International Telecommunication Union

The H.264/MPEG-4 Advanced Video Coding (AVC) Standard

Gary J. Sullivan, Ph.D.

ITU-T VCEG Rapporteur | Chair
ISO/IEC MPEG Video Rapporteur | Co-Chair
ITU/ISO/IEC JVT Rapporteur | Co-Chair

July 2005

The Advanced Video Coding Project
AVC = ITU-T H.264 / MPEG-4 part 10

- History: ITU-T Q.6/SG16 (**VCEG - Video Coding Experts Group**) "H.26L" standardization activity (where the "L" stood for "long-term")
- **August 1999**: 1st test model (TML-1)
- **July 2001:** MPEG open call for technology: H.26L demo'ed
- **December 2001**: Formation of the **Joint Video Team (JVT)** between VCEG and MPEG to finalize H.26L as a new joint project (similar to MPEG-2/H.262)
- **July 2002**: Final Committee Draft status in MPEG
- **Dec '02** technical freeze, FCD ballot approved
- **May '03** completed in both orgs
- July '04 Fidelity Range Extensions (FRExt) completed
- January '05 Scalable Video Coding launched

AVC Objectives

- **Primary technical objectives:**
  - Significant improvement in coding efficiency
  - High loss/error robustness
  - “Network Friendliness” (carry it well on MPEG-2 or RTP or H.32x or in MPEG-4 file format or MPEG-4 systems or …)
  - Low latency capability (better quality for higher latency)
  - Exact match decoding
- **Additional version 2 objectives (in FRExt):**
  - Professional applications (more than 8 bits per sample, 4:4:4 color sampling, etc.)
  - Higher-quality high-resolution video
  - Alpha plane support (a degree of “object” functionality)

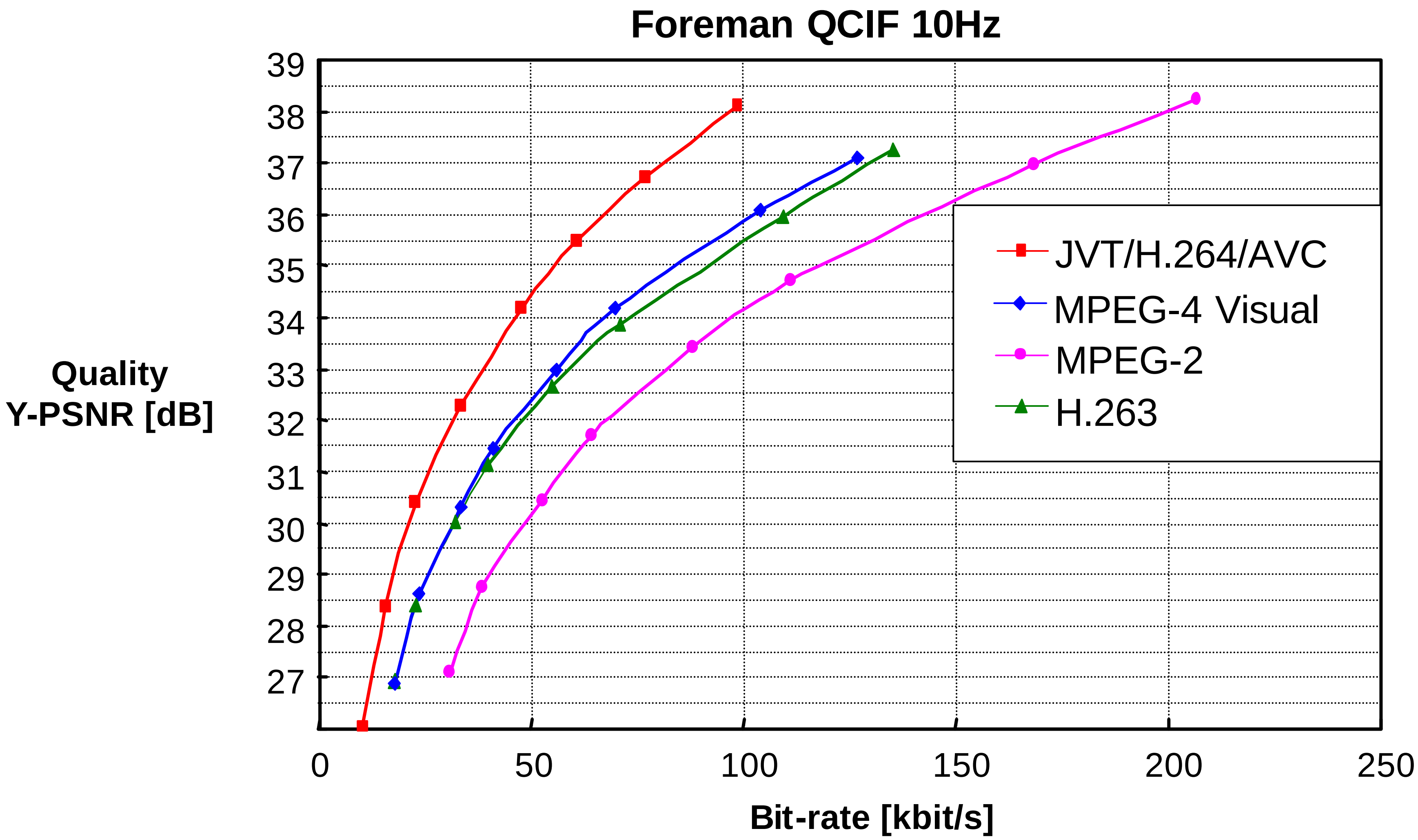
Relating to Other ITU & MPEG Standards

- **Same design to be 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) 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" position in MPEG-2 as separate codec)
  - Not backward or forward compatible with prior standards (incl. the prior MPEG-4 visual spec. – core technology is different)
  - MPEG-4 Systems / File Format supports it
- **H.222.0 | MPEG-2 Systems also supports it**

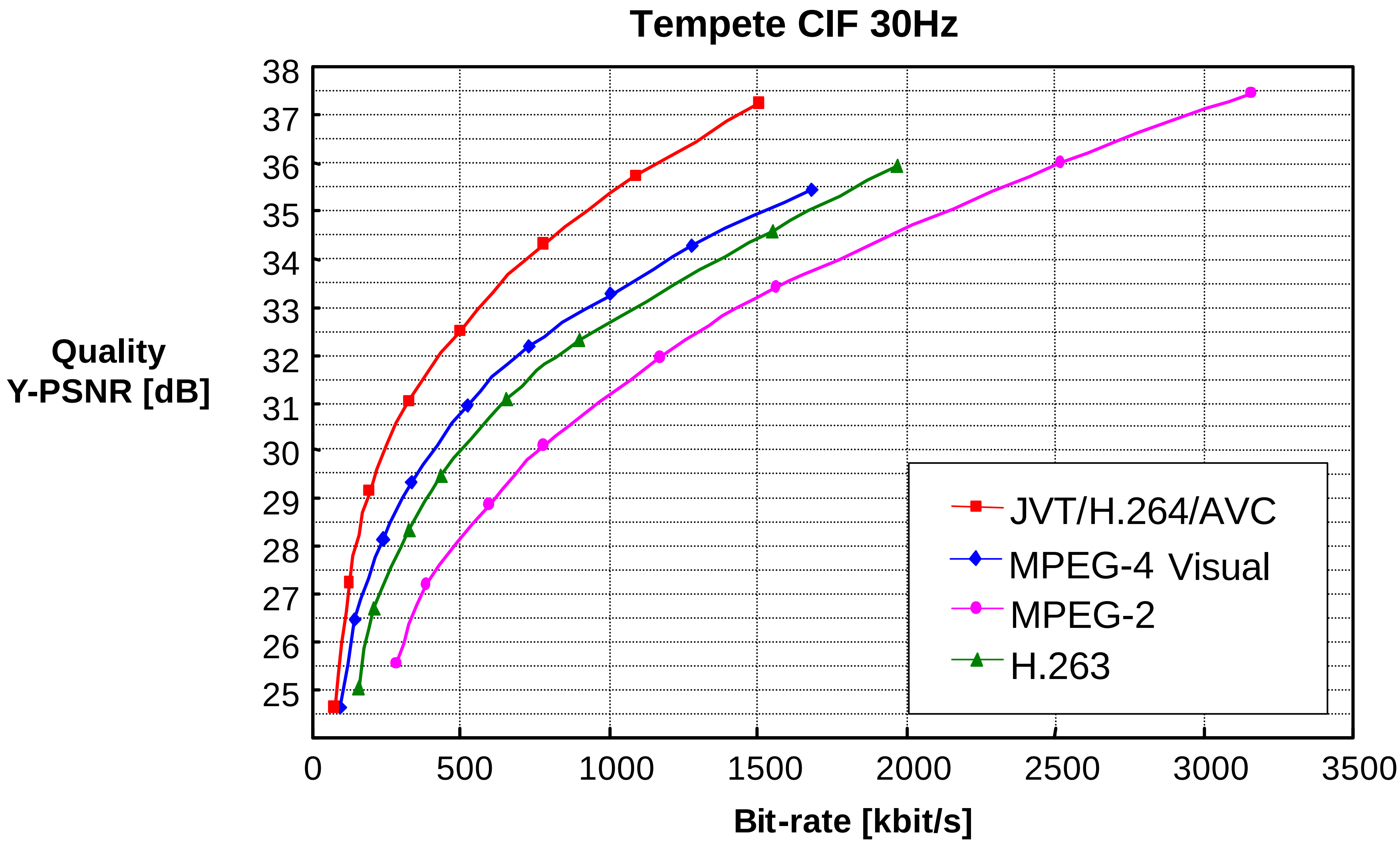
A Comparison of Performance

- Test of different standards (ICIP 2002 study)
- Using same rate-distortion optimization techniques for all codecs
- Streaming test: High-latency (included B frames)
  - Four QCIF sequences coded at 10 Hz and 15 Hz (Foreman, Container, News, Tempete) and
  - Four CIF sequences coded at 15 Hz and 30 Hz (Bus, Flower Garden, Mobile and Calendar, and Tempete)
- Real-time conversation test: No B frames
  - Four QCIF sequences encoded at 10Hz and 15Hz (Akiyo, Foreman, Mother and Daughter, and Silent Voice)
  - Four CIF sequences encoded at 15Hz and 30Hz (Carphone, Foreman, Paris, and Sean)
- Compare four codecs using PSNR measure:
  - **MPEG-2** (in high-latency/streaming test only)
  - **H.263** (high-latency profile, conversational high-compression profile, baseline profile)
  - **MPEG-4 Visual** (simple and advanced simple profiles with & without B pictures)
  - **H.264/AVC** (with & without B pictures)
- Note: These test results are from a private study and are not an endorsed report of the JVT, VCEG or MPEG organizations.

Comparison to MPEG-2, H.263, MPEG-4p2



Comparison to MPEG-2, H.263, MPEG-4p2



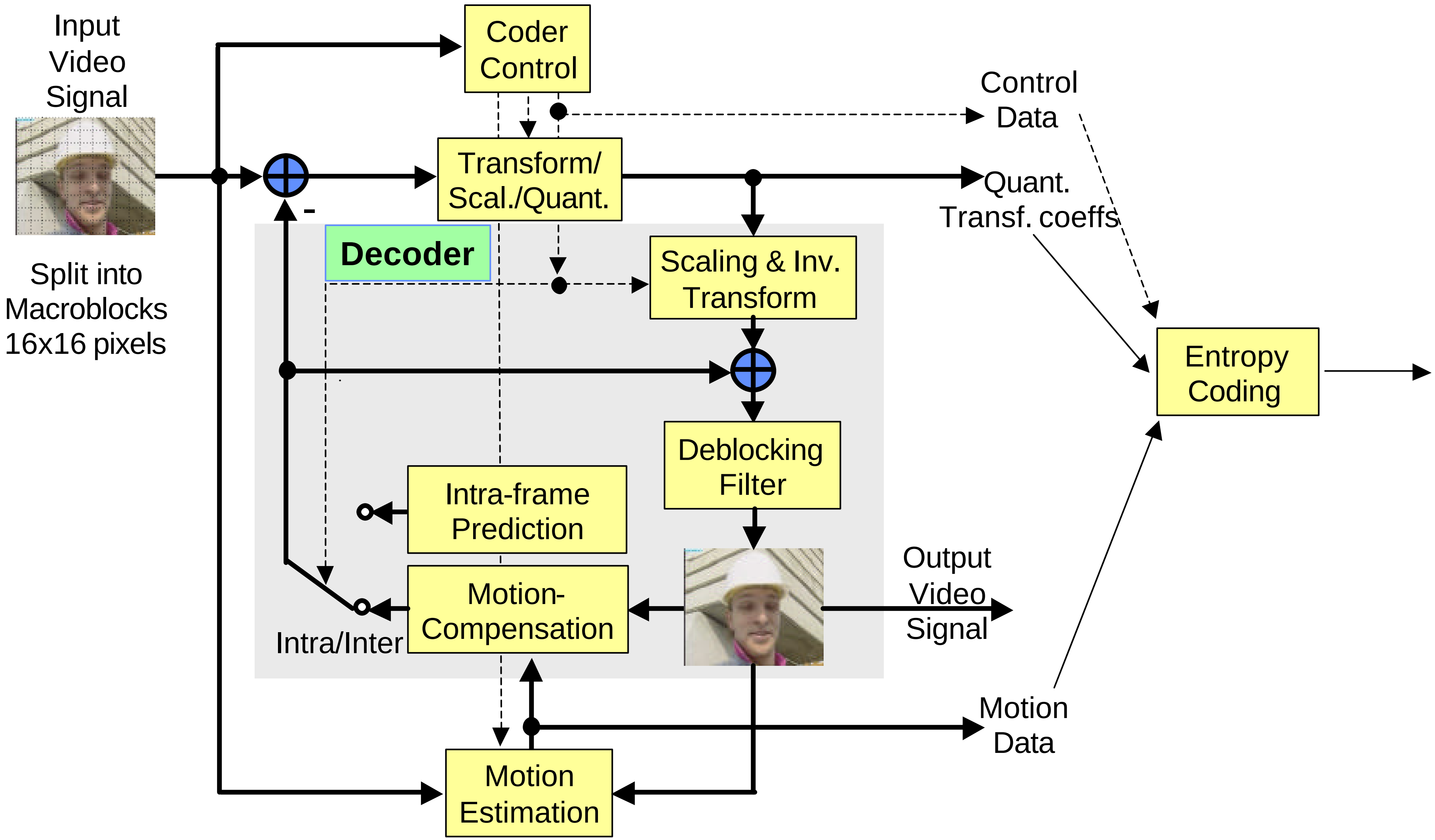
Caution: Your Mileage **Will** Vary

- This encoding software may not represent **implementation quality**
- These tests **only up to CIF** (quarter-standard-definition) resolution
- **Interlace, SDTV, and HDTV not tested** in this test
- Test sequences **may not be representative** of the variety of content encountered by applications
- These tests so far **not aligned** with profile designs
- This study reports PSNR, but **perceptual** quality is what matters

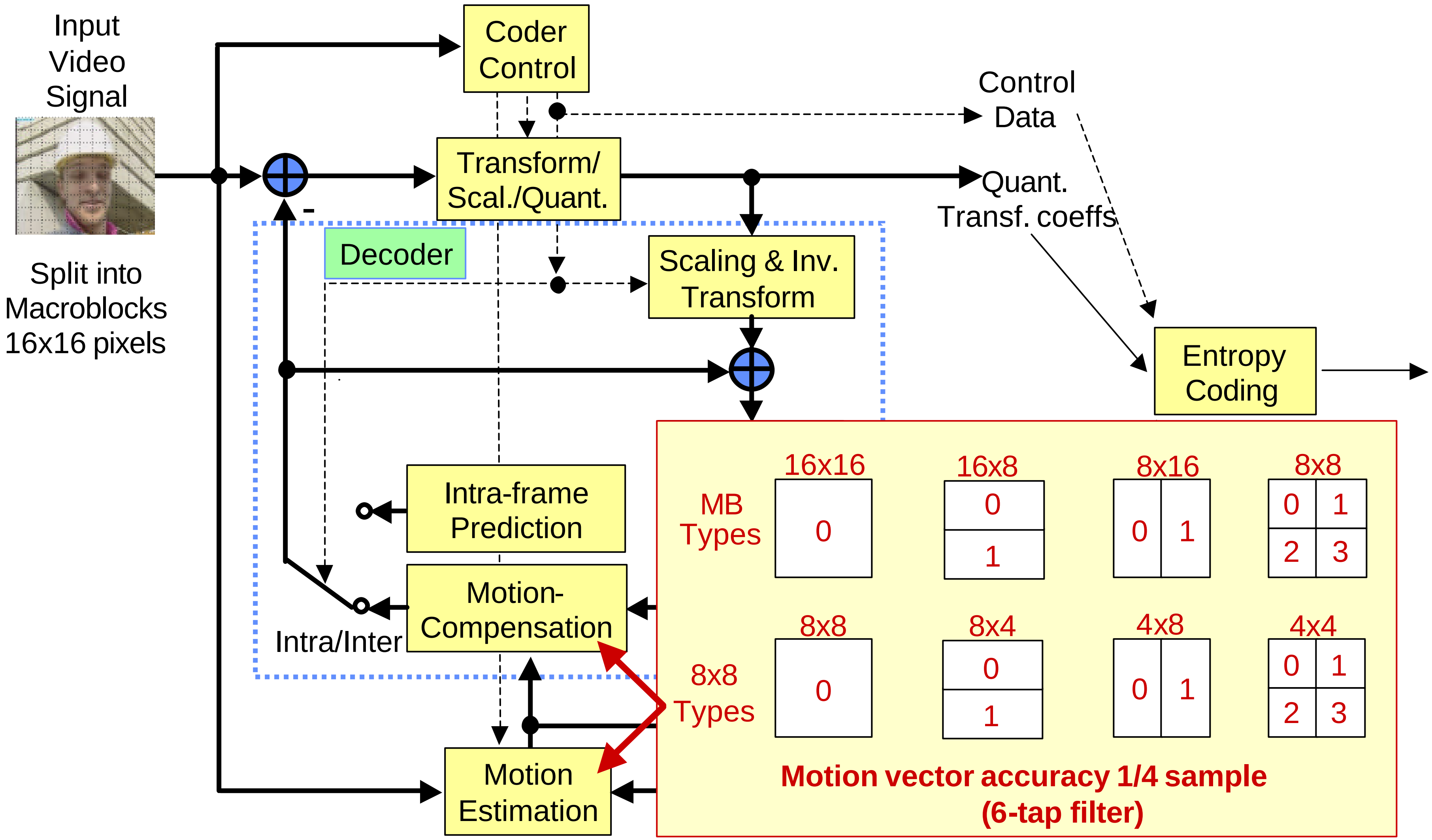
Computing Resources for the New Design

- New design includes relaxation of traditional bounds on computing resources – rough guess 2-3x the MIPS, ROM & RAM requirements of MPEG-2 for decoding, 3-4x for encoding
- Particularly an issue for low-power (e.g., mobile) devices
- Problem areas:
  - Smaller block sizes for motion compensation (cache access issues)
  - Longer filters for motion compensation (more memory access)
  - Multi-frame motion compensation (more memory for reference frame storage)
  - In-loop deblocking filter (more processing & memory access)
  - More segmentations of macroblock to choose from (more searching in the encoder)
  - More methods of predicting intra data (more searching)
  - Arithmetic coding (adaptivity, computation on output bits)

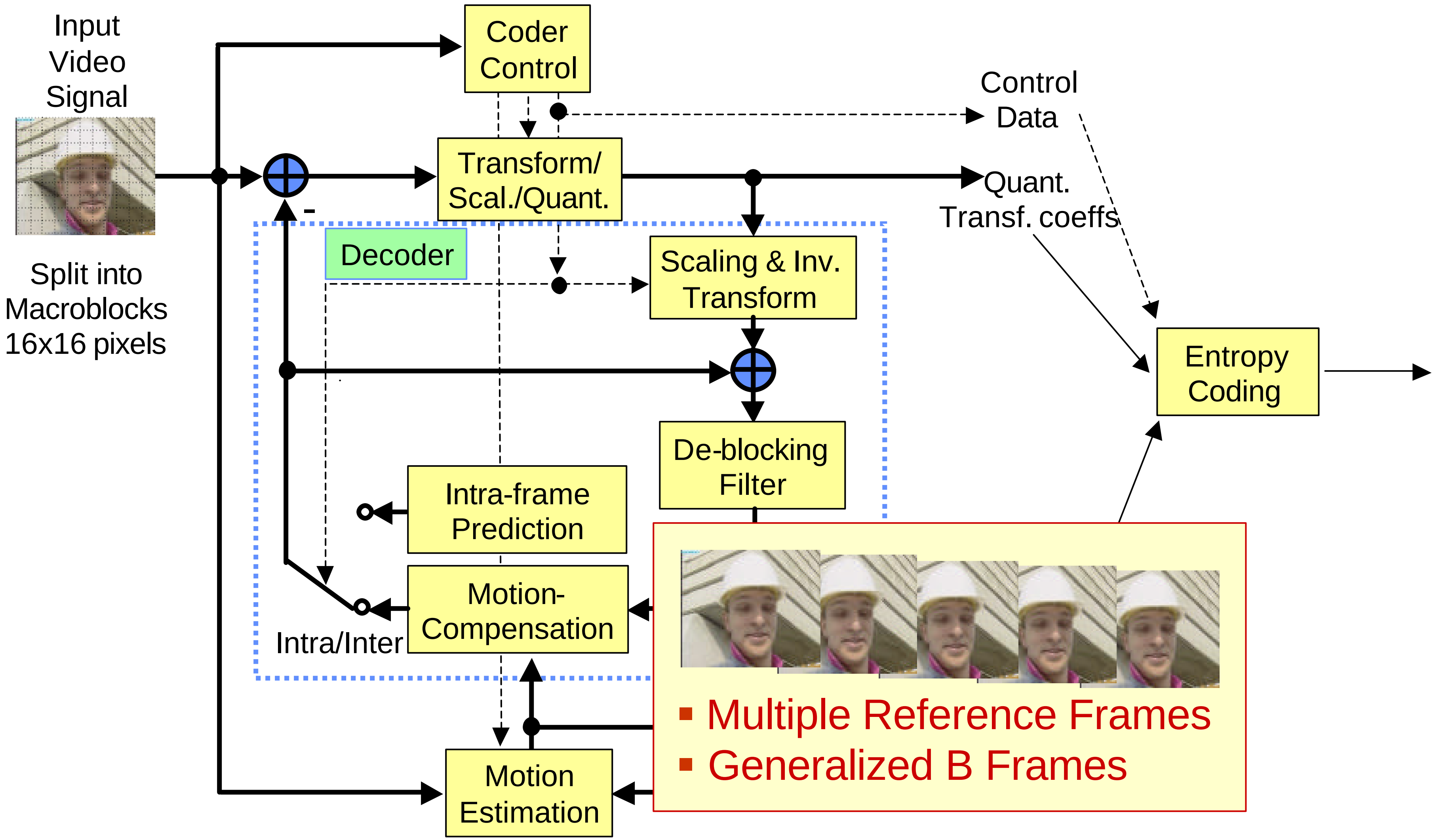
AVC Structure



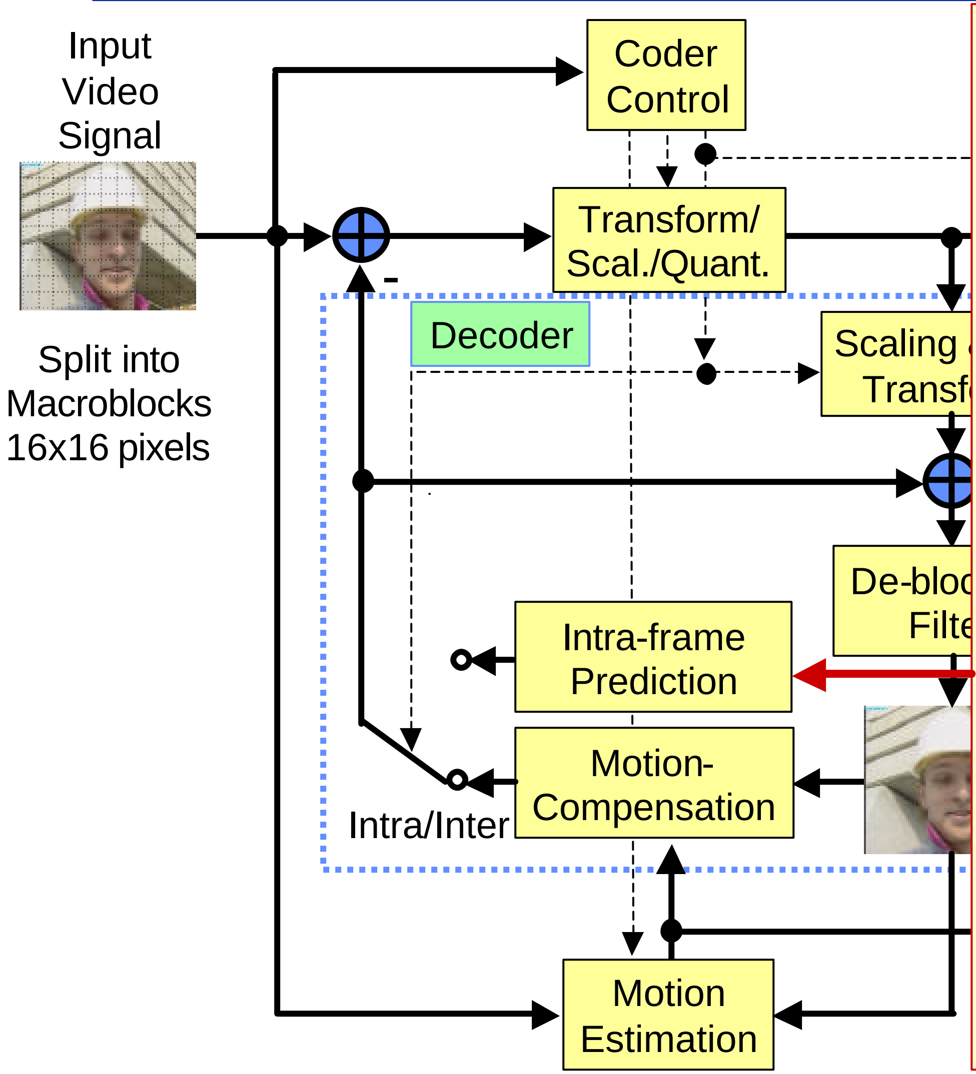
Motion Compensation Accuracy



Multiple Reference Frames

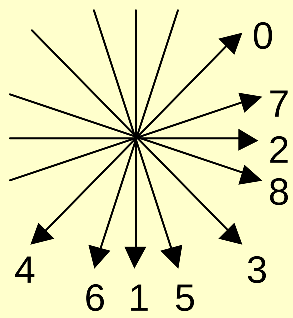


Intra Prediction



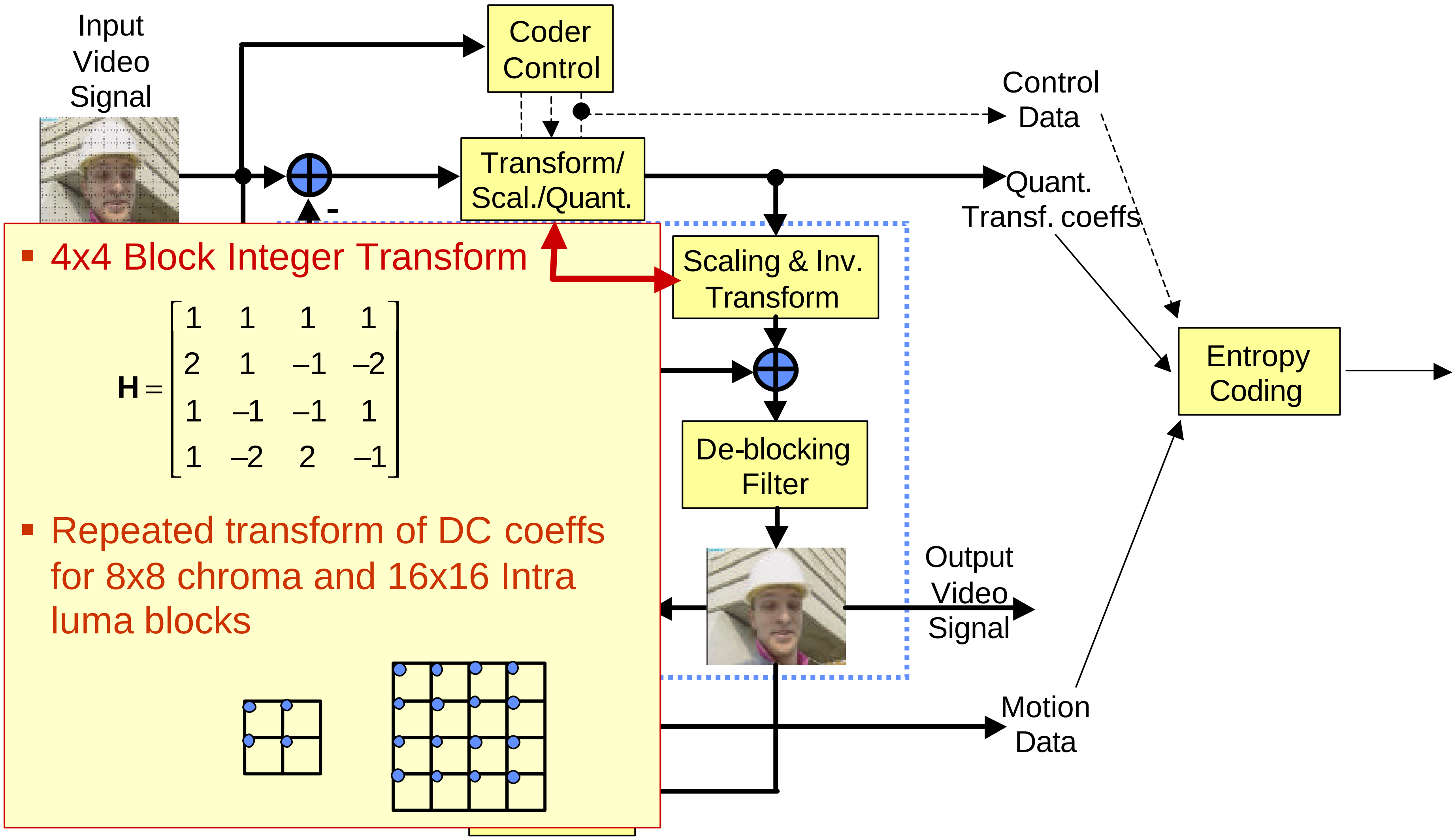
- Directional spatial prediction (9 types for luma, 1 chroma)

```
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
M
N
O
P
```



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

Transform Coding

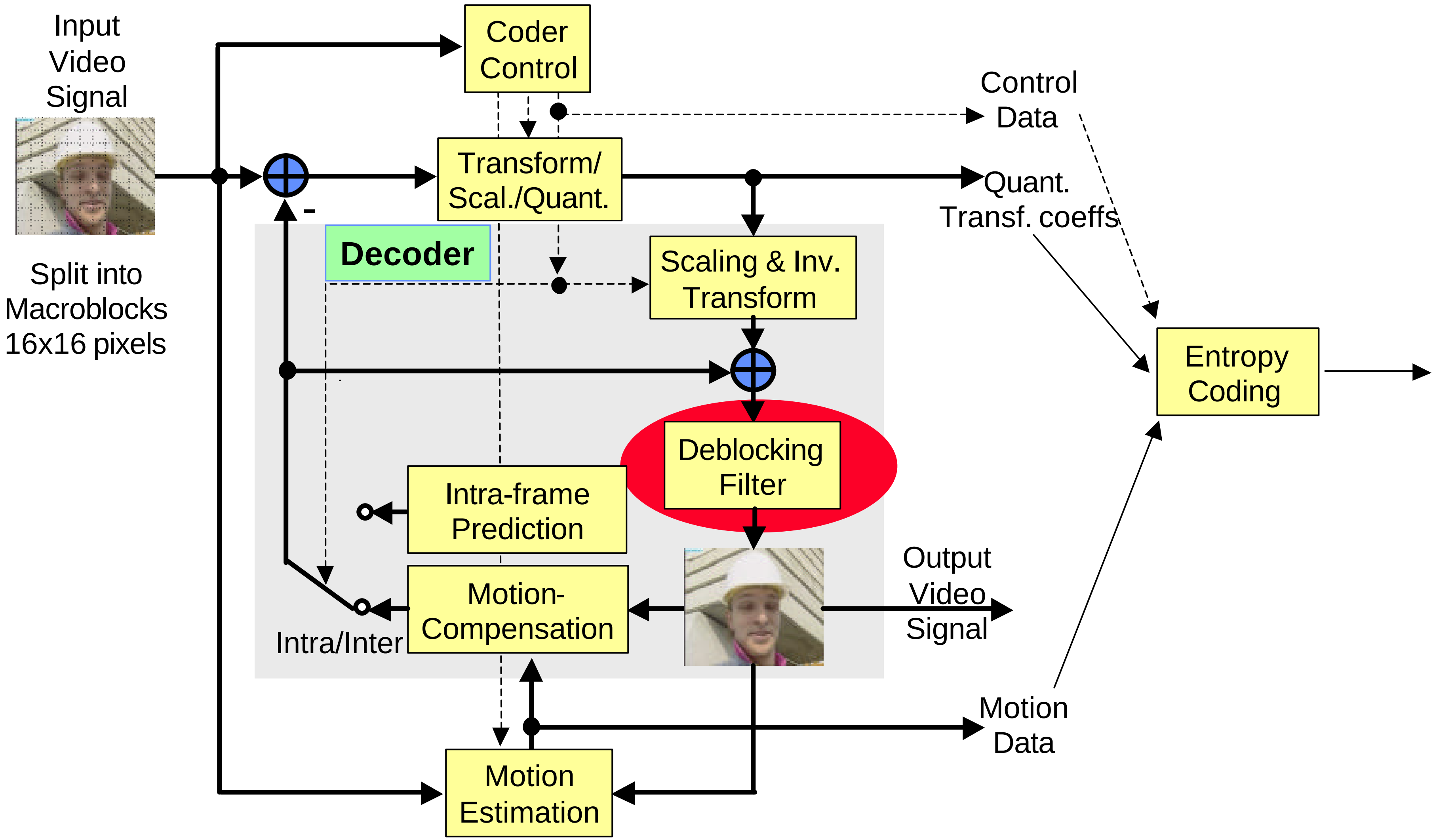


- 4x4 Block Integer Transform

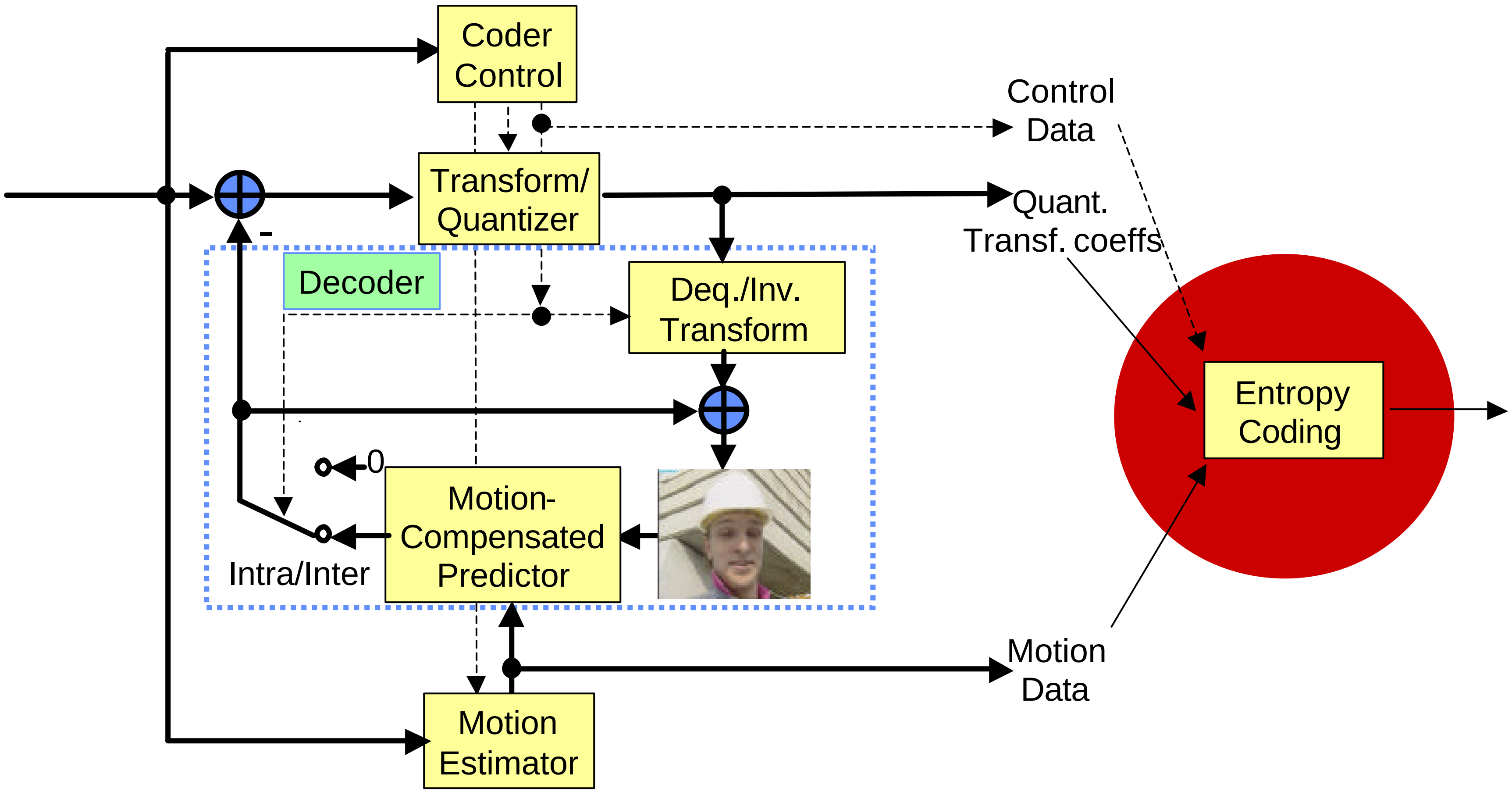
$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

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

Deblocking Filter



Entropy Coding



AVC Version 1 Profiles

- **Three profiles in version 1: Baseline, Main, and Extended**
- **Baseline (esp. Videoconferencing & Wireless)**
  - I and P progressive-scan picture coding (not B)
  - In-loop deblocking filter
  - 1/4-sample motion compensation
  - Tree-structured motion segmentation down to 4x4 block size
  - VLC-based entropy coding
  - Some enhanced error resilience features
    - Flexible macroblock ordering/arbitrary slice ordering
    - Redundant slices

Non-Baseline AVC Version 1 Profiles

- **Main Profile (esp. Broadcast)**
  - All Baseline features **except** enhanced error resilience features
  - Interlaced video handling
  - Generalized B pictures
  - Adaptive weighting for B and P picture prediction
  - ***CABAC (arithmetic entropy coding)***
- **Extended Profile (esp. Streaming)**
  - All Baseline features
  - Interlaced video handling
  - Generalized B pictures
  - Adaptive weighting for B and P picture prediction
  - ***More error resilience: Data partitioning***
  - ***SP/SI switching pictures***

Amendment 1: Fidelity-Range Extensions

- AVC standard finished 2003
  - ITU-T/H.264 finalized May, 2003
  - MPEG-4 AVC finalized July, 2003 (same thing)
  - Only corrigenda (bug fixes) since then
- Fidelity-Range Extensions (FRExt)
  - New work item initiated in July 2003
  - More than 8 bits, color other than 4:2:0
  - Alpha coding
  - More coding efficiency capability
  - Also new supplemental information

FRExt Finished July 04

- Project initiated July 2003
  - Motivations
    - Higher quality, higher rates
    - 4:4:4, 4:2:2, and also 4:2:0
    - 8, 10, or 12 bits (14 bits considered and not included)
    - Lossless
    - Stereo
- Finished in one year! (July 04)

New Things in FRExt – Part 1

- Larger transforms
  - 8x8 transform (again!)
  - Drop 4x8, 8x4, or larger, 16-point…
- Filtered intra prediction modes for 8x8 block size
- Quantization matrix
  - 4x4, 8x8, intra, inter trans. coefficients weighted differently
  - Old idea, dating to JPEG and before (circa 1986?)
  - Full capabilities not yet explored (visual weighting)
- Coding in various color spaces
  - 4:4:4, 4:2:2, 4:2:0, Monochrome, with/without Alpha
  - New integer color transform (a VUI-message item)

New Things in FRExt – Part 2

- Efficient lossless interframe coding
- Film grain characterization for analysis/synthesis representation
- Stereo-view video support
- Deblocking filter display preference

8x8 16-Bit (Bossen) Transform

$$\begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 12 & 10 & 6 & 3 & -3 & -6 & -10 & -12 \\ 8 & 4 & -4 & -8 & -8 & -4 & 4 & 8 \\ 10 & -3 & -12 & -6 & 6 & 12 & 3 & -10 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -12 & 3 & 10 & -10 & -3 & 12 & -6 \\ 4 & -8 & 8 & -4 & -4 & 8 & -8 & 4 \\ 3 & -6 & 10 & -12 & 12 & -10 & 6 & -3 \end{bmatrix}$$

8x8 Transform Advantage (JVT-K028, IBBP coding, prog. scan)

Sequence	**% BD bit-rate reduction**
Movie 1	11.59
Movie 2	12.71
Movie 3	12.01
Movie 4	11.06
Movie 5	13.46
Crawford	10.93
Riverbed	15.65
Average	**12.48**

Quantization Matrix

- Similar concept to MPEG-2 design
- Vary step size based on frequency
- Adapted to modified transform structure
- More efficient representation of weights
- Eight downloadable matrices (at least 4:2:0)
  - Intra 4x4 Y, Cb, Cr
  - Intra 8x8 Y
  - Inter 4x4 Y, Cb, Cr
  - Inter 8x8 Y

New Profiles Created by FRExt

- 4:2:0, 8-bit: "High" (HP)
- 4:2:0, 10-bit: "High 10" (Hi10)
- 4:2:2, 10-bit: "High 4:2:2" (Hi422)
- 4:4:4, 12-bit: "High 4:4:4" (Hi444)

- Effectively the same tools, but acting on different input data (with a couple of exceptions in the 4:4:4 profile)

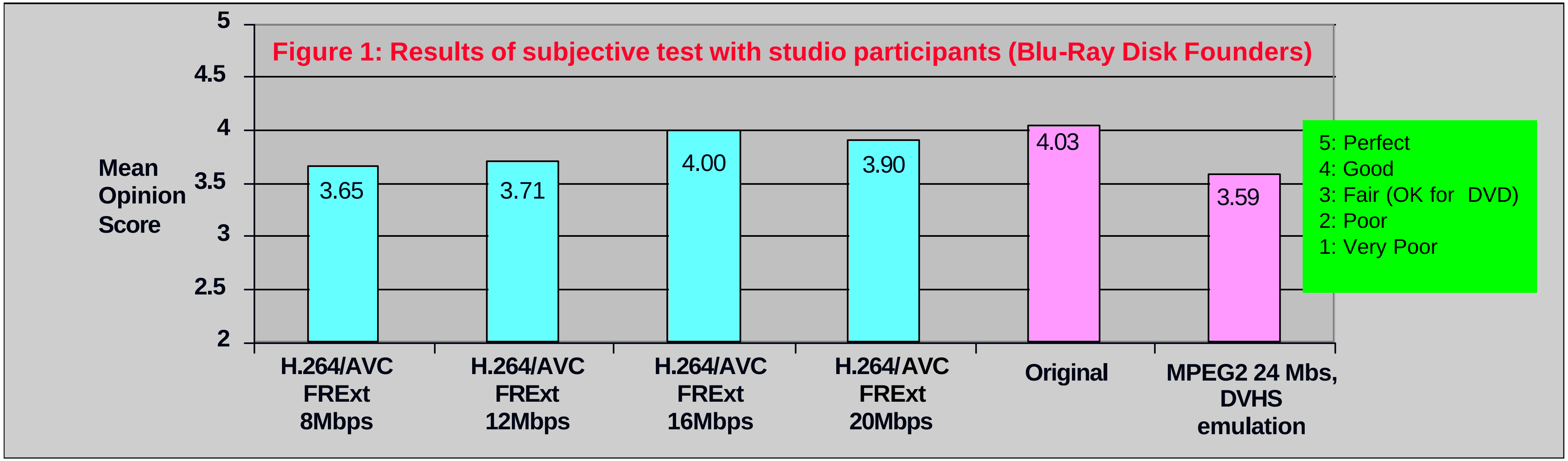
Some Notes on Quality Testing

- Use appropriate "High" profile (incl. adaptive transform)
- If testing for PSNR, use "flat" quant matrices
- Otherwise, use "non-flat" quant matrices
- Use more than 1 or 2 reference pictures
- Use hierarchical reference frames coding structure
- Use CABAC entropy coding
- If testing high-quality PSNR, use adaptive quantization*
- Use rate-distortion optimization in encoder
- Use large-range good-quality motion search

*** = See G. Sullivan & S. Sun, "On Dead-Zone…", VCIP 2005/JVT-N011**

A Performance Test for High Profile (from JVT-L033 - Panasonic)

- Subjective tests by Blu-Ray Disk Founders of FRExt HP
  - 4:2:0/8 (HP) 1920x1080x24p (1080p), 3 clips.
  - Nominal 3:1 advantage to MPEG-2
    - 8 Mbps HP scored better than 24 Mbps MPEG-2
  - Apparent **transparency** at 16 Mbps



For Further Information

- JVT, MPEG, and VCEG management team members:
  - Gary Sullivan (garysull@microsoft.com)
  - Jens Ohm (ohm@ient.rwth-aachen.de)
  - Ajay Luthra (aluthra@motorola.com)
  - Thomas Wiegand (wiegand@hhi.de)
- IEEE Transactions on Circuits and Systems for Video Technology Special Issue on H.264/AVC (July 2003)
- Paper in Proceedings of IEEE January 2005
- I. Richardson, *H.264 and MPEG-4 Video Compression*
- Overview including FRExt: SPIE Aug 2004 (Sullivan, Topiwala, and Luthra)
- Paper at VCIP 2005: Meta-overview and deployment