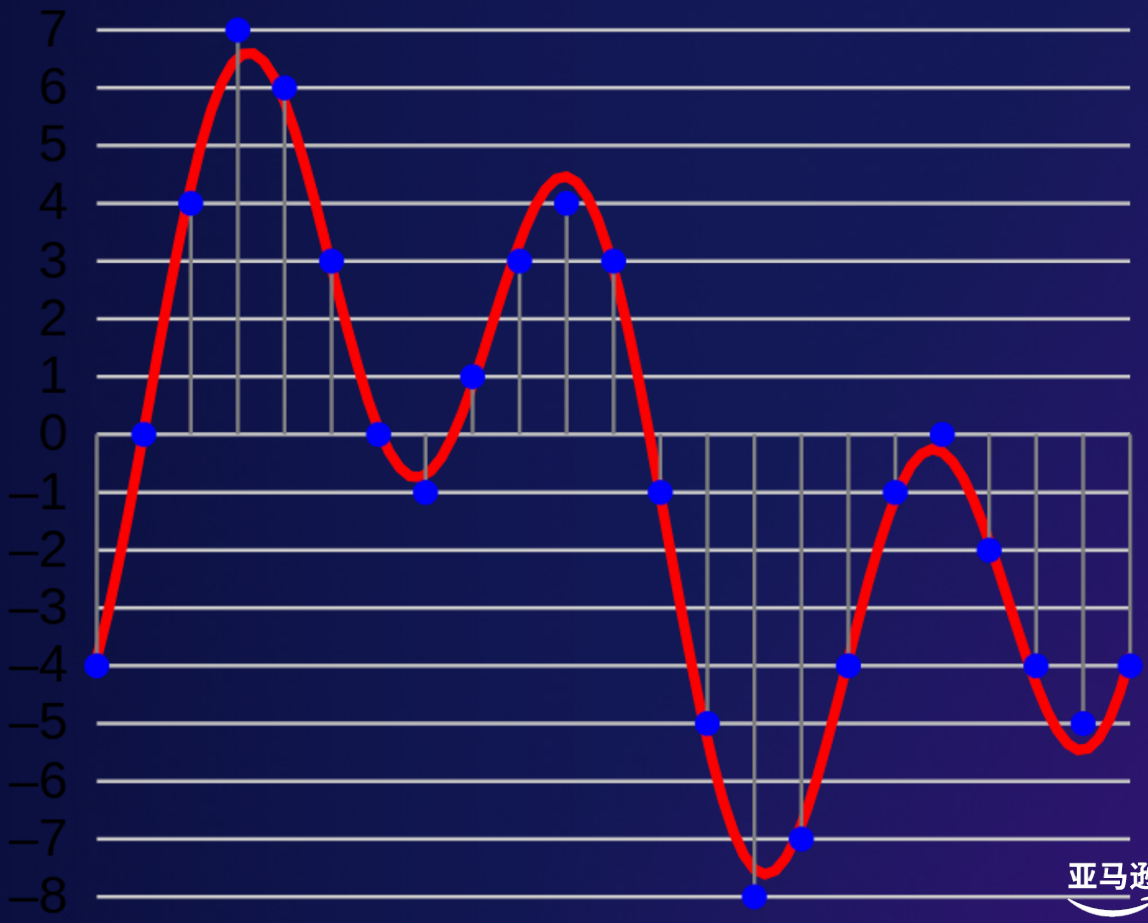
Bonus: audio codec

Sampling (8,11, 32, 44.1, 48, 50, 88, 96, 192... kHz)



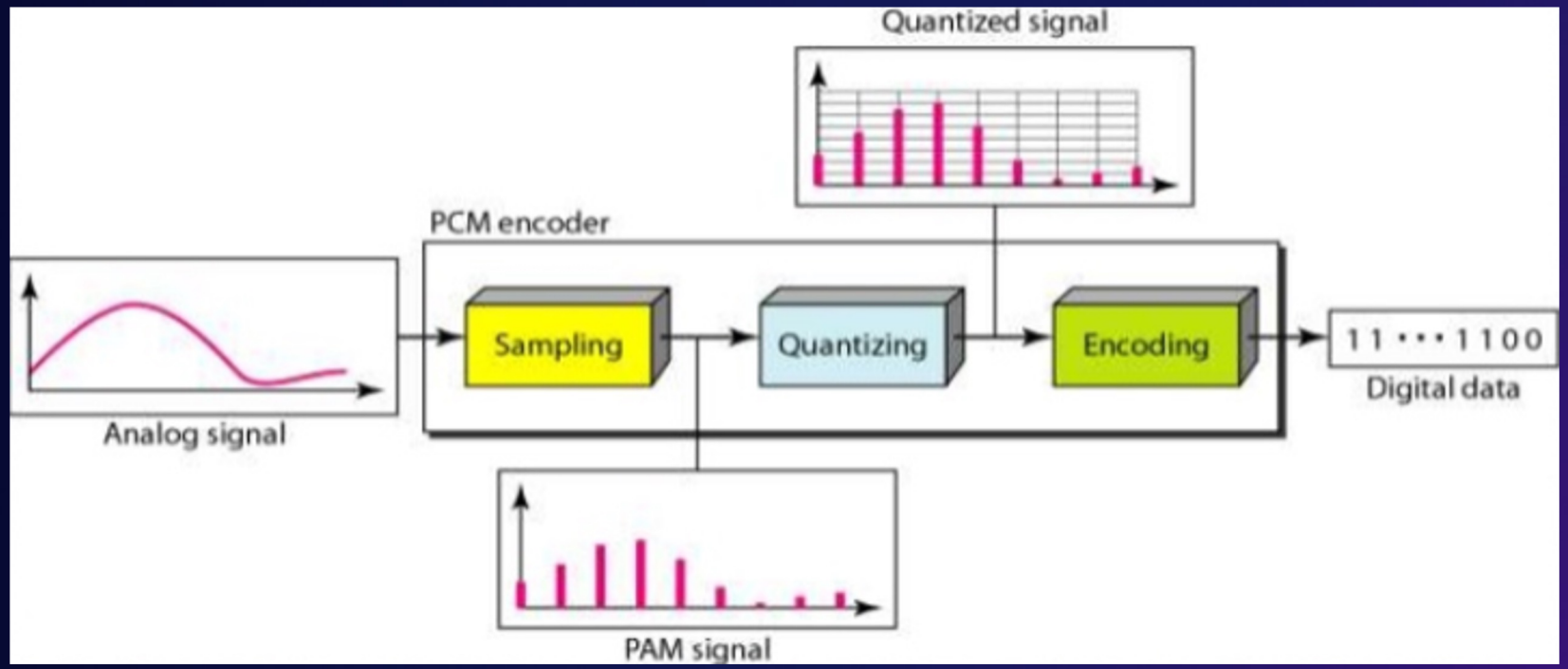


Bit depth (16, 24, 32, 64... bits) quantization

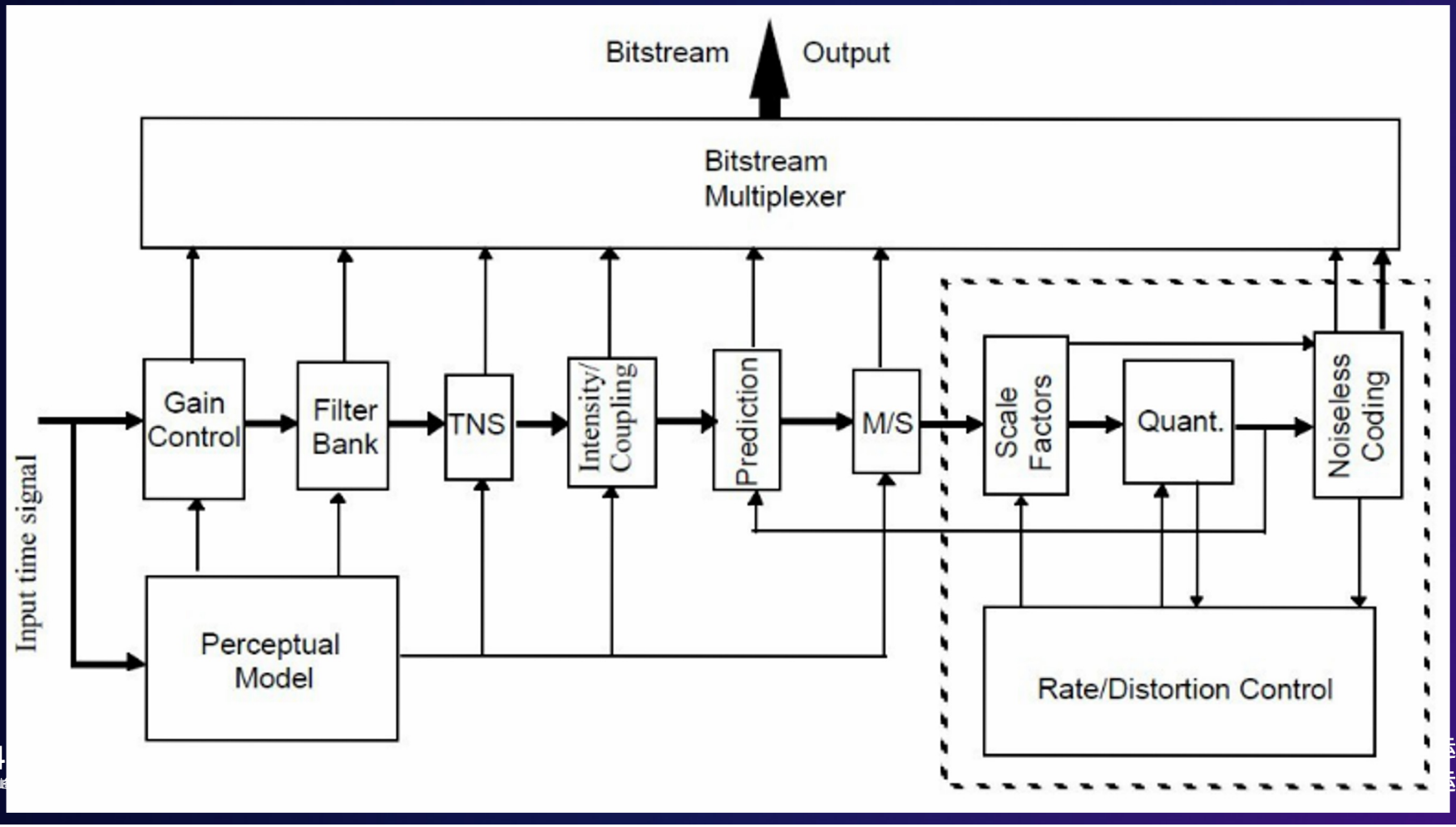


Channels 2

PCM encoder

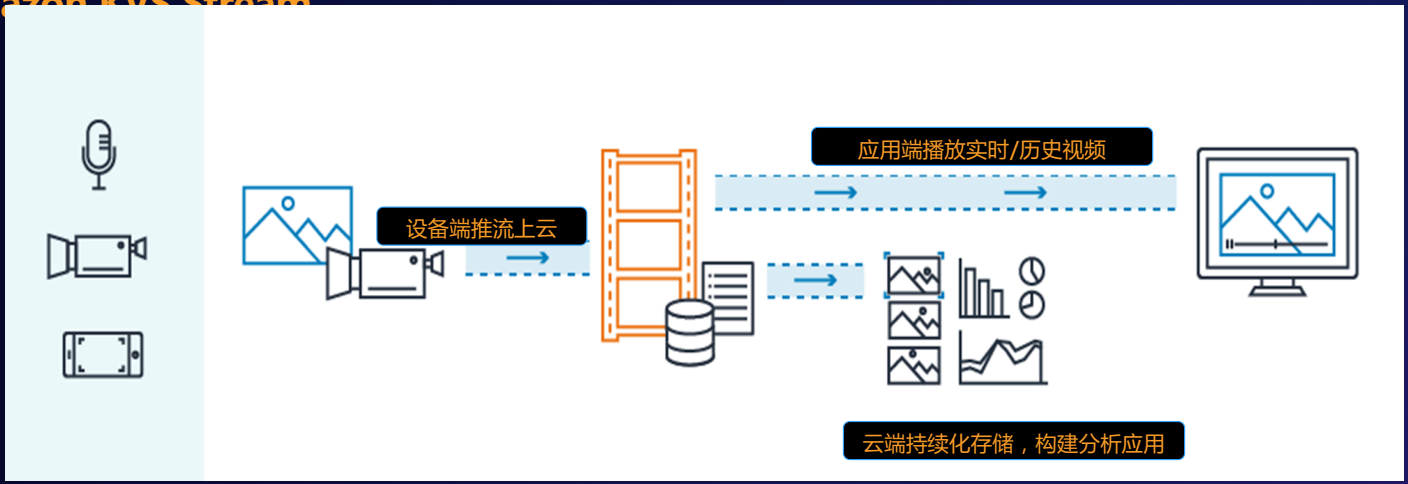

Quantized signal
PCM encoder
Sampling
Quantizing
Encoding
11 • • • 1100
Digital data
Analog signal
PAM signal

AAC CODEC block



Amazon KVS Streaming

Amazon KVS Stream



输入

设备端安装Producer SDK，将设备端产生的视频数据封装并以流的方式推入云端。

存储

Amazon Kinesis Video Stream云端服务以时间可索引的方式存储设备端推入的视频流。

消费

用户端基于请求播放实时视频流或者是云端存储的历史视频流；有些客户基于Amazon KVS流数据构建了分析应用，比如使用Amazon Rekognition，进一步提高硬件产品竞争力。

Adaptive bitrate streaming (HLS)

source: https://www.encoding.com/http-live-streaming-hls/

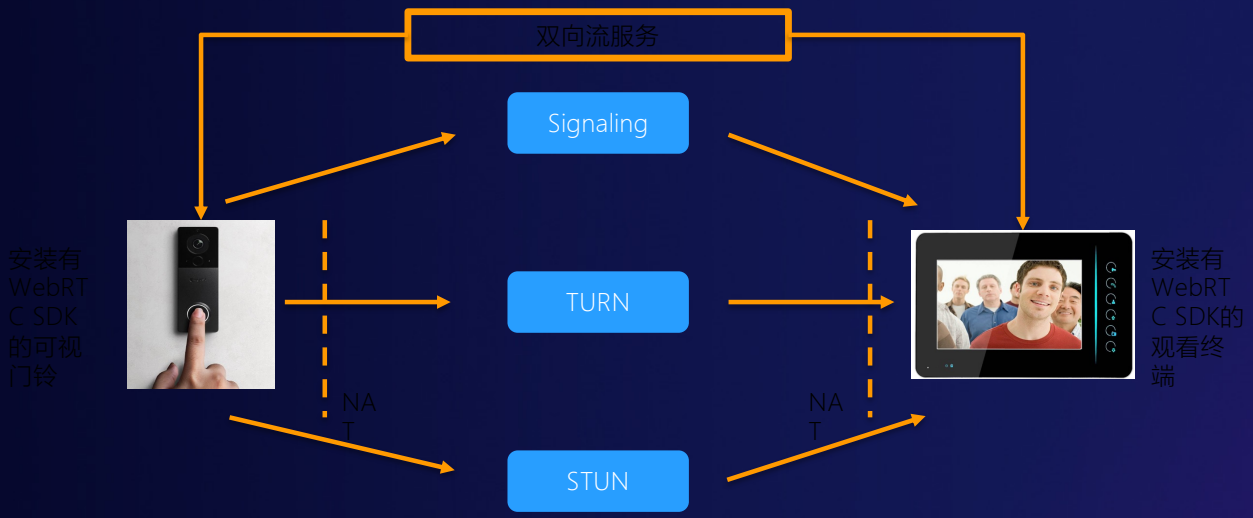
HLS vs. DASH

- **Encoding formats:** MPEG-DASH allows the use of any encoding standard. HLS, on the other hand, requires the use of H.264 or H.265.
- **Device support:** HLS is the only format supported by Apple devices. iPhones, MacBooks, and other Apple products cannot play video delivered over MPEG-DASH.
- **Segment length:** This was a larger difference between the protocols before 2016, when the default segment length for HLS was 10 seconds. Today the default length for HLS is 6 seconds, although it can be adjusted from the default. MPEG-DASH segments are usually between 2 and 10 seconds in length, although the optimum length is 2-4 seconds.
- **Standardization:** MPEG-DASH is an international standard. HLS was developed by Apple and has not been published as an international standard, even though it has wide support.

Amazon KVS WebRTC

Amazon KVS WebRTC

- Amazon KVS WebRTC为物联网视频设备提供一秒以内超低延时实时音视频通讯服务。
- Amazon KVS WebRTC是基于开源项目框架，支持客户进一步开发，透明性好。
- Amazon KVS WebRTC是基于亚马逊云科技的全球基础设施，支持全球部署并提供SLA的标准云产品



- ***信令交互 Signaling***, 信令服务器负责整个P2P连接的发现，建立，控制，和终结
- ***局域网转发 STUN***, 一旦P2P连接建立，视频流直接通过路由器从主设备(Master)到观看终端(Viewer)
- ***云转发 TURN***, 如果P2P连接尝试失败，视频流通过云端转发给观看终端(Viewer)

Thank you!

亚马逊云科技
中国峰会

亚马逊云科技
西云数据运营宁夏区域
光环新网运营北京区域