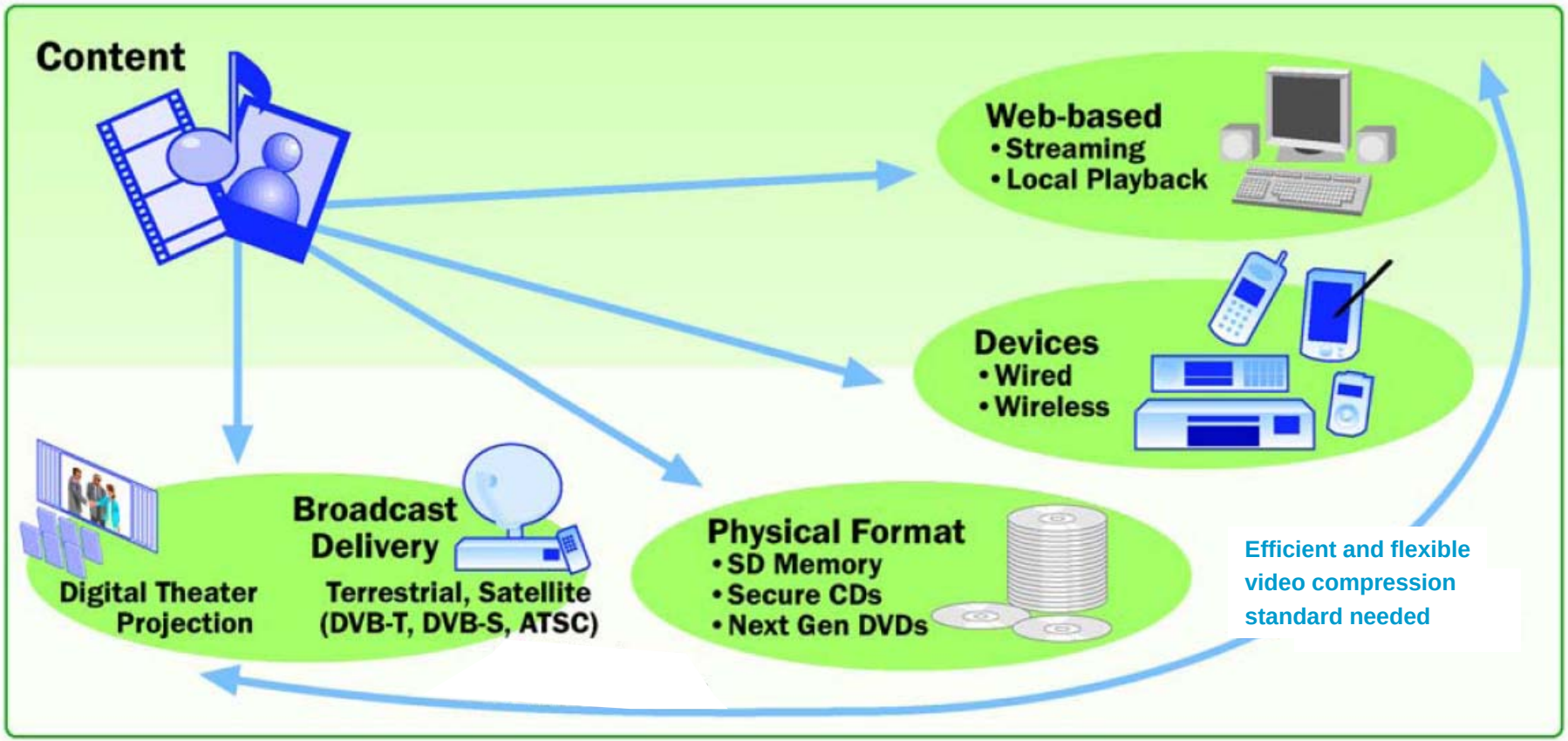
Overview: Video Coding Standards

- Video coding standards: applications and common structure
- ITU-T Rec. H.261
- (ITU-T Rec. H.263)
- ISO/IEC MPEG-1
- ISO/IEC MPEG-2
- (ISO/IEC MPEG-4)
- State-of-the-art: H.264/AVC

Applications of Video Compression



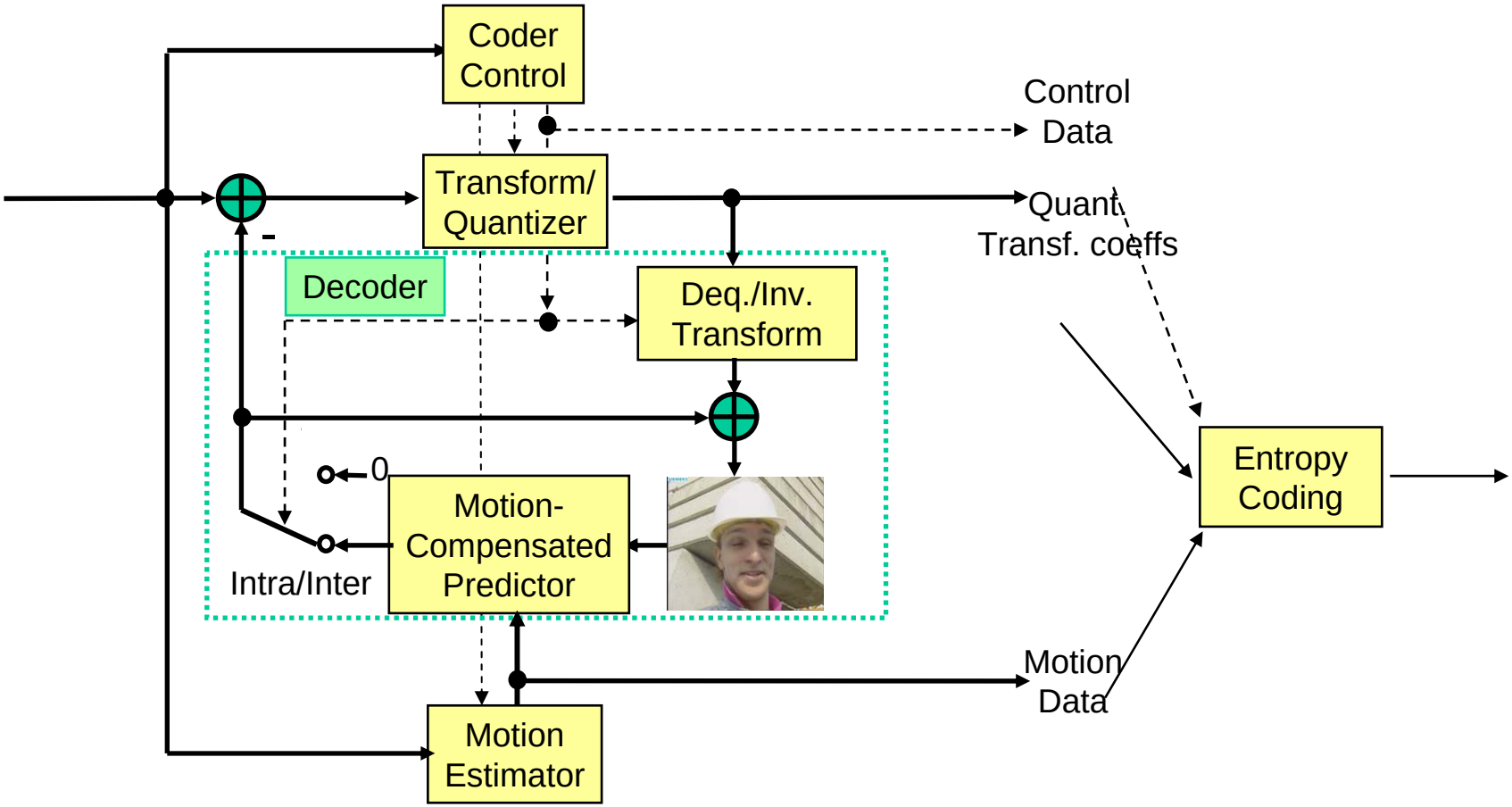
Adapted from *[Srinivasan et al., 2004]*

Applications of Video Compression

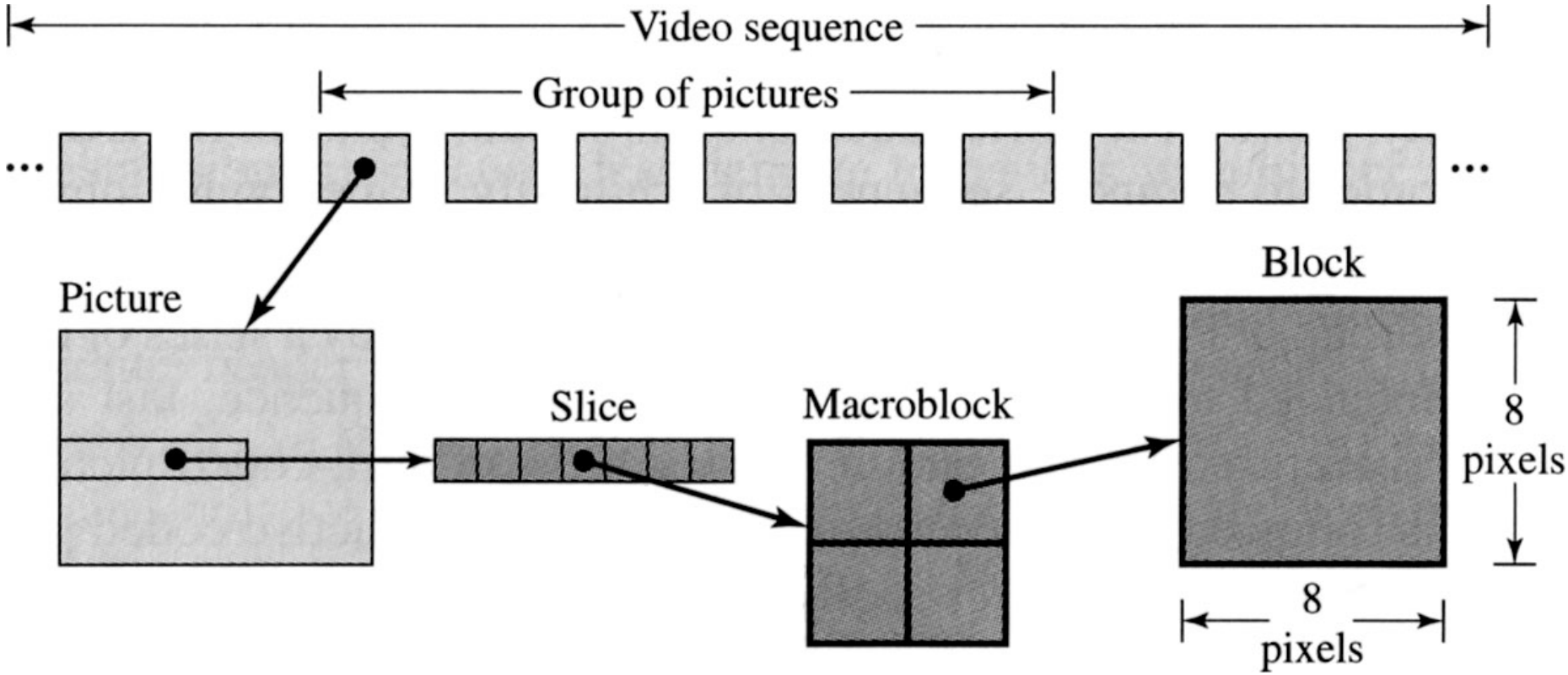
Digital television broadcasting	2 . . . 6 Mbps (10…20 Mbps for HD)	MPEG-2 (H264/AVC)
DVD video	5 . . . 8 Mbps	MPEG-2
Internet video streaming	20 . . . 300 kbps	MPEG-1, H.264/AVC, VC-1, or similar proprietary
Videoconferencing, videotelephony	20 . . . 2000 kbps	H.261, H.263, H.264/AVC
Video over 3G wireless	100 . . . 500 kbps	H.263, MPEG-4, H.264/AVC

Motion-compensated Hybrid Coding

H.261, MPEG-1, MPEG-2, H.263, MPEG-4, H.264/AVC



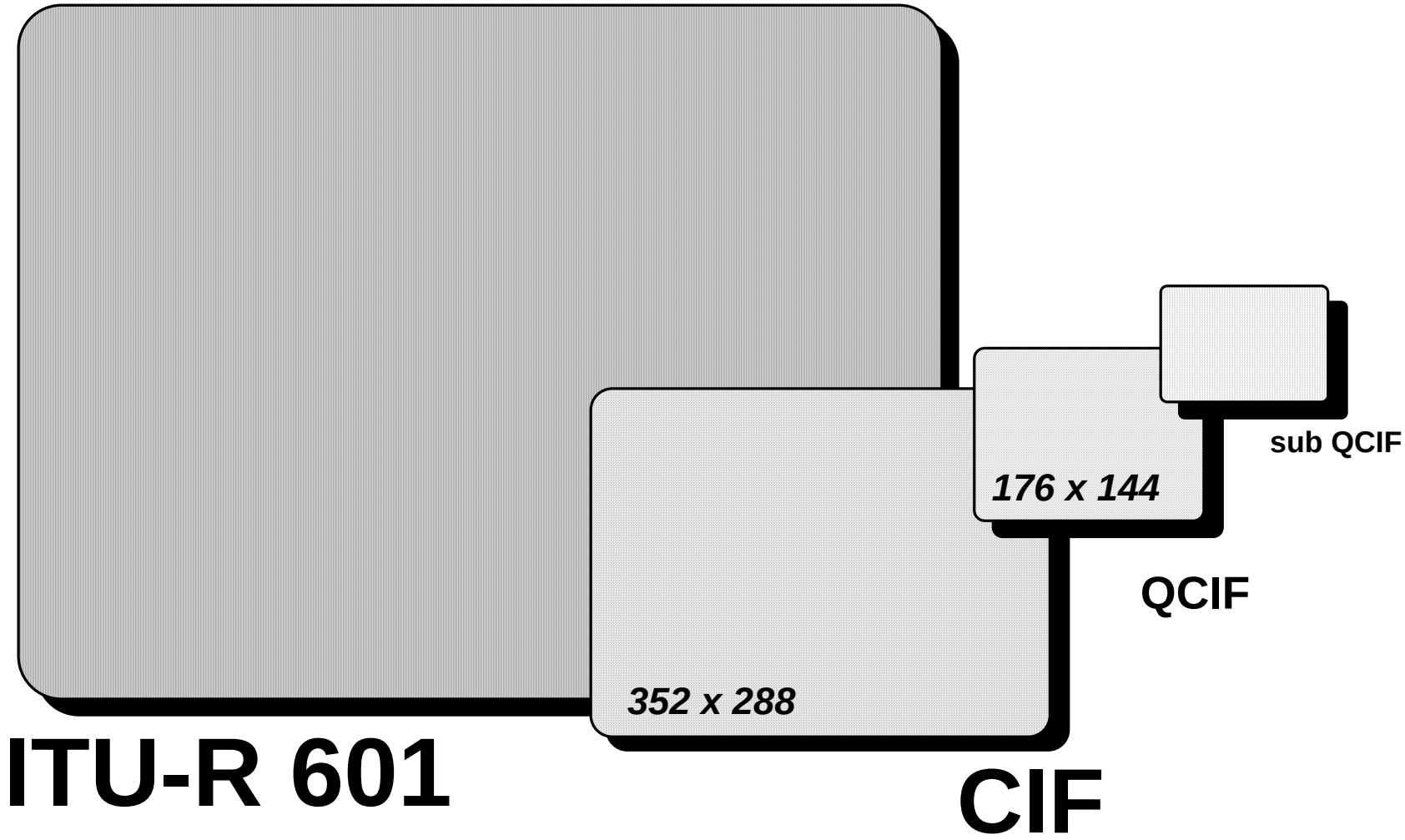
Video Compression Standards: Hierarchical Syntax



ITU-T Rec. H.261

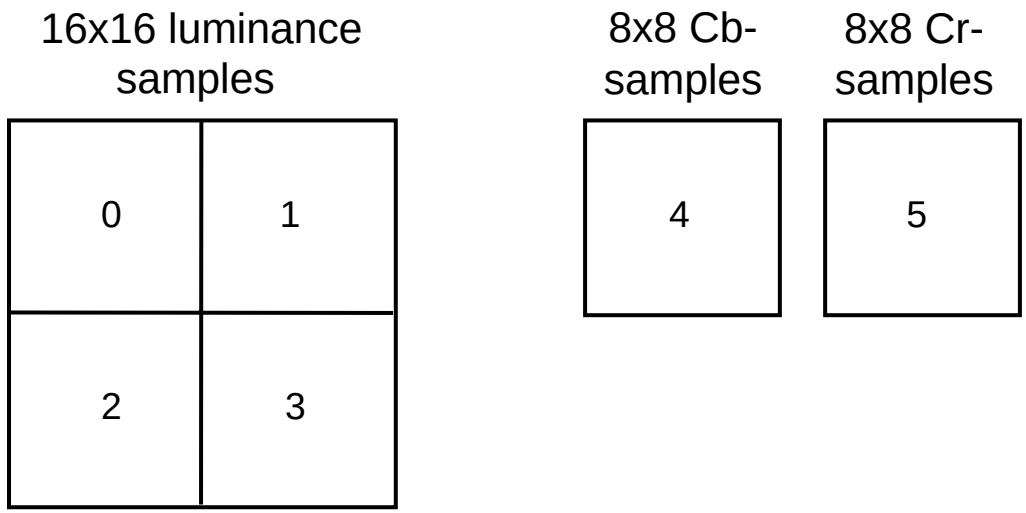
- International standard for ISDN picture phones and for video conferencing systems (1990)
- Image format: CIF (352 x 288 Y samples) or QCIF (176 x 144 Y samples), frame rate 7.5 ... 30 fps
- Bit-rate: multiple of 64 kbps (= ISDN-channel), typically 128 kbps including audio.
- Picture quality: for 128 kbps acceptable with limited motion in the scene
- Stand-alone videoconferencing system or desk-top videoconferencing system, integrated with PC

Image Formats



H.261 Macroblocks

- Macroblock (MB) of 16x16 pixels
- Sampling format: 4:2:0
- MB consists of 4 luminance and 2 chrominance blocks



H.261 Motion-Compensated Prediction

- Integer-pel accuracy
- One displacement vector per macroblock
- Maximum displacement vector range +/-16 horizontally and vertically
- Adaptive loop filter, separable in 1-D horizontal and vertical impulse response: [¼, ½, ¼]
- Differential encoding of motion vectors

H.261 Residual Coding

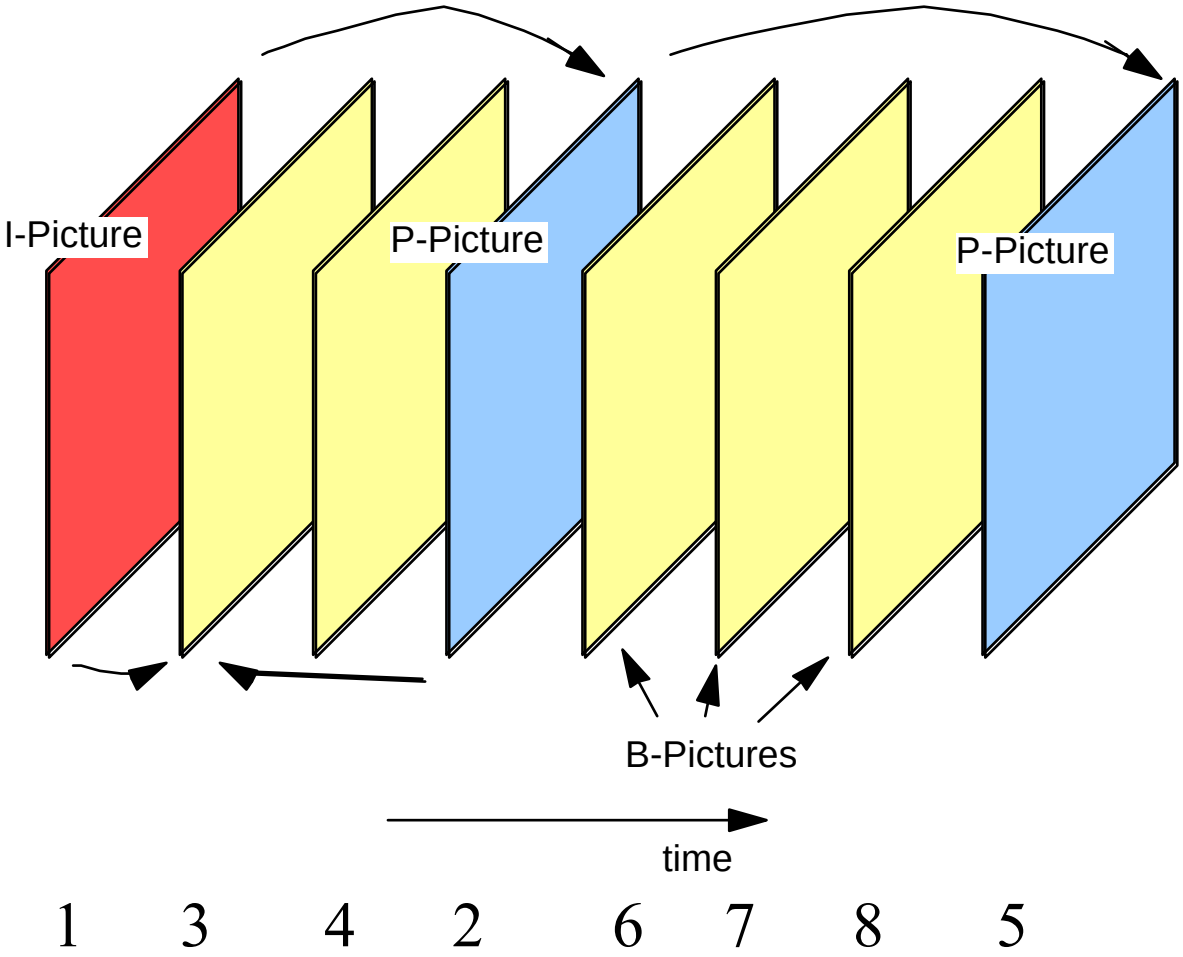
- 8x8 DCT
- Quantization
  - Uniform quantizer (Δ=8) for intra-mode DC coefficients
  - Uniform threshold quantizer (Δ=2,4,…,62) for AC coefficients in intra-mode and all coefficients in inter-mode
- Zig-zag scan
- Run-level coding for entropy coding
  - (zero-run, value) symbols
  - zero-run: the number of coefficients quantized to zero since the last nonzero coefficient
  - value: the amplitude of the current nonzero coefficient

H.261 Macroblock Types (VLC Table)

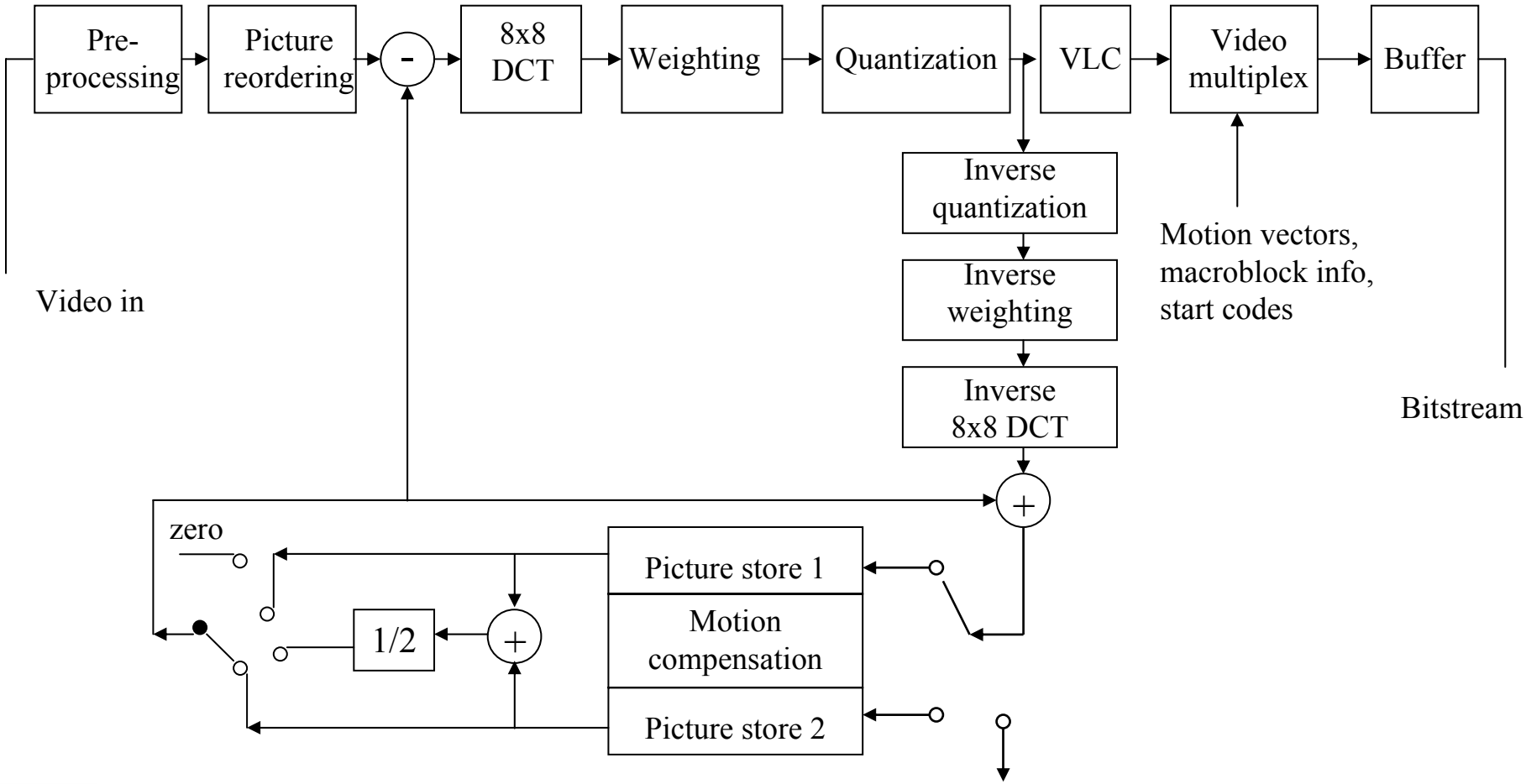
Prediction	MQUANT	MVD	CBP	TCOEFF	VLC
Intra				X	0001
Intra	X			X	0000 001
Inter			X	X	1
Inter	X		X	X	0000 1
Inter+MC		X			0000 0000 1
Inter+MC		X	X	X	0000 0001
Inter+MC	X	X	X	X	0000 0000 01
Inter+MC+FIL		X			001
Inter+MC+FIL		X	X	X	01
Inter+MC+FIL	X	X	X	X	0000 01

MPEG-1/2: GOP Structure

- "Group of Pictures" = “GOP“, GOP structure is very flexible



MPEG-1 Encoder



MPEG-1: coding of I-pictures

- I-pictures: intraframe coded
- 8x8 DCT
- Arbitrary weighting matrix for coefficients
- Differential coding of DC-coefficients
- Uniform quantization
- Zig-zag-scan, run-level-coding
- Entropy coding
- Unfortunately, not quite JPEG

MPEG-1: coding of P-pictures

- Motion-compensated prediction from an encoded I-picture or P-picture (DPCM)
- Half-pel accuracy of motion compensation, bilinear interpolation
- One displacement vector per macroblock
- Differential coding of displacement vectors
- Coding of prediction error with 8x8-DCT, uniform threshold quantization, zig-zag-scan as in I-pictures

MPEG-1: 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 = 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

MPEG-2 vs. MPEG-1

- Efficiently compress interlaced digital video at broadcast quality
  - Field/frame pictures
  - Chroma sampling
  - New prediction modes
  - Field/frame DCT
  - Additional scan patterns for DCT coefficients
  - Motion compensation with blocks of size 16x8 pels
- Improved coding efficiency by different quantization, VLC tables
- Various scalability modes

JVT Project

- ITU-T Q.6/SG16 (**VCEG - Video Coding Experts Group**) formed for ITU-T standardization activity for video compression since 1997
- **August 1999**: 1st test model (TML-1) of H.26L
- **December 2001**: Formation of the **Joint Video Team (JVT)** between **VCEG** and ISO/IEC JTC 1/SC 29/WG 11 (**MPEG)** to establish a joint standard project - **H.264 / MPEG4-AVC**
- **ITU-T Approval**: **May 2003**
- **ISO/IEC Approval**: **October 2003**

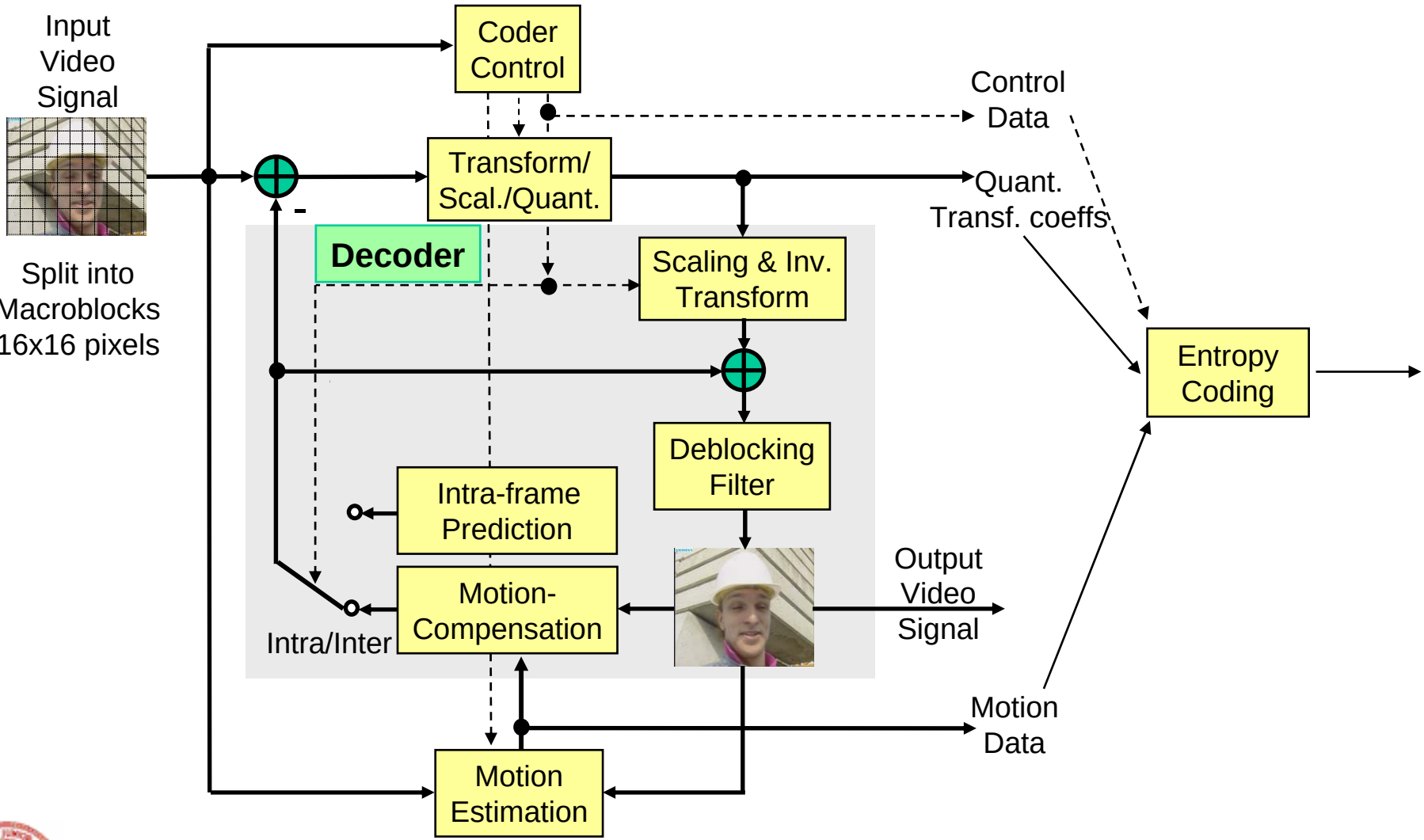
[source: G. Sullivan, VCEG]

H.264/AVC Profiles

- **Baseline:** core compression capabilities, plus error resilience, e.g., for videoconferencing, mobile video
- **Main**: high compression and quality, e.g., for broadcasting
- **Extended**: added features for efficient streaming

[source: G. Sullivan, VCEG]

H.264/AVC Coder



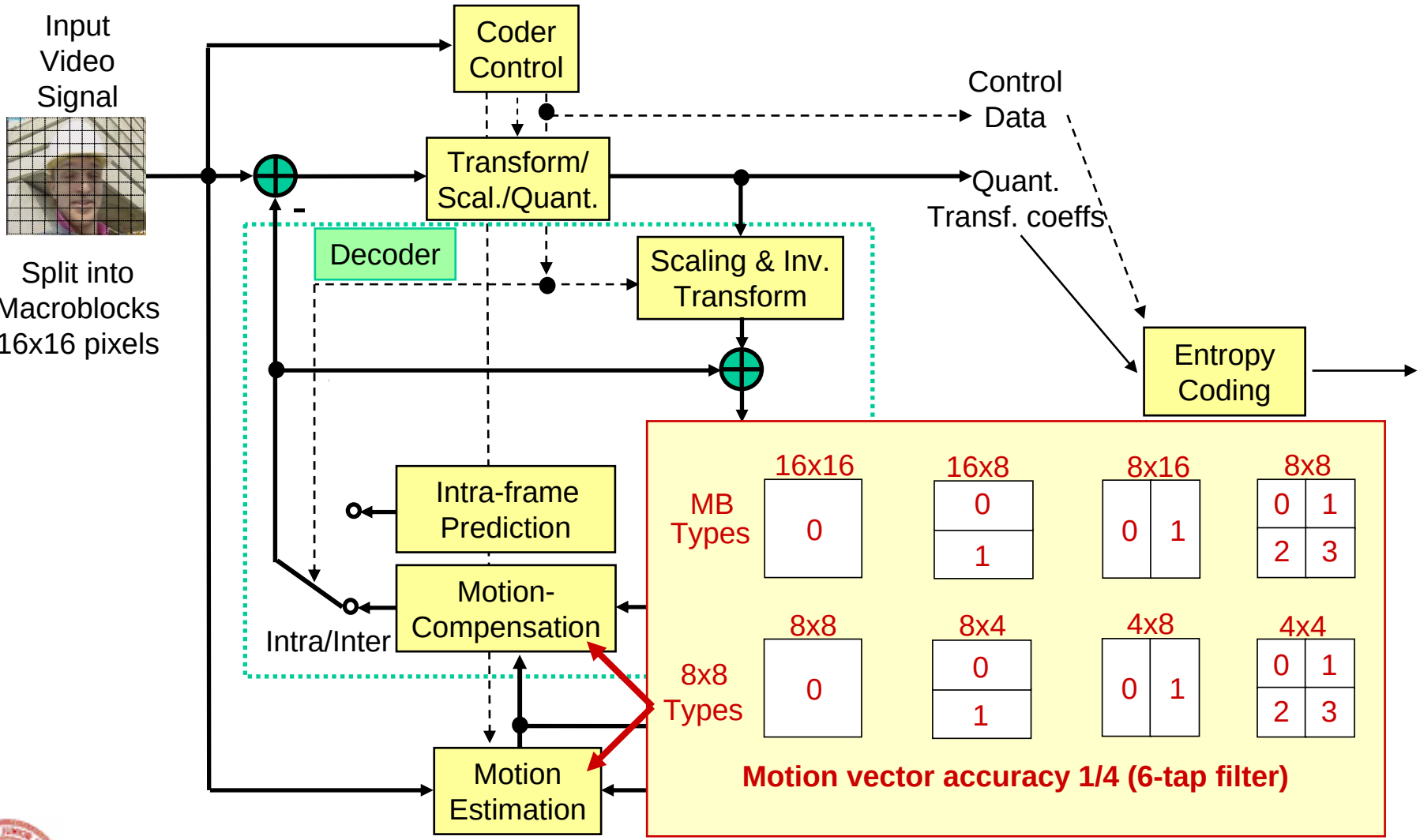
[source: G. Sullivan, VCEG]

Common Elements with other Standards

- Macroblocks: 16x16 luma + 2 x 8x8 chroma samples
- Block-wise motion compensation
- Variable block-size motion compensation
- Block transform of prediction error
- Scalar quantization
- I, P, and B coding types

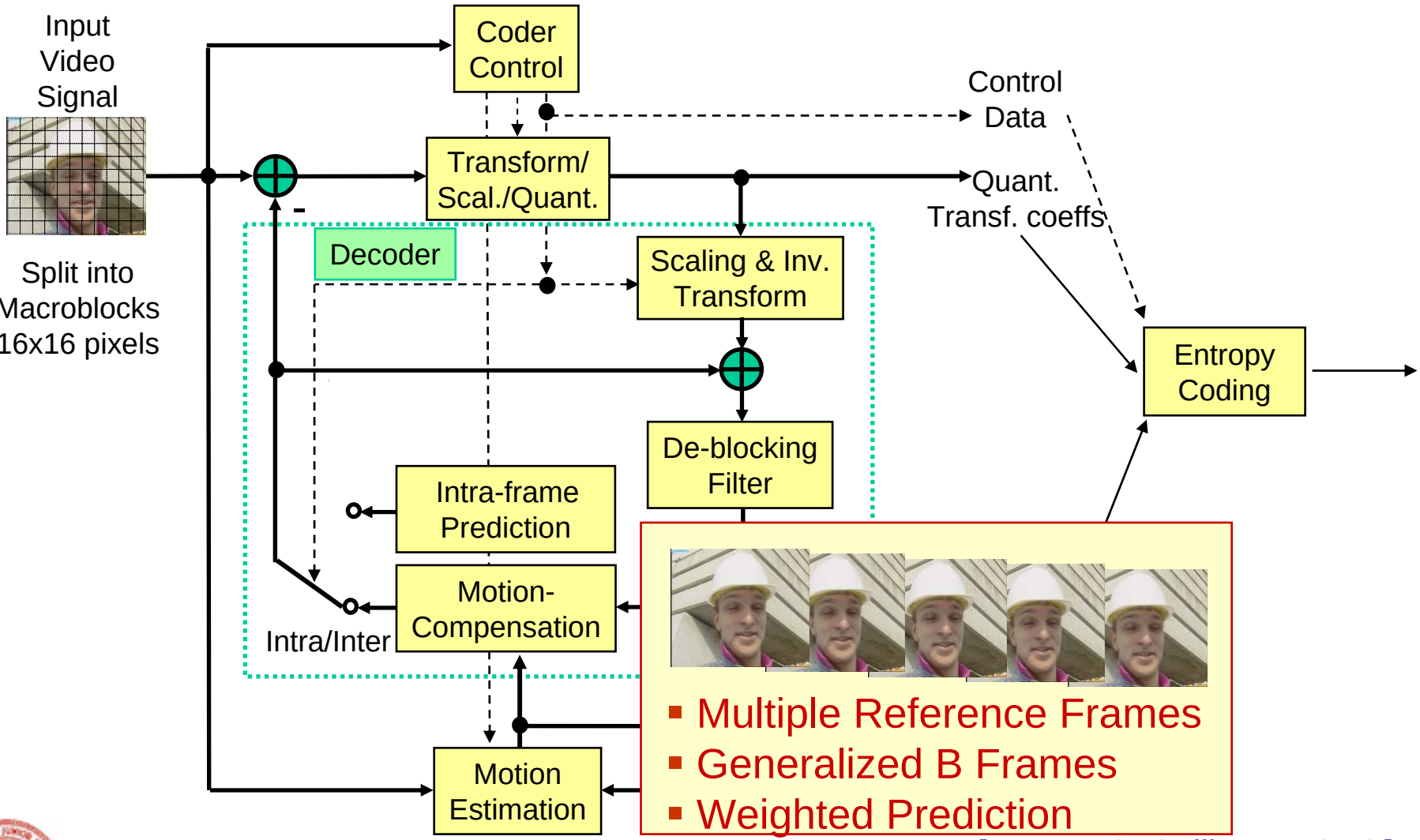
[source: G. Sullivan, VCEG]

H.264 Motion Compensation Accuracy



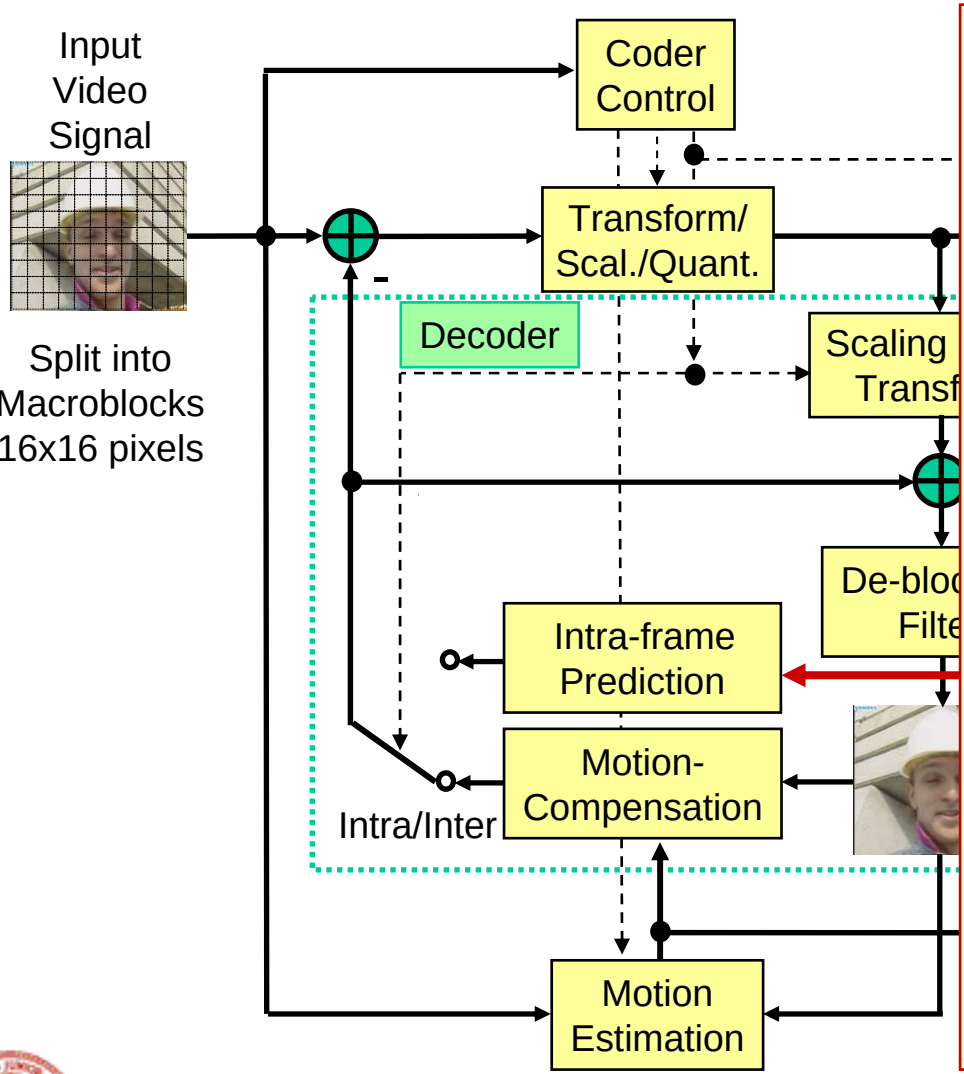
[source: G. Sullivan, VCEG]

H.264 Multiple Reference Frames



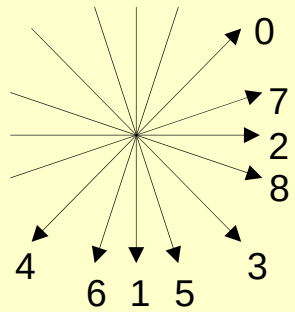
[source: G. Sullivan, VCEG]

H.264 Intra Prediction



- Directional spatial prediction (9 types for luma, 1 chroma)

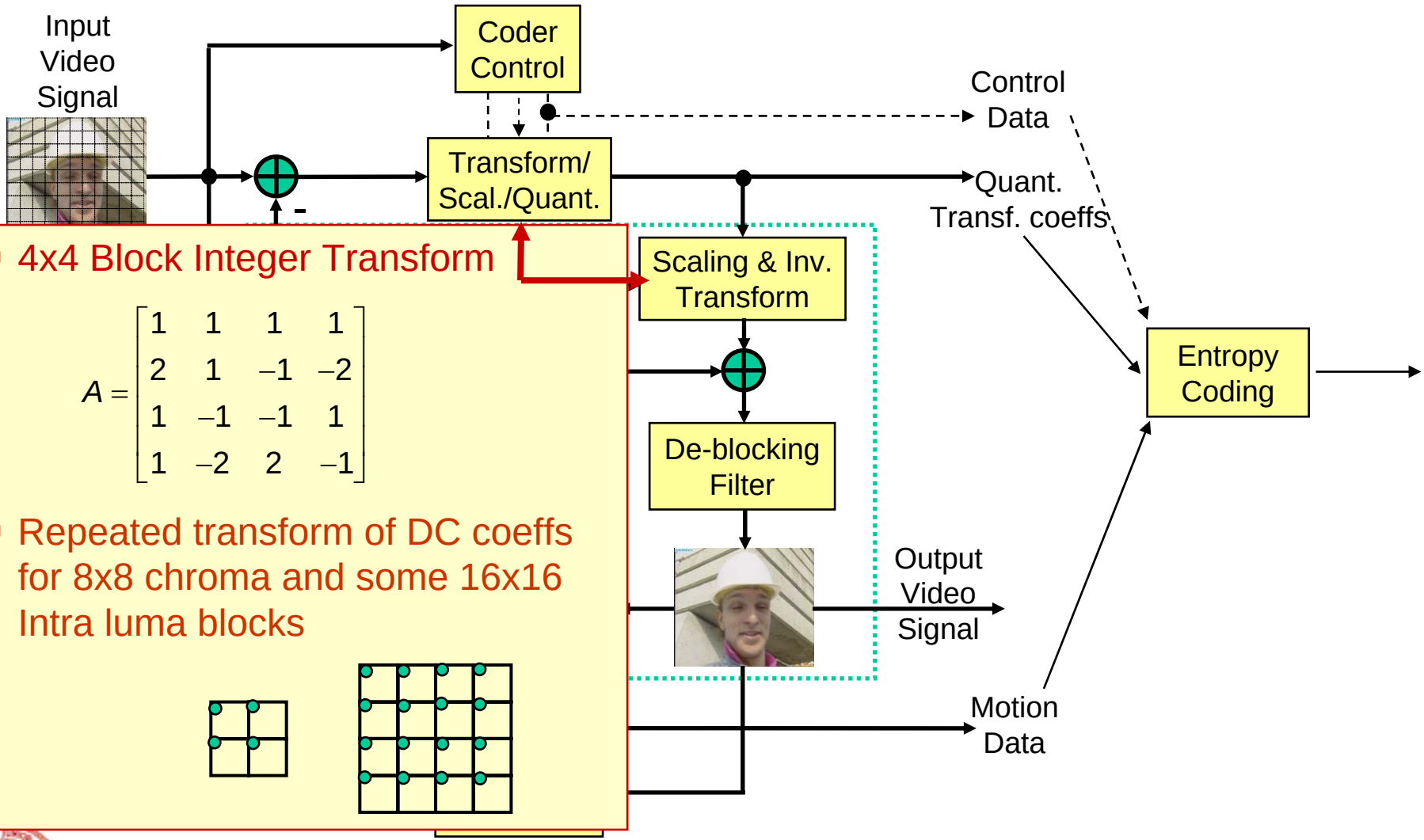
	Q	A	B	C	D	E	F	G	H
I		a	b	c	d				
J		e	f	g	h				
K		i	j	k	l				
L		m	n	o	p				



- e.g., Mode 3:
  diagonal down/right prediction
  a, f, k, p are predicted by
  (A + 2Q + I + 2) >> 2

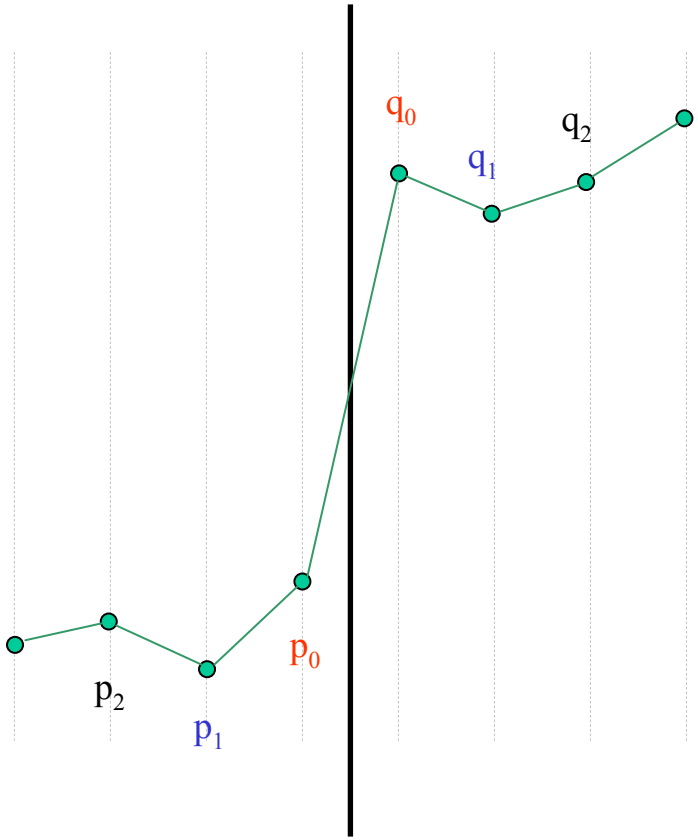
[source: G. Sullivan, VCEG]

H.264 4x4 Transform



[source: G. Sullivan, VCEG]

Deblocking Filter



4x4 Block Edge

One dimensional visualization of an edge position

Filtering of p_0 *and* q_0 only takes place if:

1. $|p_0 - q_0| < \alpha(QP)$
2. $|p_1 - p_0| < \beta(QP)$
3. $|q_1 - q_0| < \beta(QP)$

Where $\beta(QP)$ is considerably smaller than $\alpha(QP)$

Filtering of p_1 *or* q_1 takes place if additionally :

1. $|p_2 - p_0| < \beta(QP)$ *or* $|q_2 - q_0| < \beta(QP)$

(QP = quantization parameter)

[source: G. Sullivan, VCEG]

Deblocking: Subjective Result for Intra

Highly compressed first decoded intra picture at 0.28 bit/sample

Without Filter

With H264/AVC Deblocking

[source: G. Sullivan, VCEG]

Deblocking: Subjective Result for Inter

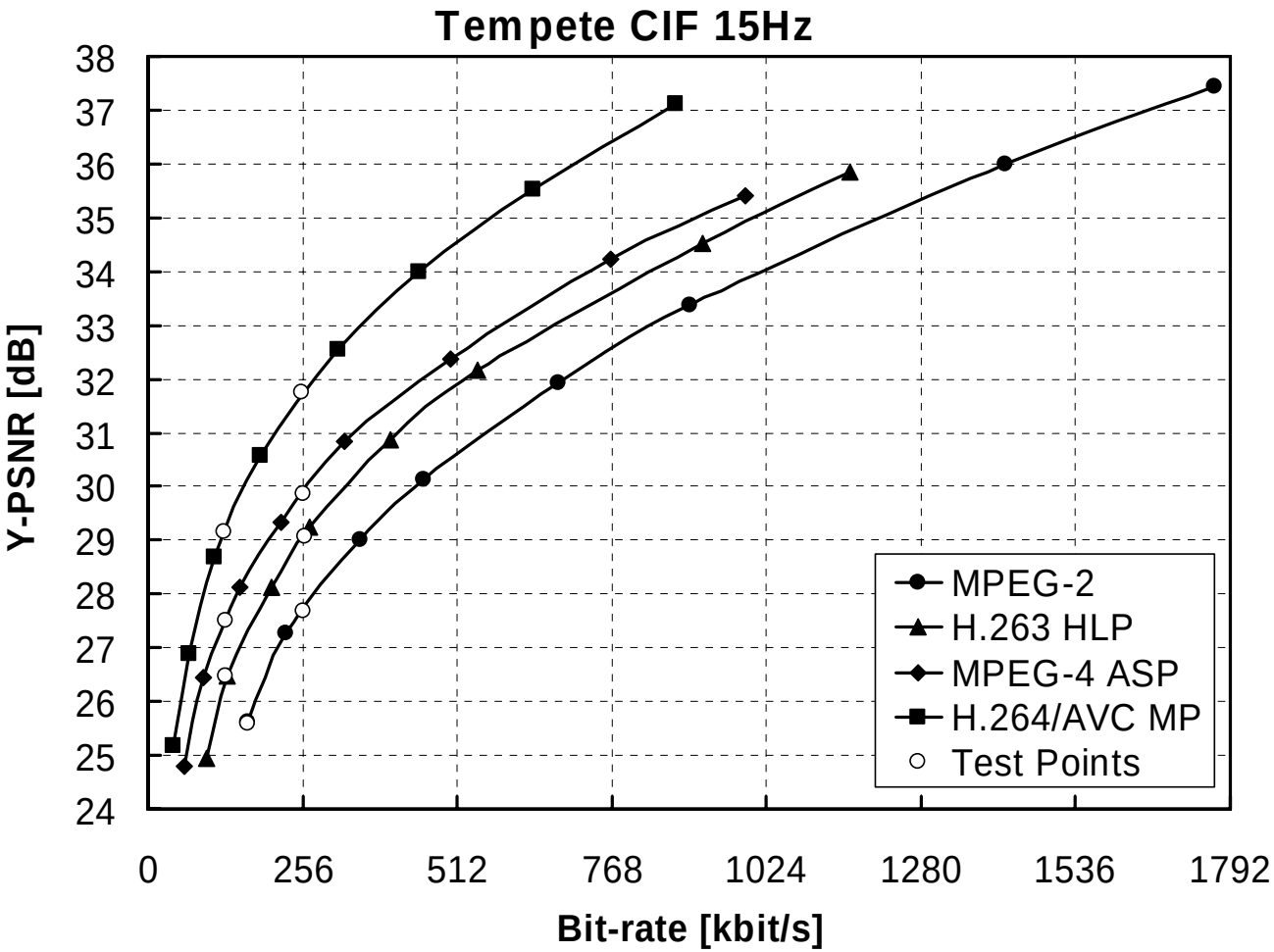
Highly compressed decoded inter picture

Without Filter

With H264/AVC Deblocking

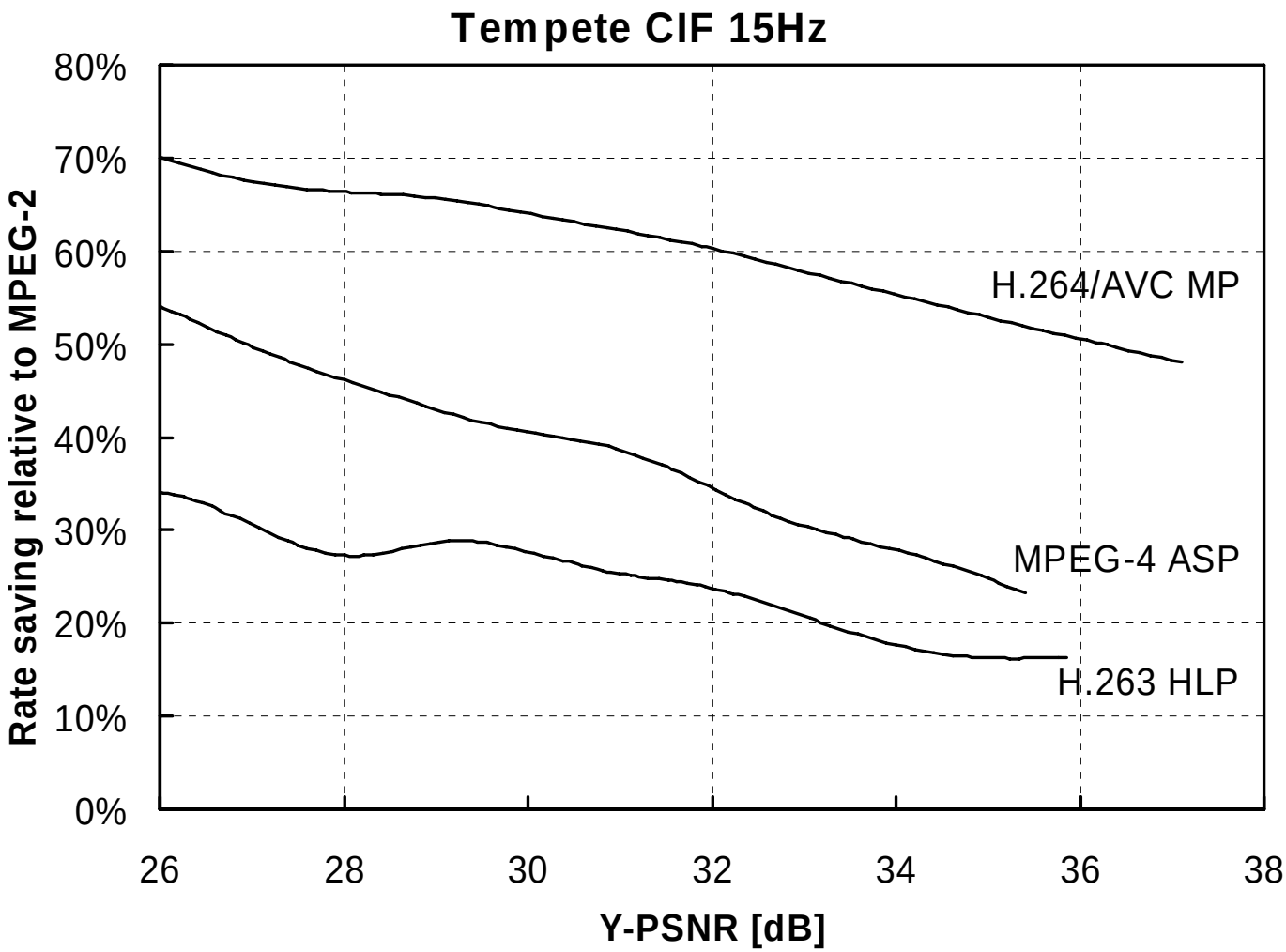
[source: G. Sullivan, VCEG]

Example Streaming Test Result



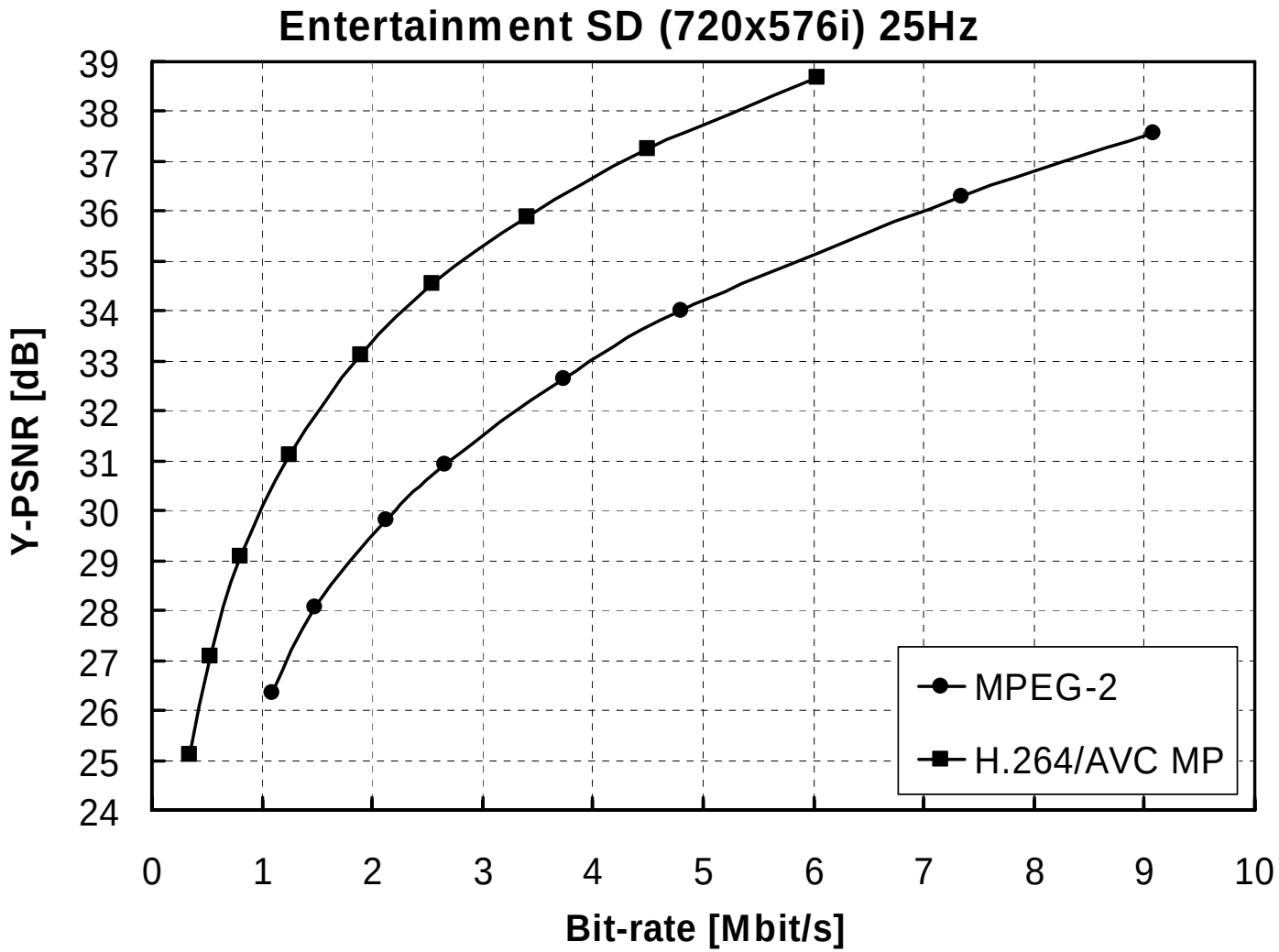
[Wiegand, et al. 2003]

Example Streaming Test Result



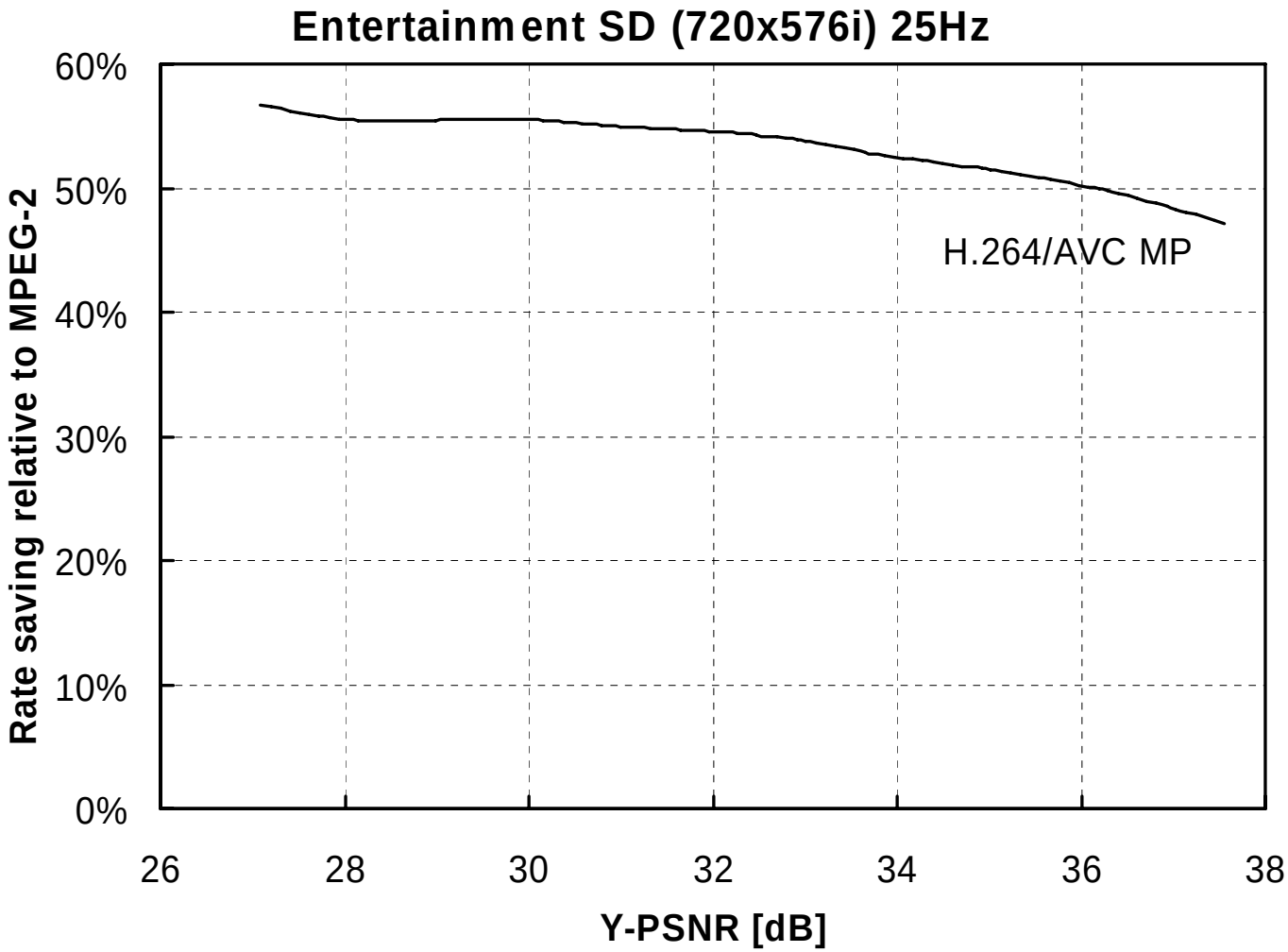
[Wiegand, et al. 2003]

Example Entertainment-Quality Applications Result



[Wiegand, et al. 2003]

Example Entertainment-Quality Applications Result



[Wiegand, et al. 2003]

Further reading

Ming Liou, "Overview of the px64 kbits/s video coding standard," Communications of the ACM, vol. 34, no. 4, pp. 59-63, April 1991.

D. LeGall, "MPEG: a video compression standard for multimedia applications," Communications of the ACM, vol. 34, no. 4, pp. 46-58, April 1991.

IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on the H.264/JVC Video Coding Standard, July 2003.