Overview: Video Coding Standards

- Video coding standards: applications and common structure
- Relevant standards organizations
- ITU-T Rec. H.261
- ITU-T Rec. H.263
- ISO/IEC MPEG-1
- ISO/IEC MPEG-2
- ISO/IEC MPEG-4
- Recent progress: H.264/AVC

The 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]

JVT Goals

- **Improved coding efficiency**
  - Average bit rate reduction of 50% given fixed fidelity compared to any other standard
  - Trade-off complexity vs. coding efficiency
- **Improved network friendliness**
  - Anticipate error-prone transport over mobile networks and the wired and wireless Internet
  - Further improve robustness techniques in H.263 and MPEG-4
- **Simple syntax specification**
  - Avoid excessive quantity of optional features
  - Minimize number of “profiles” for distinct application areas

[source: G. Sullivan, VCEG]

H.264/JVT Applications

- **Entertainment Video**
  - Broadcast: Terrestial / Satellite / Cable . . .
  - Storage: DVD / HD-DVD / PVR . . .
- **Conversational Services**
  - H.320 Conversational
  - 3GPP Conversational H.324/M
  - H.323 Conversational Internet/best effort IP/RTP
  - 3GPP Conversational IP/RTP/SIP
- **Video Streaming**
  - 3GPP Streaming IP/RTP/RTSP
  - Streaming IP/RTP/RTSP (without TCP fallback)
- **Other Applications**
  - 3GPP Multimedia Messaging Services
  - Digital camcorder

[source: G. Sullivan, VCEG]

Relationship to Other Standards

- **Identical specifications have been approved in both ITU-T / VCEG and ISO/IEC / MPEG**
- **In ITU-T / VCEG this is a new & separate standard**
  - ITU-T Recommendation H.264
  - ITU-T Systems (H.32x) will be modified to support it
- **In ISO/IEC / MPEG this is a new "part" in the MPEG-4 suite**
  - Separate codec design from prior MPEG-4 visual
  - New Part 10 called "Advanced Video Coding" (AVC – similar to "AAC" in MPEG-2 as separate audio codec)
  - MPEG-4 Systems / File Format has been modified to support it
  - **H.222.0 | MPEG-2 Systems also modified to support it**
- **IETF: RTP payload packetization**

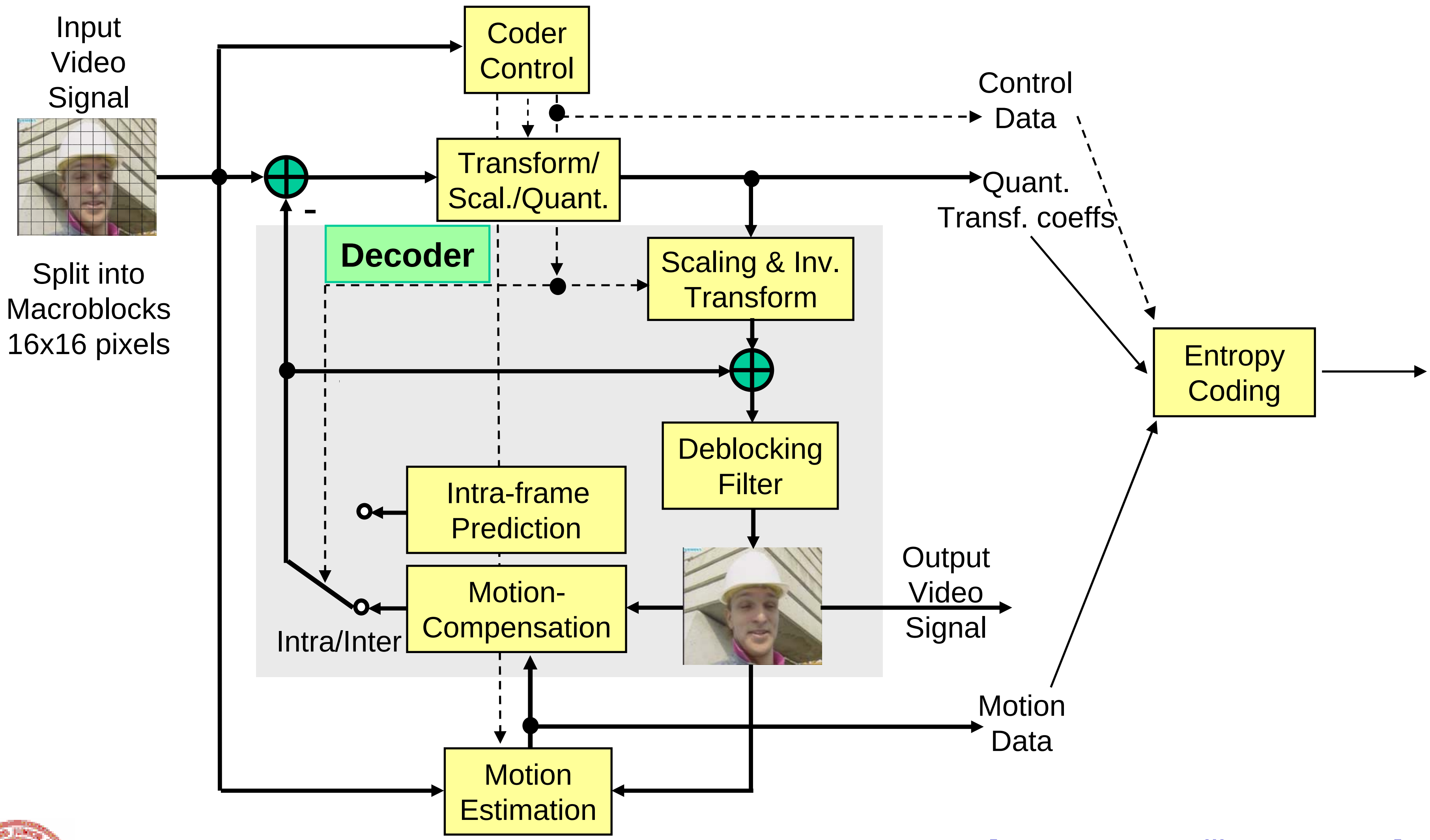
[source: G. Sullivan, VCEG]

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

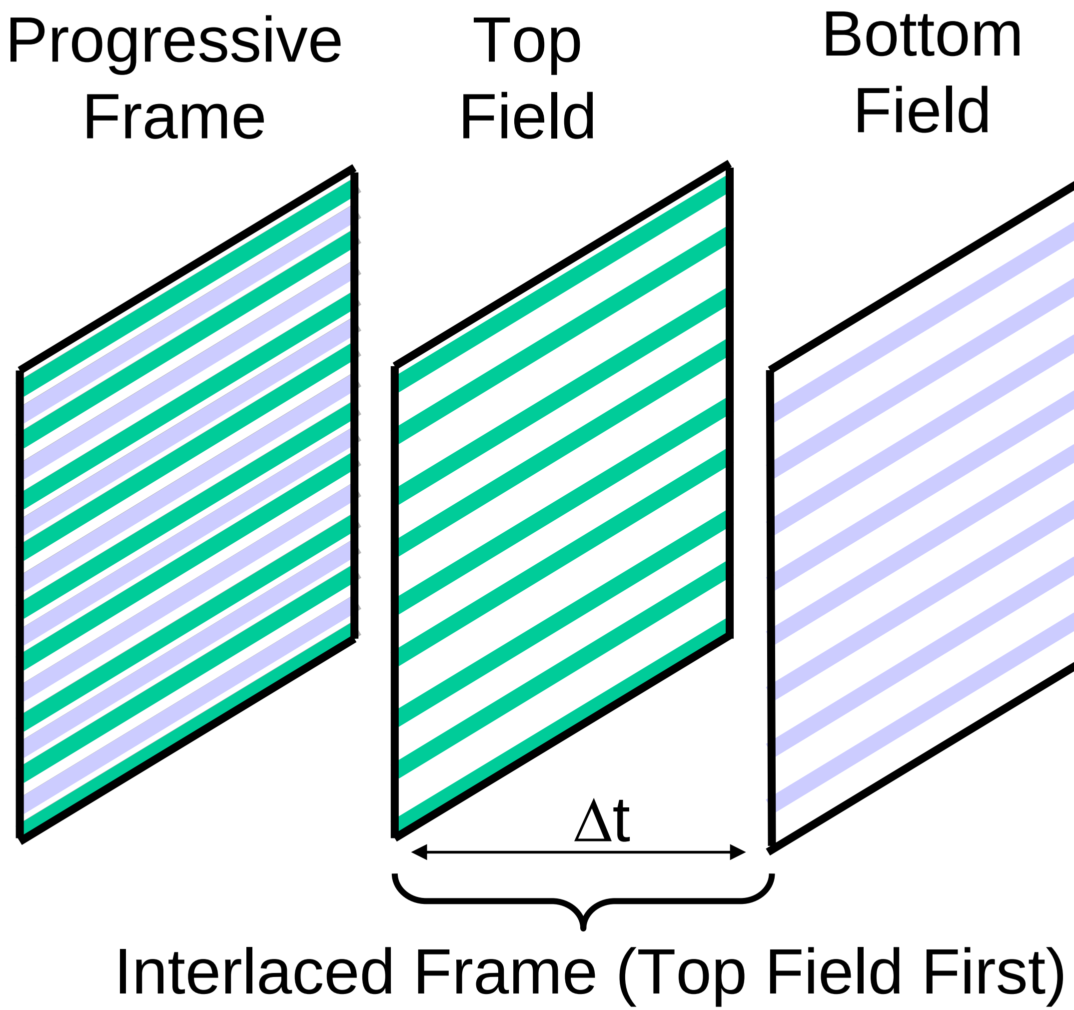
[source: G. Sullivan, VCEG]

H.264/AVC Coder



[source: G. Sullivan, VCEG]

Input Video Signal

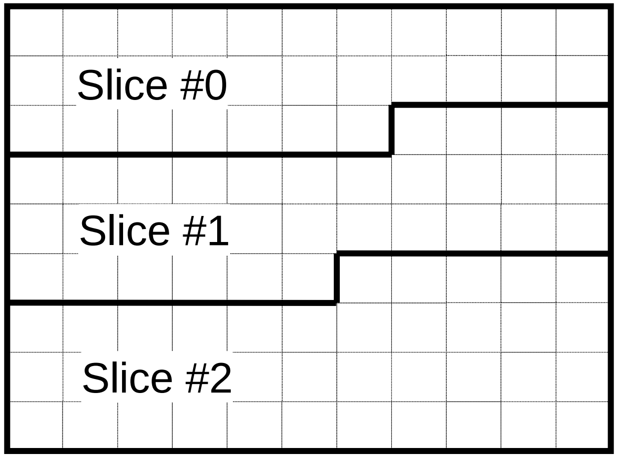


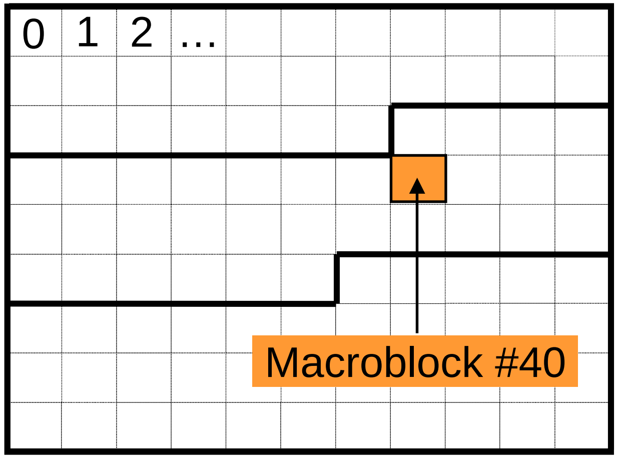
- Progressive and interlaced frames can be coded as one unit
- Progressive vs. interlace frame is signaled but has no impact on decoding
- Each field can be coded separately
- Dangling fields

[source: G. Sullivan, VCEG]

Partitioning of the Picture

- **Slices**:
  - A picture is split into 1 or several slices
  - Slices are self-contained
  - Slices are a sequence of macroblocks
- **Macroblocks**:
  - Basic syntax & processing unit
  - Contains 16x16 luma samples and 2 x 8x8 chroma samples
  - Macroblocks within a slice depend on each other
  - Macroblocks can be further partitioned

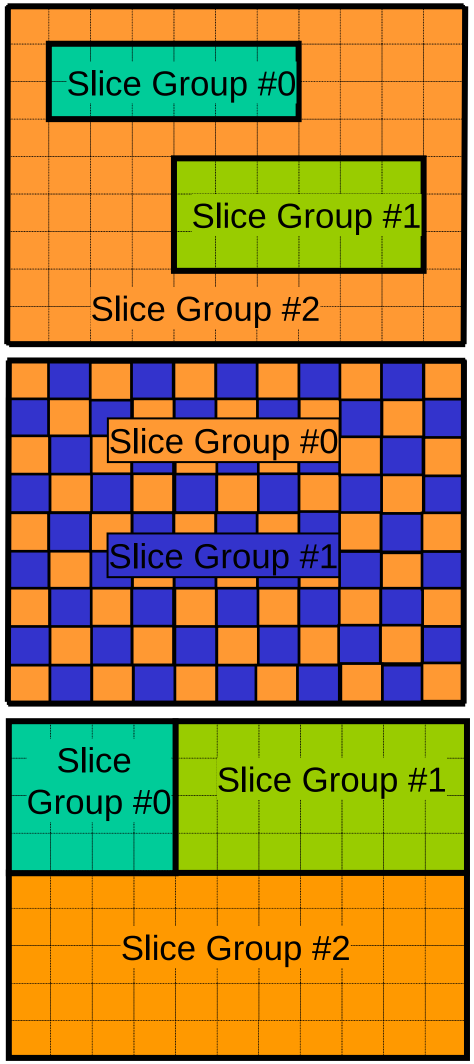




[source: G. Sullivan, VCEG]

Flexible Macroblock Ordering (FMO)

- **Slice Group**:
  - Pattern of macroblocks defined by a Macroblock allocation map
  - A slice group may contain 1 to several slices
- **Macroblock allocation map types**:
  - Interleaved slices
  - Dispersed macroblock allocation
  - Explicitly assign a slice group to each macroblock location in raster scan order
  - One or more “foreground” slice groups and a “leftover” slice group

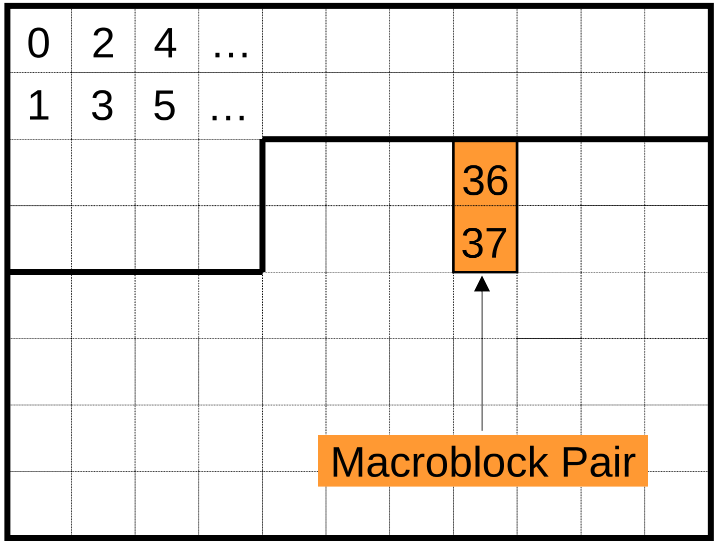


[source: G. Sullivan, VCEG]

Interlaced Processing

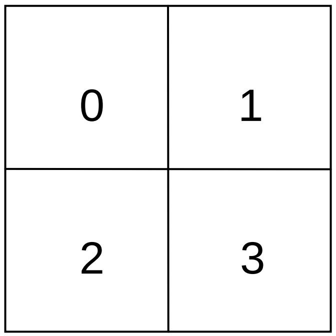
- **Field coding**: each field is coded as a separate picture using fields for motion compensation

- **Frame coding**:
  - *Type 1*: the complete frame is coded as a separate picture
  - *Type 2*: the frame is scanned as macroblock pairs, for each macroblock pair: switch between frame and field coding

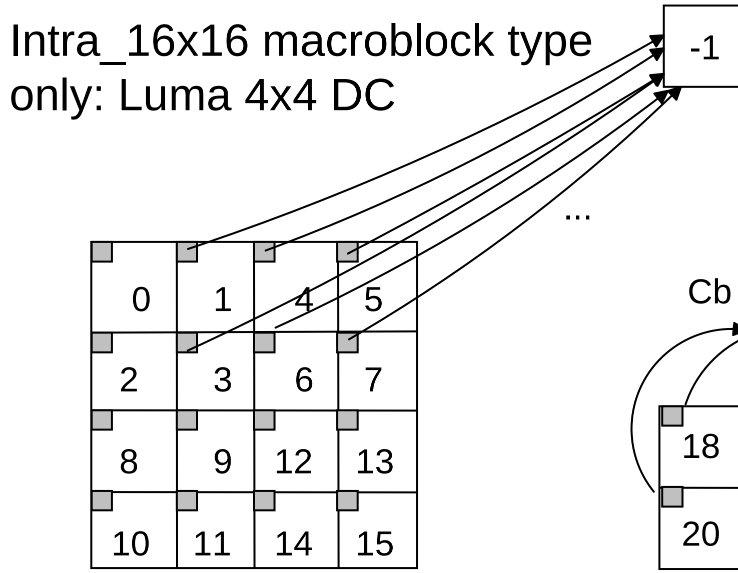


[source: G. Sullivan, VCEG]

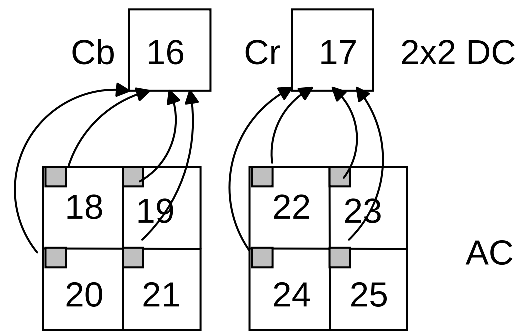
Scanning of a Macroblock



Coded Block Pattern for Luma in 8x8 block order: signals which of the 8x8 blocks contains at least one 4x4 block with non-zero transform coefficients



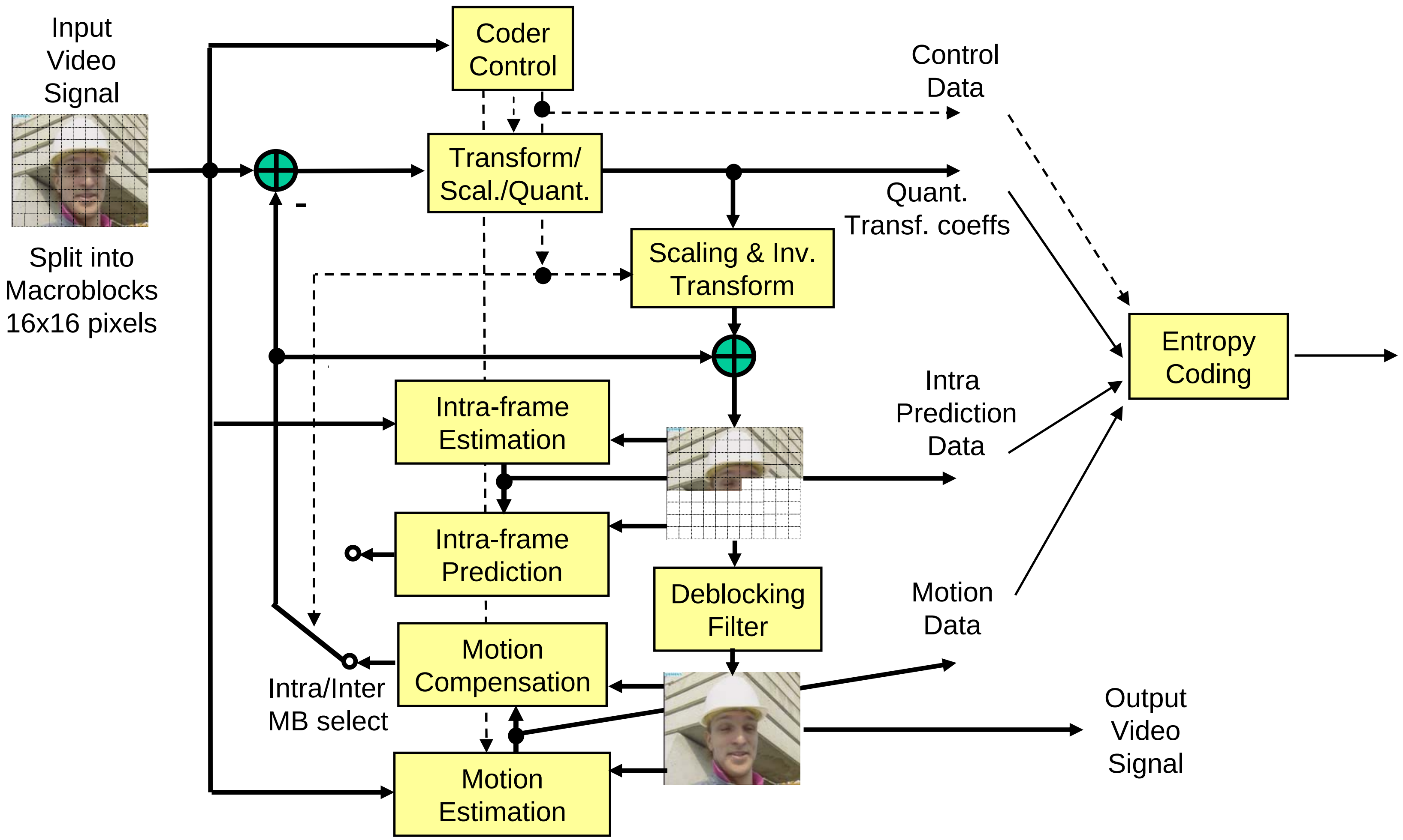
Luma 4x4 block order for 4x4 intra prediction and 4x4 residual coding



Chroma 4x4 block order for 4x4 residual coding, shown as 16-25, and intra 4x4 prediction, shown as 18-21 and 22-25

[source: G. Sullivan, VCEG]

H.264/AVC Coder



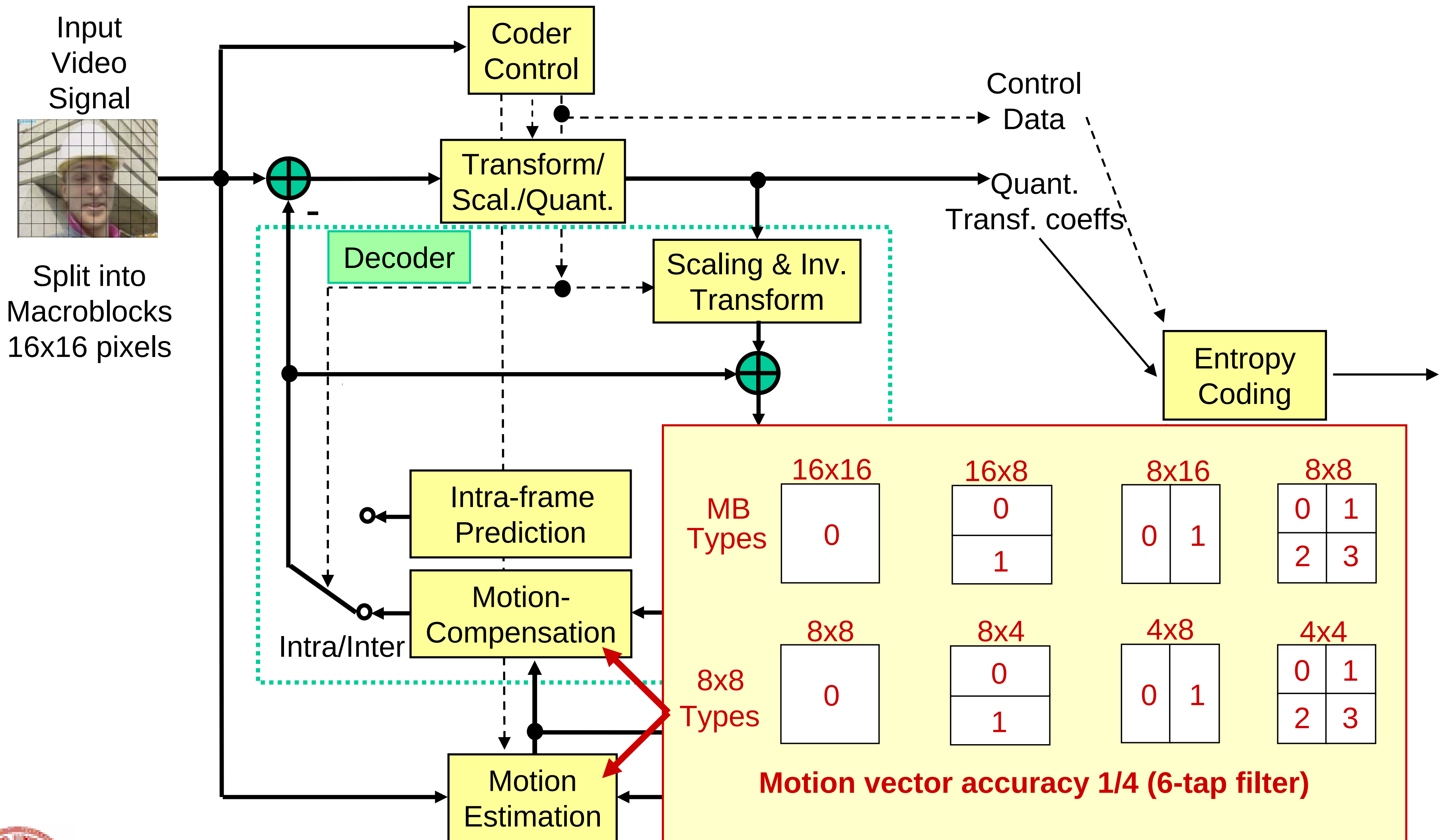
[source: G. Sullivan, VCEG]

Common Elements with other Standards

- Macroblocks: 16x16 luma + 2 x 8x8 chroma samples
- Input: Association of luma and chroma and conventional sub-sampling of chroma (4:2:0)
- Block-wise motion compensation
- Motion vectors over picture boundaries
- Variable block-size motion
- Block transforms
- Scalar quantization
- I, P, and B coding types

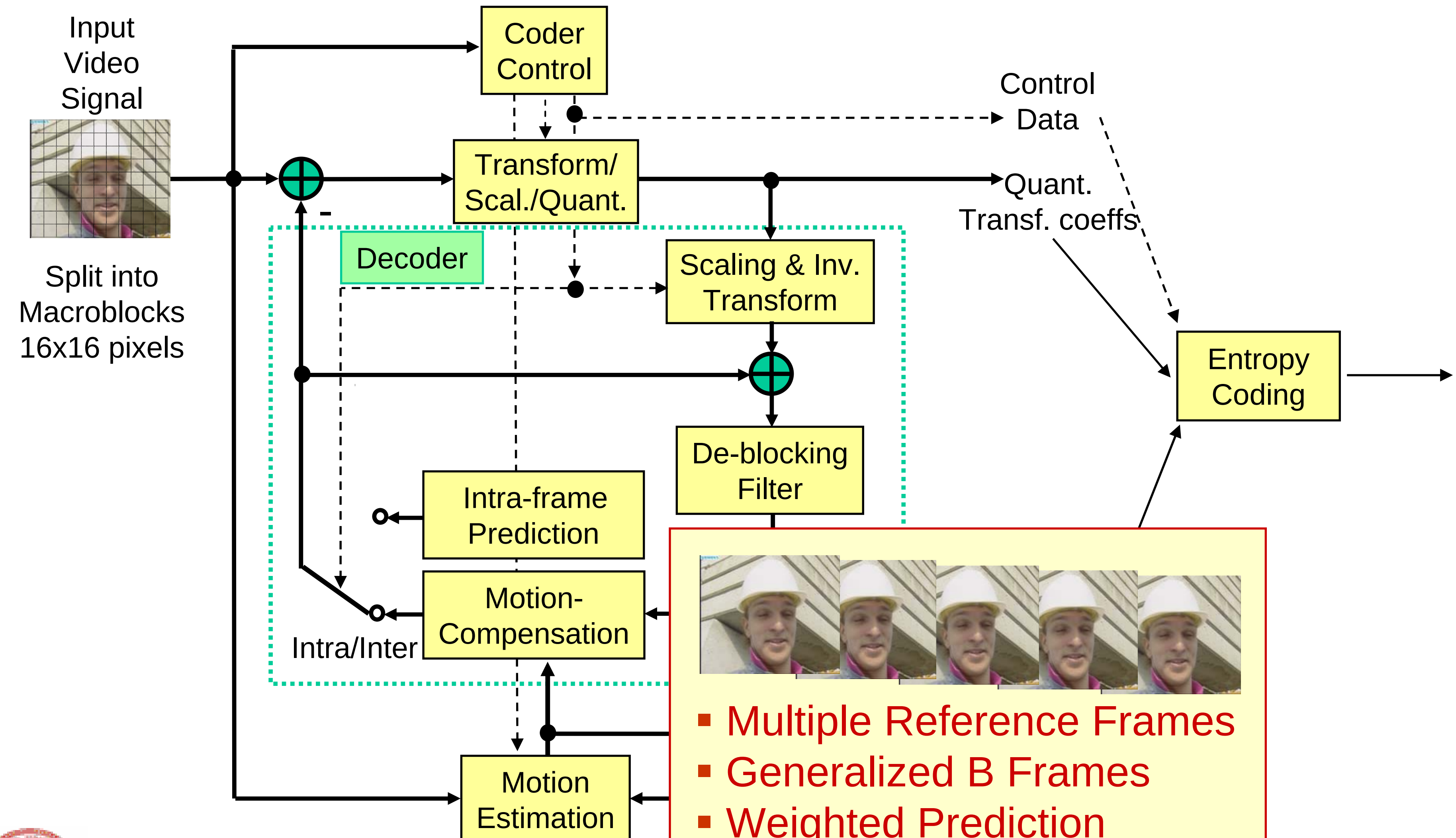
[source: G. Sullivan, VCEG]

H.264 Motion Compensation Accuracy



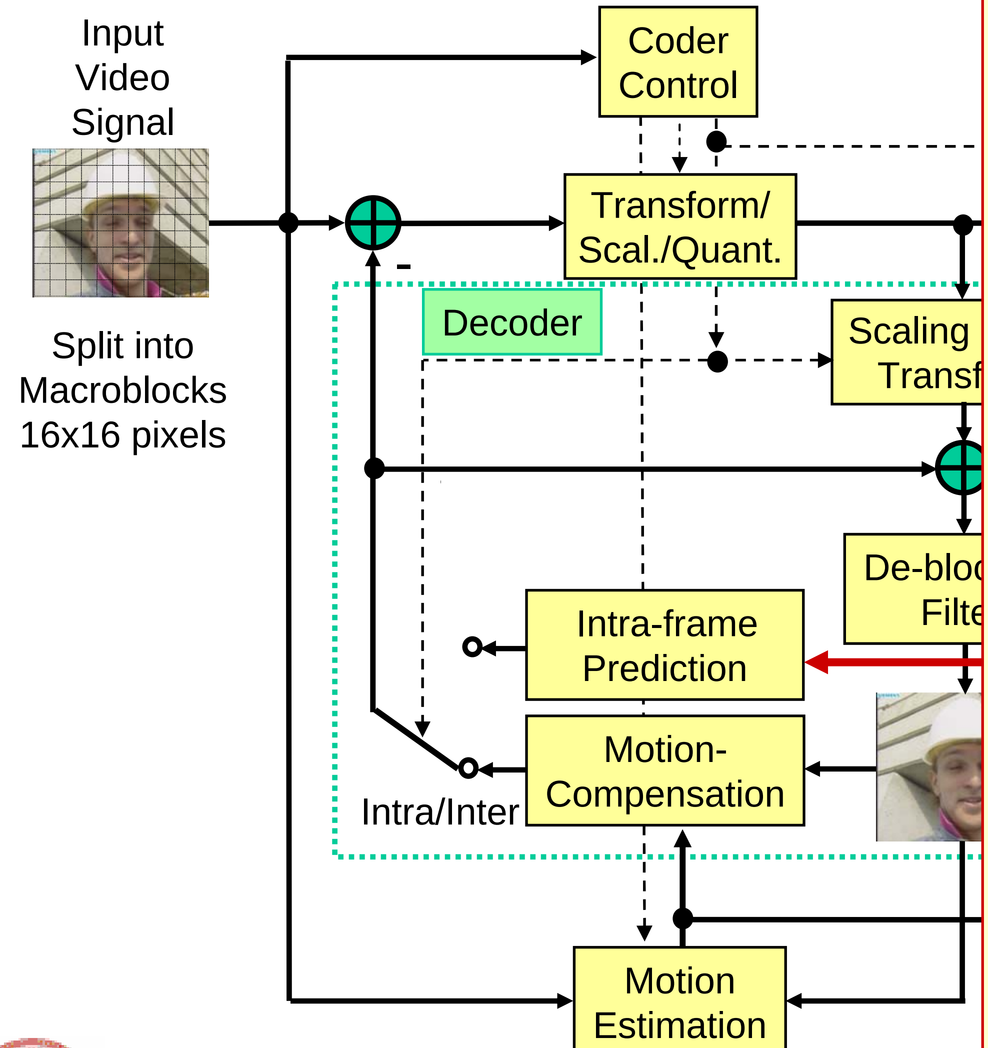
[source: G. Sullivan, VCEG]

H.264 Multiple Reference Frames



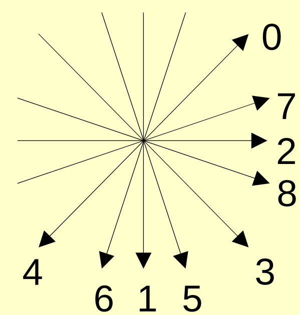
[source: G. Sullivan, VCEG]

H.264 Intra Prediction



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

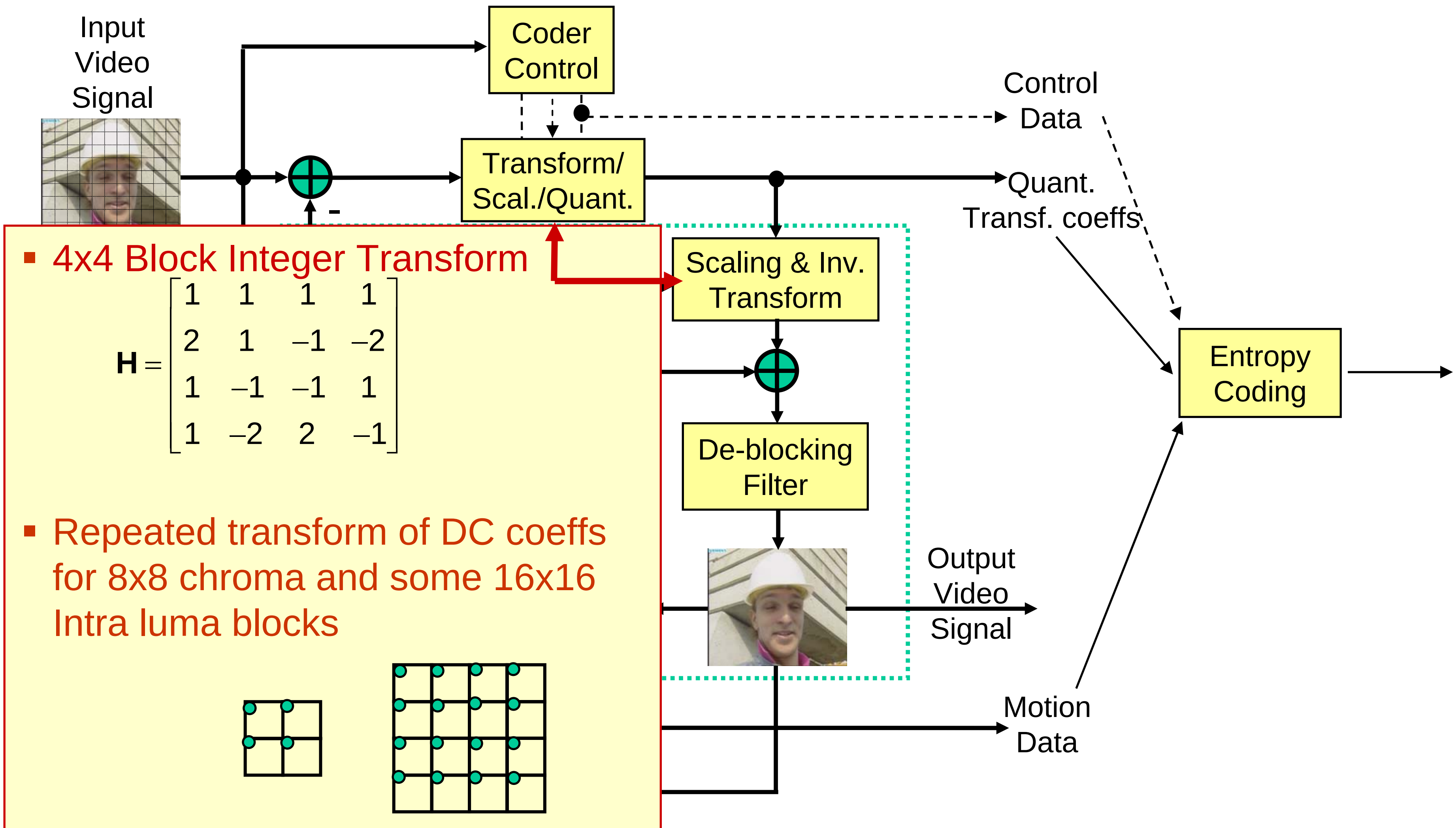
Q	A	B	C	D	E	F	G	H
I	a	b	c	d				
J	e	f	g	h				
K	i	j	k	l				
L	m	n	o	p				



- 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



- 4x4 Block Integer Transform

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

- Repeated transform of DC coeffs for 8x8 chroma and some 16x16 Intra luma blocks

[source: G. Sullivan, VCEG]

Quantization of Transform Coefficients

- Scalar quantization
- Logarithmic step size control
- Smaller step size for chroma (per H.263 Annex T)
- Extended range of step sizes
- Can change to any step size at macroblock level
- Quantization reconstruction is one multiply, one add, one shift

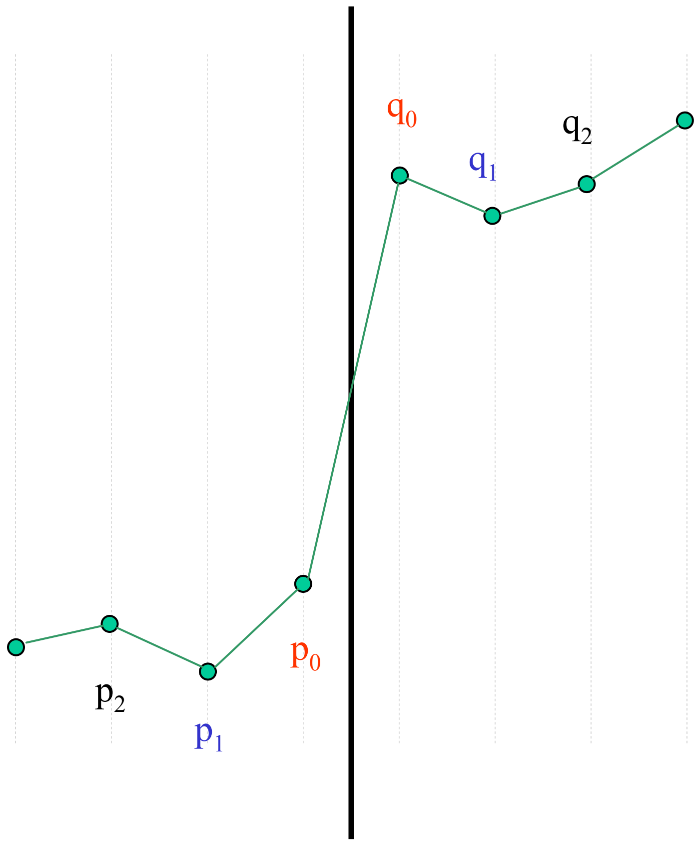
[source: G. Sullivan, VCEG]

Deblocking Filter

- Improves subjective quality *and* PSNR of the decoded picture
- Significantly superior to post filtering
- Filtering affects the edges of the 4x4 block structure
- Adaptive filtering removes blocking artifacts, but does not unnecessarily blur the visual content
  - On slice level, the global filtering strength can be adjusted to the individual characteristics of the video sequence
  - On edge level, filtering strength is made dependent on inter/intra, motion, and coded residuals
  - On sample level, quantizer dependent thresholds can turn off filtering for every individual sample
  - Specially strong filter for macroblocks with very flat characteristics almost removes "tiling artifacts"

[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

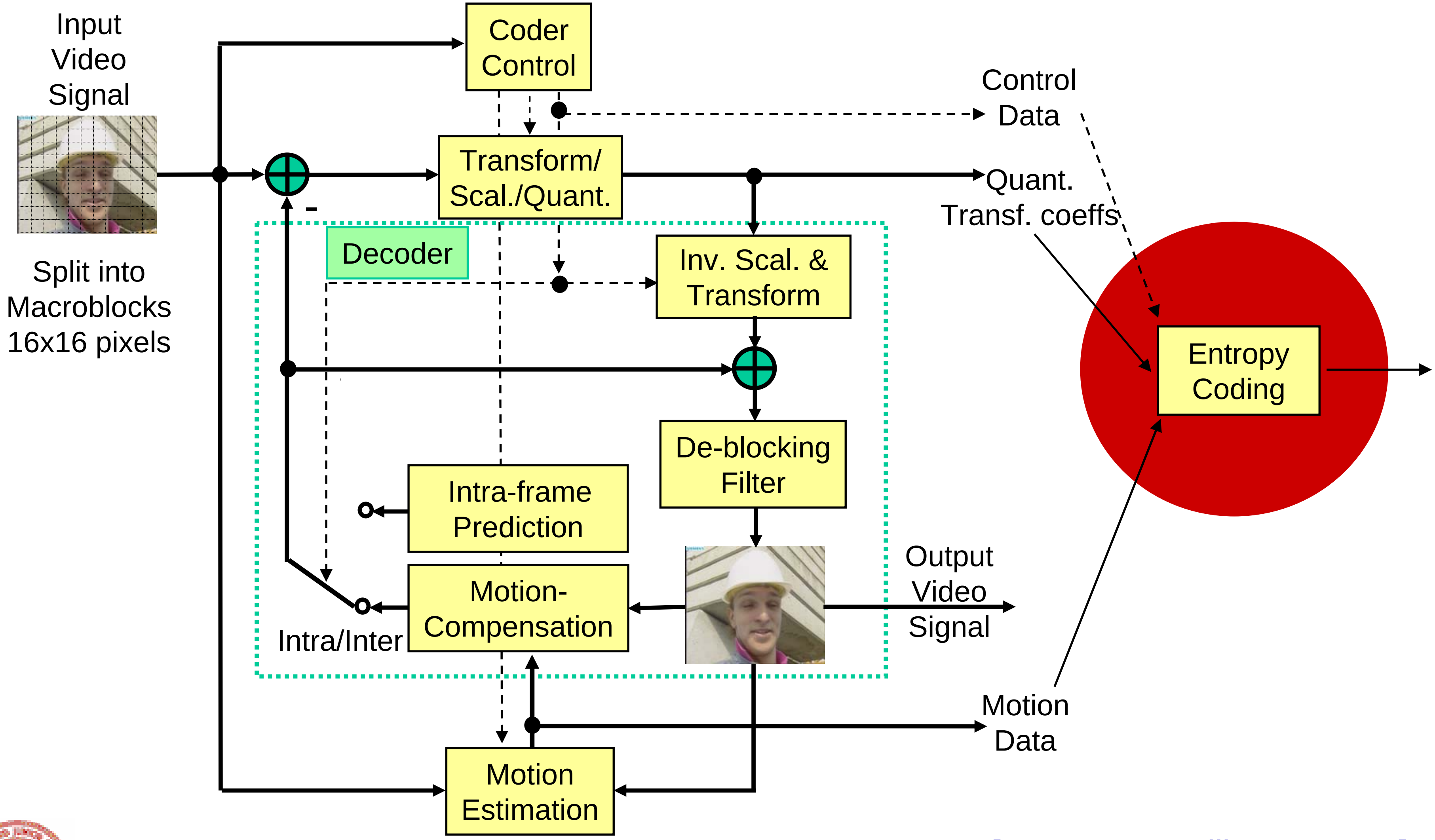
Highly compressed decoded inter picture

Without Filter

With H264/AVC Deblocking

[source: G. Sullivan, VCEG]

Entropy coding



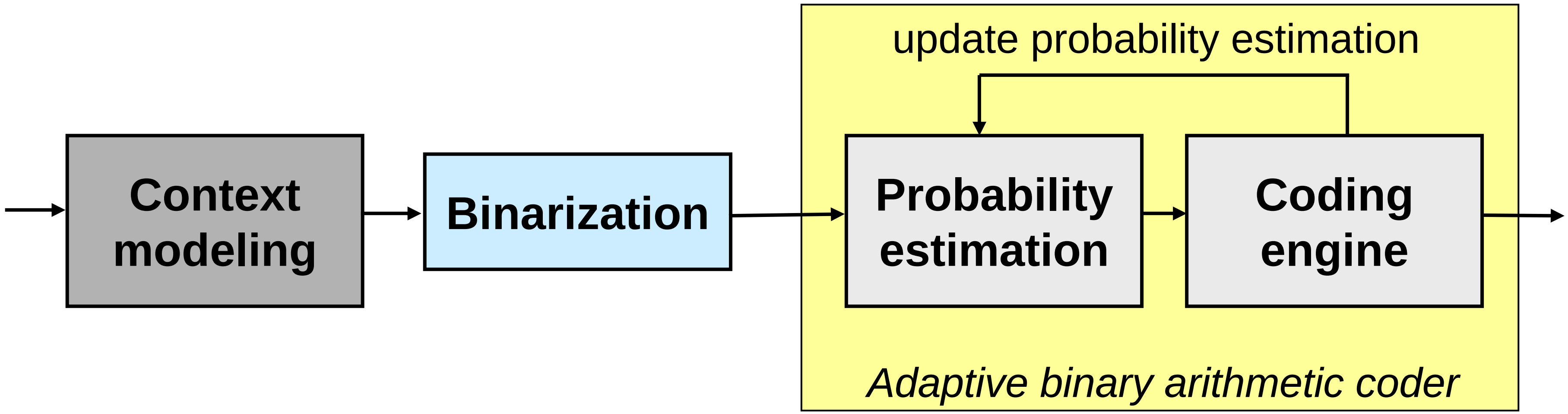
[source: G. Sullivan, VCEG]

Variable length coding

- Exp-Golomb code for almost all symbols except for transform coefficients
- Context adaptive VLCs for coding of transform coefficients
  - Number of coefficients is decoded
  - Special treatment of values +1 and -1
  - Contexts are built dependent on transform coefficients

[source: G. Sullivan, VCEG]

Context-Adaptive Arithmetic Coding (CABAC)



Chooses a model conditioned on past observations

Maps non-binary symbols to a binary sequence

Uses the provided model for the actual encoding and updates the model

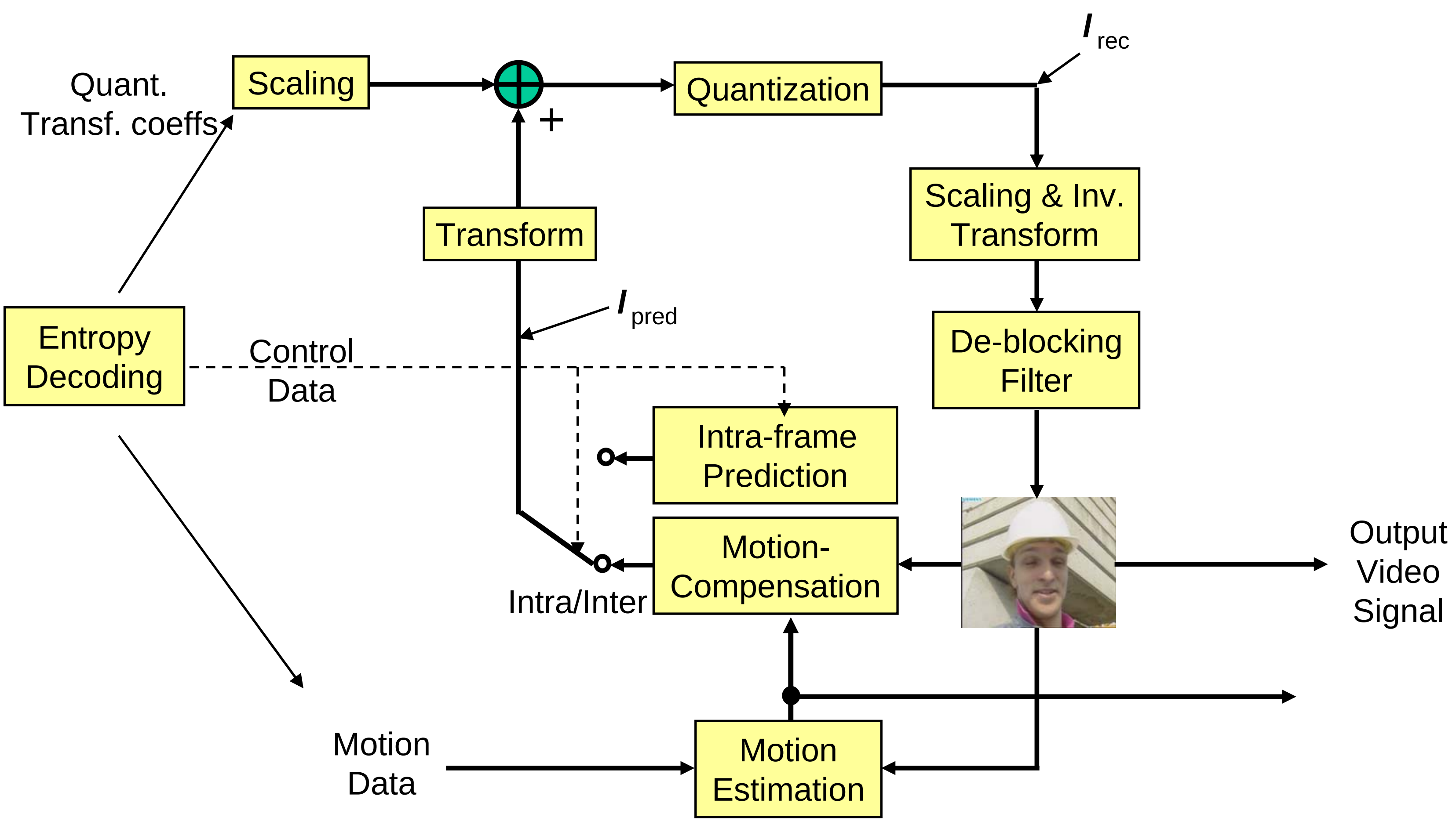
[source: G. Sullivan, VCEG]

S Pictures

- **General description**
  - Allows identical reconstruction of frames even when different reference frames are being used
  - SP pictures use of motion-compensated prediction
  - SI pictures can exactly approximate SP pictures
- **Applications**
  - Bitstream switching or splicing
  - Random access
  - Fast-forward, fast-backward
  - Error recovery and/or resiliency
  - Resynchronization such as in Video Redundancy Coding

[source: G. Sullivan, VCEG]

SP and SI Pictures



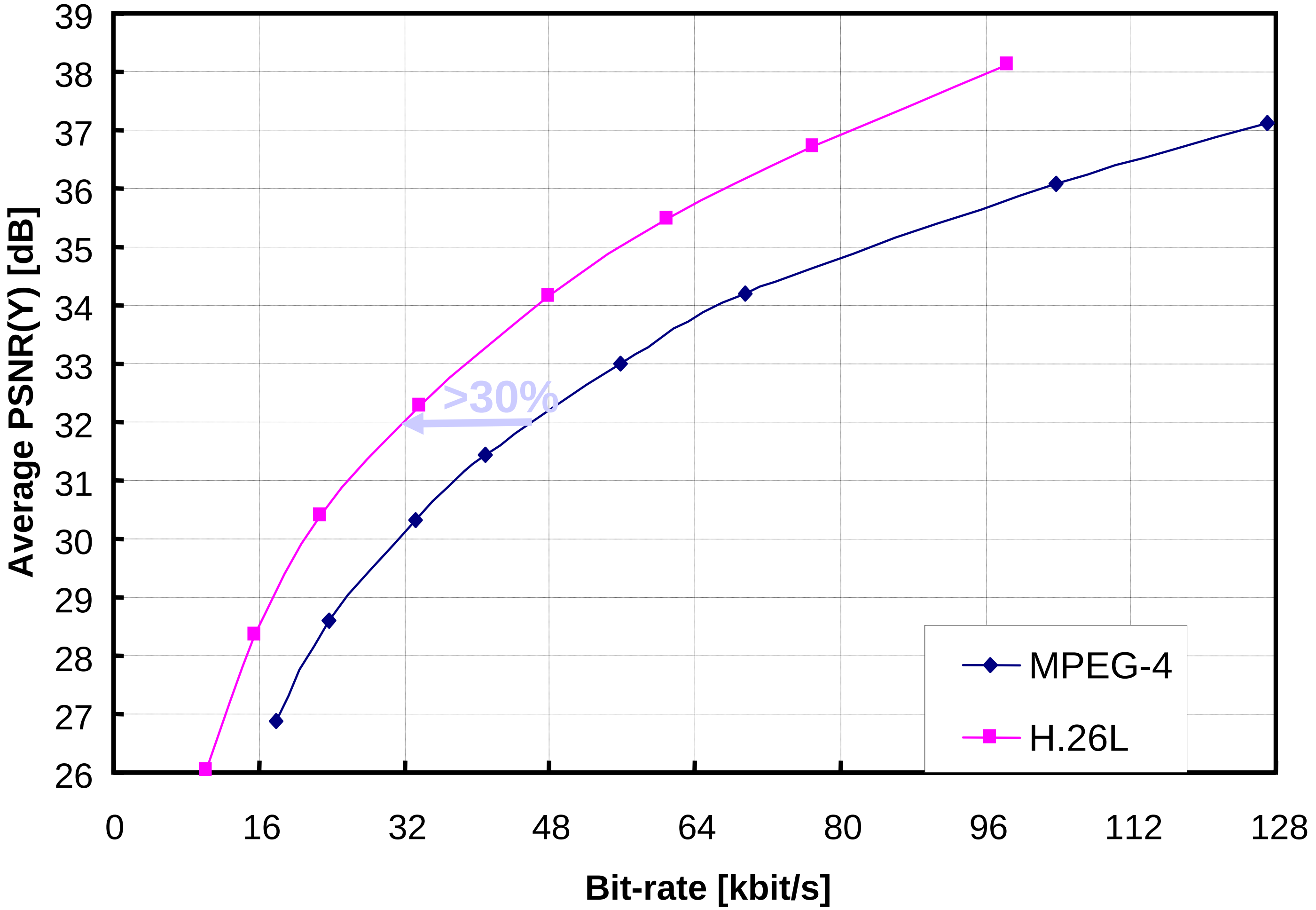
[source: G. Sullivan, VCEG]

Comparison of H.264 to MPEG-4

- MPEG-4: Advanced Simple Profile (ASP)
  - Motion Compensation: 1/4 pel
  - Global Motion Compensation
- H.264:
  - Motion Compensation: 1/4 pel
  - Using CABAC entropy coding
  - 5 reference frames (News: 17)
- Both
  - Sequence structure *IBBPBBP...*
  - $QP_B = QP_P + 2$ (step size: +25%)
  - Search range: 32x32 around 16x16 predictor
  - Lagrangian $D + \lambda R$ coder control

[source: ITU-T VCEG]

RD Curves: Foreman (QCIF, 10Hz)



[source: ITU-T VCEG]

MPEG-4 @ 32 kbit/s

H.26L @ 32 kbit/s

MPEG-4 @ 32 kbit/s

H.26L @ 32 kbit/s

MPEG-4 @ 32 kbit/s

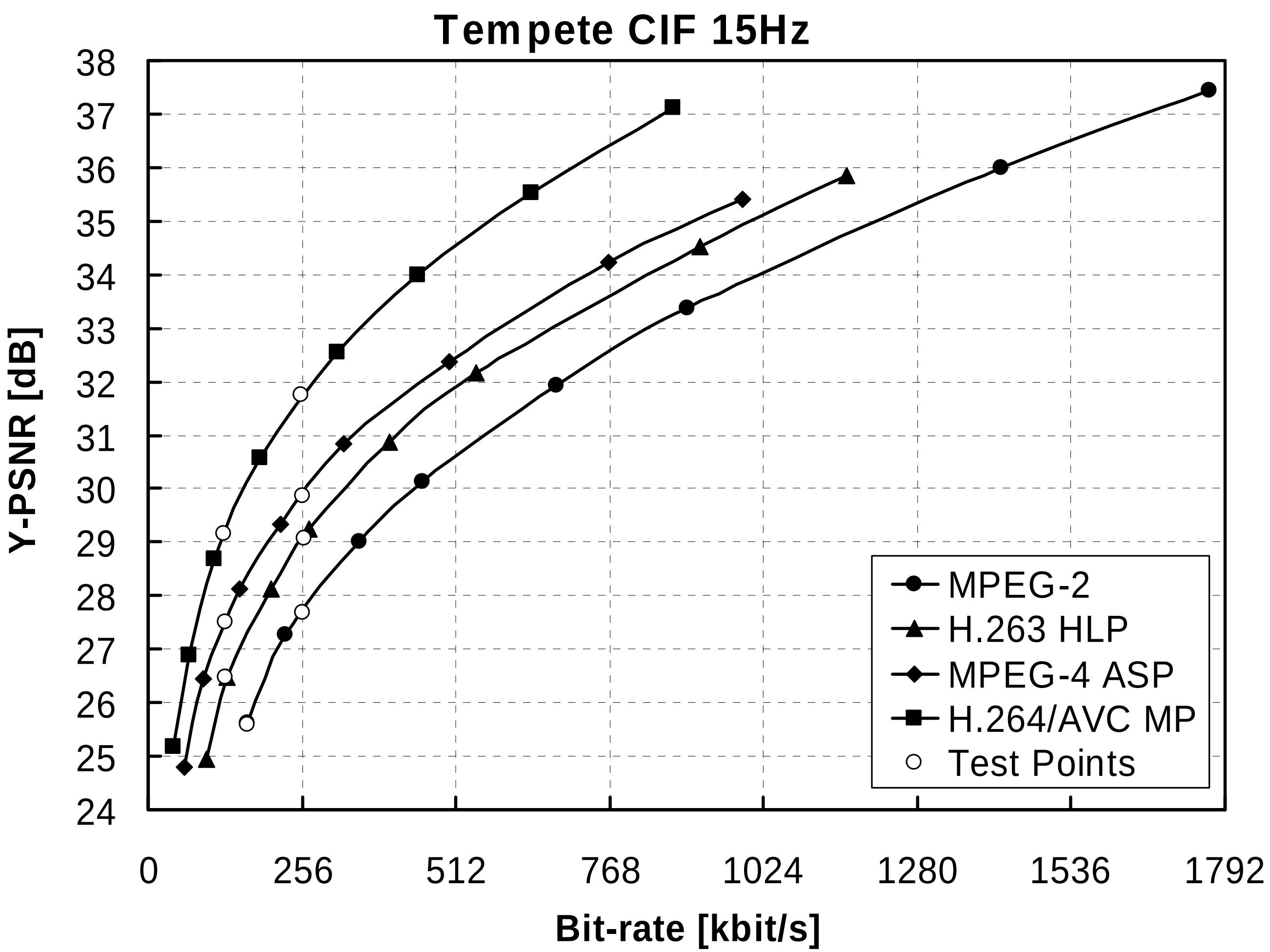
H.26L @ 32 kbit/s

Performance Streaming Application

	Average bit-rate savings relative to:		
Coder	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%

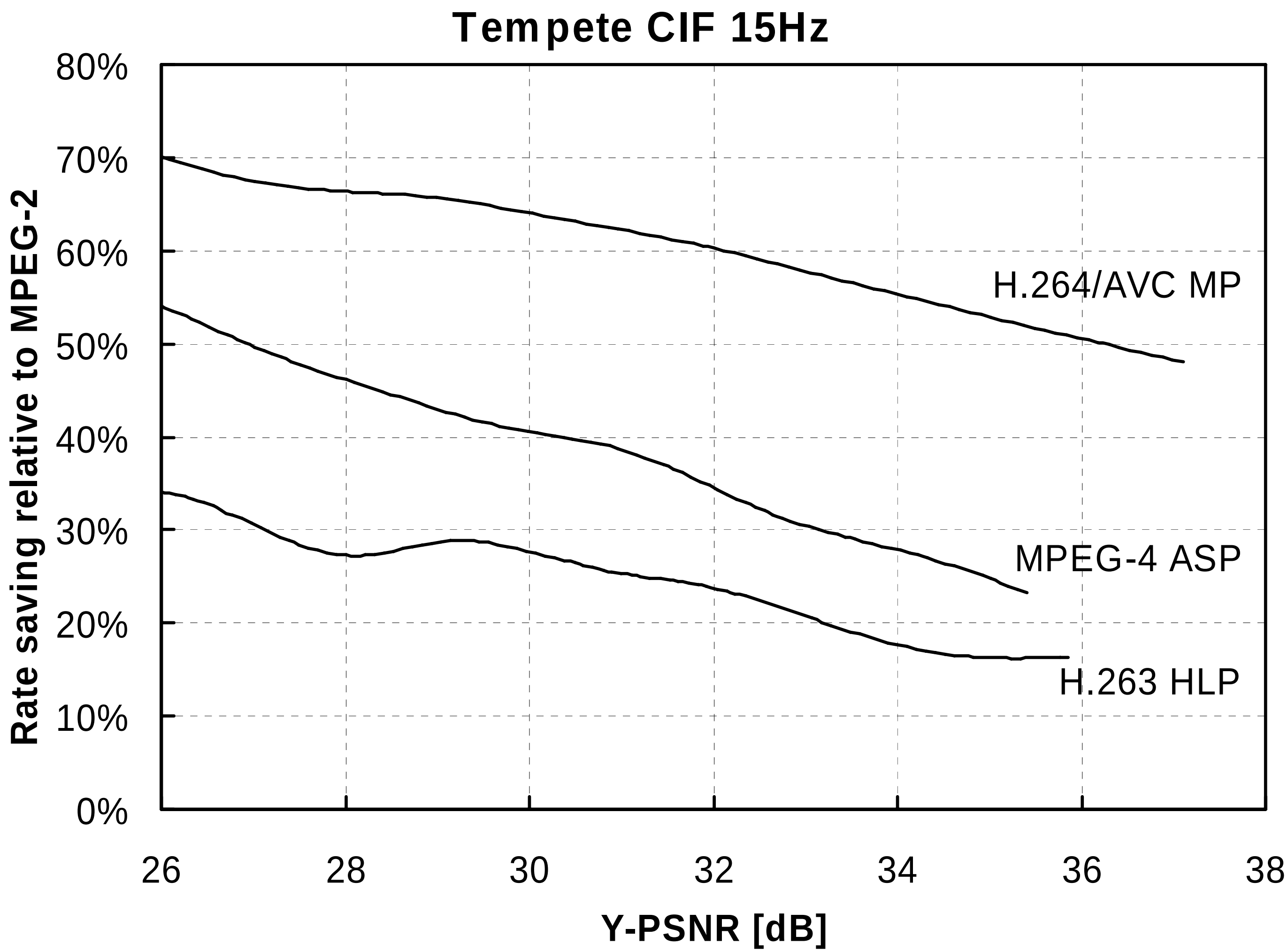
[Wiegand, et al. 2003]

Example Streaming Test Result



[Wiegand, et al. 2003]

Example Streaming Test Result



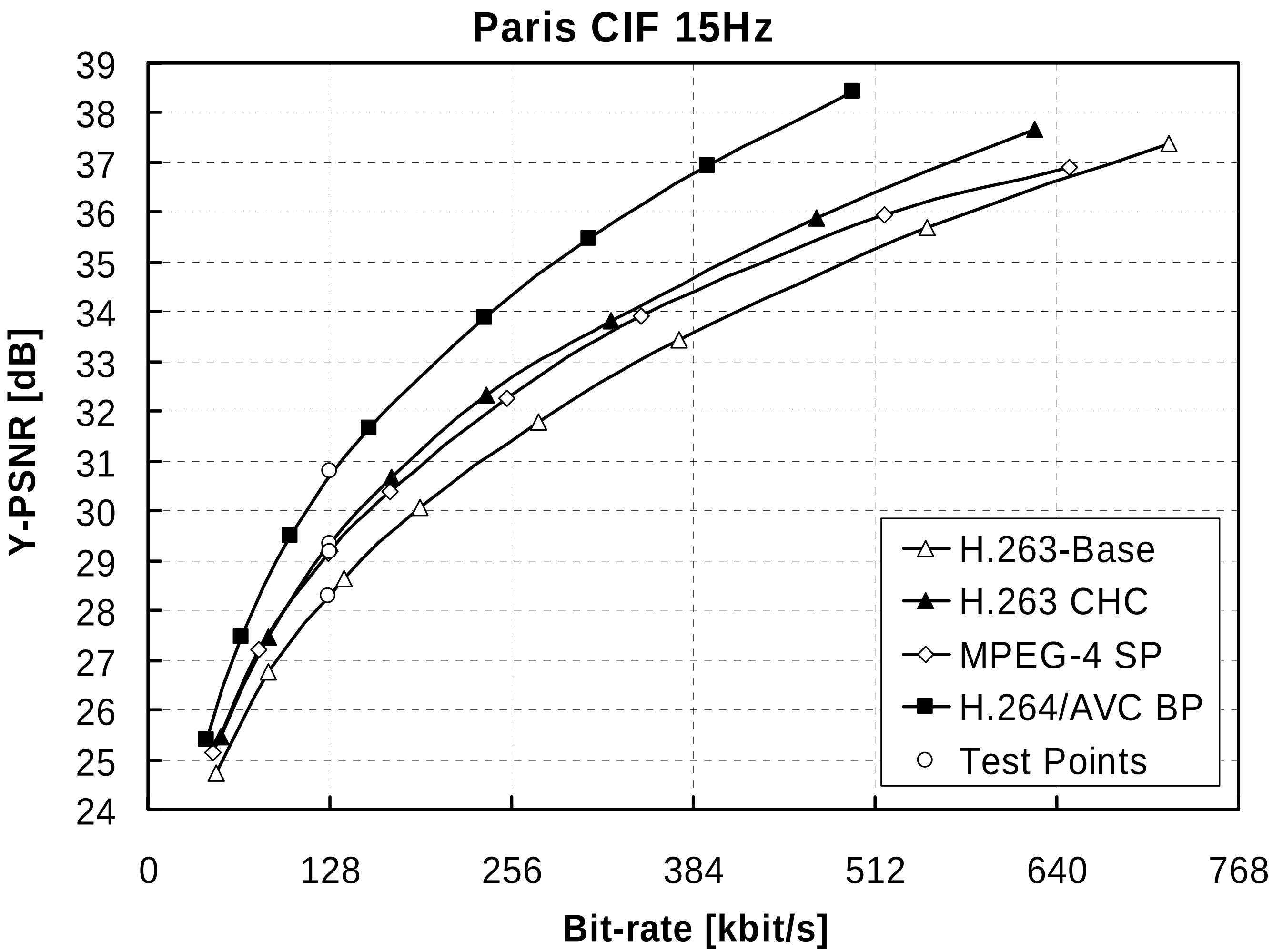
[Wiegand, et al. 2003]

Test Results for Real-Time Conversation

	Average bit-rate savings relative to:		
Coder	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%

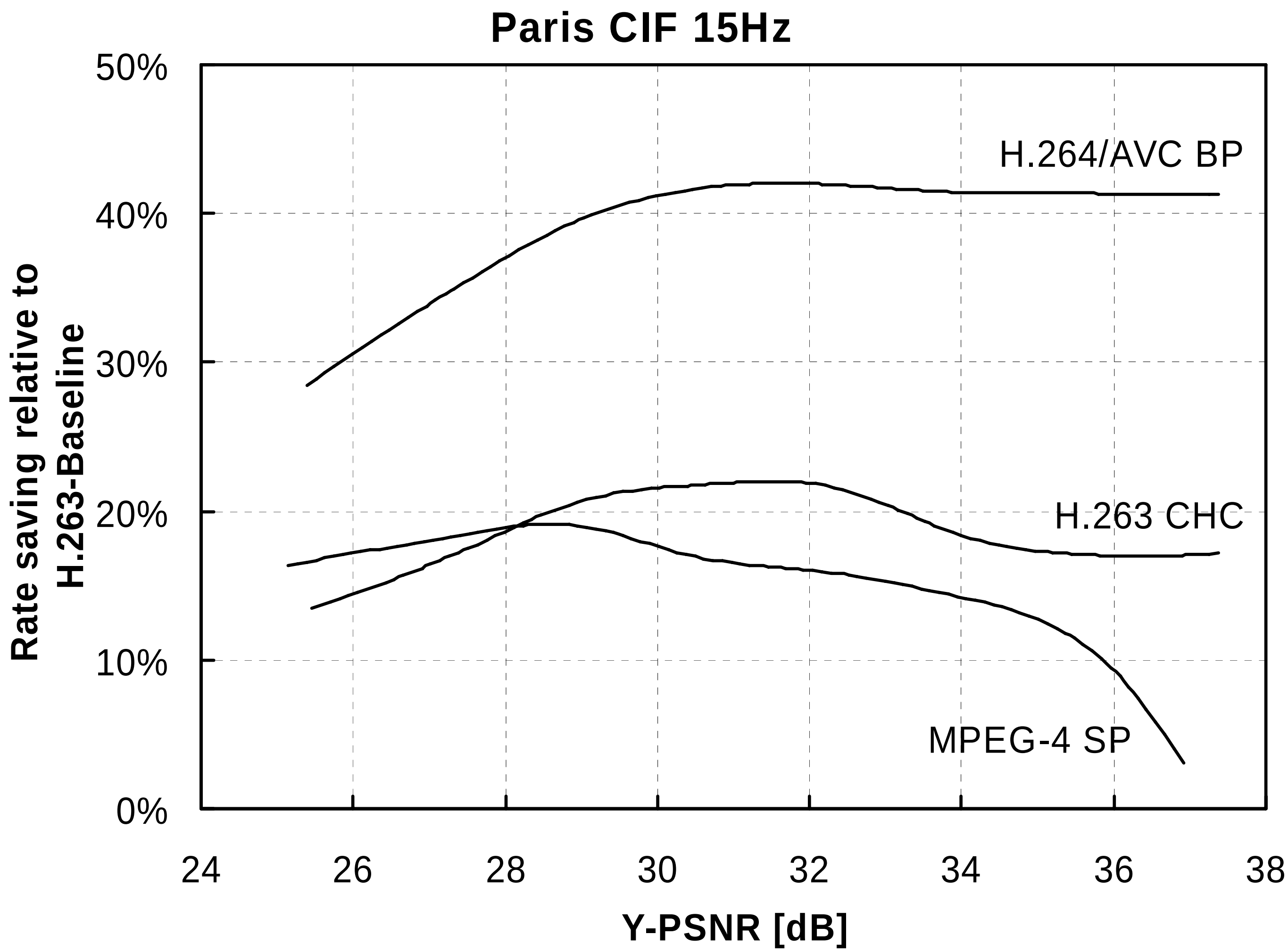
[Wiegand, et al. 2003]

Example Real-Time Conversation Result



[Wiegand, et al. 2003]

Example Real-Time Test Result



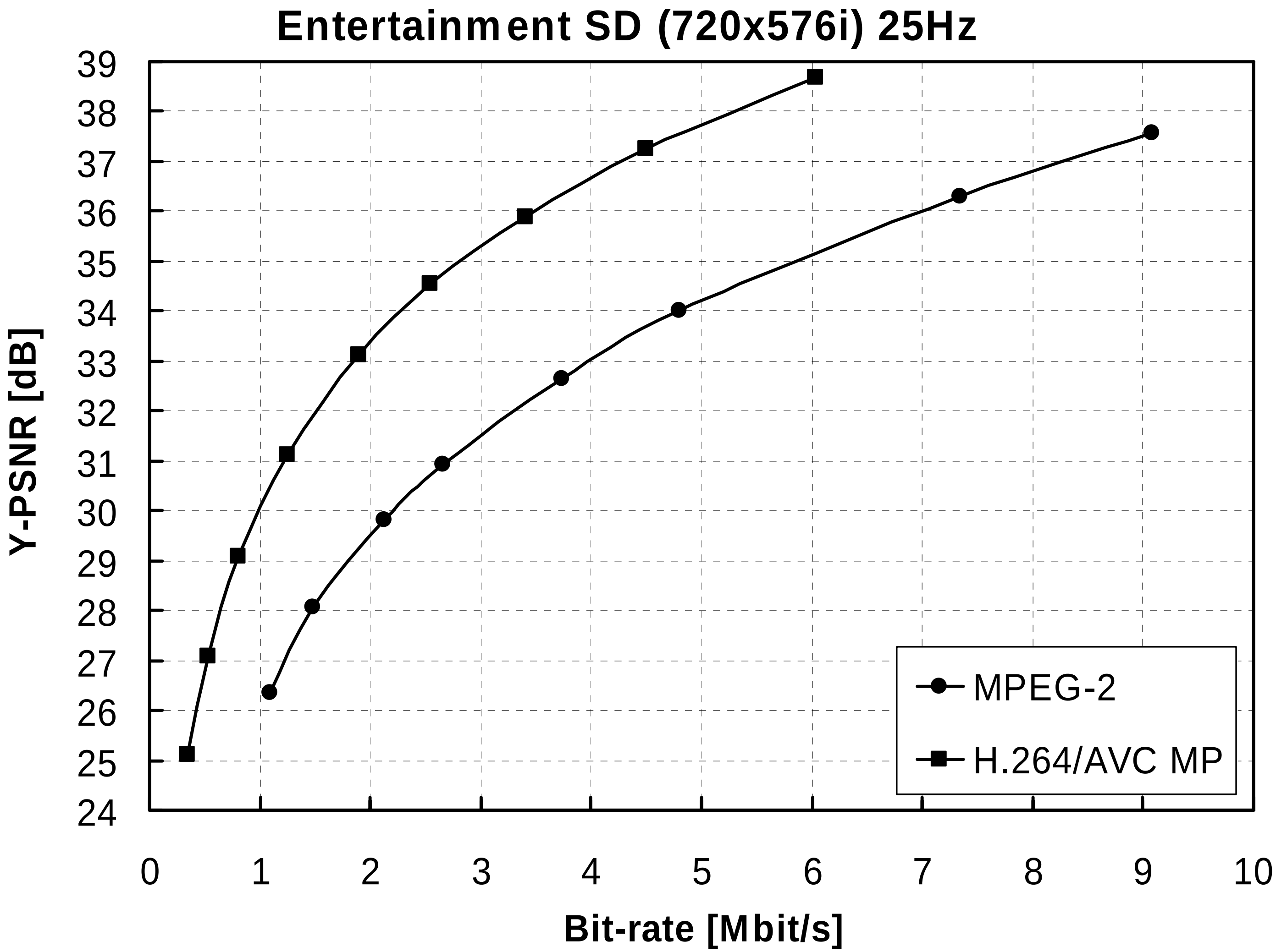
[Wiegand, et al. 2003]

Test Results Entertainment-Quality Applications

	Average bit-rate savings relative to:
Coder	MPEG-2
H.264/AVC MP	45%

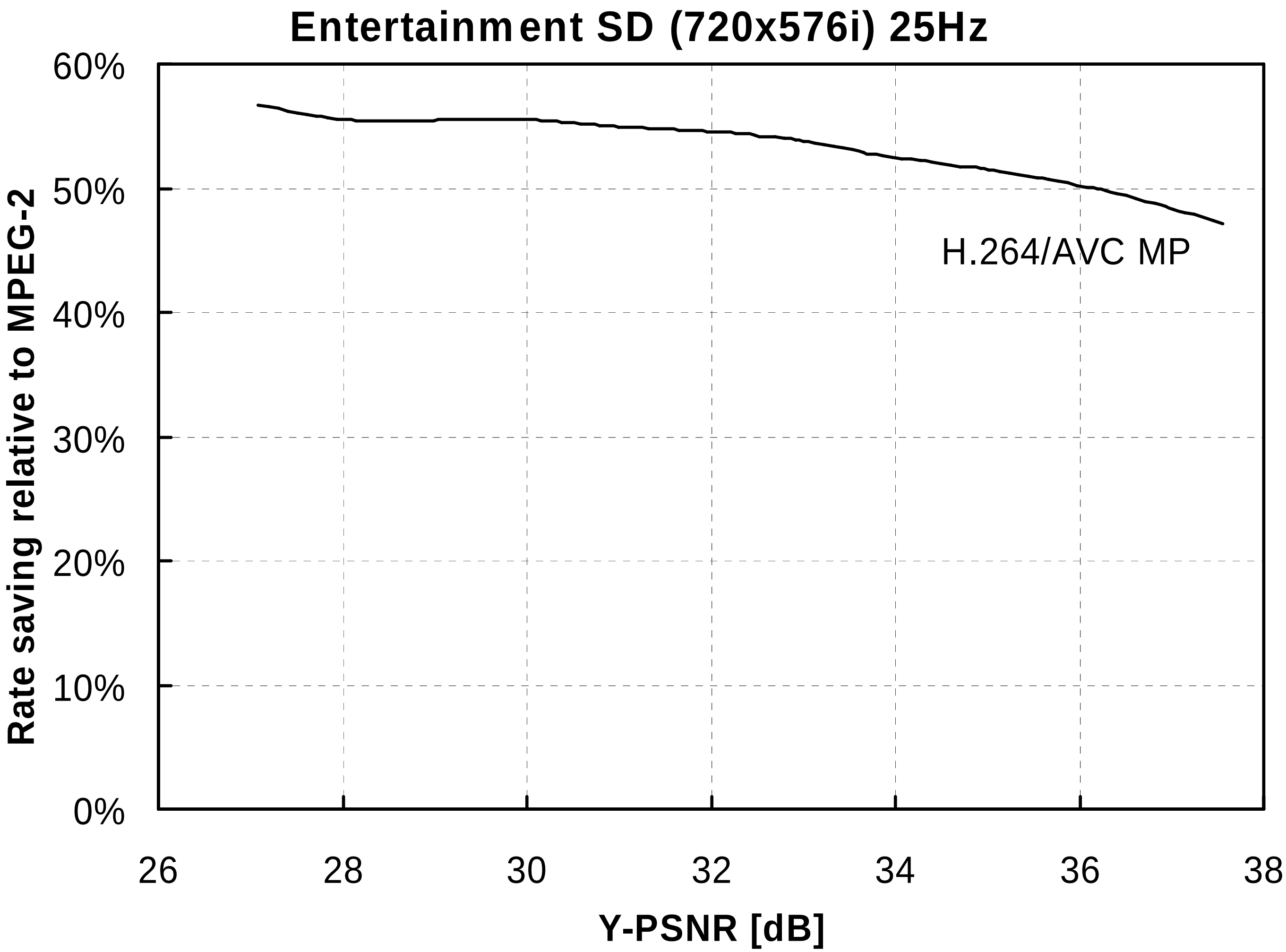
[Wiegand, et al. 2003]

Example Entertainment-Quality Applications Result



[Wiegand, et al. 2003]

Example Entertainment-Quality Applications Result



[Wiegand, et al. 2003]

Further reading

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