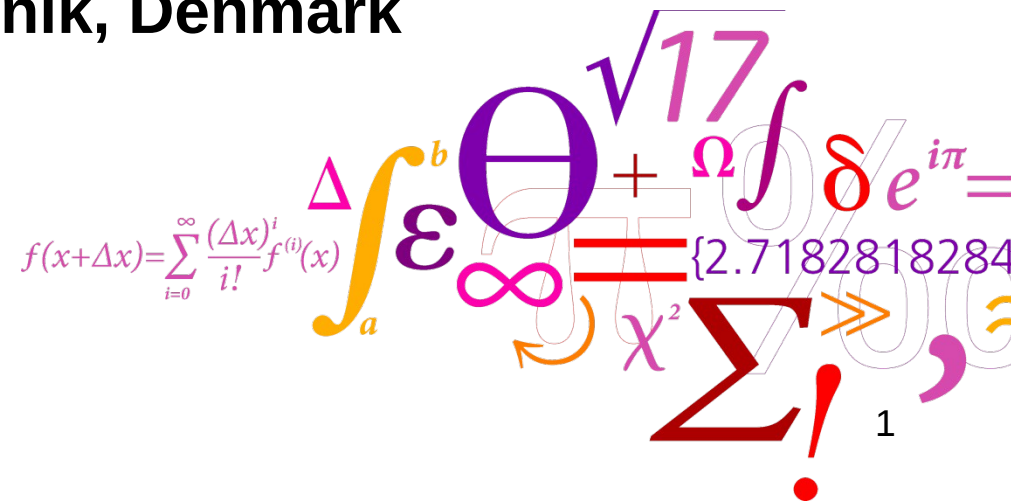
Overview of H.264/AVC

Ann Ukhanova

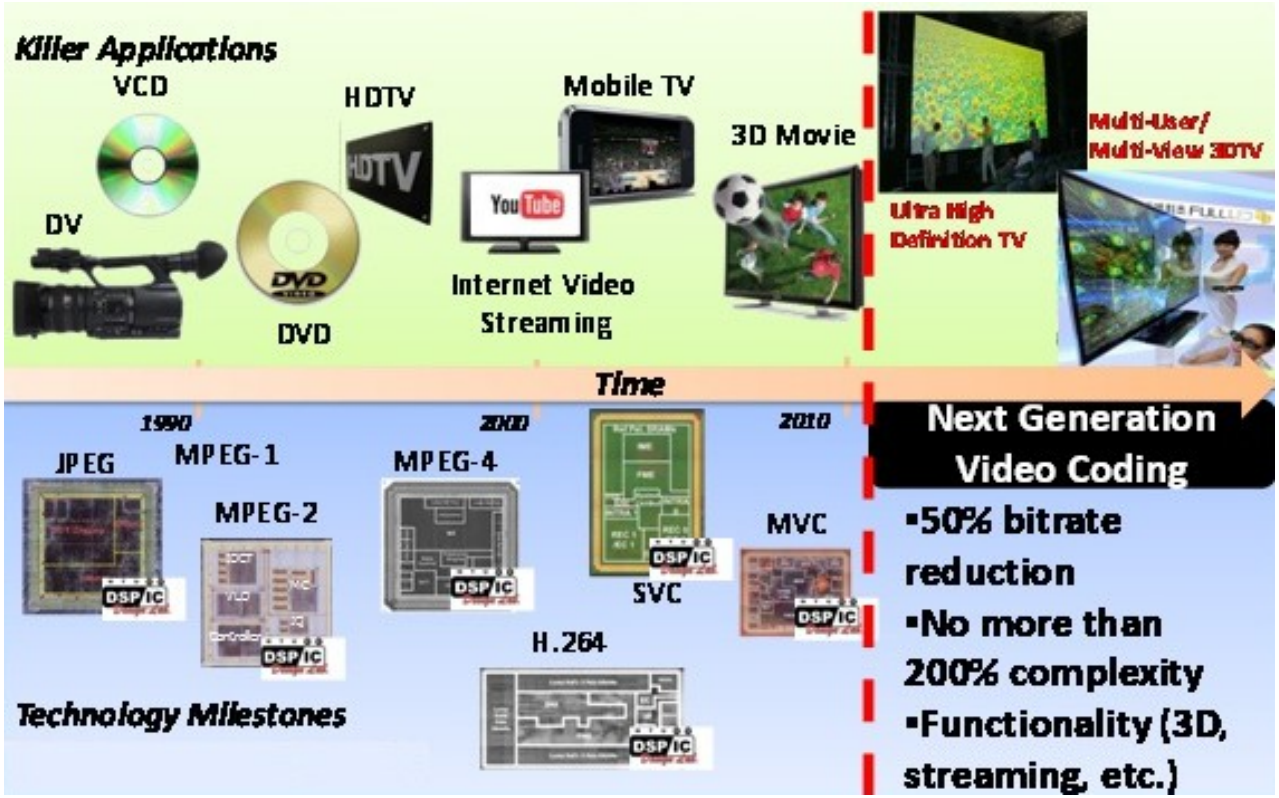
DTU Fotonik, Denmark



DTU Fotonik
Institut for Fotonik

Trend of video coding

Video coding is the process of reducing the amount of data required to represent a digital video signal, prior to transmission or storage



Video content everywhere
& not enough bandwidth

Computation increasing on every platform

Won't this be unnecessary when Megabits become free?

The need for better compression will not be reduced

- Got more bits? Give me **higher** resolution!
- Got more bits? Give me **more** channels!
- Improving worth effort? 20% of a lot is **a lot**.

Motivating video compression

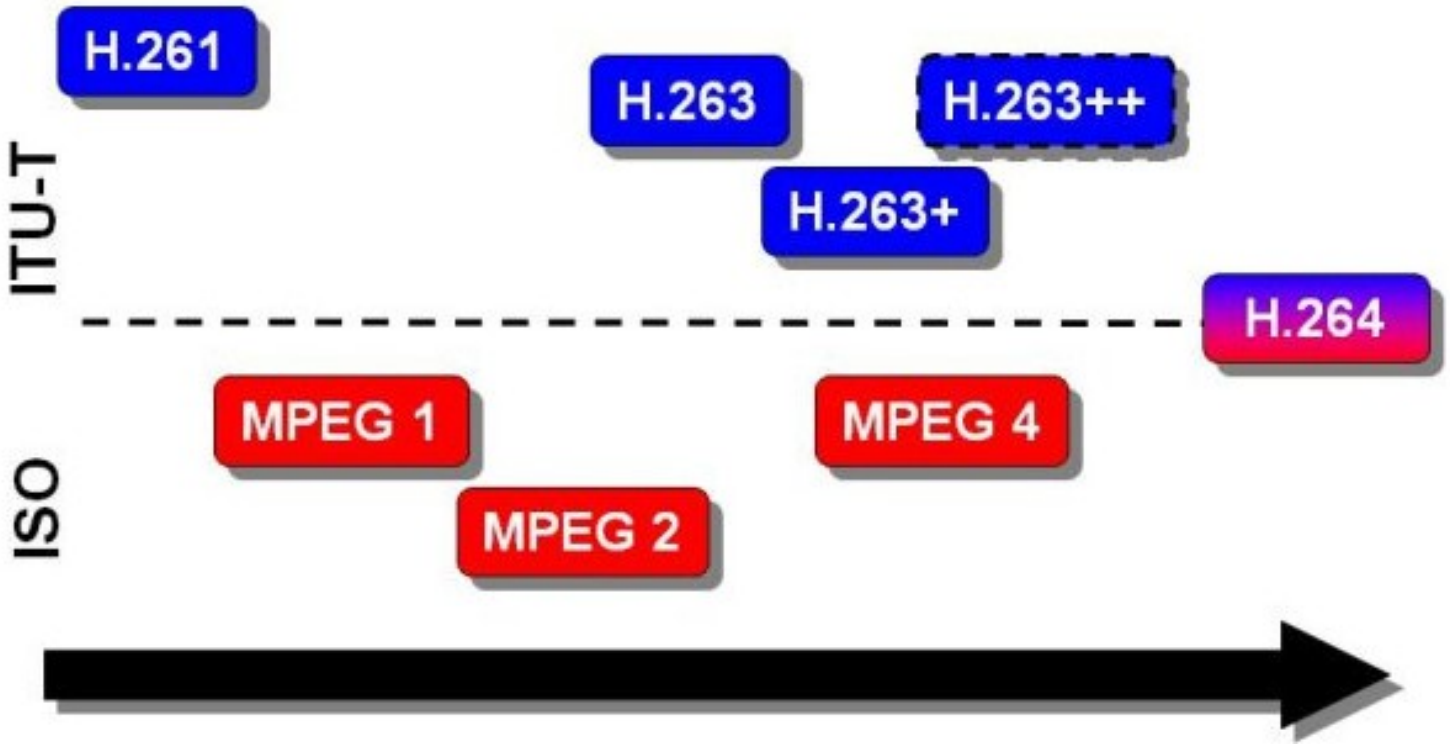
- **Digital video studio standard ITU-R Rec. 601**
  - 525-line 60 Hz or 625-line 50 Hz signals
  - 720 luminance samples and 360 chrominance samples per line.
  - The color encoding system is known as YUV 4:2:2.

Raw bitrate: 216Mbps

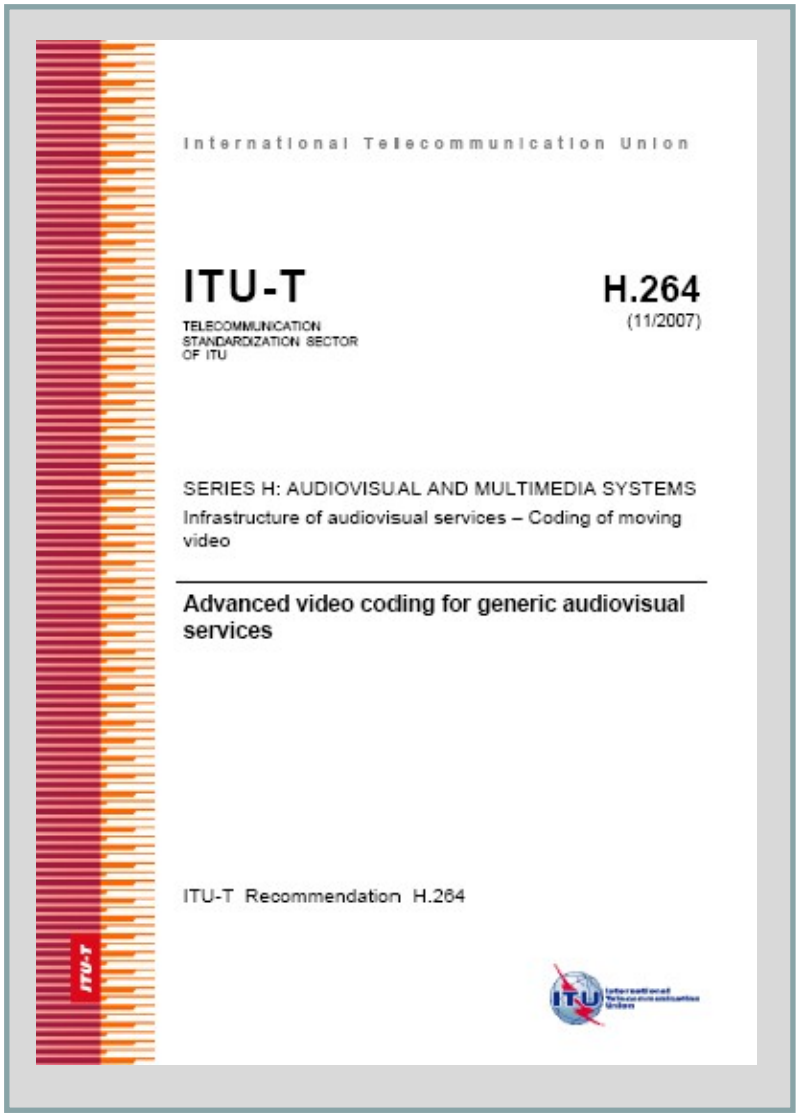
- **Some interesting bit-rates**

Terrestial TV broadcasting channel	~20 Mbps
Computer hard disk	20...40 Mbps
DVD (max. 17 GB/length of movie)	10...20 Mbps
Ethernet/Fast Ethernet	<10/100 Mbps
DSL downlink	384...2048 kbps
Dial-up modem	14.4…56 kbps
Wireless cellular data	9.6...112 kbps

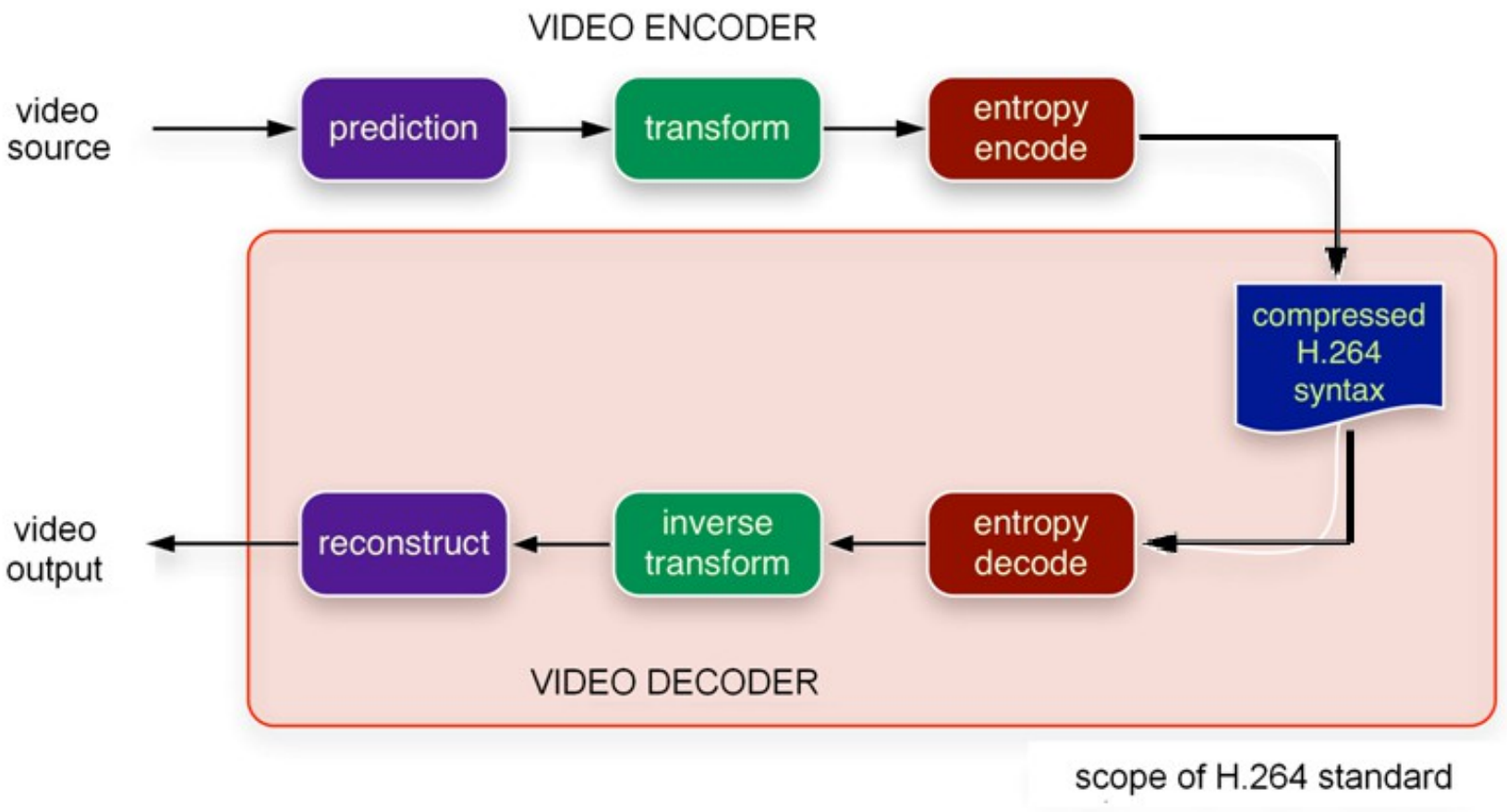
Video compression standards



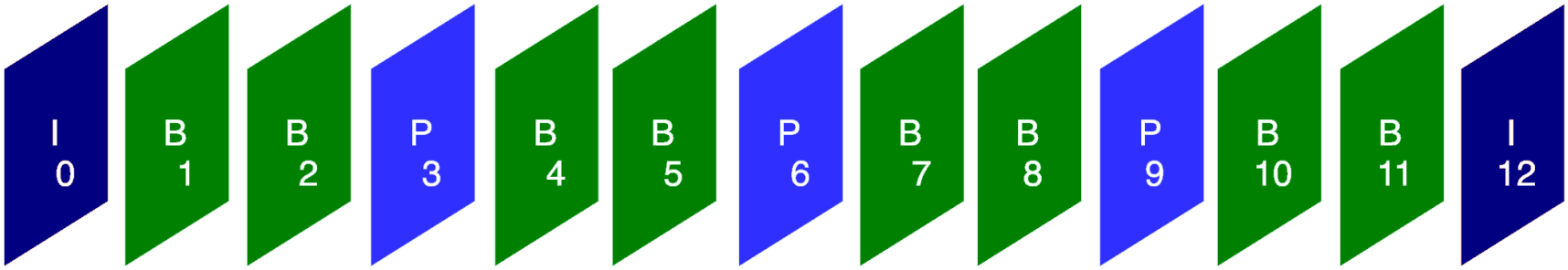
H.264 Standard



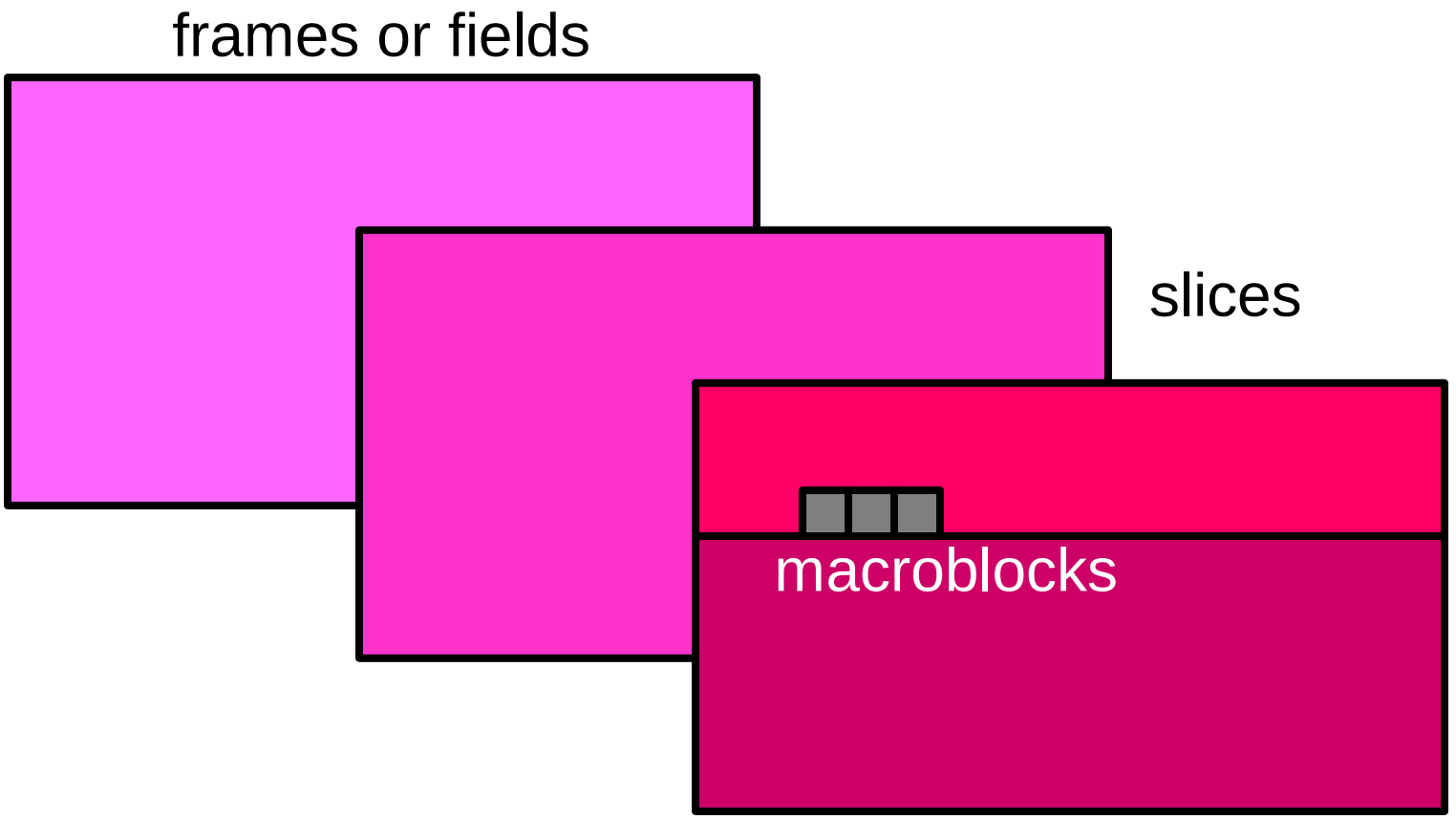
Scheme of H.264 codec



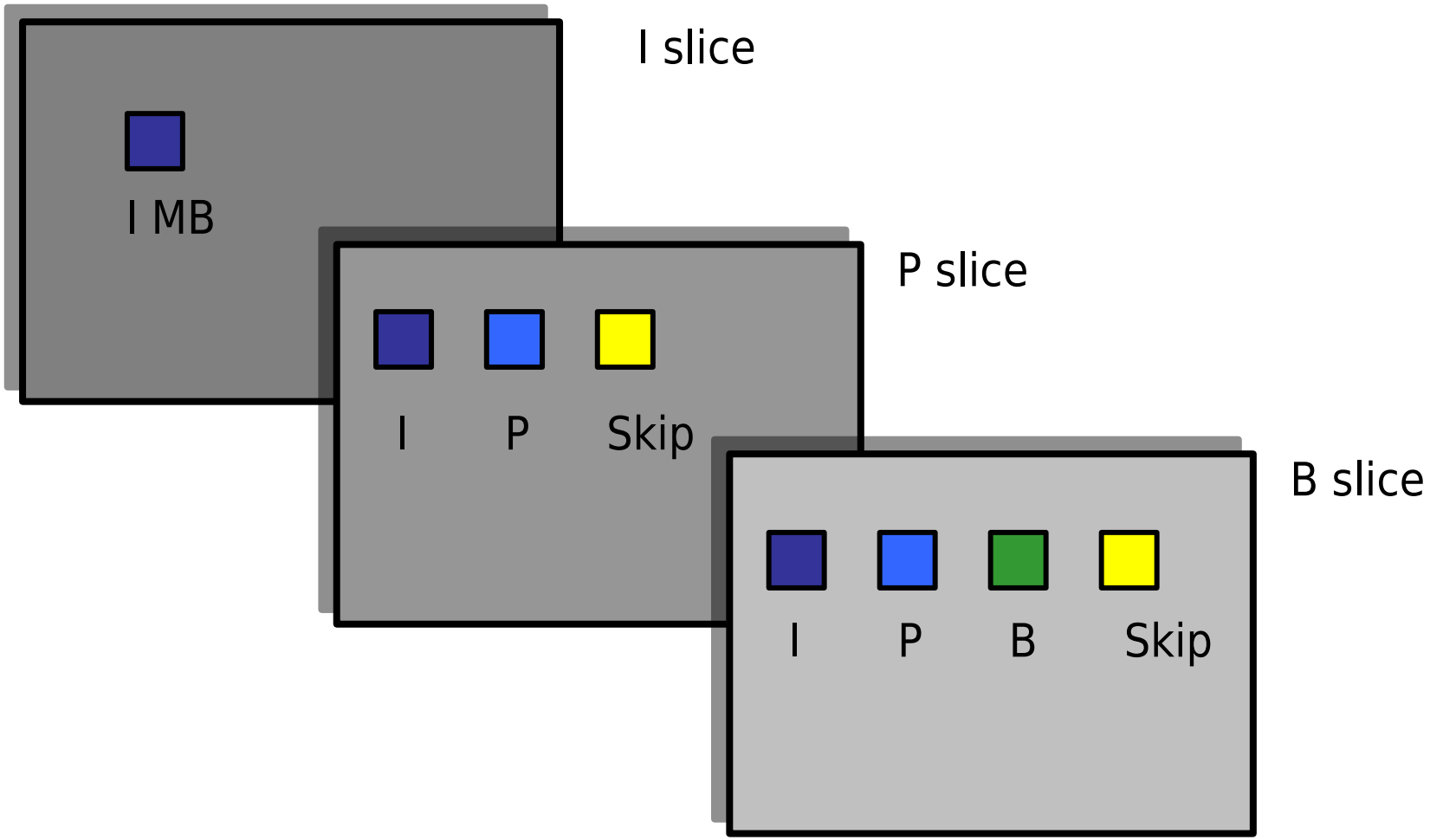
GOP structure



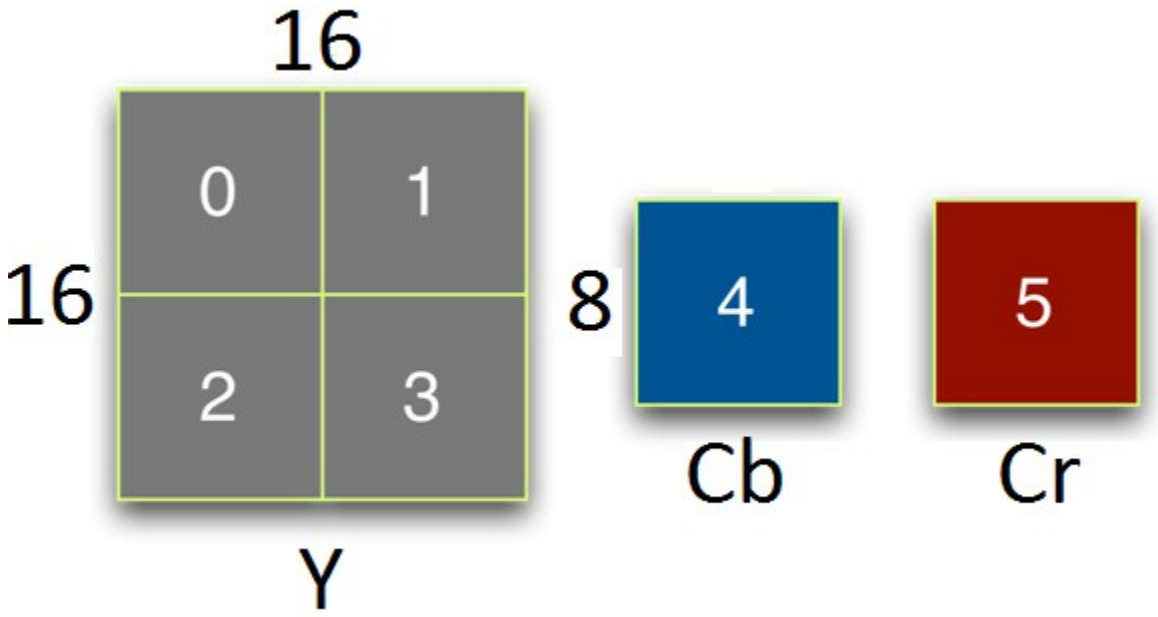
Frame, field, slice, MB



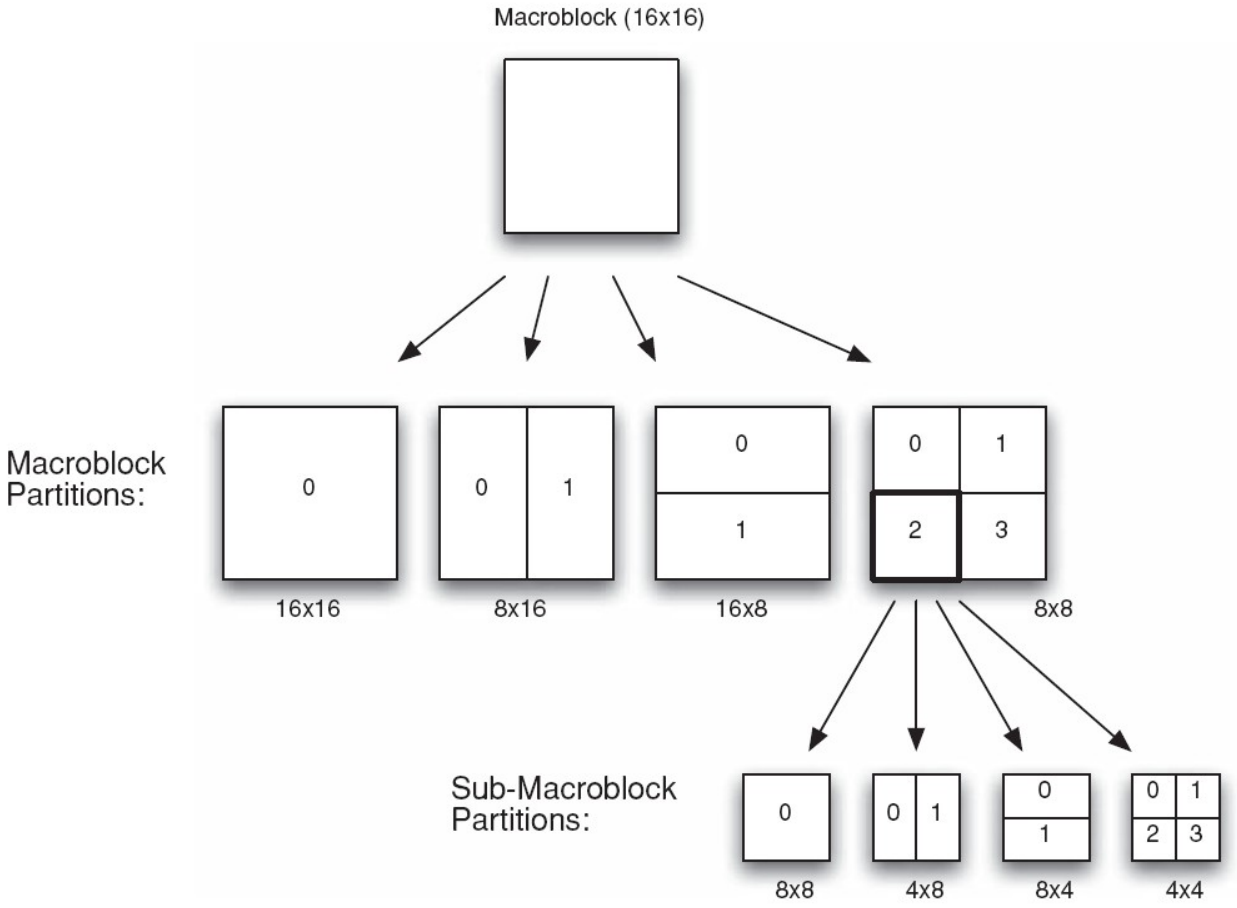
Macroblock options



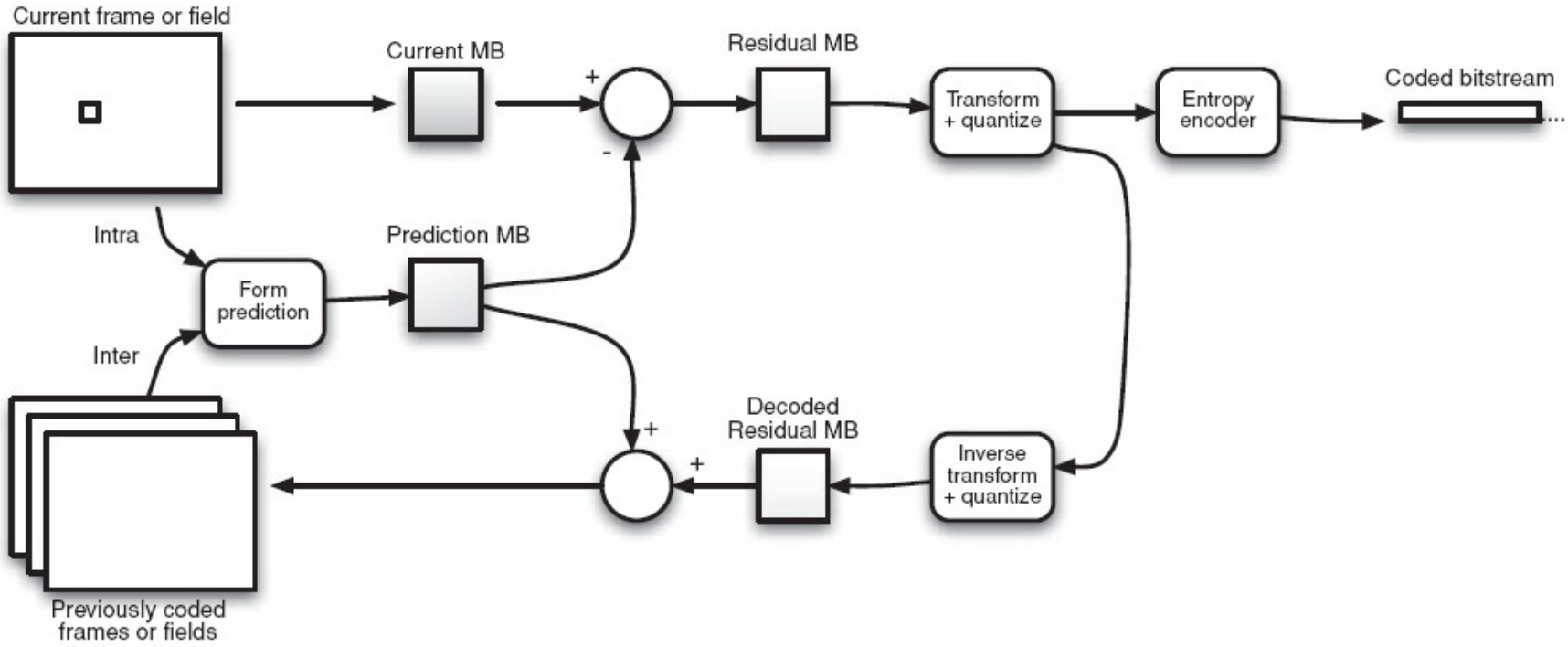
Macroblock



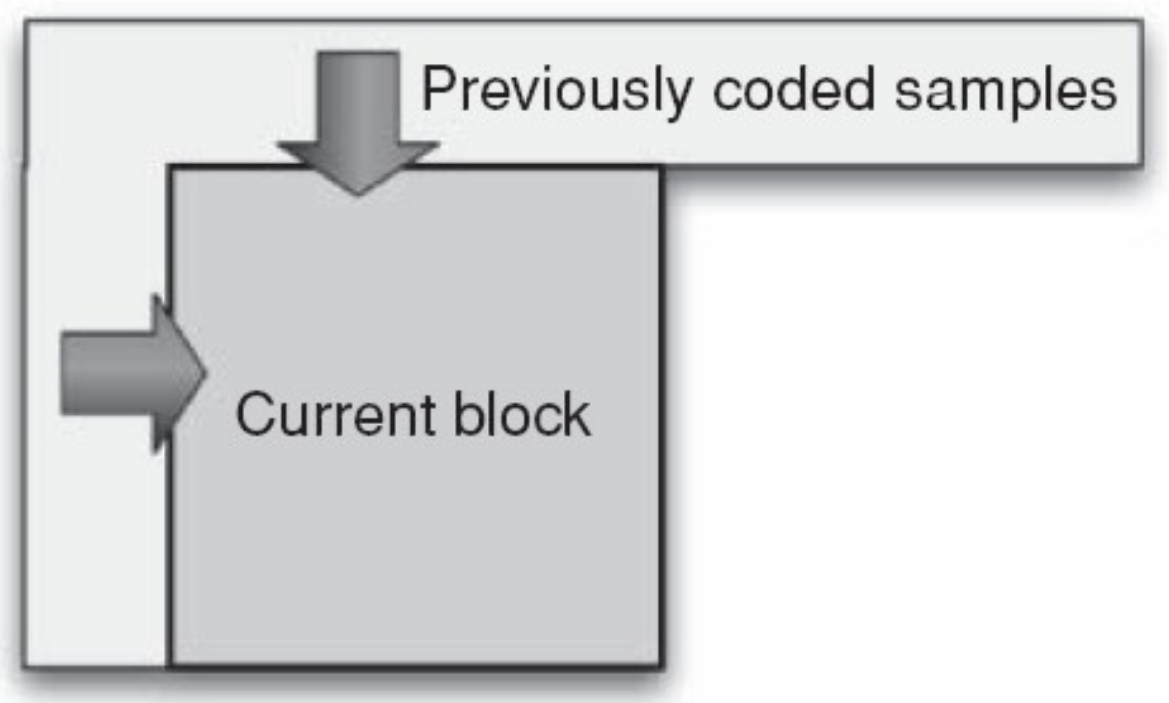
Macroblock partitions



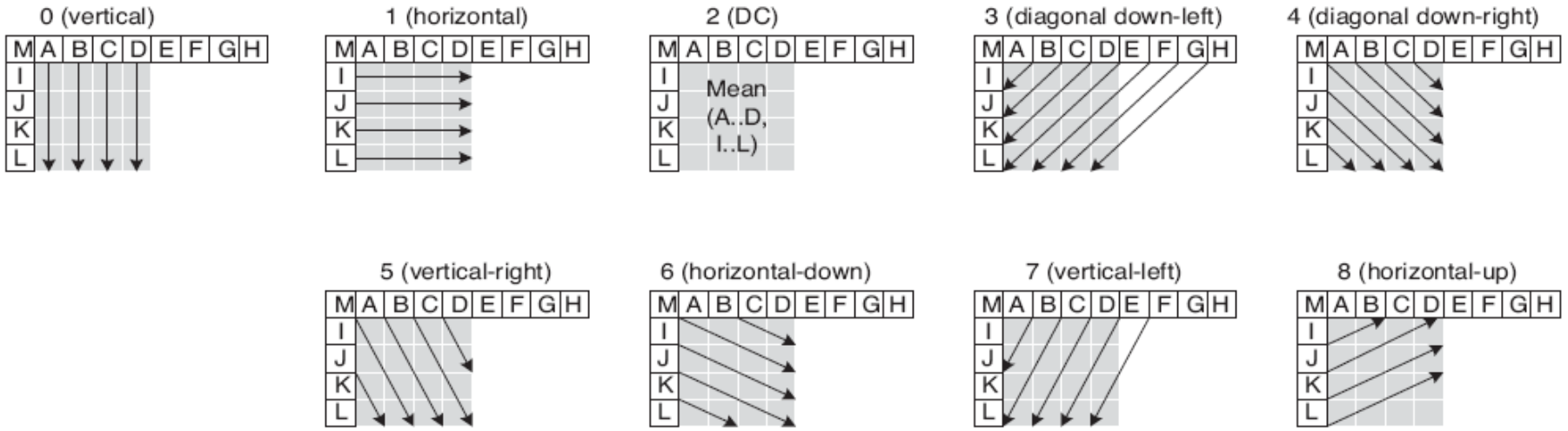
Prediction in H.264



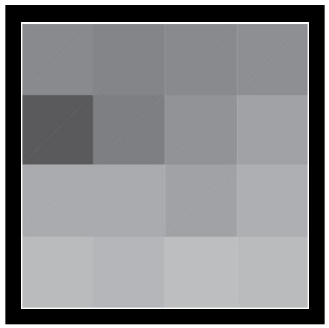
Intra prediction

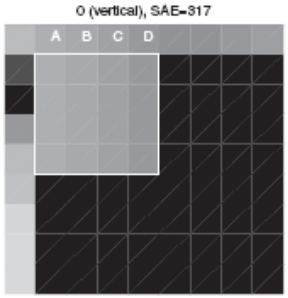


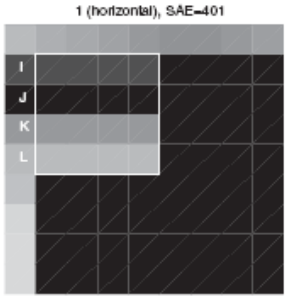
4x4 Intra (luma) prediction modes

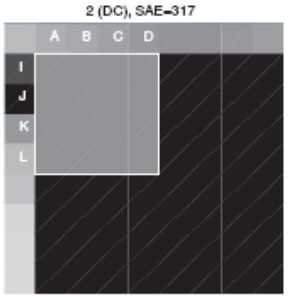


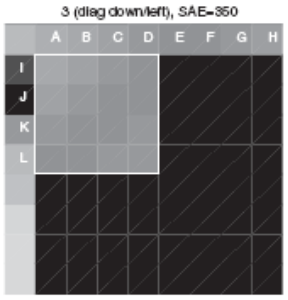
Example of prediction

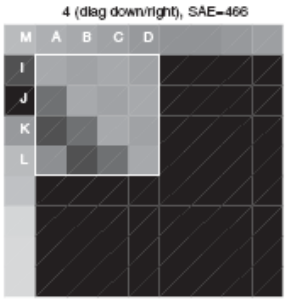


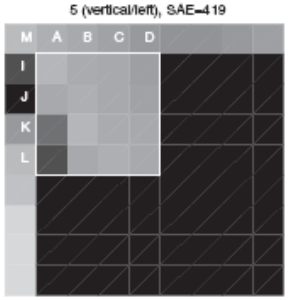


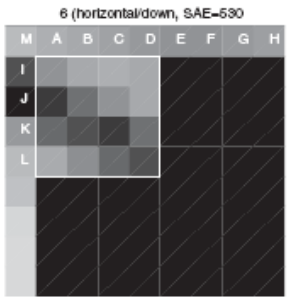


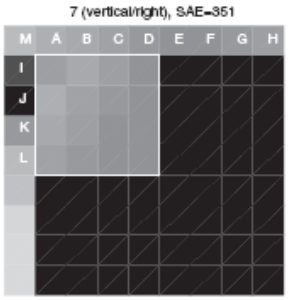


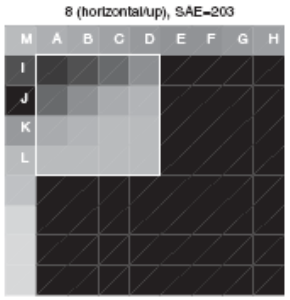




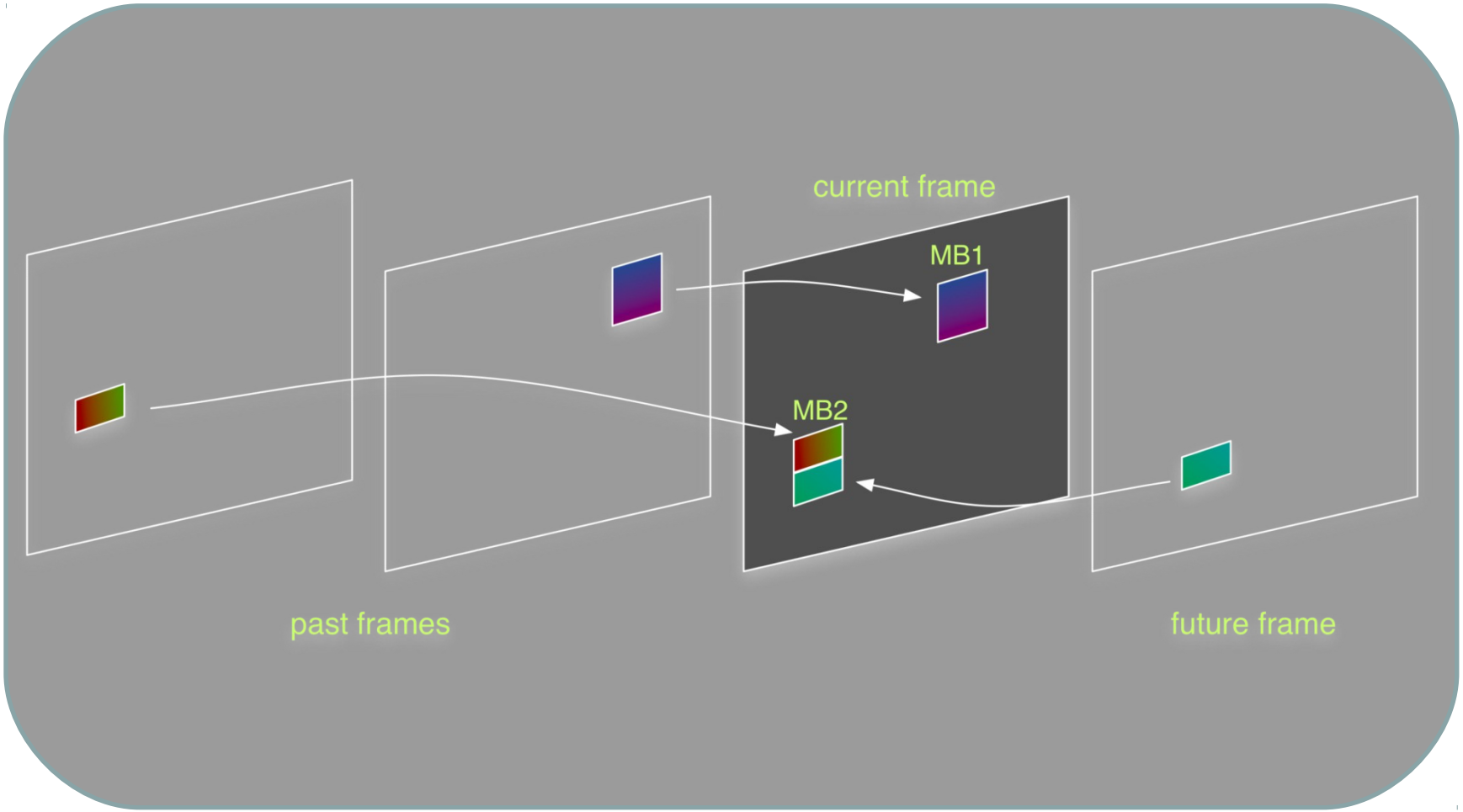








Inter prediction



Motion compensation

For each block of MxN samples in the current frame:

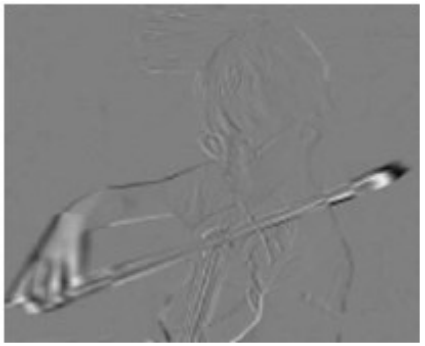
- Search a similar MxN-sample region **(motion estimation)**
- Subtract from the current block to form a residual MxN block (**motion compensation**)
- Offset between the block and the reference is a **motion vector**

Motion types:

- rigid object motion (a moving car)
- deformable object motion (a person speaking)
- camera motion (panning, tilt, zoom and rotation)

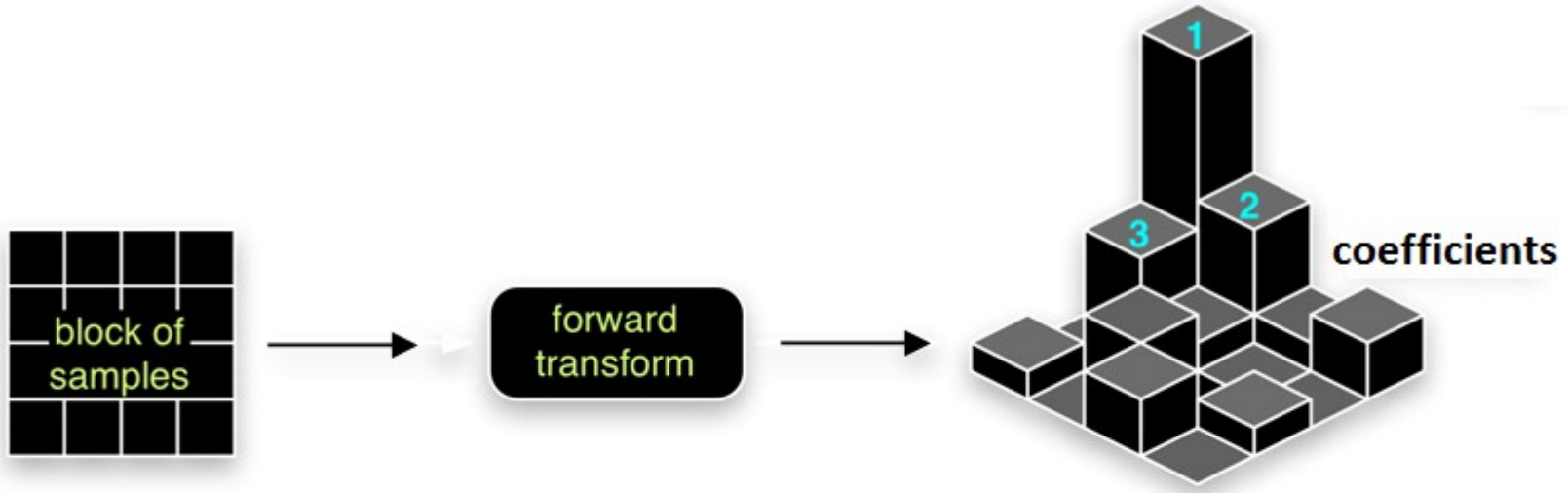
Frame 1

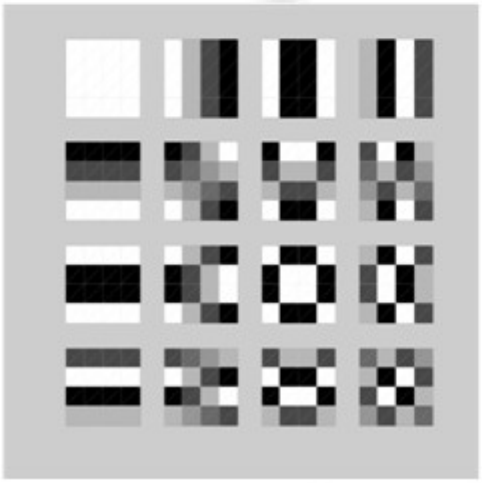
Frame 2



Difference

Forward transform

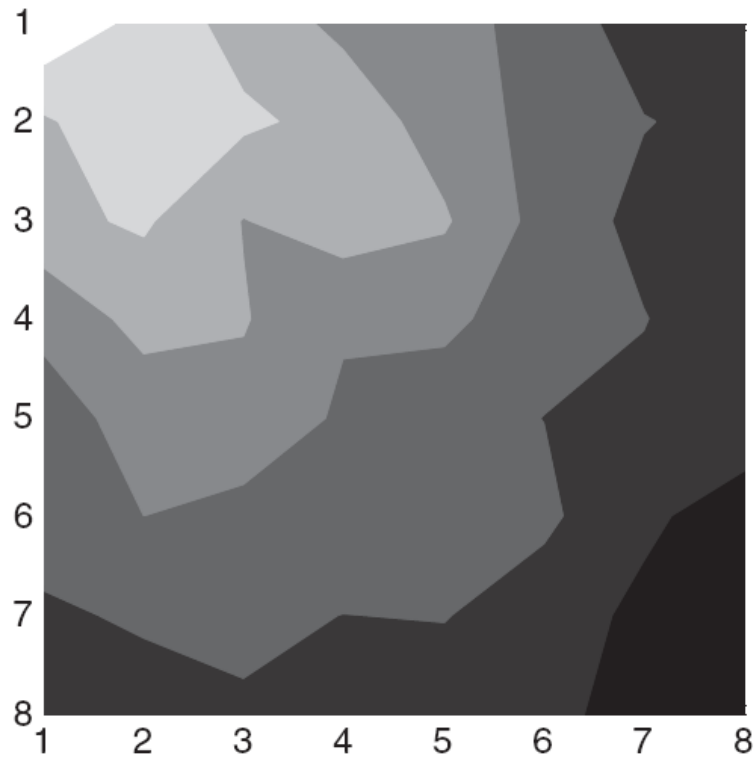




weighting and combining basis patterns would re-create the original block

DCT coefficients distribution

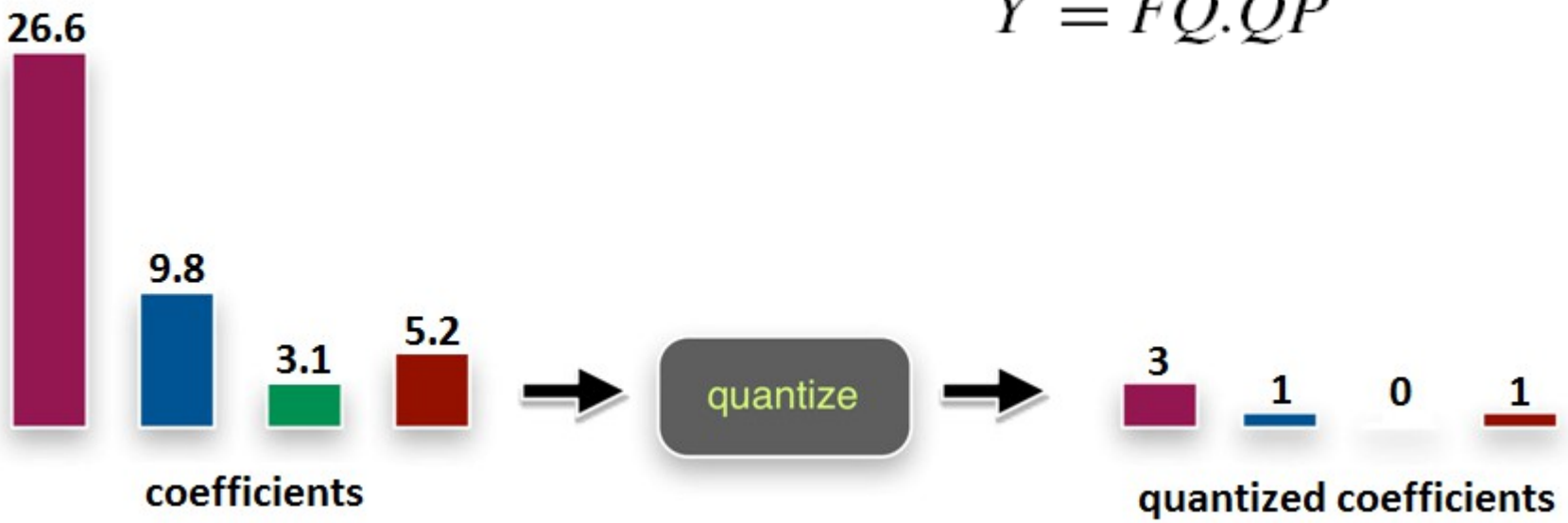
Significant DCT coefficients are typically the **‘low frequency’** positions around the DC coefficient.



Quantization

$$FQ = round\left(\frac{X}{QP}\right)$$

$$Y = FQ.QP$$



Transform + Quantize

58	64	51	58
52	64	56	66
62	63	61	64
59	51	63	69

Block of samples

961	-41	15	-48
-34	72	-30	-104
-15	3	15	24
13	81	-5	8

Integer transform

48	-1	0	-1
-1	1	-1	-2
0	0	0	0
0	1	0	0

Quantized, QP=18

Inverse Transform

58	64	51	58
52	64	56	66
62	63	61	64
59	51	63	69

Block of samples

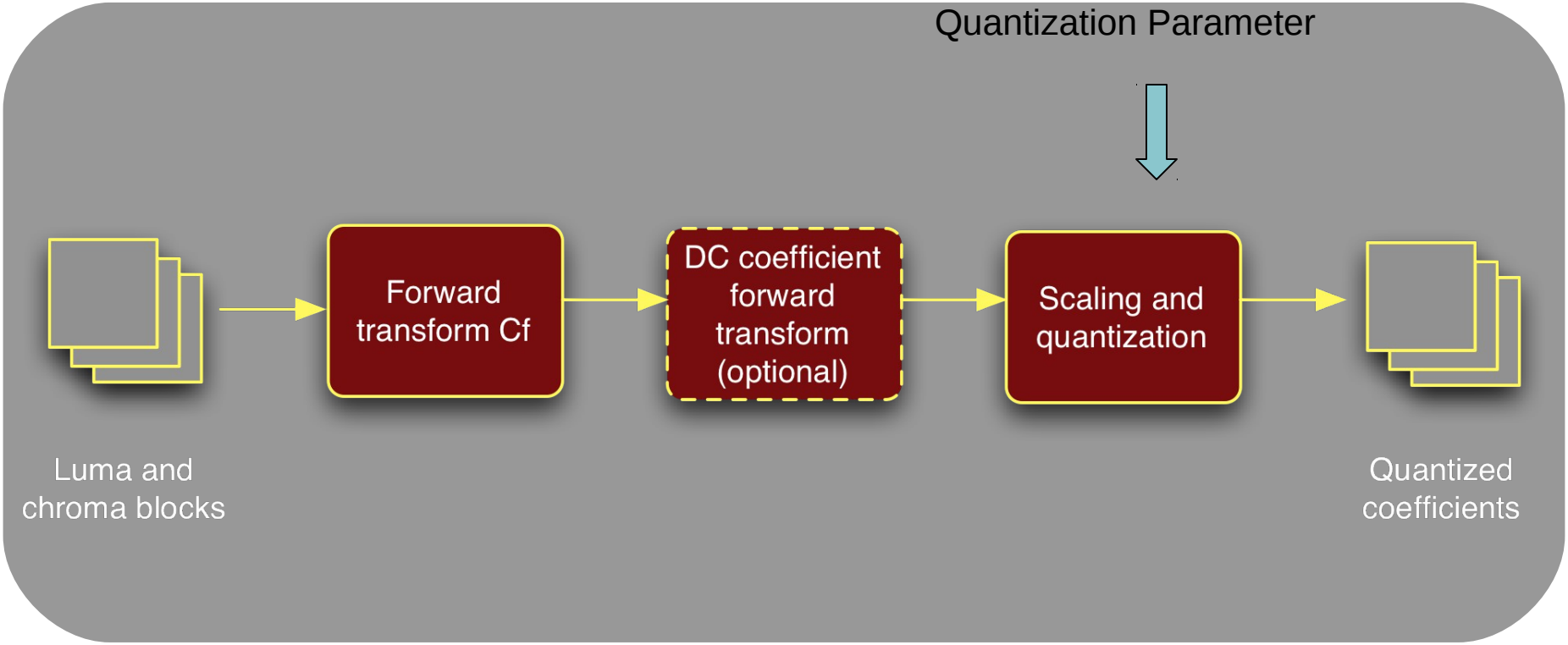
48	-1	0	-1
-1	1	-1	-2
0	0	0	0
0	1	0	0

Quantized, QP=18

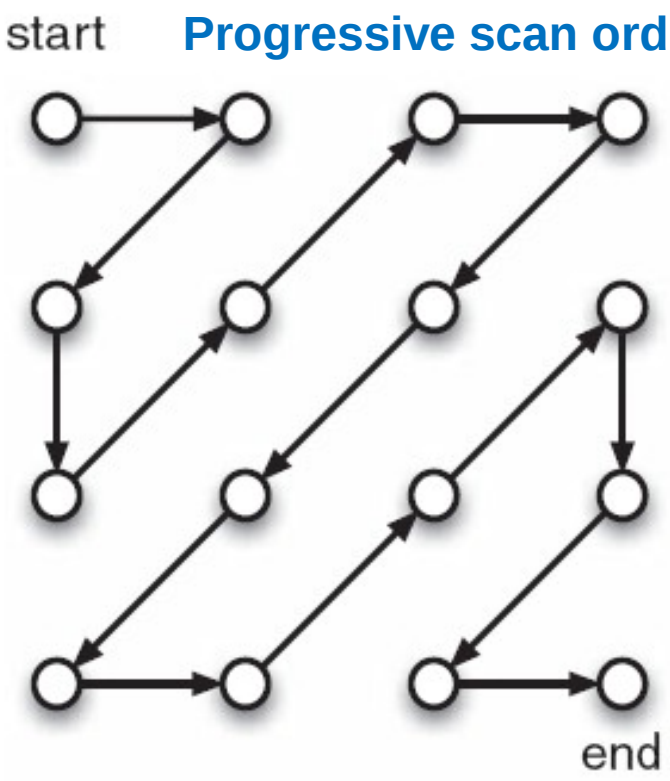
55	66	54	58
54	62	58	63
61	59	61	62
60	55	65	67

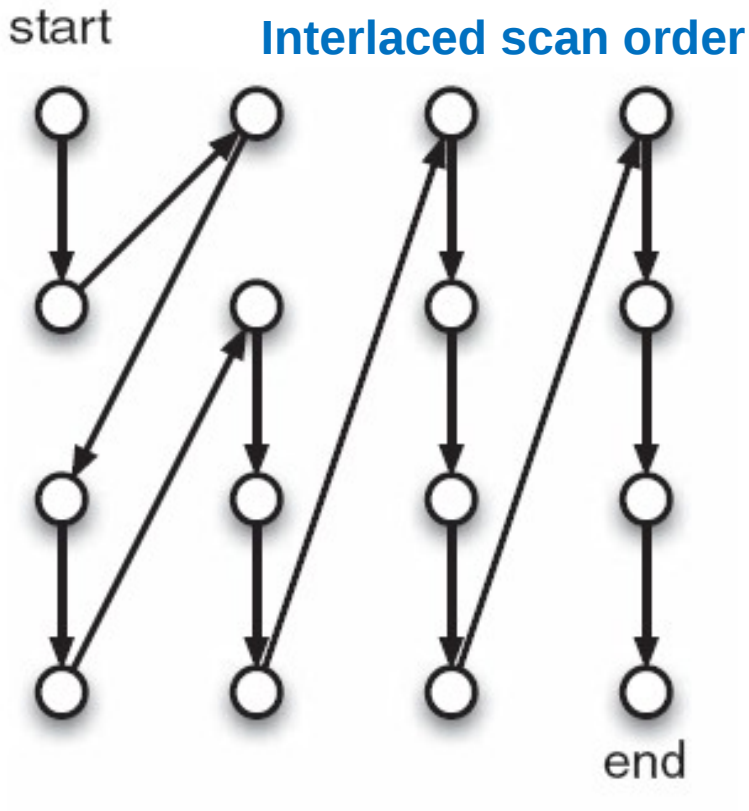
Rescaled + inverse transformed

Full scheme “T+Q”



Block scan





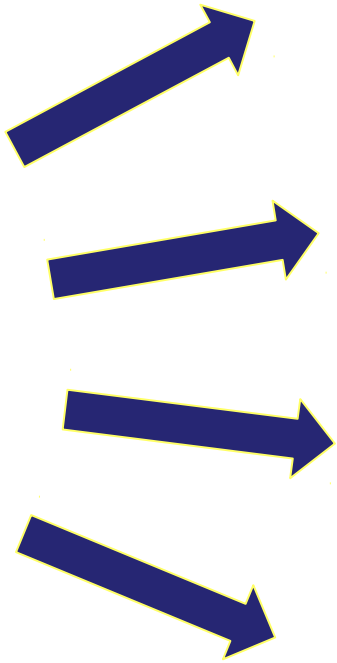
- In a **progressive frame** clustered around ‘DC’ coefficient
- In an **interlaced field** vertical frequencies tend to dominate

Entropy coding

Headers

Prediction parameters

Quantized coefficients



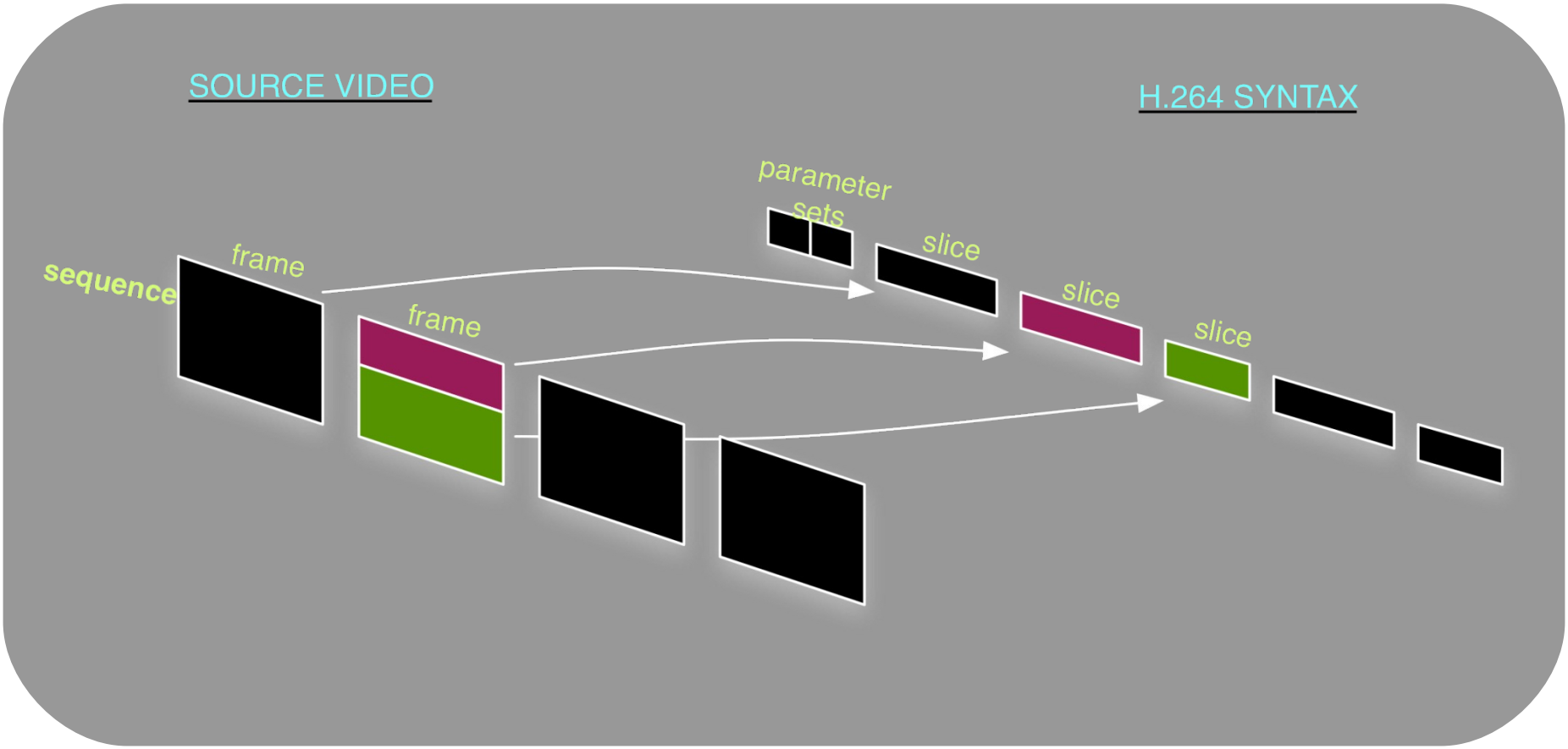
Fixed length codes

Exp-Golomb codes

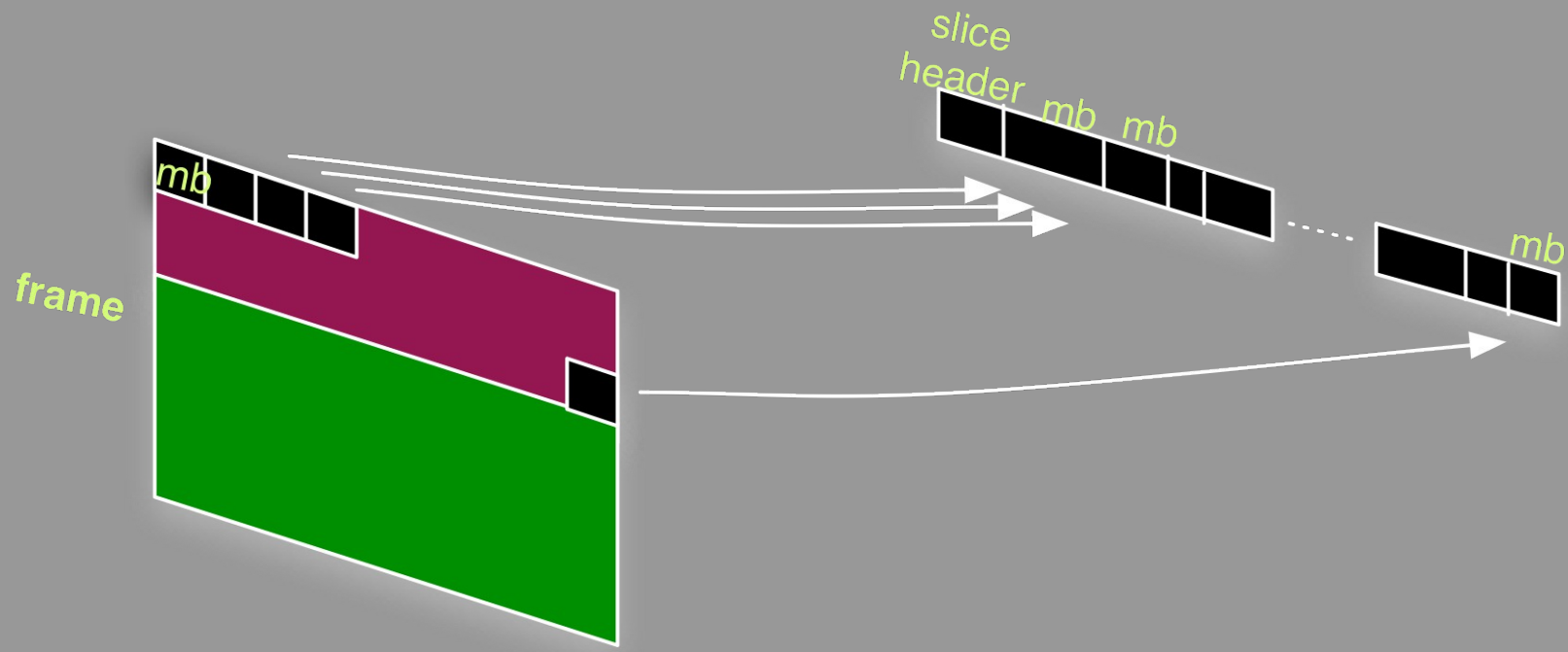
Context Adaptive Variable Length Codes

Context Adaptive Binary Arithmetic Codes

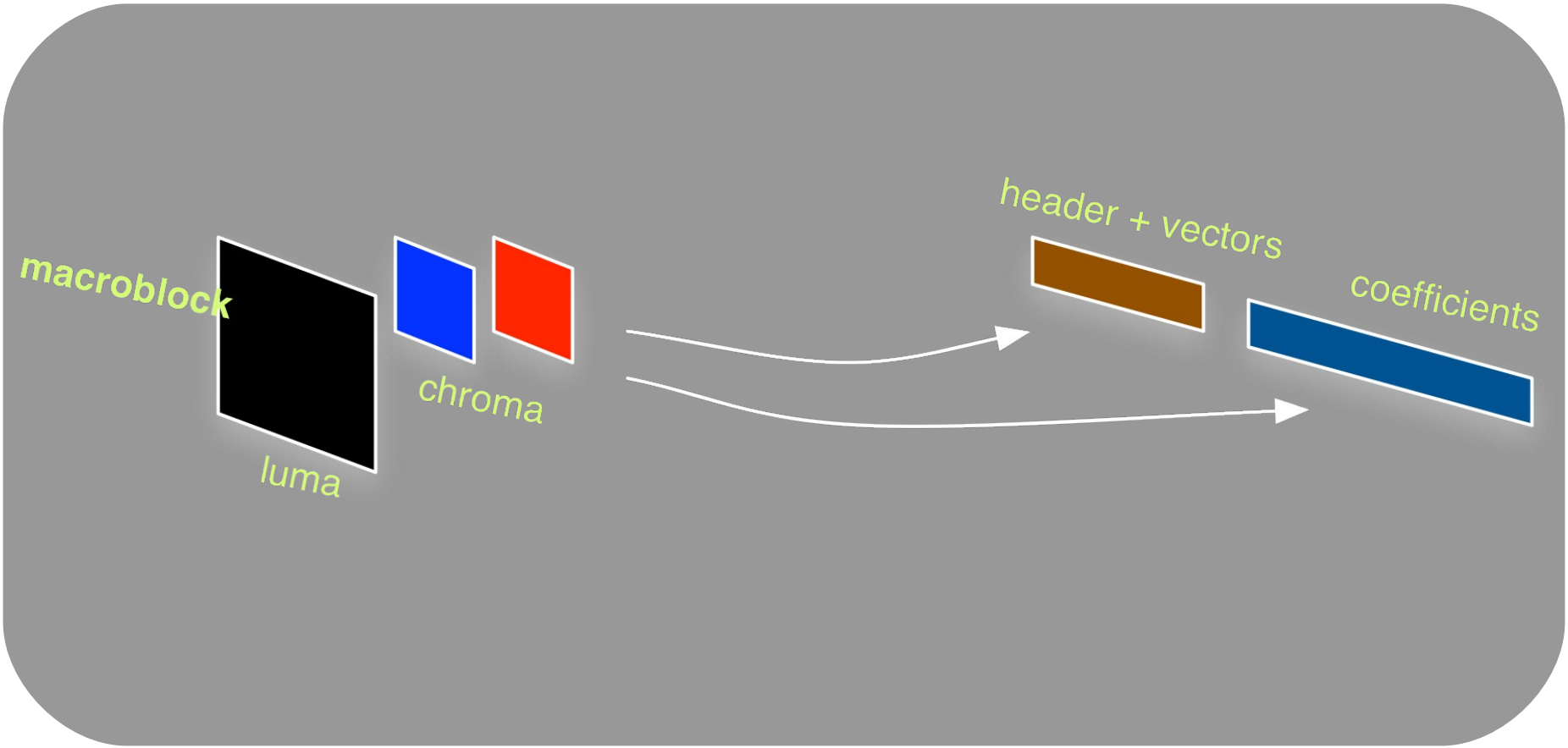
Syntax: NAL Unit Level



Syntax: Slice level



Syntax: MB level

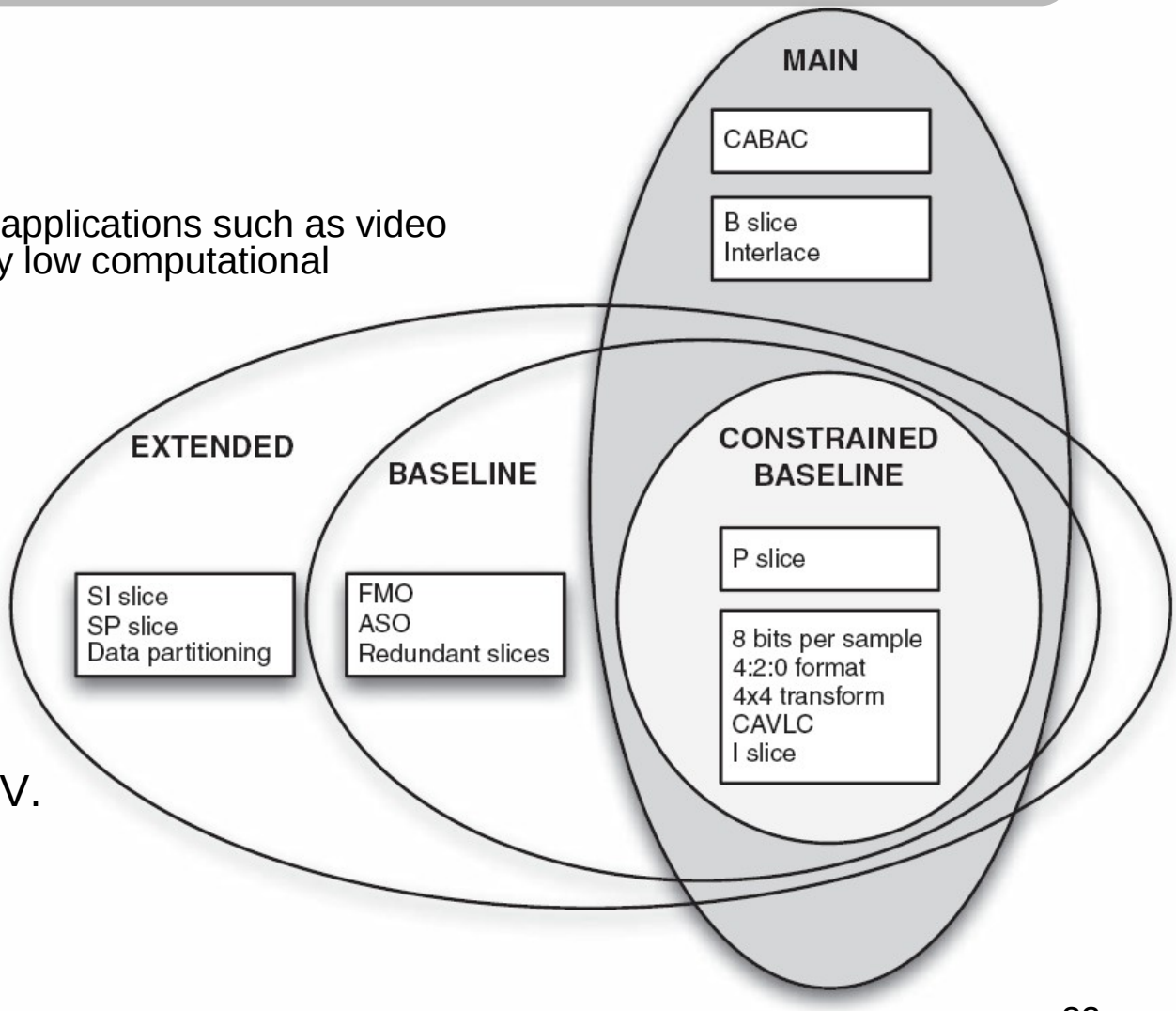


H.264 Profiles

- **Baseline Profile:**
  low-delay, 'conversational' applications such as video conferencing, with relatively low computational requirements.

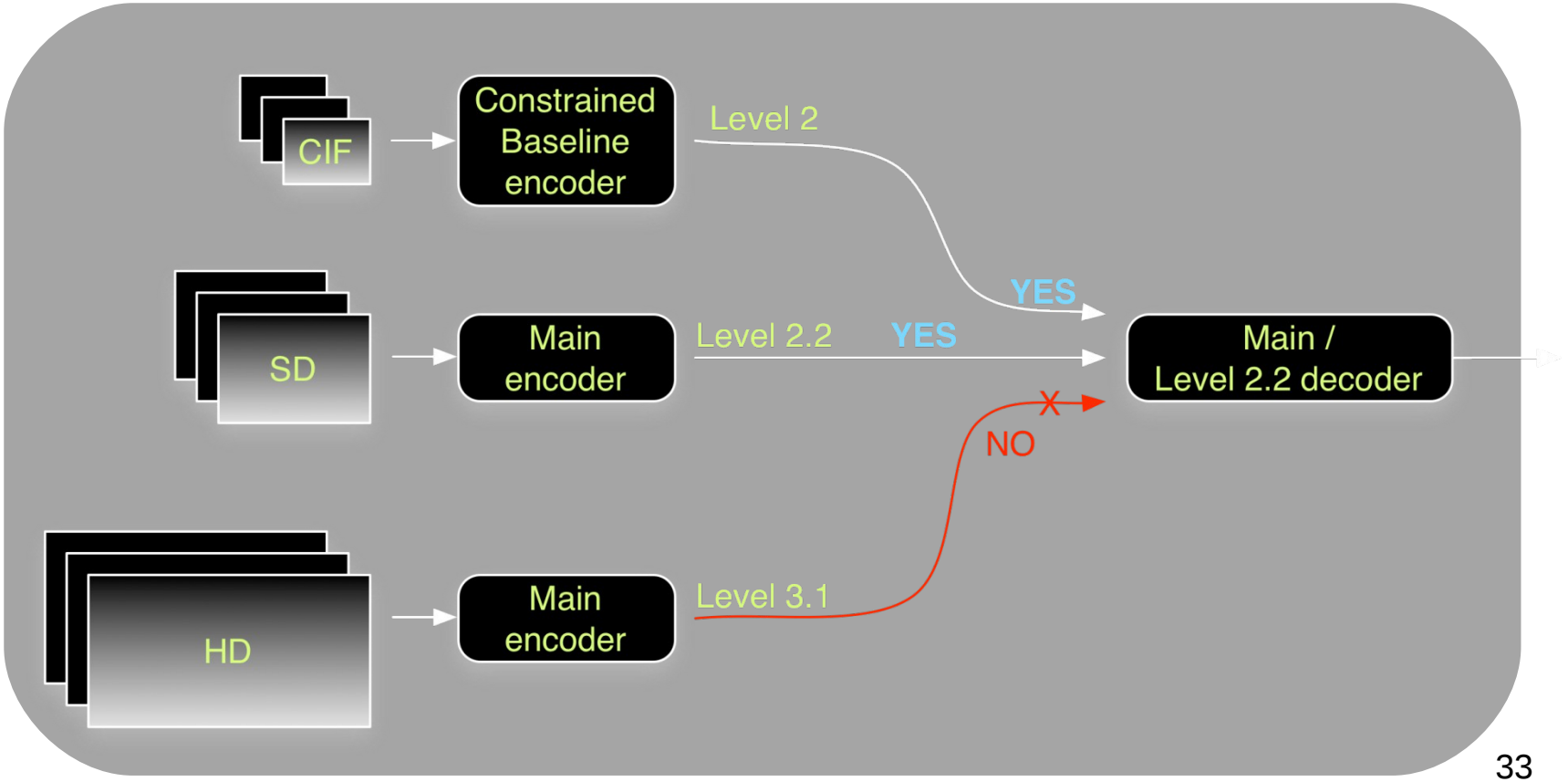
- **Main Profile:**
  basic television, entertainment app

- **High Profiles**:
  improve compression efficiency e.g. for HDTV.



H.264 Levels

H.264/AVC level specifies an upper limit on the frame size, processing rate and working memory required to decode a video sequence.



What's important?

Bitrate

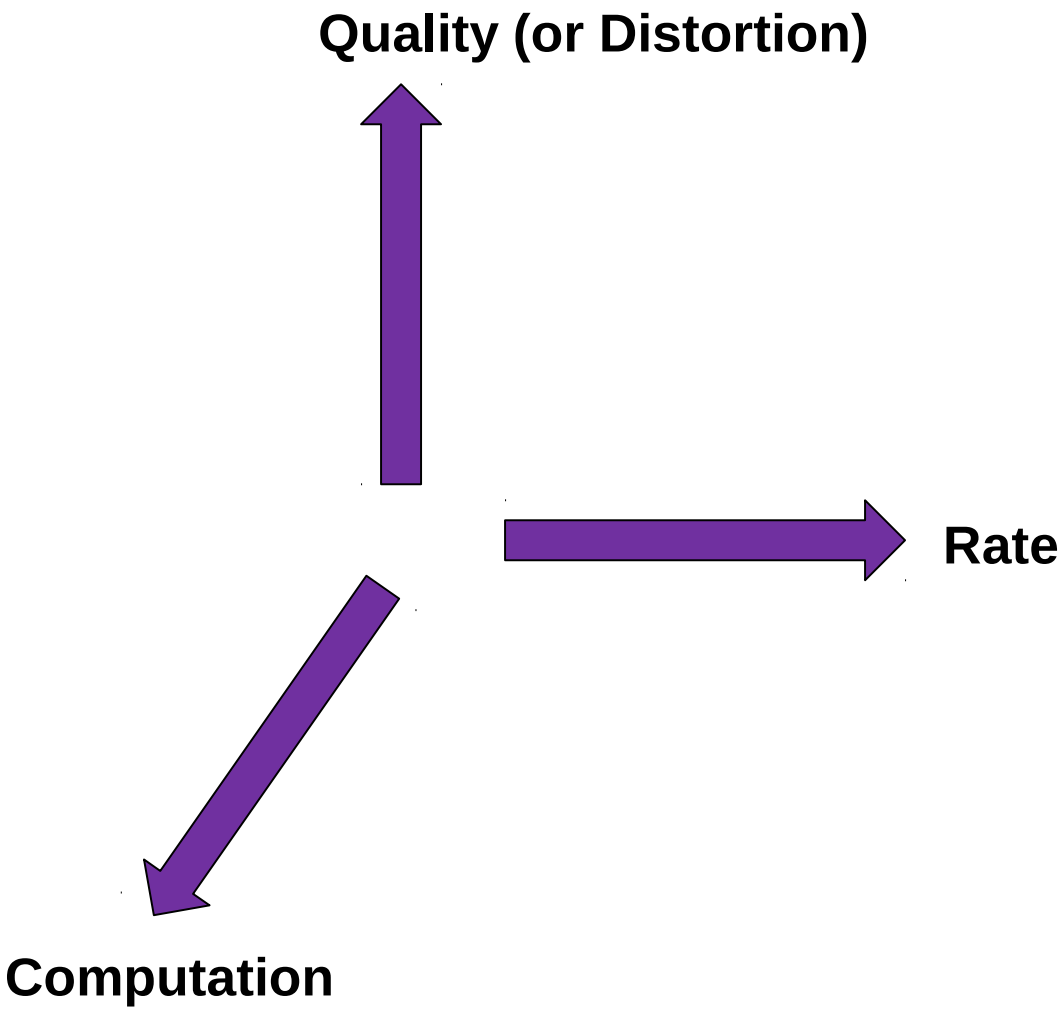
Quality

Delay

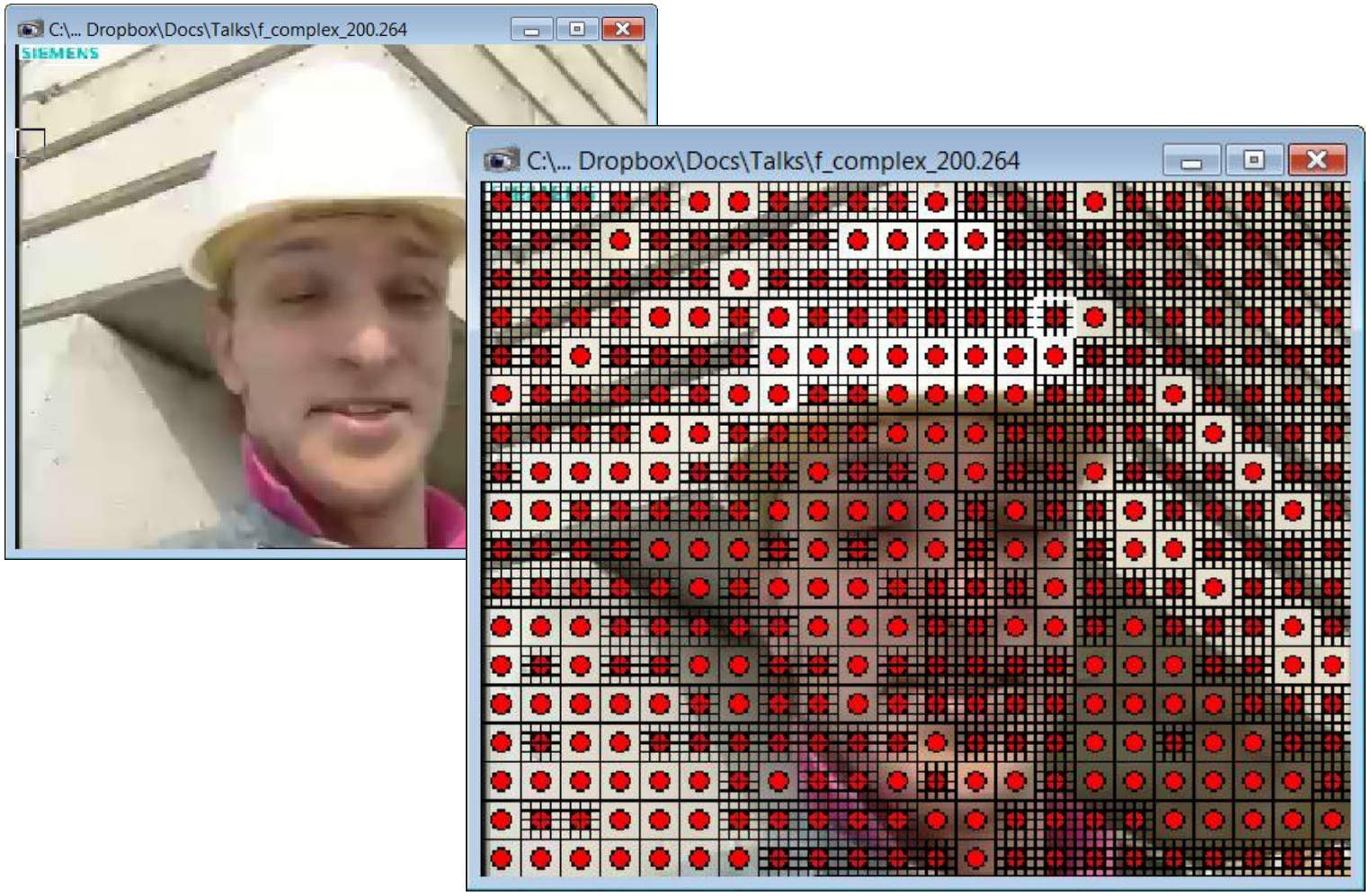
Real-time or offline

Flexibility

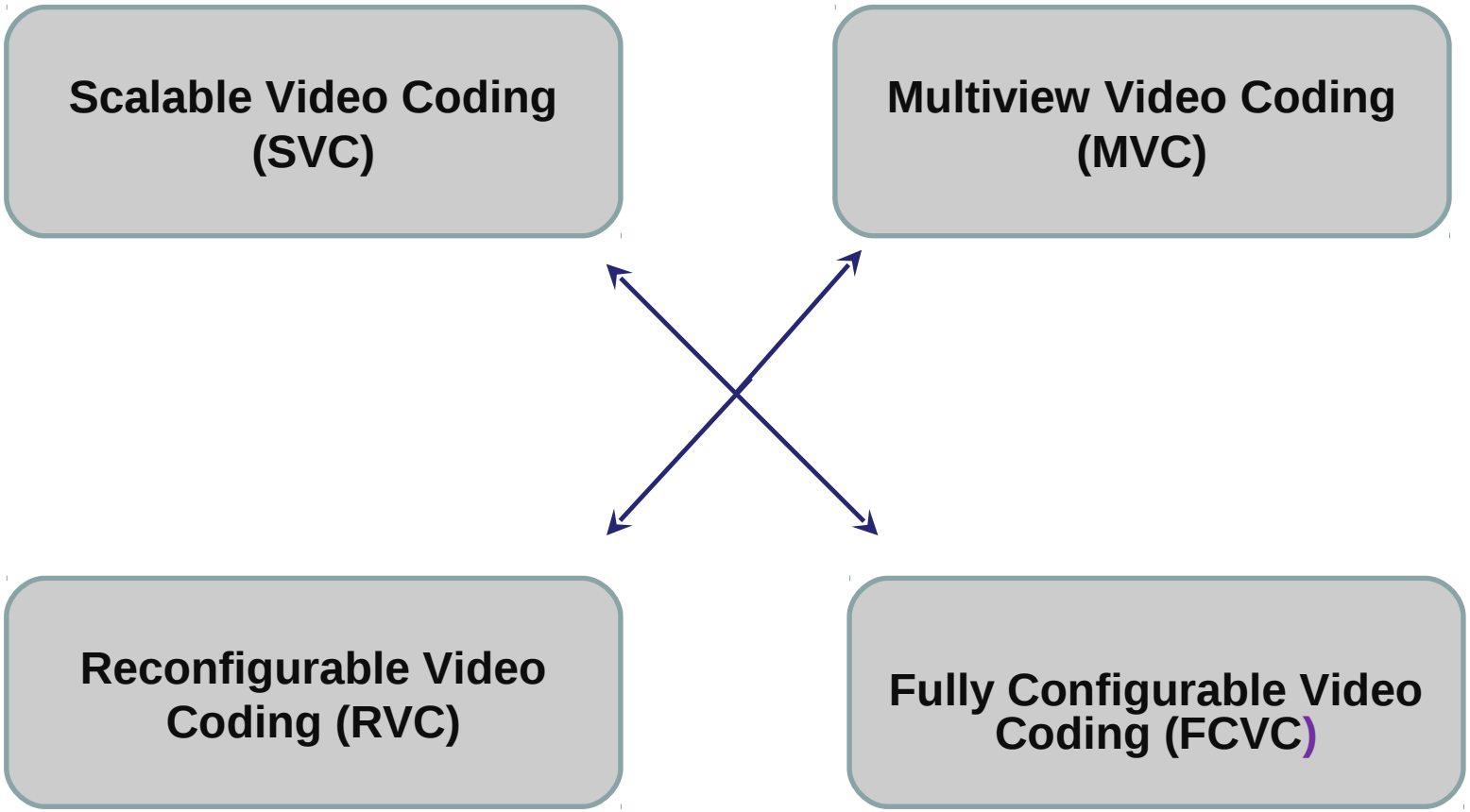
Tradeoff



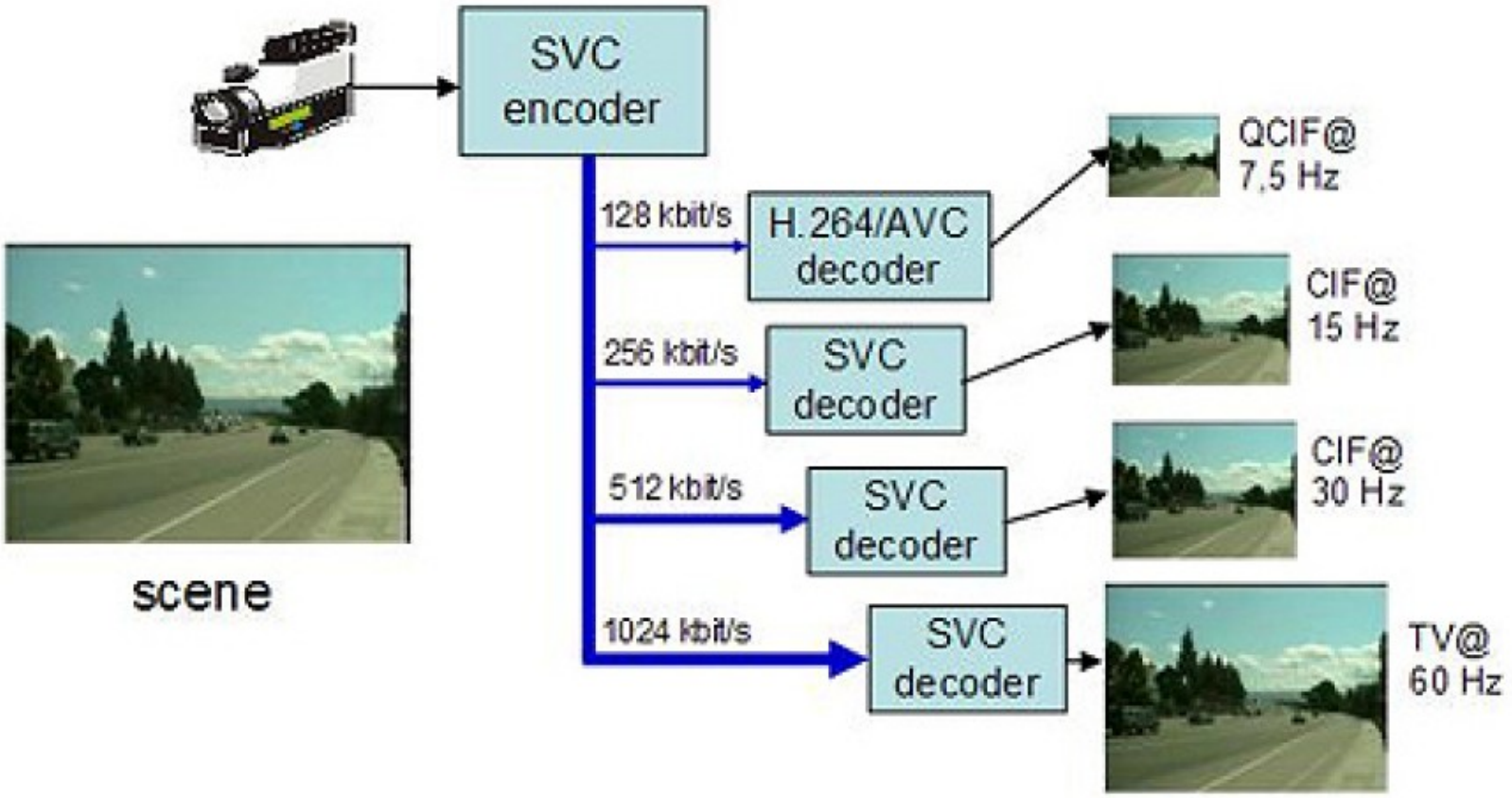
Elecard Stream Eye



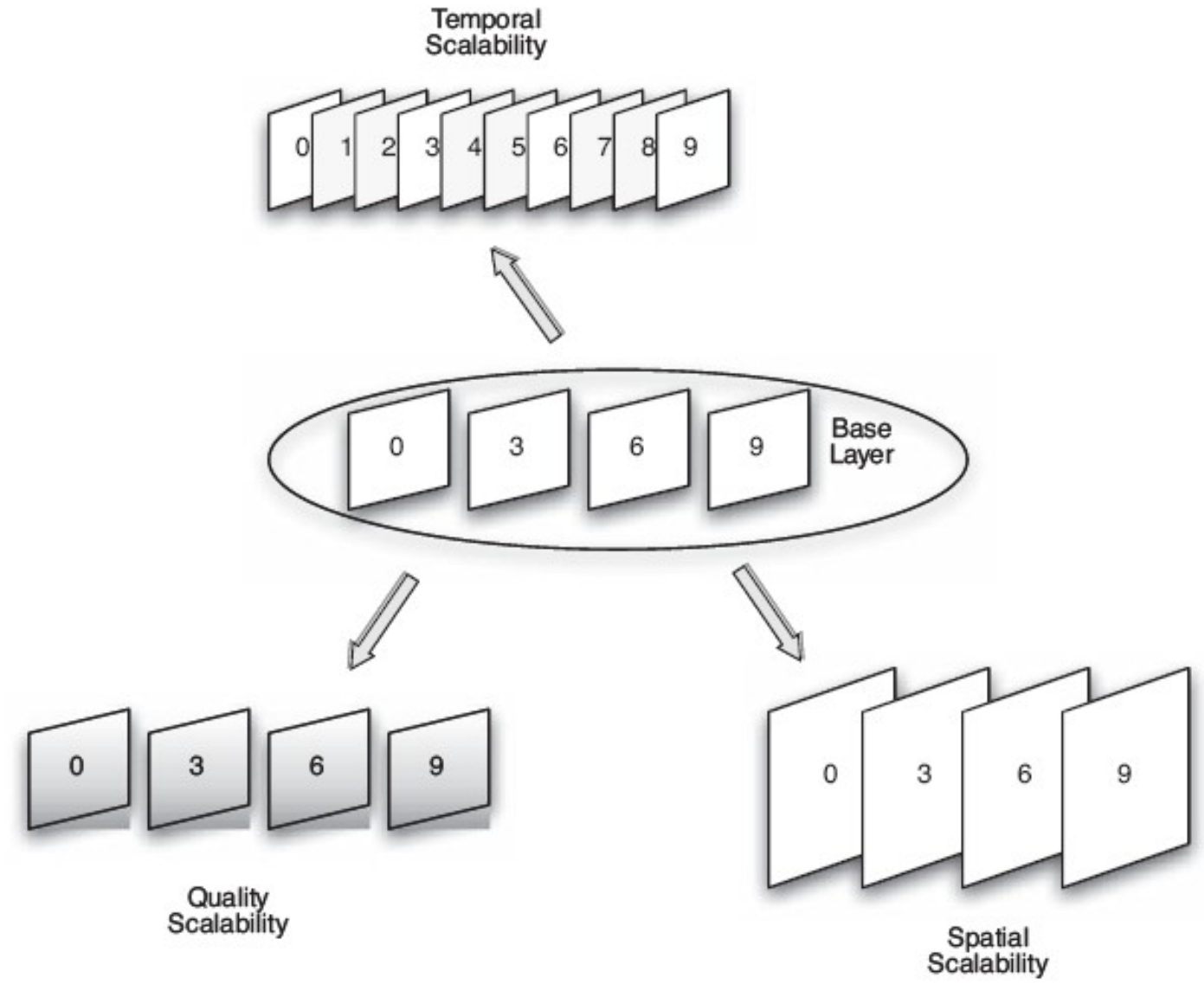
Introduction to extensions



Scalable Video Coding (SVC)

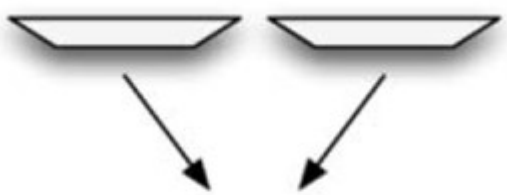


H.264/SVC

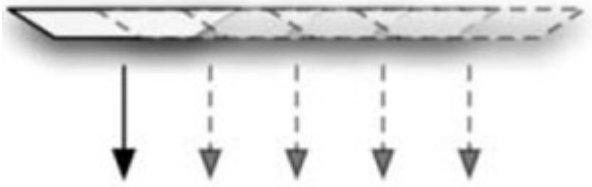

Temporal
Scalability
0 1 2 3 4 5 6 7 8 9
0 3 6 9
Base
Layer
0 3 6 9
Quality
Scalability
0 3 6 9
Spatial
Scalability

Multiview Video Coding (MVC)

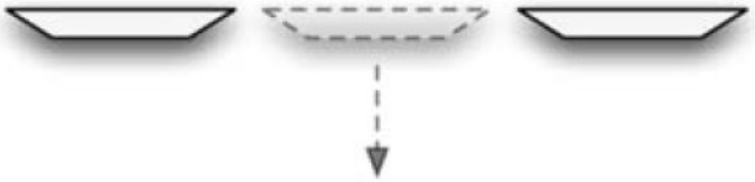
Potential implementations of multiview video include:



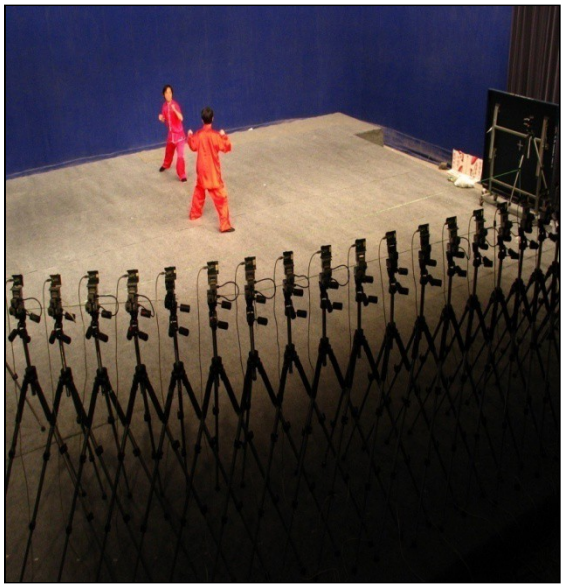
(a) Stereoscopic view



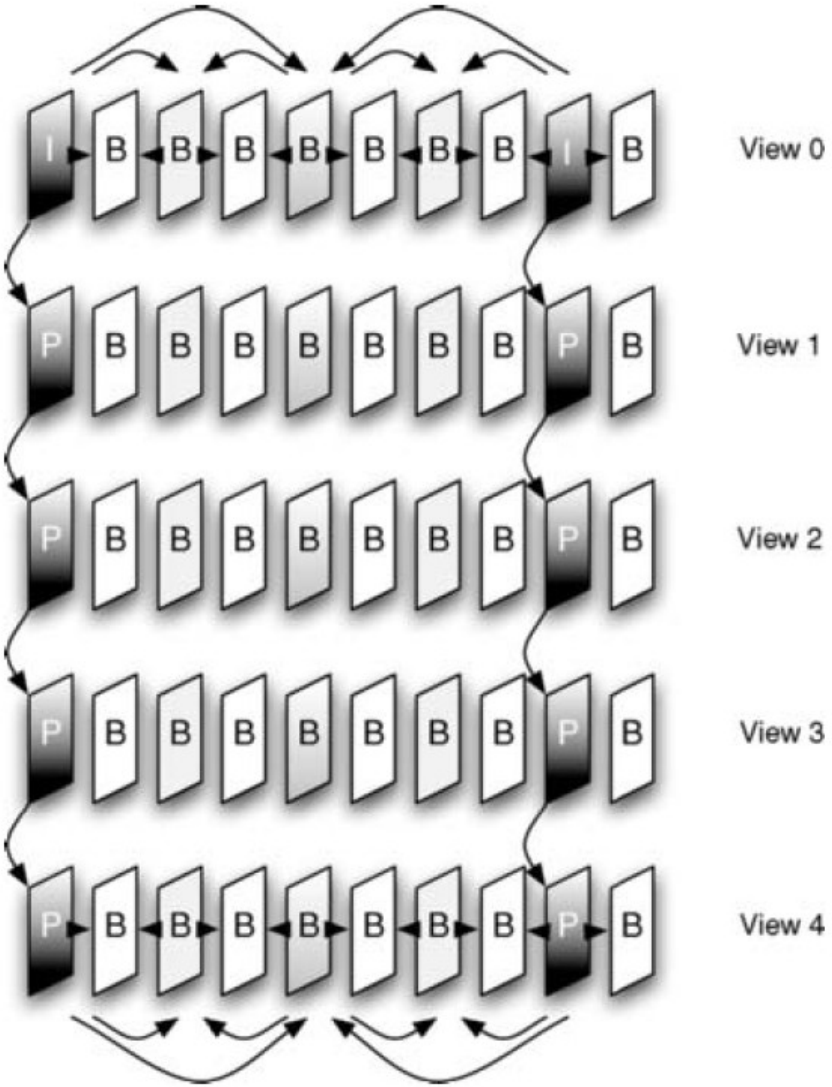
(b) 3D view



(c) Free viewpoint / rendered view



H.264 MVC


I B B B B B B B I B
View 0
P B B B B B B B P B
View 1
P B B B B B B B P B
View 2
P B B B B B B B P B
View 3
P B B B B B B B P B
View 4

Beyond H.264

High Efficiency Video Coding (HEVC/H.265) is a proposed video compression standard, a successor to H.264/AVC.

How will it compare to H.264/AVC?

- Current indications are that the new standard could provide 2x better video compression performance at the expense of significantly higher computational complexity, compared with H.264/AVC.

When will it be finished?

- February 2012: Committee Draft (complete draft of standard)
- January 2013: Final Draft International Standard (ready to be ratified as a Standard)

Useful links

- http://www.itu.int - Download the H.264 standard
- http://www.vcodex.com - Free tutorials
- http://iphome.hhi.de/suehring/tml/ - JM reference software
- http://www.elecard.com/ - StreamEye analyser
- http://tinyurl.com/h264book2 - Very useful book

The End