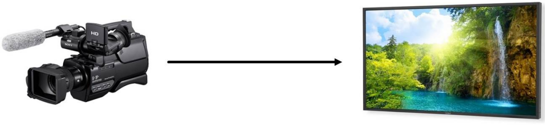
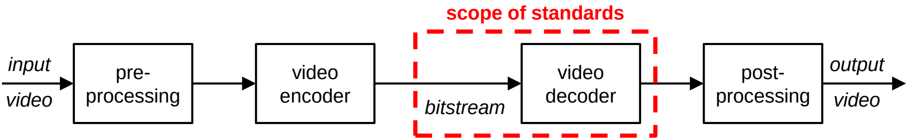
Video Coding Standards



Video Coding Standards



Why standards?

- Bitstream generated by an encoder of one manufacturer should be reliably decodable by decoders of other manufacturers
- Standards define "interface" between encoder and decoder
- Freedom for manufacturers to adapt actual implementations

Scope of video coding standards

1. Bitstream syntax and constraints for transmitted and derived parameters
2. Decoding result for conforming bitstreams (example decoding process)

➡ Encoding, pre- and postprocessing are out of the scope of standards

➡ No guarantee of video quality

Bitstream Syntax

Transmission of coding parameters

- Samples are mapped to coding parameters
  - ➡ Coding modes
  - ➡ Motion parameters
  - ➡ Transform coefficient levels
  - ➡ ...

Bitstream syntax: "Format" for transmitting coding parameters

- Order of transmission
- Conditions for presence
  - ➡ Motion vectors only for inter modes
  - ➡ Intra prediction modes only for intra modes
- Entropy coding
  - ➡ Selected codeword table
  - ➡ Binarization and probability model

Bitstream Syntax — Example

macroblock_layer( ) {	**C**	**Descriptor**
mb_type	2	ue(v) | ae(v)
if(mb_type = = I_PCM) {		
while(!byte_aligned( ))		
pcm_alignment_zero_bit	3	f(1)
for(i = 0; i < 256; i++)		
pcm_sample_luma[i **]**	3	u(v)
for(i = 0; i < 2 * MbWidthC * MbHeightC; i++)		
pcm_sample_chroma[i **]**	3	u(v)
} else {		
…		
}		
}		

Interoperability & Conformance

Interoperability is typically achieved by defining

- **Bitstream syntax**: Format for transmitting coding parameters, including constraints for transmitted and derived paramters
- **Decoding result**: Example decoding process for error-free bitstreams
- All other aspects can be adapted to application requirements
  - Preprocessing and postprocessing
  - Encoding
  - Error concealment

Conformance

- Conforming bitstream: Specified format, obeys all specified constraints
- Conforming decoder: Same result as example decoding process
- Conforming encoder: Generates conforming bitstreams

Profiles & Levels

Conformance points

- Standards are designed for broad range of potential application areas
- Some features may not be appropriate for all application scenarios (would unnecessarily increase complexity of decoder implementations)
- Define difference conformance points using **profiles** & **levels**

➡ Interoperability between applications with similar requirements

Profiles

- Set of coding features that have to be supported in conforming decoders

Levels

- Restricts values of certain key parameters

➡ Picture resolution, frame rate, bit rate, size of decoded picture buffer

Comparison of video coding standards

➡ Consider profiles that provide best coding efficiency for natural video in 4:2:0 chroma sampling format with a bit depth of 8 bit per sample

H.262 | MPEG-2 Video (Main Profile)

General

- I, P, and B pictures / conventional IPPP and IBBP coding
- Picture partitioning: 16×16 macroblocks
- Entropy coding: VLC code tables

Intra-picture prediction

- Prediction in transform domain for 8×8 transform blocks
- Predict DC coefficient from preceding intra block

Motion-compensated prediction

- Blocks of 16×16 luma samples
- Motion vectors with half-sample accuracy & bi-linear interpolation
- Motion vector prediction from left neighboring block

Transform coding

- 8×8 DCT, scalar quantization, zig-zag scan
- Run-level coding of transform coefficient levels

H.263 (High Latency Profile)

General

- I, P, and B pictures / conventional IPPP and IBBP coding
- Partitioning: 16×16 macroblocks; Entropy coding: VLC code tables
- **In-loop deblocking filter**

Intra-picture prediction

- Prediction in transform domain for 8×8 transform blocks
- DC, **horizontal**, or **vertical** prediction with **modified scanning**

Motion-compensated prediction

- Block sizes: 16×16 luma samples, **8×8 luma samples**
- Motion vectors with half-sample accuracy & bi-linear interpolation
- **Median prediction** of motion vectors & **direct mode**
- **Multiple reference pictures**

Transform coding

- 8×8 DCT, scalar quantization, zig-zag scan
- **Run-level-last coding** of transform coefficient levels

MPEG-4 Visual (Advanced Simple Profile)

General

- I, P, and B pictures / conventional IPPP and IBBP coding
- Picture partitioning: 16×16 macroblocks
- Entropy coding: VLC code tables

Intra-picture prediction

- Prediction in transform domain for 8×8 transform blocks
- DC, horizontal, or vertical prediction with modified scanning

Motion-compensated prediction

- Block sizes: 16×16 luma samples, 8×8 luma samples
- **Quarter-sample** motion vector precision, **8-tap + bi-linear interpolation**
- Median prediction of motion vectors & direct mode
- No multiple reference pictures, but global motion compensation (rarely used)

Transform coding

- 8×8 DCT, scalar quantization, zig-zag scan
- Run-level-last coding of transform coefficient levels

H.264 | MPEG-4 AVC (High Profile)

General

- I, P, and B slices / **arbitrary coding structures** / 16×16 macroblocks
- Entropy coding: VLC or **adaptive binary arithmetic coding (CABAC)**
- In-loop deblocking filter

Intra-picture prediction

- **Spatial intra prediction** with block sizes from 4×4 to 16×16
- **9 spatial intra prediction modes** (4 modes for 16×16 blocks)

Motion-compensated prediction

- Variable block sizes: **16×16 to 4×4**, including **non-square blocks**
- Quarter-sample motion vector precision, **6-tap** + bi-linear interpolation
- Median prediction of motion vectors & direct mode
- Multiple reference pictures

Transform coding

- **Integer transform** with **variable block sizes**: 4×4, 8×8, cascaded 16×16
- Improved coding of transform coefficient levels (**CAVLC** or **CABAC**)

H.265 | MPEG-H HEVC (Main Profile)

General

- I, P, and B slices / arbitrary coding structures
- Picture partitioning: **Variable-sized coding units** (64×64 to 8×8)
- CABAC / in-loop deblocking filter & **sample adaptive offset filter**

Intra-picture prediction

- Spatial intra prediction with block sizes from 4×4 to **32×32**
- **35 spatial intra prediction modes** (for all block sizes)

Motion-compensated prediction

- Variable block sizes: **64×64** to 4×4, including non-square blocks
- Quarter-sample motion vector precision, **7/8-tap filter** for all positions
- **Switched motion vector prediction** & **merge mode**
- Multiple reference pictures

Transform coding

- Integer transform with variable block sizes: 4×4 to **32×32**
- Improved coding of transform coefficient levels (CABAC)

Comparison of Syntax Features & Coding Tools

Increased number of supported coding options

- Transform sizes
- Partitioning modes for intra-picture prediction
- Partitioning modes for motion-compensated prediction
- Intra prediction modes
- Selectable reference pictures
- Precision of motion vectors
- Selectable motion vector predictors
- Options for inferring motion parameters
- Supported coding structures

Additional improvements

- Longer and better interpolation filters for sub-sample MCP
- More advanced entropy coding techniques
- Suitable in-loop filters

Comparison of Coding Efficiency

Encoder control

- Standard only specify bitstream syntax and decoding process
- During development of standards: Example encoding process (Test Model)
- Encoding technology has been improved over time

➞ Use unified, highly efficient Lagrangian encoder control

➞ Fair comparison of achievable coding efficiency

➞ Ignore: Real-time operation, error robustness, ...

Encoder configuration

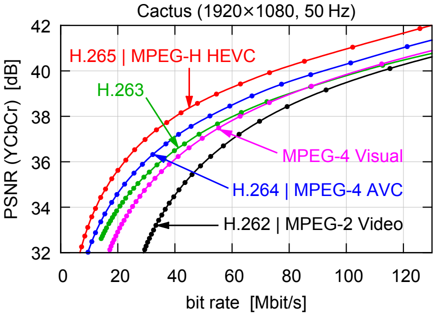
- Enable all coding tools that contribute to coding efficiency
- Choose best known configuration (e.g., coding structure) for each standard

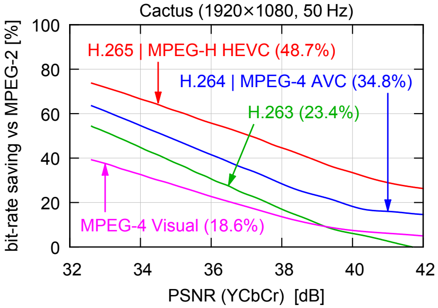
Measuring of coding efficiency

- Bit rate: Average bit rate
- Quality: Average PSNR with weighting of color components

$$\mathrm{PSNR}_{\mathrm{YCbCr}} = \left(6 \cdot \mathrm{PSNR}_{\mathrm{Y}} + \mathrm{PSNR}_{\mathrm{Cb}} + \mathrm{PSNR}_{\mathrm{Cr}} \right) / 8$$

Intra-Only Video Applications





Intra-only coding

- All pictures are coded separately (without inter-picture prediction)
- Rarely used in practice: Professional editing applications
- Shows improvements of intra-picture coding
- Test set: Five entertainment-quality sequences (resolution of 1920 × 1080)

Intra-Only Video Applications

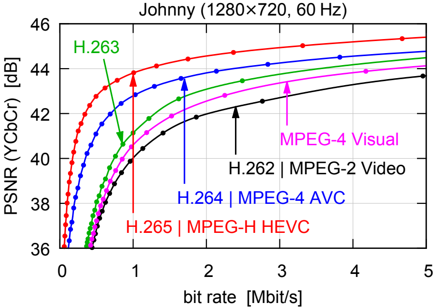
video coding standard	average bit-rate savings relative to ...			
	H.262 | MPEG-2 Video	MPEG-4 Visual	H.263	H.264 | MPEG-4 AVC
MPEG-4 Visual	18.5 %			
H.263	23.1 %	7.0 %		
H.264 | MPEG-4 AVC	34.2 %	20.7 %	15.0 %	
H.265 | MPEG-H HEVC	47.3 %	36.9 %	32.9 %	21.2 %

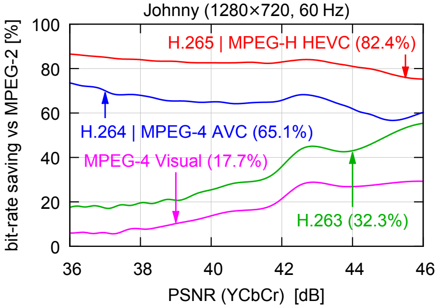
Average bit-rate savings

- Averaged over all sequences in test set

➡ Improvements from one generation of standards to the next

Interactive Video Applications





Low-delay configuration

- Targets interactive applications that require low delay
- Video telephony, video conferencing
- Restriction: All pictures are coded in acquisition/display order
- Test set: Six video conferencing sequences (resolution of 1280 × 720)

Interactive Video Applications

video coding standard	average bit-rate savings relative to ...			
	H.262 | MPEG-2 Video	MPEG-4 Visual	H.263	H.264 | MPEG-4 AVC
MPEG-4 Visual	17.7 %			
H.263	29.3 %	14.3 %		
H.264 | MPEG-4 AVC	61.4 %	52.8 %	44.2 %	
H.265 | MPEG-H HEVC	76.5 %	71.1 %	65.9 %	39.6 %

Average bit-rate savings

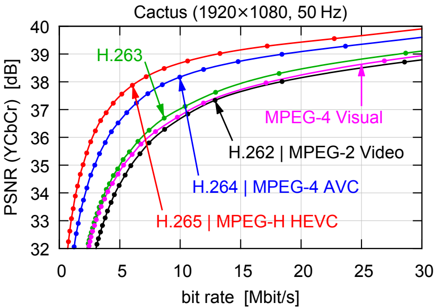
- Averaged over all sequences in test set

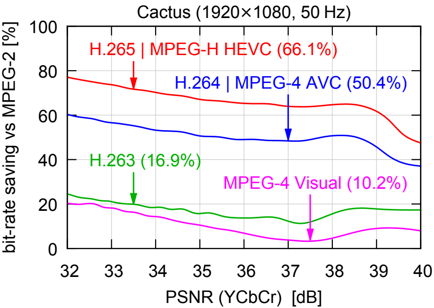
➡ Improvements from one generation of standards to the next

➡ Large improvement from H.263 to H.264 | MPEG-4 AVC

➡ Large improvement from H.264 | MPEG-4 AVC to H.265 | MPEG-H HEVC

Entertainment-Quality Video Applications





Random access configuration

- Targets entertainment applications that require regular random access points
- Television broadcast, Internet or wireless streaming, optical discs
- Restriction: Random access point every second (but higher delay)
- Test set: Five entertainment-quality sequences (resolution of 1920×1080)

Entertainment-Quality Video Applications

video coding standard	average bit-rate savings relative to ...			
	H.262 | MPEG-2 Video	H.263	MPEG-4 Visual	H.264 | MPEG-4 AVC
H.263	6.0 %			
MPEG-4 Visual	9.5 %	2.9 %		
H.264 | MPEG-4 AVC	49.1 %	45.2 %	43.8 %	
H.265 | MPEG-H HEVC	67.4 %	64.8 %	64.1 %	37.1 %

Average bit-rate savings

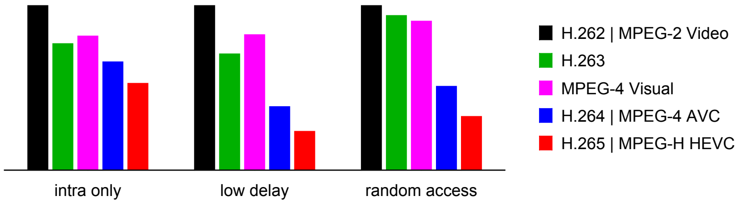
- Averaged over all sequences in test set

➡ Improvements from one generation of standards to the next

➡ Large improvement from H.263 to H.264 | MPEG-4 AVC

➡ Large improvement from H.264 | MPEG-4 AVC to H.265 | MPEG-H HEVC

Part Summary



Video coding standards

- Follow basic approach of block-based hybrid video coding
- Specify bitstream syntax and decoding process
- Main improvement: More coding options for representing blocks of samples

Coding efficiency comparison

- Unified and highly efficient Lagrangian coder control
- Consider different application requirements

→ Continuous improvement of coding efficiency