Next-Generation High Efficiency Video Coding (HEVC) Standard

Gary J. Sullivan

Co-chair JCT-VC, Chair VCEG, Co-Chair Video in MPEG,
Video & Image Technology Architect, Microsoft

15 February 2011

Presentation for Hollywood Post Alliance (HPA) Technology Retreat /
ATSC Next Generation Broadcast Television Symposium

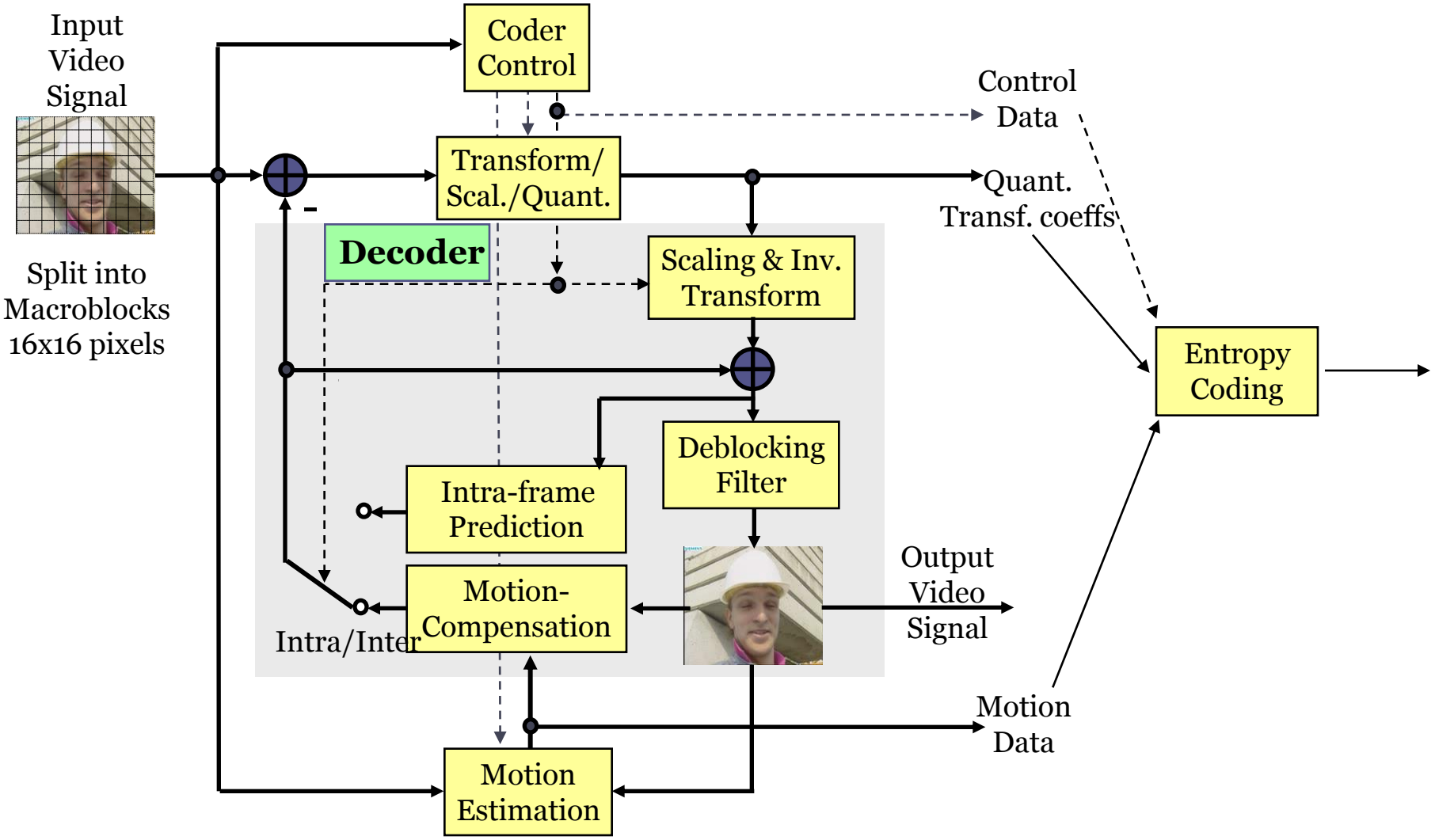
Video Coding Standards Organizations

- **ISO/IEC MPEG** = "**Moving Picture Experts Group**"
  (ISO/IEC JTC 1/SC 29/WG 11 = International Standardization Organization and International Electrotechnical Commission, Joint Technical Committee 1, Subcommittee 29, Working Group 11)
- **ITU-T VCEG** = "**Video Coding Experts Group**"
  (ITU-T SG16/Q6 = International Telecommunications Union – Telecommunications Standardization Sector (ITU-T, a United Nations Organization, formerly CCITT), Study Group 16, Working Party 3, Question 6)
- **JVT** = "**Joint Video Team**" collaborative team of MPEG & VCEG
- **SMPTE** (**Society of Motion Picture and Television Engineers**)
- **New**: **JCT-VC** = "**Joint Collaborative Team on Video Coding**" team of MPEG & VCEG, continuing the collaborative relationship for a new project (established January 2010)

Some standardization activities in ISO/IEC MPEG & ITU-T VCEG & JCT-VC

- AVC (H.264 / MPEG-4 pt 10) and its recent extensions
  - SVC emerging (esp. in real-time communication & 1080p60 on 720p60)
  - Stereo / MVC coding (esp. for 3D Blu-ray)
  - Frame-compatible frame packing SEI (esp. for broadcast)
- Work formally in progress
  - Dynamic Adaptive Streaming over HTTP (DASH)
    - At DIS (formerly called FCD) stage, completion 2011-2012
  - High Efficiency Video Coding (HEVC)
    - Joint MPEG & VCEG new team: JCT-VC
    - Schedule CD Jan 2012, DIS July 2012, FDIS & Consent Jan 2013
- Explorations
  - 3DV for enhanced 3D with motion parallax
  - SSV / SREFS / MFC frame-compatible stereo video scalable enhancement

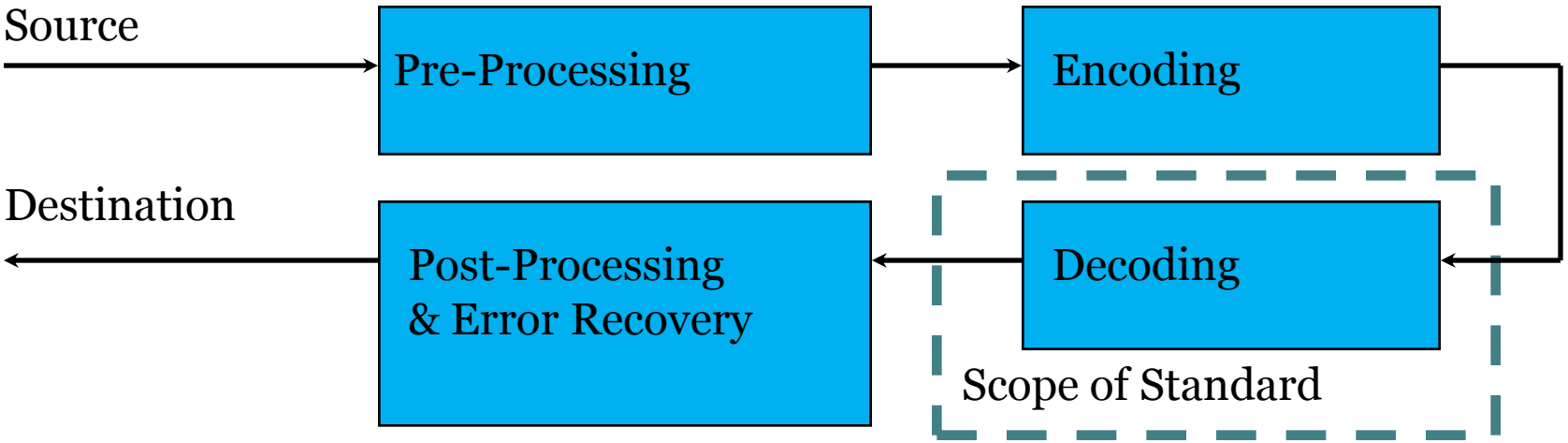
H.264 / 14496-10 AVC Structure



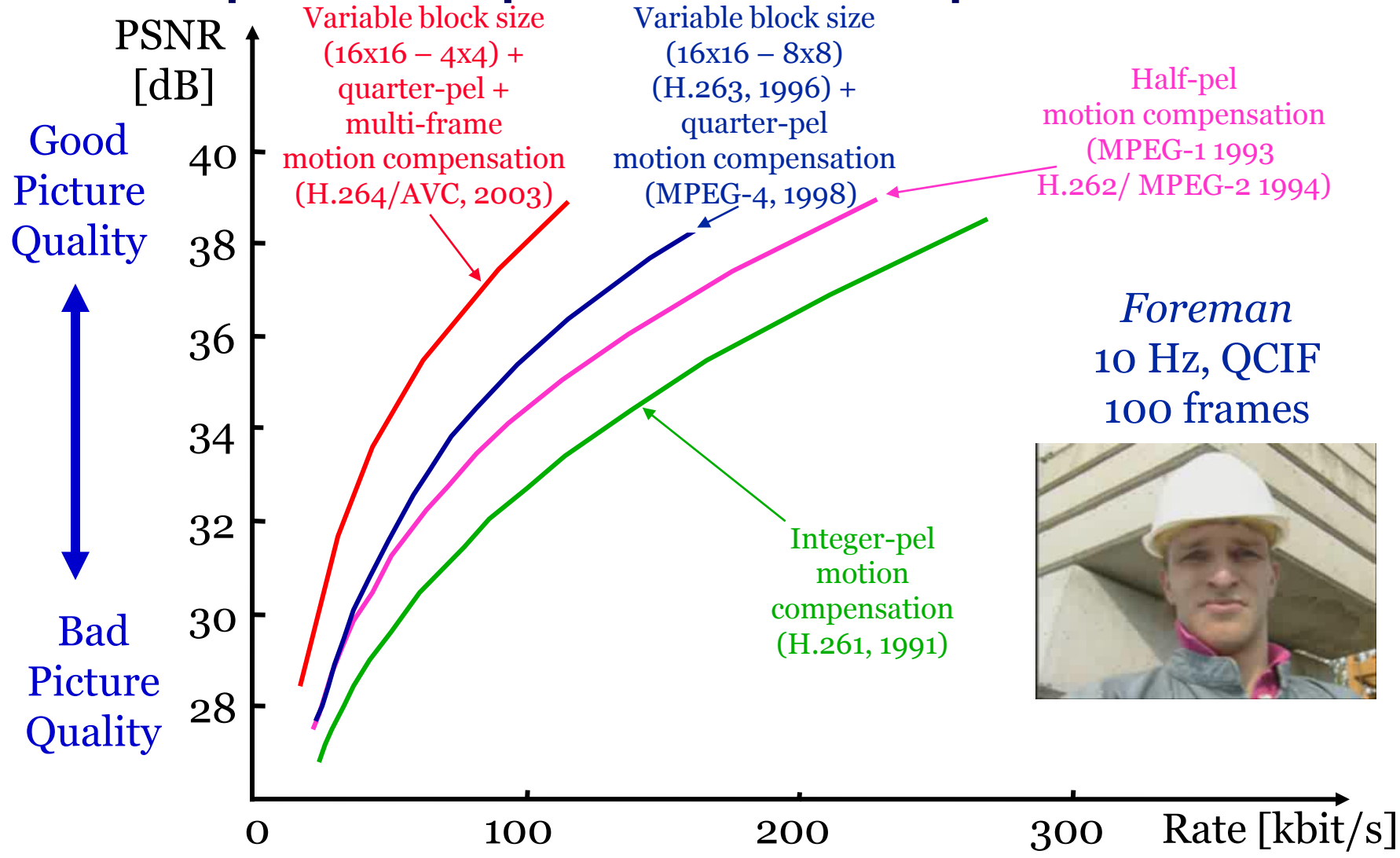
Based on a diagram originally produced by Thomas Wiegand (Fraunhofer HHI)

The *Scope* of Video Coding Standardization

- Only the *Syntax* and *Decoder* are standardized:
  - Permits optimization beyond the obvious
  - Permits complexity reduction for implementability
  - Provides some capability, but *no guarantee* of Quality



Example Compression Comparison



Based on a diagram originally produced by Thomas Wiegand (Fraunhofer HHI)

HEVC and the new JCT-VC Partnership

- Initial groundwork in VCEG and MPEG
- Agreement in January 2010 to form new team **VCEG-AM90** / **N11112**
- Joint Call for Proposals on Video Coding Technology issued January 2010 **VCEG-AM91** / **WG 11 N11113**
- **Joint Collaborative Team (JCT) on Video Coding (JCT-VC)**
- Chairs: Gary Sullivan (Microsoft) & Jens-Rainer Ohm (RWTH Aachen Univ.)
- Meetings so far
  - First meeting: Dresden Germany 15-23 April 2010
  - Second meeting: Geneva, Switzerland 21-28 July 2010
  - Third meeting: Guangzhou, China 7-15 October 2010
  - Fourth meeting: Daegu, Korea 20-28 January 2011
  - **Next meeting**: Geneva, Switzerland 16-23 March 2011
- Project name **High Efficiency Video Coding (HEVC)** agreed April 2010
- Document archives are publicly accessible
  - http://phenix.it-sudparis.eu/jct
  - http://ftp3.itu.ch/av-arch/jctvc-site
  - http://www.itu.int/ITU-T/studygroups/com16/jct-vc/index.html
- Email reflector http://mailman.rwth-aachen.de/mailman/listinfo/jct-vc

Call for Proposals Testing

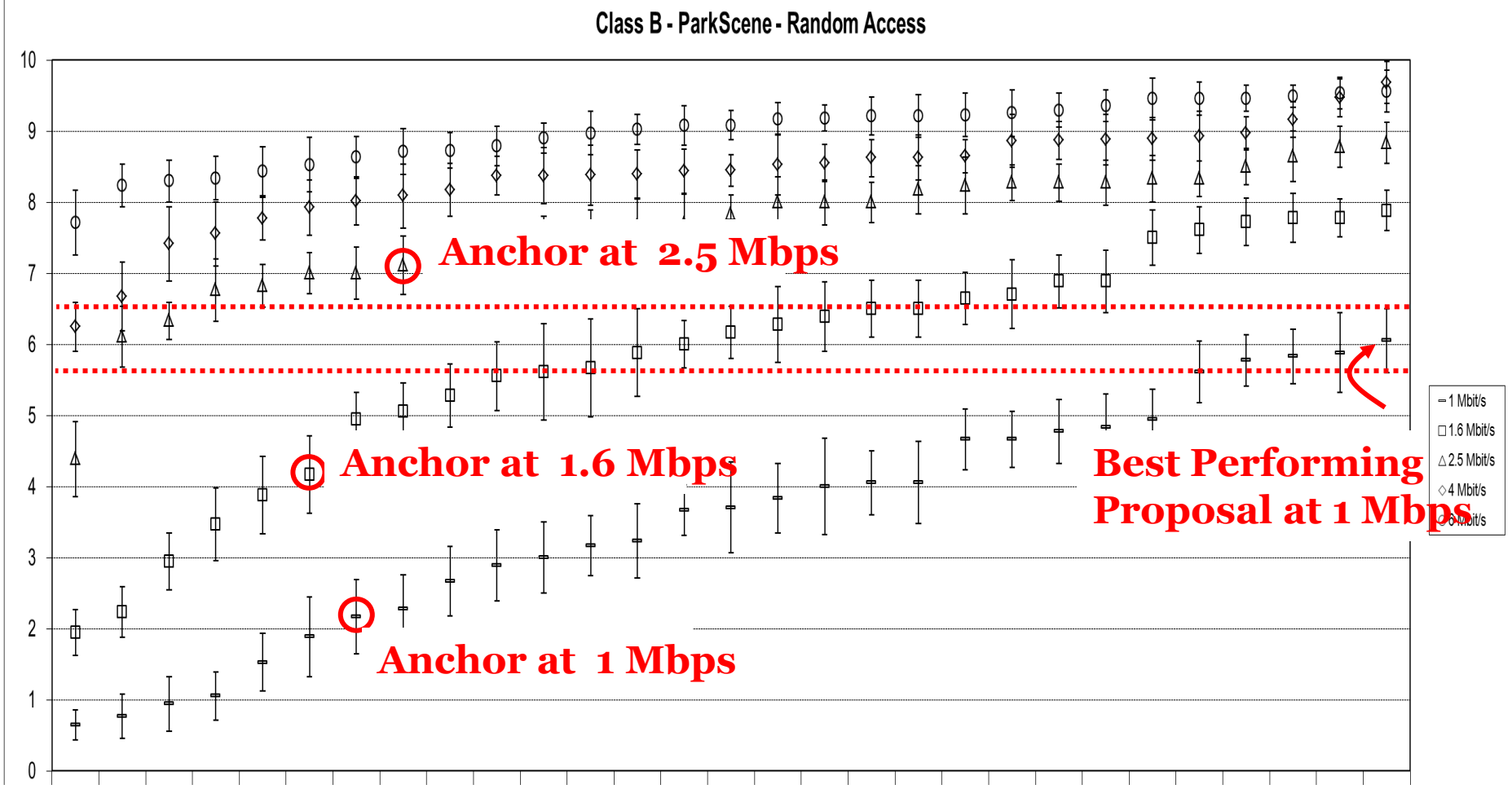
- 27 complete proposals submitted (some multi-organizational)
- Each proposal was a major package – lots of encoded video, extensive documentation, extensive performance metric submissions, sometimes software, etc.
- Extensive subjective testing (3 test labs, 4 200 video clips evaluated, 850 human subjects, 300 000 scores)
- Quality of proposal video was compared to AVC (ITU-T Rec. H.264 | ISO/IEC 14496-10) anchor encodings
- Test report issued **JCTVC-A204**
- In a number of cases, comparable quality at half bit rate

Test Classes and Bit Rates

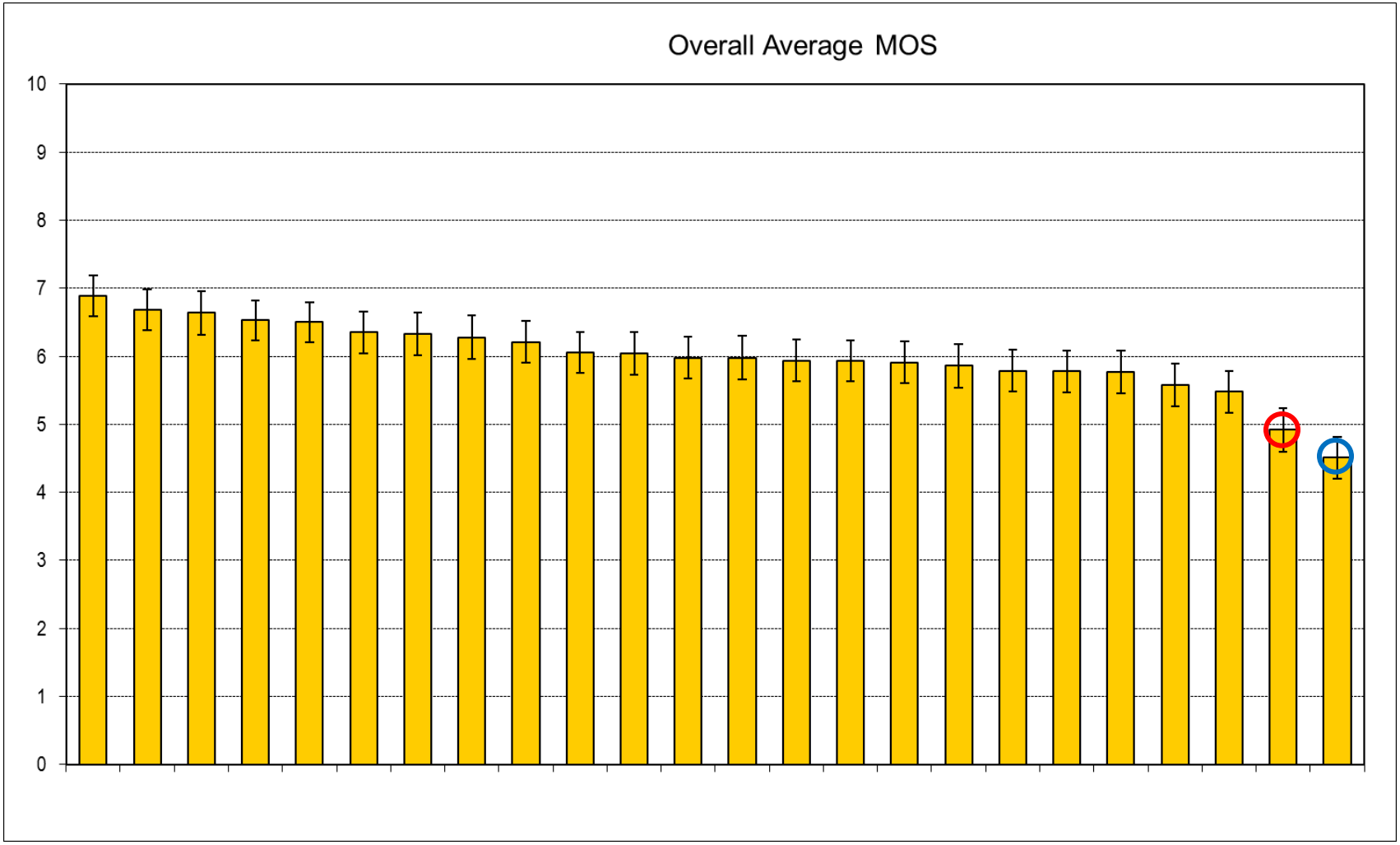
Class	Bit Rate 1	Bit Rate 2	Bit Rate 3	Bit Rate 4	Bit Rate 5
A: 2560x1600p30	2.5 Mbit/s	3.5 Mbit/s	5 Mbit/s	8 Mbit/s	14 Mbit/s
B1: 1080p24	1 Mbit/s	1.6 Mbit/s	2.5 Mbit/s	4 Mbit/s	6 Mbit/s
B2: 1080p50-60	2 Mbit/s	3 Mbit/s	4.5 Mbit/s	7 Mbit/s	10 Mbit/s
C: WVGAp30-60	384 kbit/s	512 kbit/s	768 kbit/s	1.2 Mbit/s	2 Mbit/s
D: WQVGAp30-60	256 kbit/s	384 kbit/s	512 kbit/s	850 kbit/s	1.5 Mbit/s
E: 720p60	256 kbit/s	384 kbit/s	512 kbit/s	850 kbit/s	1.5 Mbit/s

- 3-5 video clips subjectively tested in Classes B-E
- Testing for both “random access” (1 sec) and “low delay” (no picture reordering) conditions
- Complexity also considered in anchor encodings

Example Proposal Results Graph



Overall Average Mean Opinion Score



Basic Technology Architecture

- All proposals basically conceptually similar to AVC (and prior standards)
  - Block-based
  - Variable block sizes
  - Block motion compensation
  - Fractional-pel motion vectors
  - Spatial intra prediction
  - Spatial transform of residual difference
  - Integer-based transform designs
  - Arithmetic or VLC-based entropy coding
  - In-loop filtering to form final decoded picture
- Lots of variations at the individual "tool" level
- Proposal survey output documents:
  - Decoder speed **JCTVC-A201**
  - Architectural outline **JCTVC-A202**
  - Table of design elements **JCTVC-A203**

HEVC Model (HM 2) Technical Overview

High Efficiency	Low Complexity
Coding unit tree structure (8x8 up to 64x64 luma samples)	
Prediction units (N=4, 8, 16, 32, shapes: 2Nx2N, NxN for smallest; for inter also 2NxN & Nx2N)	
Transforms can cross prediction unit boundaries for Intra; not for Inter	
Transform unit tree structure (maximum of 3 levels)	Transform unit tree structure (maximum of 2 levels)
Transform block size of 4x4 to 32x32 samples (always square)	
Angular intra prediction (17 directions for 4x4, 3 directions for 64x64, 34 directions for others)	
Luma motion compensation: 1/4 sample precision, 8x8 separable with 6 bit tap values	
Chroma motion compensation: 1/8 sample precision, 4x4 separable with 6 bit tap values	
Advanced motion vector prediction with motion vector “competition” and “merging”	
Context adaptive binary arithmetic entropy coding	Low complexity entropy coding
Increased bit depth capability (10 bits)	Transform precision extension (2 bits)
Deblocking filter	
Adaptive loop filter	X

HEVC Timeline

- 2011 Meetings
  - Jan 20-28, 2011 Daegu
  - March 16-23, 2011 Geneva
  - July 14-22, 2011 Torino
  - Nov 22-30, 2011 Geneva
- CD Feb 1-10, 2012 San Jose
- Meeting April 2-9, 2012 Geneva
- DIS July 11-20, 2012 Stockholm
- Meeting Oct 10-19, 2012 Suzhou
- FDIS & Consent [Final spec] Jan 2013 Geneva

HEVC Expectations & Final Words

- Very active project (200+ documents & people per meeting)
- Very diverse company & university participation
- Significant technical advance over prior standards
- Computational/implementation complexity is a big challenge
- Parallelism is an increased theme
- Deliverables
  - Video coding specification
  - Reference software
  - Conformance testing specification
- "Profiles" for various applications (mobile, high-quality entertainment, etc.)
- Likely multiple versions and extensions (SVC, MVC, etc)
- Contact: JVT, JCT-VC, VCEG, MPEG video chairs:
  - Gary J. Sullivan (garysull@microsoft.com)
  - Jens-Rainer Ohm (ohm@ient.rwth-aachen.de)