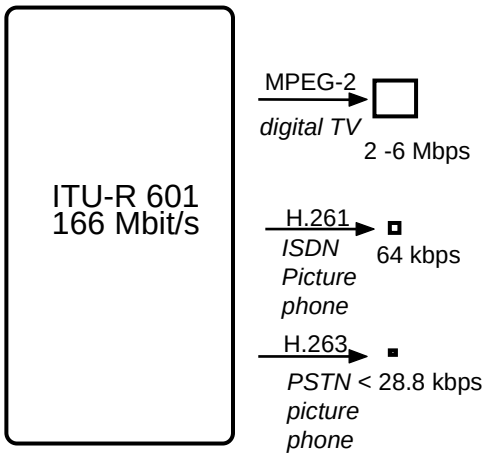
Video Coding Standards

- H.120
- H.261
- MPEG-1 and MPEG-2/H.262
- H.263
- MPEG-4

Video Coding Standards



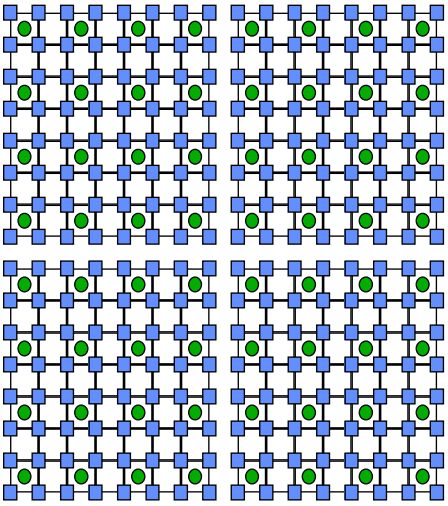
H.120: The First Digital Video Coding Standard

- ITU-T (ex-CCITT) Rec. H.120: The first digital video coding standard (1984)
- v1 (1984) had conditional replenishment, DPCM, scalar quantization, variable-length coding, switch for quincunx sampling
- v2 (1988) added motion compensation and background prediction
- Operated at 1544 (NTSC) and 2048 (PAL) kbps
- Few units made, essentially not in use today

H.261: The Basis of Modern Video Compression

- ITU-T (ex-CCITT) Rec. H.261: The first widespread practical success
- First design (late ‘90) embodying typical structure that dominates today: 16x16 macroblock motion compensation, 8x8 DCT, scalar quantization, and variable-length coding
- Other key aspects: loop filter, integer-pel motion compensation accuracy, 2-D VLC for coefficients
- Operated at 64-2048 kbps
- Still in use, although mostly as a backward-compatibility feature – overtaken by H.263

H.261&3 Macroblock Structure



= luminance pixel

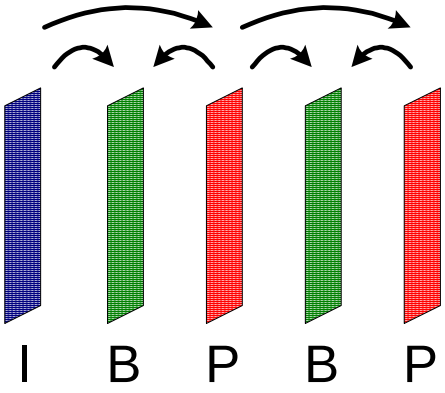
= chrominance pixel (two chroma fields)

- Intra/Inter Decisions: 16x16 macroblocks
- DCT of 8x8 blocks
- H.261: 16x16 1-pel motion
- H.263: 16x16 1/2-pel motion or
- H.263 AP mode: 8x8 1/2-pel motion with overlapping

MPEG-1

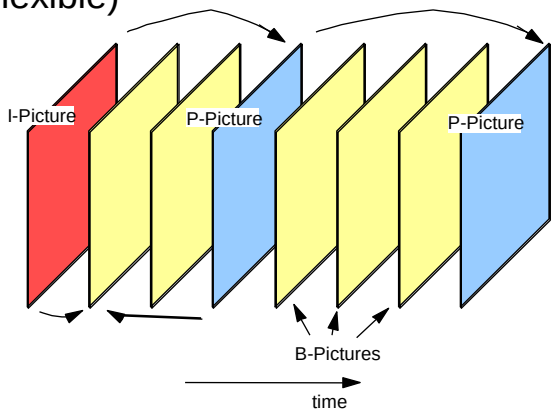
- Formally ISO/IEC 11172-2 (‘93), developed by ISO/IEC JTC1 SC29 WG11 (MPEG) – use is fairly widespread, but mostly overtaken by MPEG-2
- Superior quality to H.261 when operated at higher bit rates (≥ 1 Mbps for CIF 352x288 resolution)
- Can provide approximately VHS quality between 1-2 Mbps using SIF 352x240/288 resolution
- Technical features: Adds bi-directional motion prediction and half-pixel motion to H.261 design

Predictive Coding with B-Pictures



Hierarchical Syntax

- "Video Sequence“
- "Group of Pictures“ = “GOP“ (GOP structure is very flexible)



MPEG: Coding of B-Pictures

- Motion compensated prediction from two consecutive P- or I-pictures

  either

  only forward prediction (1 vector/macroblock)

  or

  only backward prediction (1 vector/macroblock)

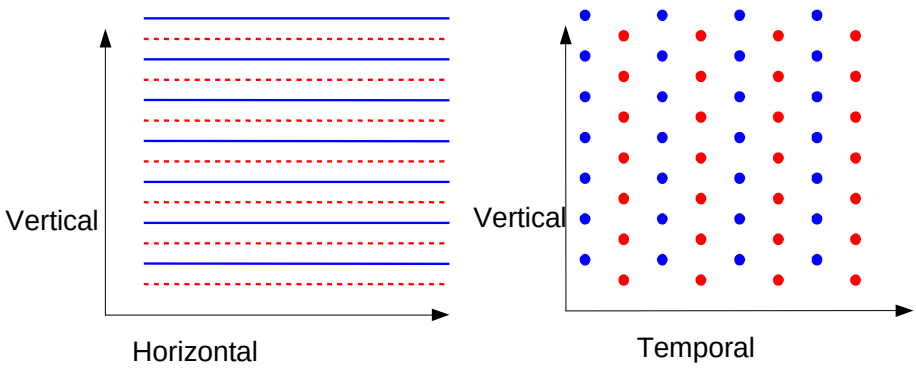
  or

  average of forward and backward prediction = bidirectional prediction = interpolation (2 vectors/macroblock)
- Half-pel accuracy of motion compensation: bilinear interpolation
- Coding of prediction error with 8x8-DCT, uniform quantization, zig-zag-scan as in I-pictures, VLC

MPEG-2/H.262

- Formally ISO/IEC 13818-2 & ITU-T H.262, developed ('94) jointly by ISO/IEC SC29 WG11 (MPEG) and ITU-T
- Now in wide use for DVD standard and DTV (the most commonly used video coding standard)
- Primary new technical features: support for interlaced-scan pictures and scalability
- Essentially the same as MPEG-1 for progressive-scan pictures, and MPEG-1 forward compatibility required
- Not especially useful below 2 Mbps (range of use normally 2-20 Mbps)

Interlaced Video



H.263: The Next Generation

- ITU-T Rec. H.263 (v1: 1995): The next generation of video coding performance, developed by ITU-T
- Has overtaken H.261 as dominant videoconferencing codec
- Superior to H.261 at all bit rates
- Wins by a factor of two at very low rates
- Versions 2: H.263+ (late ’97/early ’98) and 3 (about now) later developed

Technical Features of H.263

- "Baseline" Algorithm Features beat H.261
  - Half-pel motion compensation
  - 3-D variable length coding of DCT coefficients
  - Median motion vector prediction
  - More efficient coding pattern signaling
  - Deletable GOB header overhead
- Optional Enhanced Modes
  - Increased motion vector range with picture extrapolation
  - Advanced prediction mode (AP mode): Variable-size and overlapped block motion compensation (OBMC) with picture extrapolation
  - PB-frames (bi-directional prediction)
  - Arithmetic entropy coding
  - Continuous-presence multipoint / video mux

Motion Compensation H.261 vs. H.263

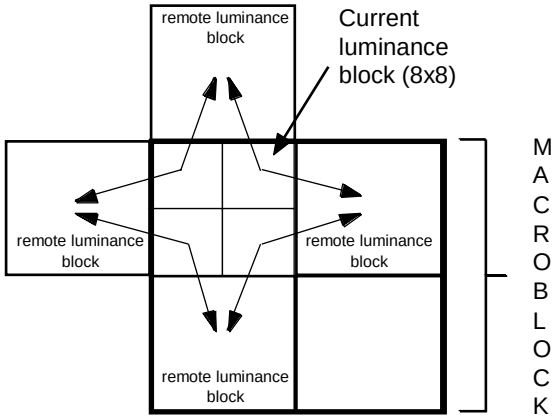
- H.261 (1990): integer-pel accuracy, loop filter, 1 motion vector per MB

- H.263 (1995): half-pel accuracy, no loop filter, 1 motion vector per MB (already in MPEG-1,2)

- H.263 option „Advanced Prediction Mode"
  - overlapped block motion compensation (OBMC),
  - switch between 1 or 4 motion vectors per MB

- H.263 option „PB frames"

H.263 AP mode: OBMC



H.263 AP mode: OBMC Weights

4	5	5	5	5	5	5	4
5	5	5	5	5	5	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	5	5	5	5	5	5
4	5	5	5	5	5	5	4

for MV of current luminance block

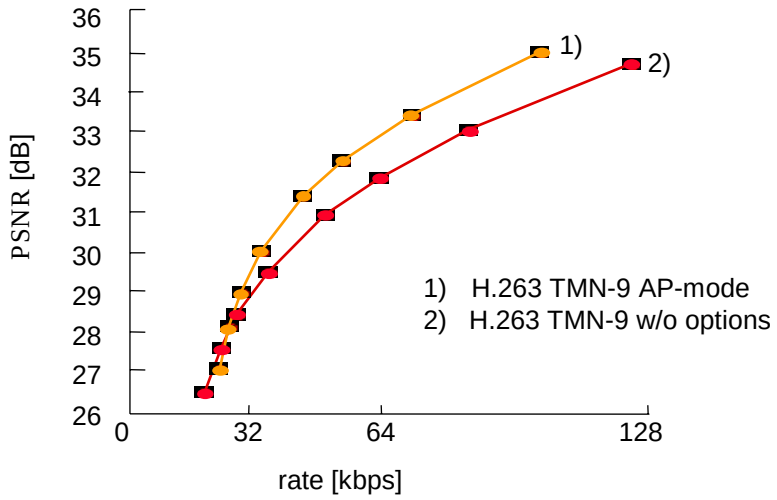
2	2	2	2	2	2	2	2
1	1	2	2	2	2	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	2	2	2	2	2	2
2	2	2	2	2	2	2	2

for remote MV of top/bottom luminance block

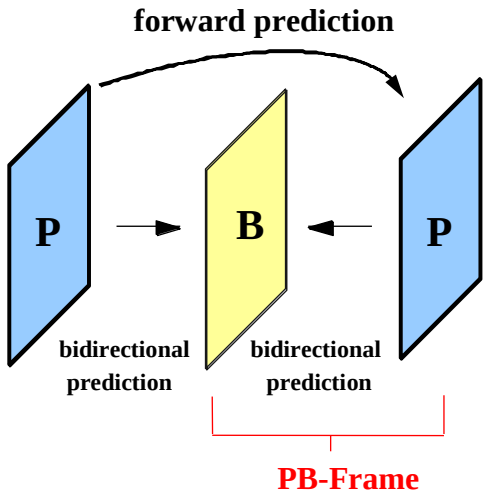
2	1	1	1	1	1	1	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	1	1	1	1	1	1	2

for remote MV of left/right luminance block

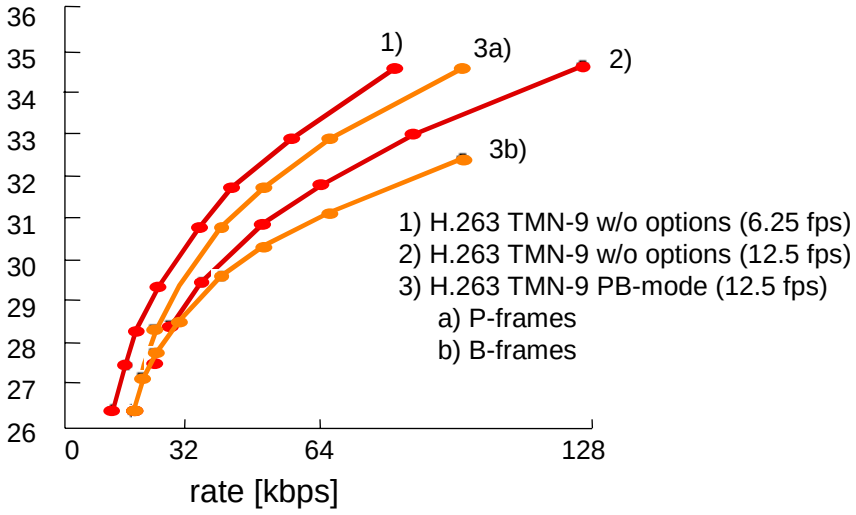
Performance of H.263 AP Mode



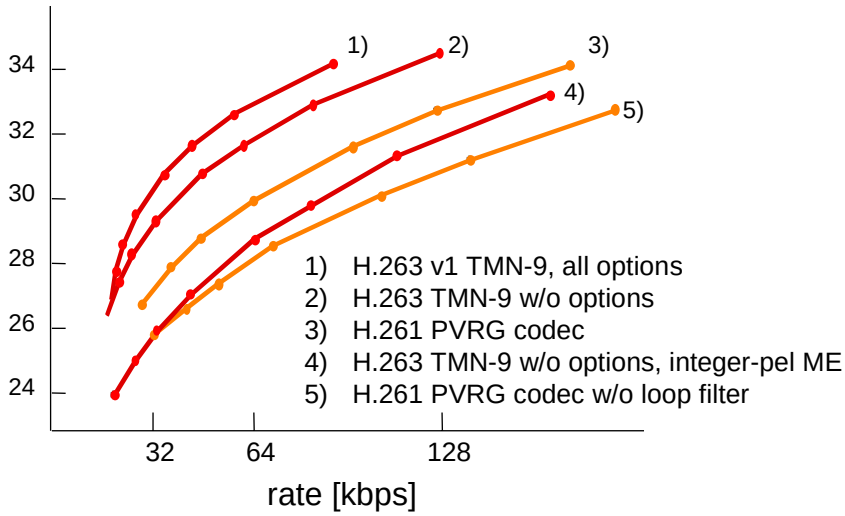
H.263: PB-frames



Performance of H.263 PB-Mode



Performance of H.263 and H.261



H.263+ Feature Categories

- Improved compression efficiency (e.g., 15-25% overall improvement over H.263v1)
- Error resilience (1st resilient video standard)
- Custom and Flexible Video Formats
- Scalability for resilience and multipoint
- Supplemental enhancement information

H.263+: Better Ways of Coding

- Efficiency & Perceptual Enhancement
  - Advanced intra-coding
  - Deblocking filter (in the loop)
  - Alternate inter VLC for heavy motion
  - Modified quantization/coefficient range
  - Improved PB-frames (and B frames too)
  - Tweaks of prior features (RC, UMV, AP, CPM)
- Dynamic Resolution Features
  - Reference-picture resampling (also global motion)
  - Reduced-resolution update

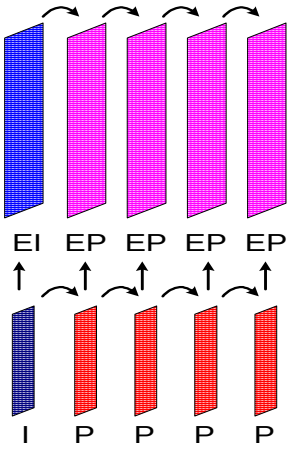
H.263+: Improved Error Resilience

- Slice-structured coding (packetization, low delay, object-oriented coding)
- Independent segment decoding (packetization, encoder parallelization)
- Reference picture selection (with optional back-channel)
- Scalability Features (see part 3)

H.263+: New Kinds of Pictures

- Custom Video Source Formats
  - Custom picture resolutions
  - Custom pixel/picture aspect ratios
  - Custom picture clock rates
- Scalability Pictures
  - SNR scalability (fidelity enhancement layer)
  - Temporal scalability (B pictures)
  - Spatial scalability (resolution enhancement layer)

Spatial Scalable Coding



What New Applications Does H.263+ Enable?

- Error/packet-loss resilient video
- Scalable multiple bit rate video at low bit rates
- Flexible-format low bit rate video
- High frame rate video (e.g., 72 frames / sec)
- Progressive picture quality refinement
- Dynamic video picture resolution
- Graphics/still-picture snapshot coding
- Object-layered video using chroma key
- Special effects / global motion

MPEG-4: H.263 + Additions + Variable Shape Coding

- MPEG-4 (v1: early 1999), formally ISO/IEC 14496-2: Roughly follows H.263 design and adds all prior features and various “trick modes” and (most important) shape coding
- Includes zero-tree wavelet coding of still pictures, segmented coding of shapes, coding of synthetic content
- v2 (early 2000) developed, v3 and v4 in progress

MPEG-4 v1: Simple Profile

- H.263v1 Baseline (exact compatibility to H.263, plus custom picture format with a different picture header)
- Variable block-size and picture-extrapolating MC (parts of H.263v1 Annexes D & F & H.263+ Annex J)
- Spatial-Predictive Intra Coding (similar to H.263 Annex I)
- Altered step size for DC coefficients
- Relative placement of rectangular pictures (sort of H.263+ Annex R)
- Slice-structured coding (like H.263+ Annex K)
- Data partitioning and reversible VLC (not in H.263 until v3)

MPEG-4 v1: Core Profile

- Binary shape coding (shape coding is the most unique new feature of MPEG-4 – different than H.263+ Annex L chroma keying)
- B-picture temporal scalability (a subset of H.263+ Annex O)
- P-picture temporal scalability (a subset of H.263+ Annex N)
- MPEG-2-style inverse quantization (not in H.263)

MPEG-4 v1: Main Profile

- Grey-scale shape coding (soft representation of shapes)
- Interlace (backward compatibility for 1940's-era analog compression technique)
- Static sprite coding (warpable pictures, different than H.263+ Annex P)
- Scalable still pictures

MPEG-4 v1: Other Technical Features

- Temporal & spatial scalability (a subset of H.263+ Annex O)
- Overlapped block motion compensation (part of H.263v1 Annex F)
- 12-bit video
- Dynamic 2D mesh coding
- Face animation modeling

MPEG-4 v2

- Fidelity
  - Quarter-pel Motion Compensation
  - Global MC (somewhat different than H.263+ Annex P)
  - Shape-Adaptive DCT (for shape-based coding)
  - Reduced-Resolution Update (H.263+ Annex Q)
- Error Resilience
  - Reference Picture Selection (H.263+ Annex N)

MPEG-4 v2

- Object based spatial scalability (spatial scalability for arbitrary shaped coding)
- Multiple auxiliary components (specialized apps)
- Mesh coding for Body Mesh and 3-D Mesh (synthetic and semi-synthetic content)
- Still-Picture Coding
  - Wavelet tiling for still picture (random access within a still picture)
  - Error resilience for still pictures (recovery from errors)
  - Scalable arbitrary shape for still picture coding (scalability with shape coding for still pictures)

H.263++ New Version 3 Features, I

- **Annex U**: Fidelity enhancement by macroblock and block-level reference picture selection
  - Improved compression performance
  - Improved error resilience
- **Annex V**: Packet Loss & Error Resilience using data partitioning with reversible VLCs (roughly similar to MPEG-4 data partitioning, but improved by using reversible coding of motion vectors rather than coefficients)

H.263++ New Version 3 Features, II

- **Annex W**: Additional Supplemental Enhancement Information
  - IDCT Mismatch Elimination (specific fixed-point fast IDCT)
  - Arbitrary binary user data
  - Text messages (arbitrary, copyright, caption, video description, and URI)
  - Error Resilience:
    - Picture header repetition (current, previous, next+TR, next-TR)
    - Spare reference pictures for error concealment
  - Interlaced field indications (top & bottom)