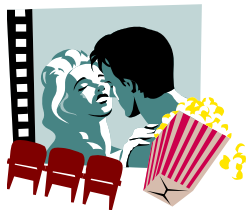
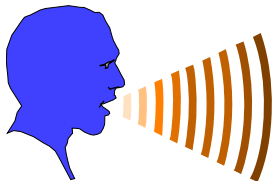





Video Processing & Communications

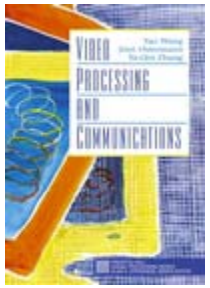
Video Coding Standards – Part II

Yao Wang
Polytechnic University, Brooklyn, NY11201
http://eeweb.poly.edu/~yao

Based on: Y. Wang, J. Ostermann, and Y.-Q. Zhang, Video Processing and Communications, Prentice Hall, 2002.

Outline

- Overview of Standards and Their Applications
- ITU-T Standards for Audio-Visual Communications
  - H.261
  - H.263
  - H.263+, H.263++
- ISO Standards for
  - MPEG-1
  - MPEG-2
  - MPEG-4
- H.264/AVC

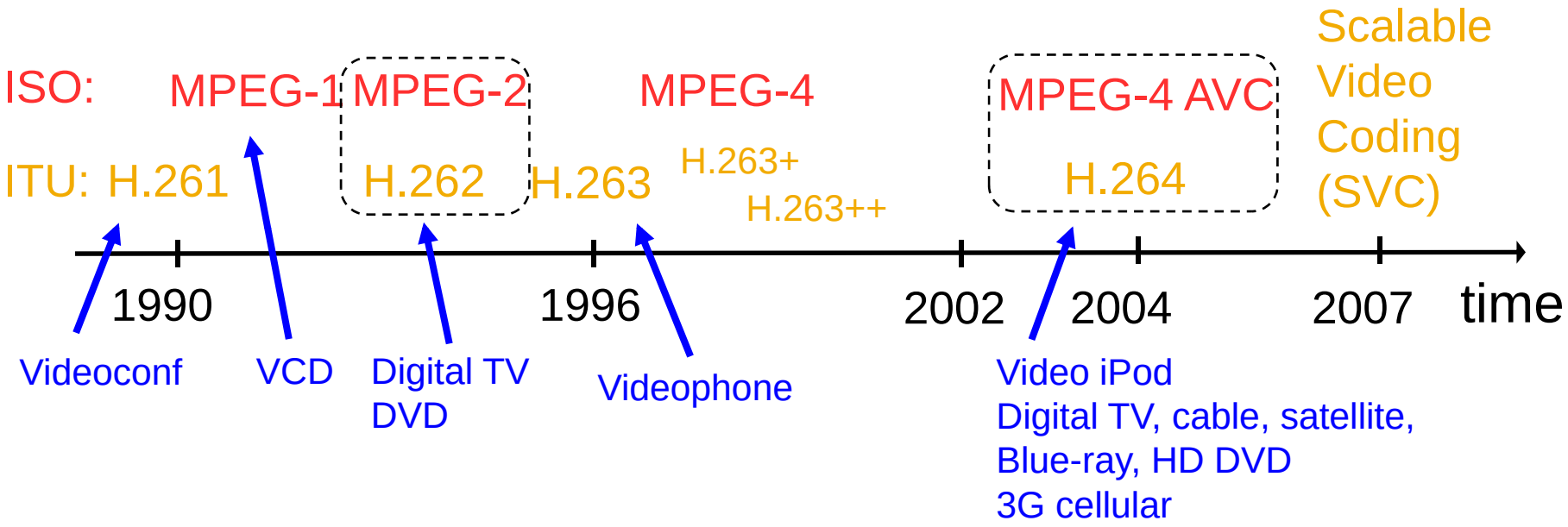


Current Image and Video Compression Standards

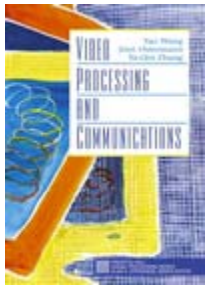
Standard	Application	Bit Rate
JPEG	Continuous-tone still-image compression	Variable
H.261	Video telephony and teleconferencing over ISDN	p x 64 kb/s
MPEG-1	Video on digital storage media (CD-ROM)	1.5 Mb/s
MPEG-2	Digital Television	2-20 Mb/s
H.263	Video telephony over PSTN	33.6-? kb/s
MPEG-4	Object-based coding, synthetic content, interactivity	Variable
JPEG-2000	Improved still image compression	Variable
H.264 / MPEG-4 AVC	Improved video compression	10's kb/s to Mb/s

MPEG and JPEG: International Standards Organization (ISO)
H.26x family: International Telecommunications Union (ITU)

History of Video Coding Standards



- Recent development:
  - HEVC, 2012



H.264/AVC Standards

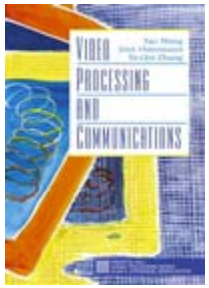
- Developed by the joint video team (JVT) including video coding experts from the ITU-T and the ISO MPEG
- Finalized March 2003
- Improved video coding efficiency, up to 50% over H.263++/MPEG4
  - Half the bit rate for similar quality
  - Significantly better quality for the same bit rate
- Reference & figures for this section are from
  - *Ostermann et al., Video coding with H.264/AVC: Tools, performance, and complexity, IEEE Circuits and Systems Magazine, First Quarter, 2004*

New Video Coding Tools

- Intra-prediction
- Integer DCT with variable block sizes
- Adaptive deblocking filtering
- Multiple reference frame prediction

Spatial prediction

- H.261
  - Motion vector prediction using previously encoded MV
- MPEG-1
  - DC coefficients coded predictively
- H.263
  - MV prediction using the median of three neighbors
  - Optional: Intra DC prediction (10-15% improvement)
- MPEG-4
  - DC prediction: can predict DC coefficient from *either* the previous block or the block above
  - AC prediction: can predict one column/row of AC coefficients from *either* the previous block or the block above
- H.264
  - Pixel domain directional intra prediction



H.264 Intra prediction

•Instead of the simple DC coefficient prediction to exploit the correlation between nearby pixels in the same frame, more sophisticated spatial prediction is used

•Apply prediction to the entire 16*16 block (INTRA_16x16), or apply prediction separately to sixteen 4*4 blocks (INTRA_4x4)

•Adaptive directional prediction

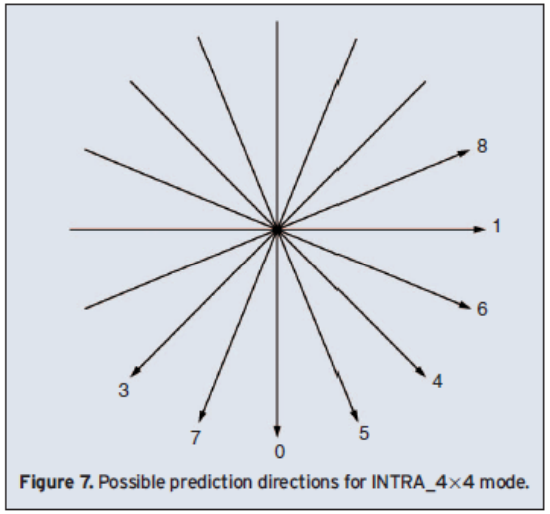


Figure 7. Possible prediction directions for INTRA_4×4 mode.

8 possible directions

Sample Intra Prediction Modes

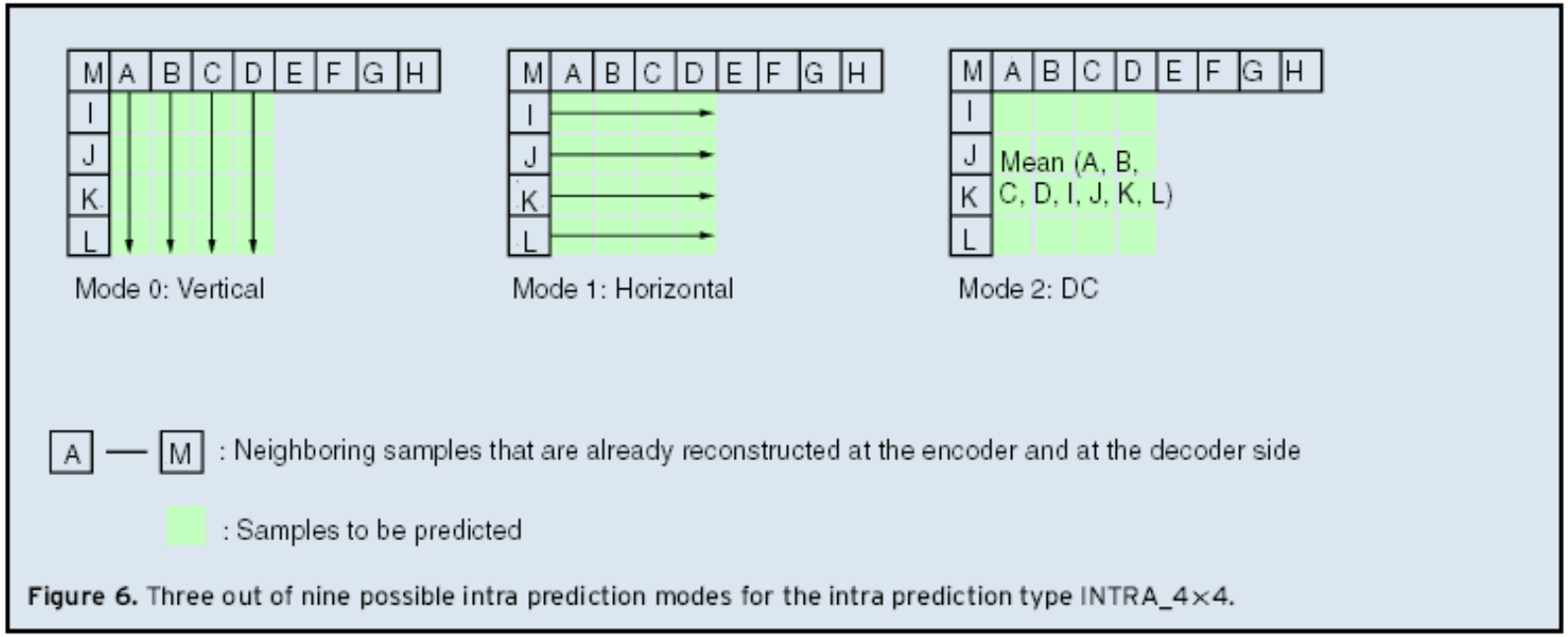


Figure 6. Three out of nine possible intra prediction modes for the intra prediction type INTRA_4×4.

From [Ostermann04]

Motion Compensation

- Quarter-pel accuracy
- Variable block size
- Multiple reference frames
  - Generalized B-picture
- Weighted prediction (fade in, fade out, etc)

Variable Blocksize Motion Compensation

- Use variable size block-based motion compensation
  - 16x16, 16x8, 8x16, 8x8, 8x4, 4x8, 4x4
  - H.263/MPEG4 use only 16x16 and 8x8

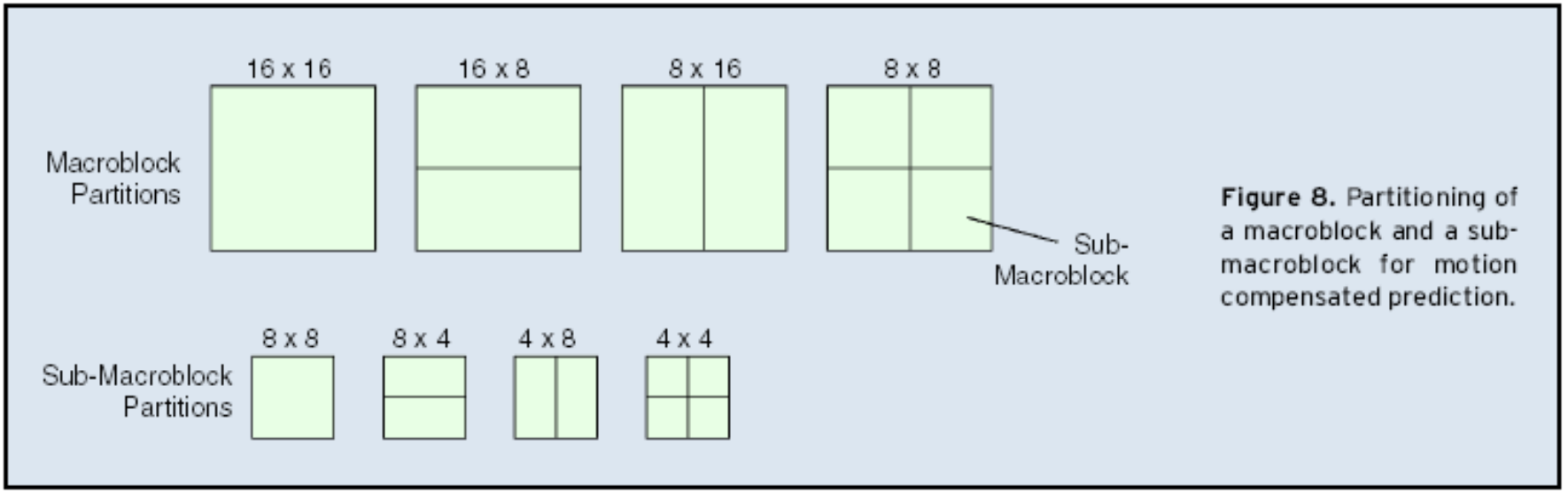


Figure 8. Partitioning of a macroblock and a sub-macroblock for motion compensated prediction.

From [Ostermann04]

Multiple Reference Frames for Motion Compensation

- Can use one or two from several possible reference frames
- When two reference frames are used, arbitrary weights can be used to combine them – Generalized B-picture

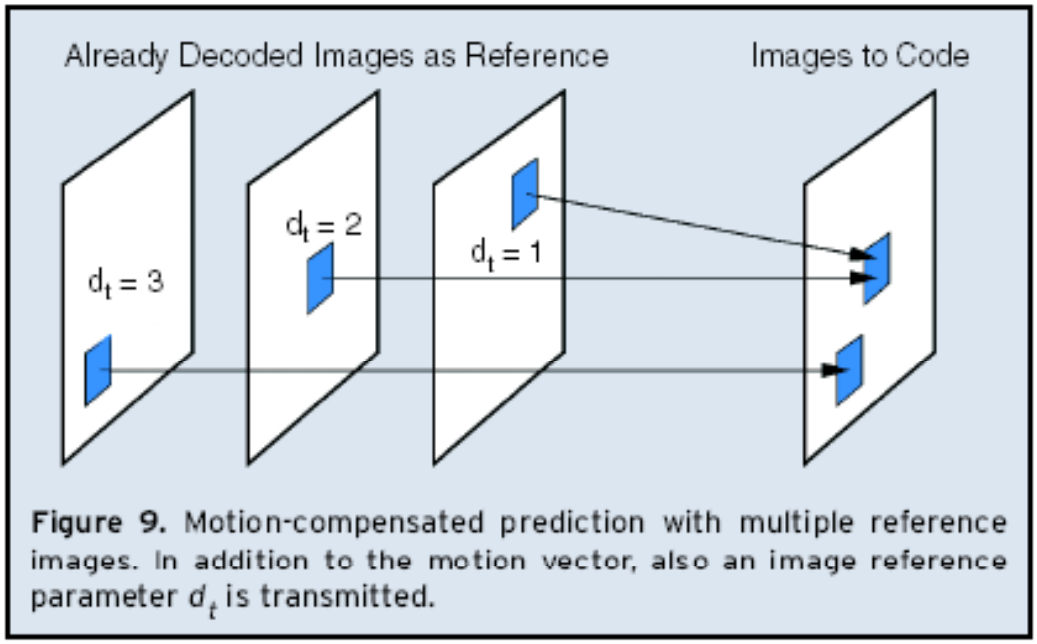
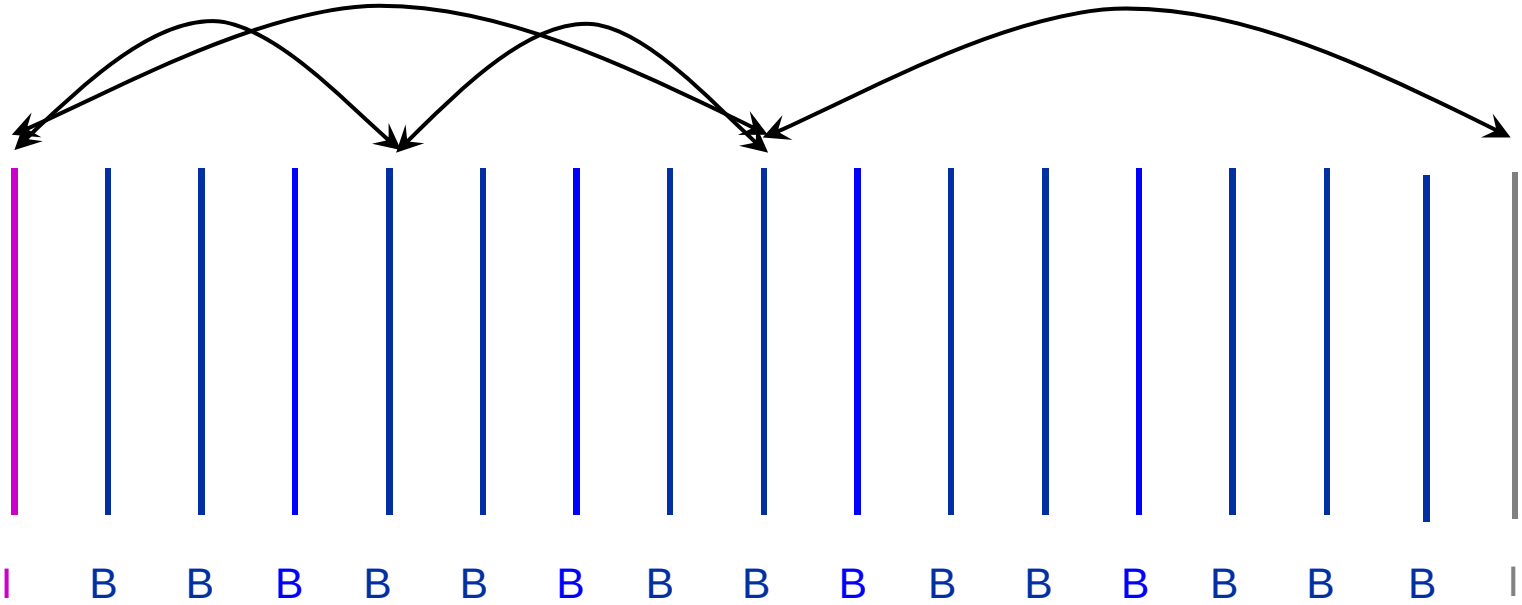


Figure 9. Motion-compensated prediction with multiple reference images. In addition to the motion vector, also an image reference parameter d_t is transmitted.

From [Ostermann04]

Generalized B-frames

In H.264, B frames can be used for prediction



Display order: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Transform

- 8x8 DCT
  - H.261
  - MPEG-1
  - H.263
  - MPEG-2
  - MPEG-4
  - DCT is non-integer; the result depends on the implementation details
- H.264:
  - Integer transforms, variable size (2x2, 4x4, 16x16)

Integer Transform

- Smaller block size (4x4 or 2x2) can better represent boundaries of moving objects, and match prediction errors generated by smaller block size motion compensation
- Integer transform can be implemented more efficiently and no mismatch problem between encoder and decoder

Primary transform

$$H_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad H_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad H_3 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Figure 10. Matrices H_1, H_2 and H_3 of the three different transforms applied in H.264/AVC.

From [Ostermann04]

Variable Length Coding

- H.261
  - DCT coefficients are converted into runlength representations and then coded using VLC (Huffman coding for each pair of symbols)
    - Symbol: (Zero run-length, non-zero value range)
  - Other information are also coded using VLC (Huffman coding)
- H.263
  - 3-D VLC for DCT coefficients (runlength, value, EOB)
  - Syntax-based arithmetic coding (option)
    - 4% savings in bit rate for P-mode, 10% saving for I-mode, at 50% more computations
- MPEG-4
  - 3-D VLC similar to H.263

H.264 Entropy Coding

- Baseline technique: CAVLC (context adaptively switched sets of variable length codes)
- A more complex technique called CABAC: context-based adaptive binary arithmetic coding
- Both offer significant improvement over Huffman coding which uses pre-designed coding tables based on some assumed statistics

Loop Filter

- In-Loop filtering can be applied to suppress propagation of coding noise temporally
- H.261
  - Separable filter [1/4,1/2,1/4]
  - Loop filter can be turned on or off
- MPEG-1
  - No loop filter (half-pel motion compensation provides some)
- H.263
  - Optional deblocking filter included in H.263+
  - Overlapped block motion effectively smoothes block boundaries
  - Decoder can choose to implement out-of-loop deblocking filter
- H.264
  - Deblocking filter adapts to the strength of the blocking artifact
- H.265
  - More elaborate inloop filtering

Adaptive Deblocking

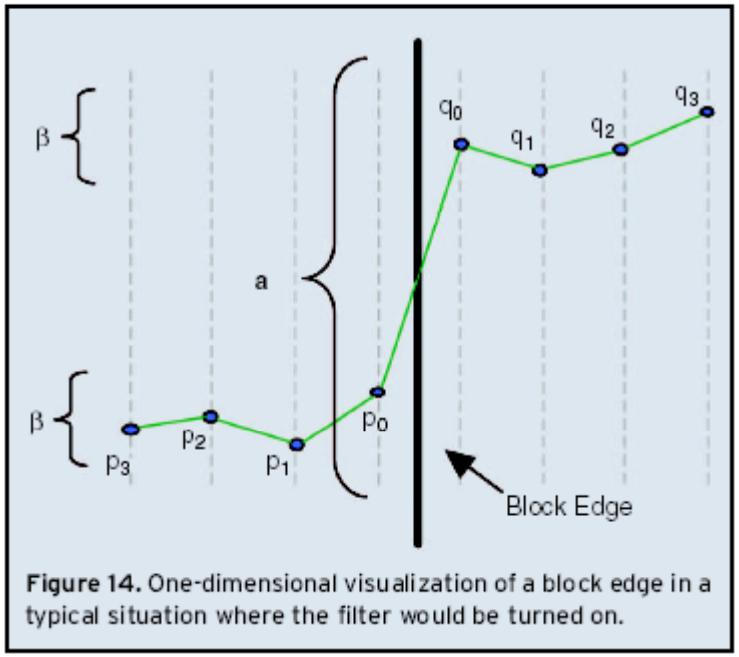


Figure 14. One-dimensional visualization of a block edge in a typical situation where the filter would be turned on.

From [Ostermann04]

- Whether filtering will be turned on depends on the pixel differences involving pixels p0,…, q0,…, and the filter depends on block characteristics and coding mode.
- Deblocking results in bit rate savings of 6-9% at medium qualities, and more remarkable subjective improvements,

Profiles and Levels

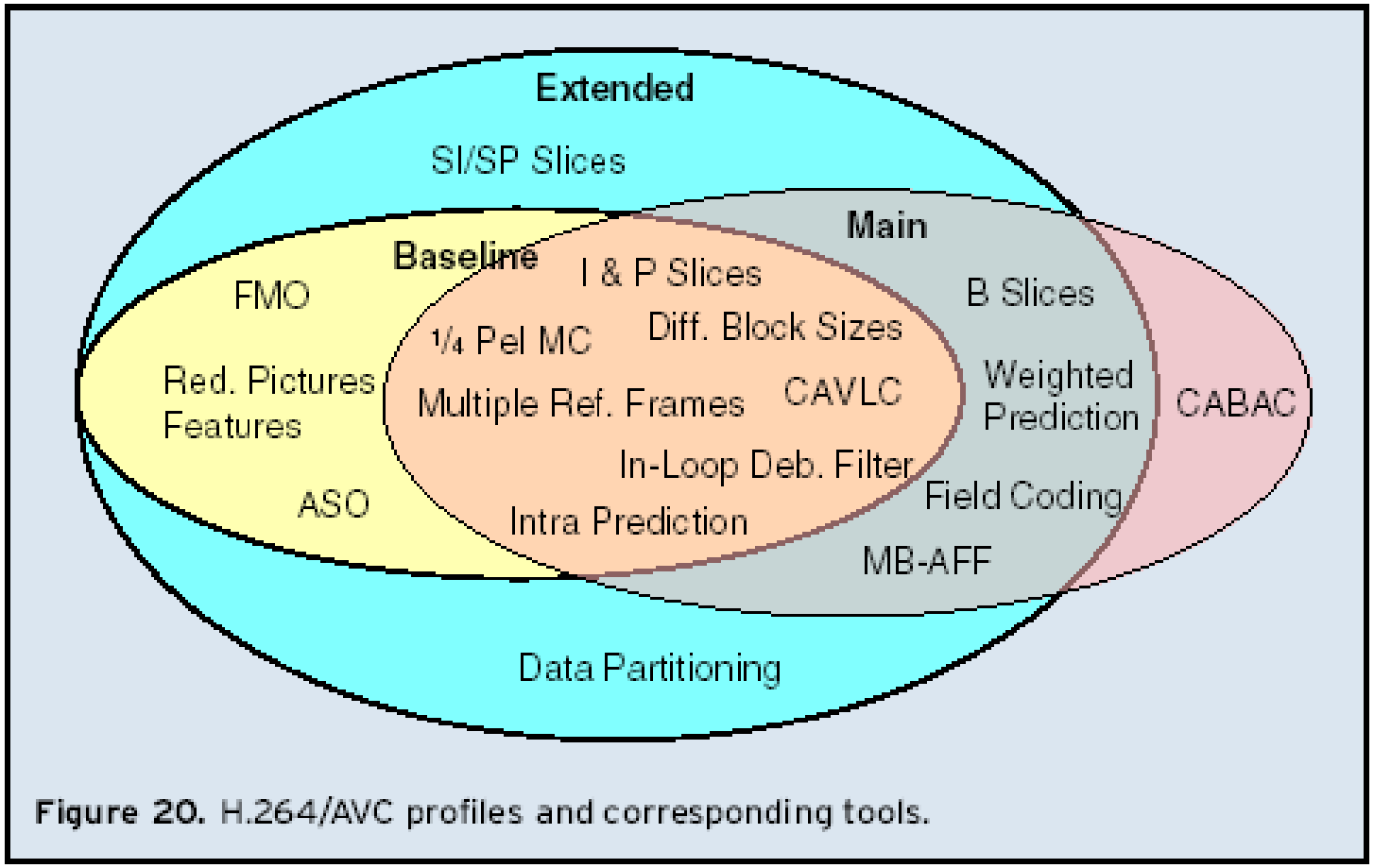


Figure 20. H.264/AVC profiles and corresponding tools.

From [Ostermann04]

Comparison with Previous Standards

- Coding efficiency: in terms of achievable rates for target video quality (PSNR)
  - Video streaming application
  - Video conferencing application
- Complexity:
  - Encoder
  - Decoder

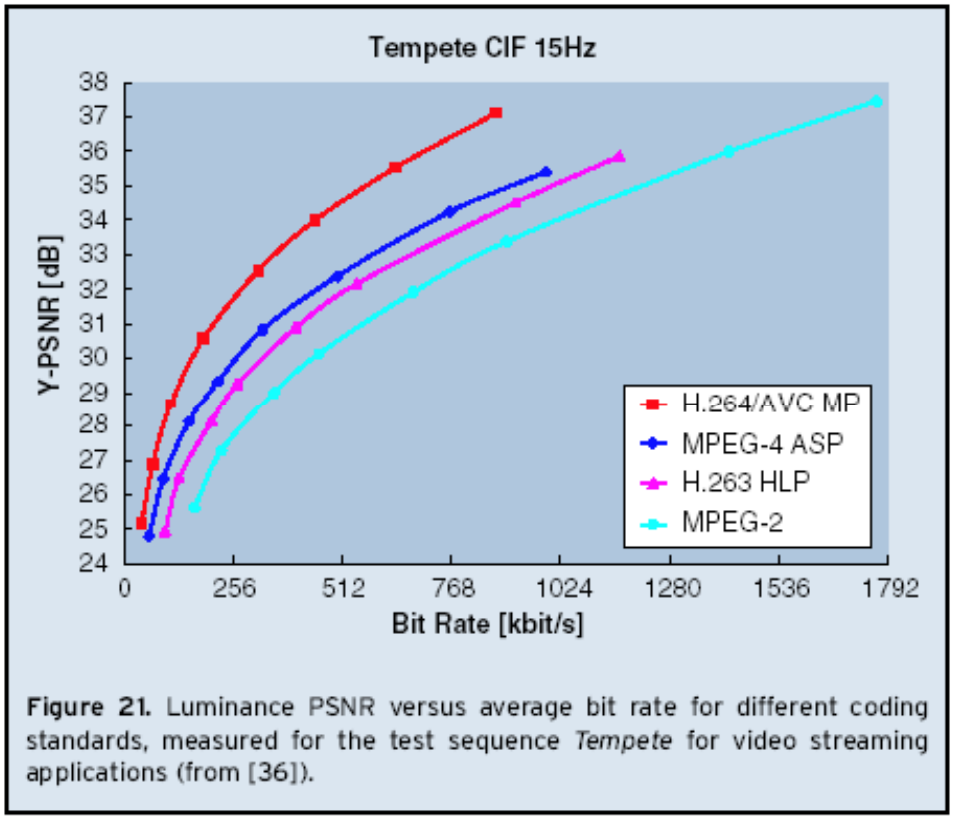


Figure 21. Luminance PSNR versus average bit rate for different coding standards, measured for the test sequence *Tempete* for video streaming applications (from [36]).

Coding efficiency for video streaming

Table 1.
Average bit rate savings for video streaming applications (from [10]).

Coder	Average Bit Rate Savings Relative To: MPEG-4 ASP	H.263 HLP	MPEG-2
H.264/AVC MP	37.44%	47.58%	63.57%
MPEG-4 ASP	–	16.65%	42.95%
H.263 HLP	–	–	30.61%

From [Ostermann02]

Coding efficiency for conferencing

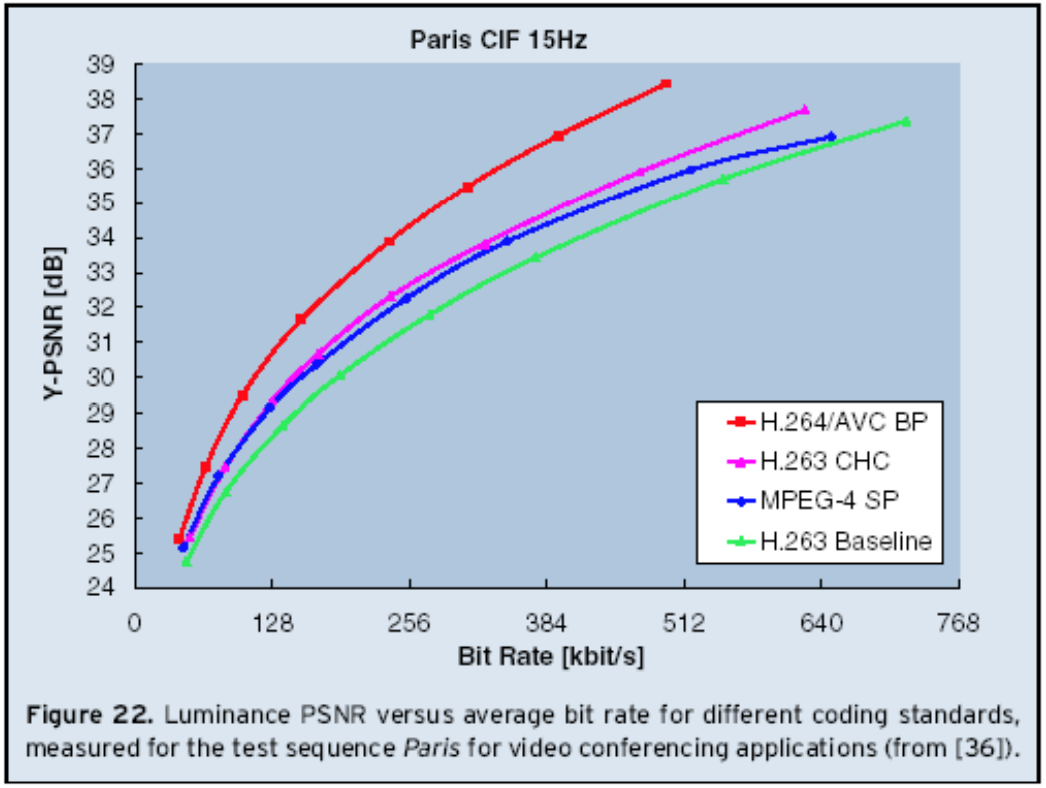


Figure 22. Luminance PSNR versus average bit rate for different coding standards, measured for the test sequence *Paris* for video conferencing applications (from [36]).

Table 2.
Average bit rate savings for video conferencing applications (from [10]).

Coder	Average Bit Rate Savings Relative To: H.263 CHC	MPEG-4 SP	H.263 Base
H.264/AVC BP	27.69%	29.37%	40.59%
H.263 CHC	–	2.04%	17.63%
MPEG-4 SP	–	–	15.69%

From [Ostermann02]

What about complexity ?

- H.264 decoder is about 2 times as complex as an MPEG-4 Visual decoder for the Simple profile
- H.264 encoder is about 10 times as complex as a corresponding MPEG-4 Visual encoder for the Simple profile
- The H.264/AVC main profile decoder suitable for entertainment applications is about 4 times more complex than MPEG-2

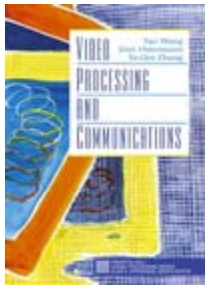
AVS (Audio Visual Coding Standard) Overview

- Chinese standard; 2002-2003 (Video)
- Licensing fees for all ISO and ITU standards after (not including) MPEG-1
- China produces more than 30 million Set Top Boxes
- Interlaced pictures, SDTV and HDTV
- Similar (slightly less) compression efficiency as H.264
- Interlaced pictures
- Intra prediction
- Variable block-size MC
- ¼ resolution motion, 4-tap interpolation filter
- 8x8 Integer Transform
- Deblocking

High Efficiency Video Coding (HEVC) The latest video coding standard

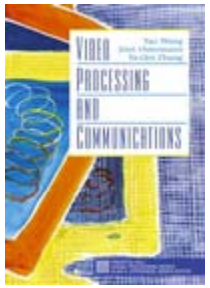
- Targeting for high resolution videos: HD (1920x1080) to ultra HD (7680x4320), progressive only (60p)
- Two targeted applications
  - Random access
  - Low delay
- Two categories of profile
  - High efficiency (HE)
  - Low complexity (LC)
- Performance: 2x better video compression performance compared to H.264/AVC.
  - Half the bit rate for similar quality

- Committee draft: Feb 2012.
- Target Standadization: Early 2013

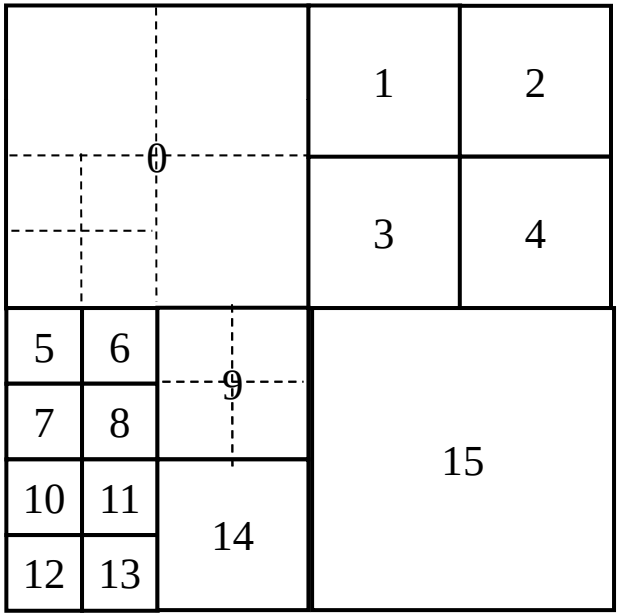


New Coding Tools in HEVC

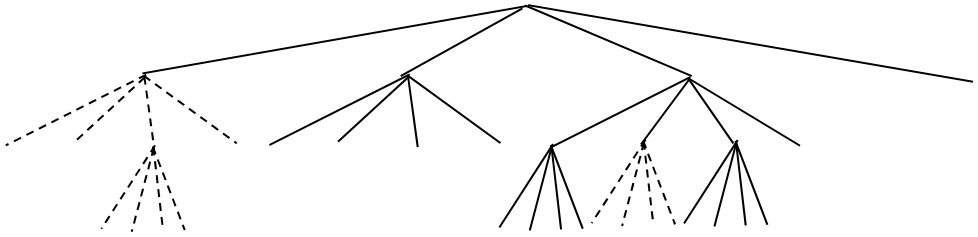
- Quadtree partition in 64x64 blocks: Block sizes from 8x8 to 64x64
- Up to 34 directions for intra-prediction
- For sub-pel motion estimation (down to ¼ pel), use 6- or 12-tap interpolation filter
- Advanced motion vector prediction
- CABAC or Low Complexity Entropy Coding
- Deblocking filter or Adaptive Loop Filter
- Extended precision options



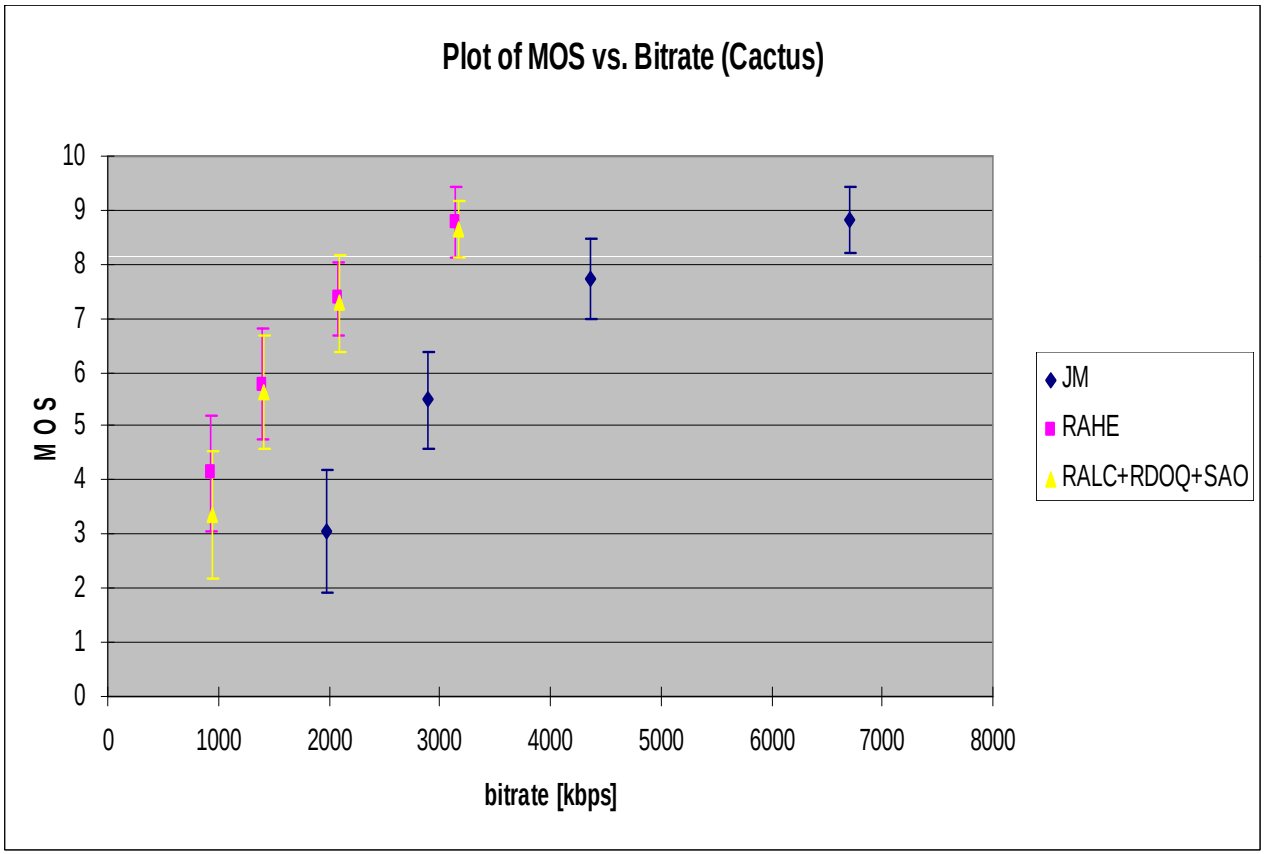
Tree Structure for block partition



Processing order



HEVC vs. H.264 Performance (sample sequence)

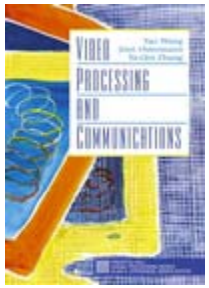


JM: H264 reference code

RAHE: HEVC high efficiency

RALC: HEVA low complexity

Better visual quality at half of the bit rate!

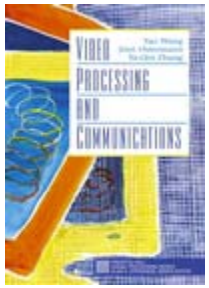


Summary

- H.261:
  - First video coding standard, targeted for video conferencing over ISDN
  - Uses block-based hybrid coding framework with integer-pel MC
- H.263:
  - Improved quality at lower bit rate, to enable video conferencing/telephony below 54 bkps (modems or internet access, desktop conferencing)
  - Half-pel MC and other improvement
- MPEG-1 video
  - Video on CD and video on the Internet (good quality at 1.5 mbps)
  - Half-pel MC and bidirectional MC
- MPEG-2 video
  - TV/HDTV/DVD (4-15 mbps)
  - Extended from MPEG-1, considering interlaced video

Summary (Cnt'd)

- MPEG-4
  - To enable object manipulation and scene composition at the decoder -> interactive TV/virtual reality
  - Object-based video coding: shape coding
  - Coding of synthetic video and audio: animation
- H.264:
  - Significant improvement in coding efficiency over H.263/MPEG4
  - Fundamentally similar ideas but with more adaptive/optimized implementation, feasible only with recent advance in computation power.
- Other MPEG standards
  - MPEG-7
    - To enable search and browsing of multimedia documents
  - MPEG-21
    - beyond MPEG-7, considering intellectual property protection, etc.



References

- Chap. 13
- H.264:
  - J. Ostermann et al., Video coding with H.264/AVC: Tools, performance, and complexity, IEEE Circuits and Systems Magazine, First Quarter, 2004
  - IEEE Trans. Circuits and Systems for Video Technology, special issue on H.264, July 2003.
- AVS
  - http://vspc.ee.cuhk.edu.hk/~ele5431/AVS.pdf (King Ngan, Chinese University of Hong Kong)
- HEVC
  - IEEE Trans. Circuits and Systems for Video Technology, Special Section on the Joint Call for Proposals on High Efficiency Video Coding (HEVC) Standardization. Dec. 2010