Video Coding: Recent Developments for HEVC and Future Trends

Initial overview section by Gary Sullivan

Video Architect, Microsoft Corporate Standards Group

30 March 2016

Presentation for *Data Compression Conference*, Snowbird, Utah

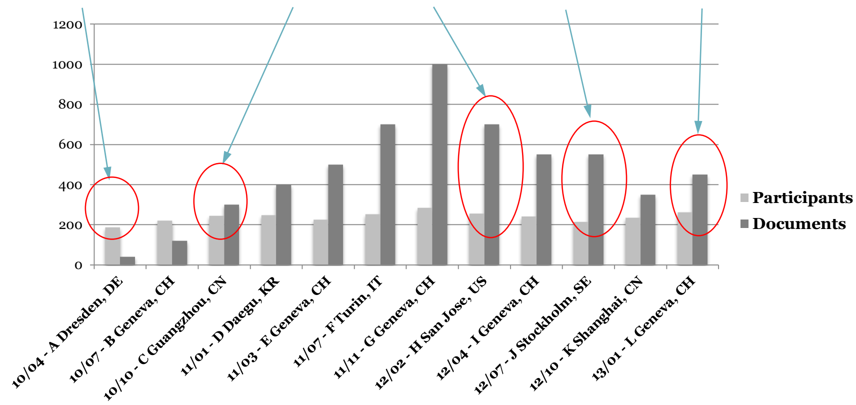
Major Video Coding Standards

Mid 1990s: MPEG-2 → Mid 2000s: H.264/MPEG-4 AVC → Mid 2010s: HEVC

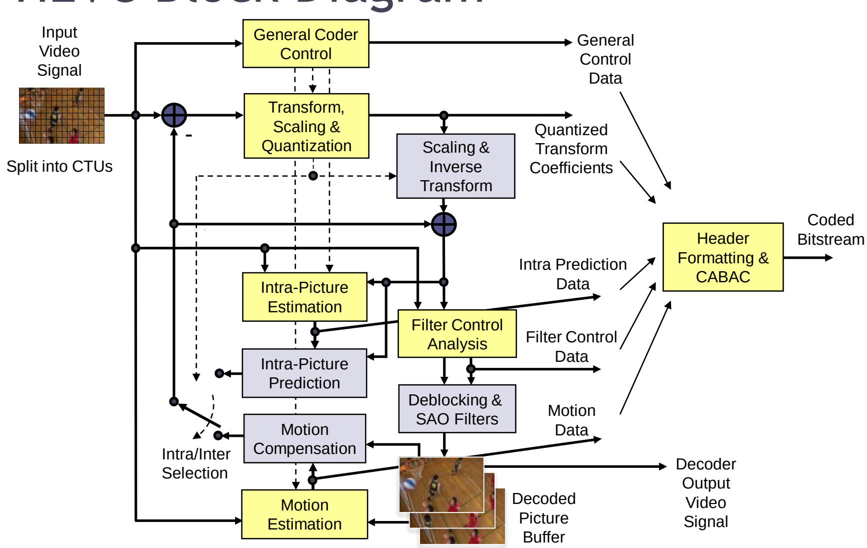
- These are the joint work of the same two bodies
  - ISO/IEC Moving Picture Experts Group (MPEG)
  - ITU-T Video Coding Experts Group (VCEG)
  - Most recently working on **High Efficiency Video Coding (HEVC)** as Joint Collaborative Team on Video Coding (JCT-VC)
- HEVC version 1 was completed in January 2013
  - Standardized by ISO/IEC as **ISO/IEC 23008-2** (MPEG-H Part 2)
  - Standardized by ITU-T as **H.265**
  - 3 profiles: Main, 10 bit, and still picture

Project Timeline and Milestones

First Test Model under Consideration (TMuC) | First Working Draft and Test Model (Draft 1 / HM1) | ISO/IEC CD (Draft 6 / HM6) | ISO/IEC DIS (Draft 8 / HM8) | ISO/IEC FDIS & ITU-T Consent (Draft 10 / HM10)



HEVC Block Diagram



HEVC version 2 extensions (2014): Format range extensions ("RExt")

- 4:4:4, 4:2:2, monochrome, all-intra
- Increased bit depths
- Completion April 2014 (a few aspects one meeting later)
- 21 profiles
- New technologies:
  - Color-related metadata (SMPTE 428 XYZ & DCDM VUI, 2084 PQ VUI, 2086 MDCV SEI, knee SEI, adaptation mapping SEI)
  - High-precision weighted prediction
  - Extended-precision processing
  - High bit depths up to 16 bits per sample
  - High-throughput entropy coding with bypass byte alignment
  - Cross-component prediction
  - Entropy coding persistent Rice parameter adaptation
  - Transform-skip enhancements (residual DPCM, rotation, block sizes, entropy coding context)
  - Enhanced chroma QP control

Gary J. Sullivan 2016-03-30

HEVC version 2 extensions (2014): Scalable HEVC (SHVC)

- Hooks for extensions built into version 1
- Temporal scalability also included in version 1
- New spatial, SNR, color gamut, & bit depth enhancements
- Completion July 2014
- 2 profiles
- New technologies
  - Architecturally simple multi-loop "reference index" design
  - Lots of work on high-level syntax properties and generality
  - Independent non-base layer coding
  - Alpha channel, overlays, other metadata
  - Colour gamut scalability mapping
  - AVC / external base layer possible

Gary J. Sullivan 2016-03-30

HEVC extensions (2014 & 5): Multiview (MV-HEVC) & 3D-HEVC

- Developed in another team: the Joint Collaborative Team on 3D video coding (JCT-3V)
- Two new profiles added
- New technologies
  - High-level syntax design harmonized with SHVC extensions
  - Completion July 2014 for multiview & depth map encoding
  - Combined view and depth map coding (in version 3 finalized in Feb 2015)
- Also some new 3D video extensions to AVC
  - Joint coding of depth and texture
  - Multi-resolution frame-compatible coding
  - Further use of depth maps

Gary J. Sullivan 2016-03-30

HEVC Screen Content Coding (SCC) Extensions (2016)

- For video containing a significant portion of rendered (moving or static) graphics, text, or animation
- Use for wireless displays, remote computer desktop access, real-time screen sharing for videoconferencing, news with scrolling graphics, etc.
- Call for Proposal issued January 2014
- 7 responses evaluated (JCTVC-Q0031 – JCTVC-Q0037)
  - Qualcomm, ITRI, MediaTek, Huawei, Microsoft, Mitsubishi Electric, Interdigital
- Schedule:
  - First Test Model April 2014
  - PDAM Feb 2015
  - DIS June 2015
  - ISO/IEC FDIS Feb 2016
  - ITU-T Consent June 2016

Gary J. Sullivan 2016-03-30

SCC Extensions Technical features (2016)

- Palette mode
  - Directly code values of samples (no transform) to create palette
  - Palette prediction from region to region
  - Scan to copy from above, left, palette index or escape
- Current-picture referencing (a.k.a. Intra block copy)
  - Like motion compensation, but within the current picture
  - Displacement vectors are integer-valued
  - Essentially otherwise identical to inter-picture referencing
- Adaptive MV resolution
  - Displacement vectors for inter-picture prediction of sequence or slice can be restricted to integer values
- Adaptive color transform (based on YCoCg)
  - Cross-component colour-transform within the decoding process
  - Especially helpful for coding RGB-format video
  - Based on YCoCg-R (a lifting-based reversible colour transform)
- Intra boundary filtering disabling
  - Very small modification to prevent blurring of predictors

Gary J. Sullivan 2016-03-30

PSNR-Based Study of SCC – SPIE 2015 (with Bin Li and Jizheng Xu of Microsoft)

- Objective PSNR-based measurements
- Software used
  - Joint Model (JM) with version 18.6 for H.264/MPEG-4 AVC
  - HEVC Model (HM) with version 16.4 for HEVC version 1 and its format range extensions
  - HM-16.4+SCM-4.0 version for Draft 3 of the HEVC screen content coding extensions
- Configurations
  - RA = Random Access
  - LD = Low Delay
  - AI = All-Intra

Note 1: For recent *subjective* test results (not including RExt and SCC), see the article in the January 2016 issue of *IEEE Trans. CSVT*

Note 2: Subjective gains tend to be larger than PSNR-based gains, when comparing AVC to HEVC

Gary J. Sullivan 2016-03-30

HEVC SCC versus AVC FRExt: Lossy (percentage bit-rate savings)

	AI			RA			LD		
	Y/G	U/B	V/R	Y/G	U/B	V/R	Y/G	U/B	V/R
RGB TGM	86.1%	83.5%	84.1%	80.4%	76.1%	77.4%	77.7%	73.0%	74.4%
RGB mixed content	80.1%	76.2%	76.0%	74.3%	68.2%	67.3%	69.7%	60.8%	59.8%
RGB Animation	52.4%	45.0%	40.1%	54.8%	49.5%	43.2%	56.4%	51.0%	43.1%
RGB Camera captured	58.4%	35.6%	44.3%	63.3%	42.6%	51.5%	60.1%	36.5%	48.2%
YUV TGM	74.6%	75.0%	77.0%	68.1%	70.4%	73.3%	65.4%	67.6%	70.5%
YUV Mixed content	63.6%	64.8%	64.8%	56.9%	63.2%	63.1%	51.5%	60.5%	60.6%
YUV Animation	23.4%	35.5%	29.3%	32.2%	48.0%	41.8%	39.0%	59.6%	54.6%
YUV Camera captured	26.5%	18.5%	25.4%	40.0%	42.2%	42.6%	39.8%	51.5%	53.6%
YUV 4:2:0 TGM	69.9%	64.4%	65.3%	62.3%	61.6%	62.6%	60.3%	58.7%	59.5%
YUV 4:2:0 Mixed content	57.4%	52.2%	53.1%	51.8%	51.7%	52.7%	47.0%	46.0%	47.7%
YUV 4:2:0 Animations	31.7%	33.3%	33.5%	36.4%	45.2%	44.2%	39.5%	45.0%	43.9%
Average	**66.5%**	**63.0%**	**64.3%**	**62.9%**	**62.2%**	**63.3%**	**60.6%**	**59.8%**	**61.2%**

HEVC SCC versus HEVC FRExt: Lossy (percentage bit-rate savings)

	AI			RA			LD		
	Y/G	U/B	V/R	Y/G	U/B	V/R	Y/G	U/B	V/R
RGB TGM	64.5%	60.9%	62.0%	56.9%	51.0%	53.1%	50.8%	42.9%	45.3%
RGB mixed content	54.8%	49.7%	49.5%	50.2%	42.4%	42.0%	41.7%	29.2%	28.2%
RGB Animation	26.3%	19.5%	16.8%	26.2%	17.3%	12.9%	24.4%	11.9%	5.5%
RGB Camera captured	25.6%	5.5%	10.4%	28.3%	5.8%	14.4%	26.1%	1.6%	11.9%
YUV TGM	57.4%	61.3%	62.8%	48.0%	52.6%	55.3%	40.5%	44.9%	47.4%
YUV Mixed content	45.2%	50.9%	50.8%	36.7%	45.1%	44.8%	23.8%	33.6%	33.3%
YUV Animation	1.2%	10.9%	7.6%	0.4%	10.1%	6.8%	0.0%	7.0%	4.9%
YUV Camera captured	0.4%	0.0%	0.2%	0.6%	0.2%	0.3%	0.6%	0.3%	0.2%
YUV 4:2:0 TGM	49.0%	49.3%	50.5%	39.4%	40.6%	42.2%	32.7%	33.2%	34.4%
YUV 4:2:0 Mixed content	36.6%	37.6%	37.6%	29.4%	31.2%	31.5%	18.0%	18.7%	19.5%
YUV 4:2:0 Animations	7.3%	11.7%	10.7%	3.8%	12.4%	9.9%	2.0%	8.2%	5.7%
Average	**44.7%**	**44.2%**	**45.0%**	**38.1%**	**37.6%**	**38.9%**	**31.5%**	**29.7%**	**31.0%**

HEVC SCC versus AVC FRExt: Lossless (percentage bit-rate savings)

	AI	RA	LD
RGB TGM	66.7%	59.1%	58.9%
RGB mixed content	44.4%	28.4%	26.6%
RGB Animation	20.9%	14.3%	13.0%
RGB Camera captured	6.9%	2.9%	2.8%
YUV TGM	53.5%	47.7%	47.4%
YUV Mixed content	29.3%	16.7%	14.9%
YUV Animation	5.0%	7.6%	6.2%
YUV Camera captured	0.7%	1.4%	1.4%
YUV 4:2:0 TGM	44.2%	39.8%	36.6%
YUV 4:2:0 Mixed content	26.9%	16.9%	14.1%
YUV 4:2:0 Animations	4.5%	9.6%	7.7%
Average	**39.5%**	**33.2%**	**31.8%**

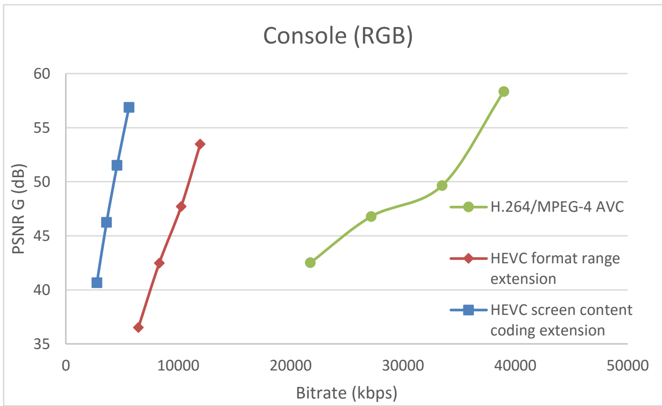
Gary J. Sullivan 2016-03-30

HEVC SCC versus HEVC FRExt: Lossless (percentage bit-rate savings)

	AI	RA	LD
RGB TGM	45.8%	35.2%	32.2%
RGB mixed content	24.3%	6.3%	3.9%
RGB Animation	4.4%	1.1%	1.1%
RGB Camera captured	0.2%	0.4%	0.4%
YUV TGM	46.7%	36.4%	33.3%
YUV Mixed content	23.9%	6.3%	3.8%
YUV Animation	1.7%	0.3%	0.3%
YUV Camera captured	0.0%	0.0%	0.0%
YUV 4:2:0 TGM	34.1%	23.9%	21.1%
YUV 4:2:0 Mixed content	21.6%	6.0%	3.5%
YUV 4:2:0 Animations	0.7%	0.2%	0.1%
Average	**29.0%**	**19.1%**	**16.9%**

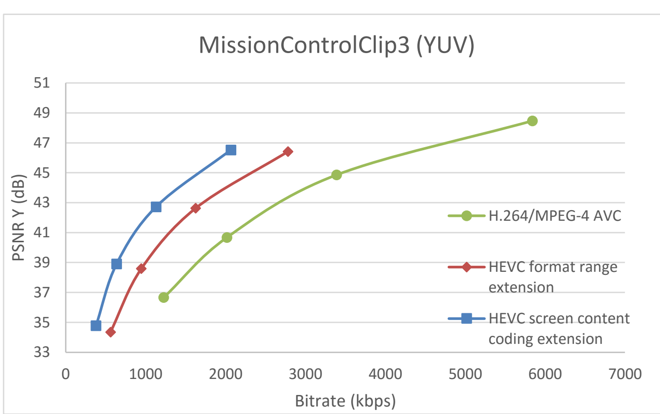
Gary J. Sullivan 2016-03-30

Example R-D Curve from test: for RGB purely screen-rendered sequence



Gary J. Sullivan 2016-03-30

Example R-D Curve from test: for YUV mixed camera & rendered sequence



Gary J. Sullivan 2016-03-30

14 New Profiles in 2016 version

- 4 Basic Screen-Extended profiles
  - 4:2:0 8 and 10 bit
  - 4:4:4 8 and 10 bit
- 3 Additional High-Throughput profiles using "wavefronts"
  - 4:4:4 8, 10, and 14 bit
- 3 Screen-Extended High-Throughput profiles
  - Corresponding to each of the added high-throughput profiles
- 4 Scalable Format Range Extensions profiles combining scalability and RExt
  - Scalable monochrome 8, 12, and 16 bit
  - Scalable 4:2:0 8 bit

High Dynamic Range and Wide Colour Gamut (2016/2017)

- Lots of recent interest
- Possibly more compelling than beyond-HD resolution
- Some key aspects and supplemental information:
  - SMPTE ST 2085 Perceptual Quantization (PQ) Transfer Function VUI
  - SMPTE ST 2086 Mastering Display Colour Volume SEI message
  - CEA 861.3 Content Light Level SEI message
  - ARIB B67 Hybrid Log-Gamma (HLG) Transfer Function VUI
  - New ITU-R Rec. BT.[HDR-TV] with PQ, HLG, and IC_TC_P
  - Alternative Transfer Characteristics SEI message
  - Colour Remapping Information SEI message
  - Ambient Viewing Environment SEI message
- Conclusion reached in February 2016: No new profiles needed for HDR (without backward compatibility considerations)
- Guideline development for 4:2:0 10 bit with PQ
- Further work on backward compatibility

Beyond today's HEVC: Further Compression Exploration

- Background work:
  - Brainstorming workshop October 2014
  - 4th-gen ITU-T VCEG "Key Technical Areas" (KTA) study Jan-Oct 2015
  - Seminar October 2015
- Formation of new "Joint Video Exploration Team" (JVET) Oct 2015
- New test model "Joint Exploration Model" (JEM) JEM 1
- JEM 2 Feb 2016
- Each aspect at most about 5% (most contributing <1%)
- Some have substantial increases in encoding and decoding time
- These are basically well-understood and straightforward techniques
- Varying tradeoffs of compression versus complexity

Gary J. Sullivan 2016-03-30

JEM 1 Compression Gain over HEVC

Random Access Configuration JEM 1.0 versus HM-16.6 Main 10			
Class	Y	U	V
Class A	21%	30%	24%
Class B	21%	13%	9%
Class C	21%	15%	18%
Class D	21%	10%	12%
Overall	**21%**	**17%**	**15%**

(JEM 2 perhaps 1% more.)

(This does *not* include the KLT feature.)

Gary J. Sullivan 2016-03-30

JEM 2 Enhanced Techniques (1 of 2)

- Larger coding tree units and larger transform blocks
- Quadree plus binary tree segmentation (separate branch)
- Intra prediction improvement
  - Intra mode coding with 67 prediction modes
  - Four-tap intra interpolation filter
  - Boundary prediction filters
  - Cross-component prediction
  - Position-dependent prediction combination
  - Adaptive reference sample smoothing
- Inter prediction improvement
  - Sub-PU based motion vector prediction
  - Adaptive motion vector resolution
  - Higher-precision motion vector storage
  - Overlapped block motion compensation
  - Local illumination compensation
  - Affine motion compensation prediction
  - Pattern-matched motion vector derivation
  - Bi-directional optical flow

Gary J. Sullivan 2016-03-30

JEM 2 Enhanced Techniques (2 of 2)

- Transform improvement
  - Adaptive multiple core transforms
  - Secondary transforms
  - Signal-dependent KLT transform (separate branch)
- Entropy coding improvement
  - Context model selection for transform coefficient levels
  - Multiple adaption rate probability estimation
  - Initialization for context models

Gary J. Sullivan 2016-03-30

“Type 1” (Royalty-Free Goal) Video in MPEG

- “Web Video Coding” (a.k.a. AVC Constrained Baseline)
  - MPEG-4 Part 29
  - Spec is derived from AVC
  - Finished (twice, sort of) in April 2014
- “Video Coding for Browsers” (a.k.a. Google’s VP8)
  - MPEG-4 Part 31
  - 2nd DIS ballot on hold pending resolution of patent declaration
- “Internet Video Coding”
  - CD ballot passed, now at DIS ballot stage
  - Coding efficiency was lagging but has improved
  - Tested as roughly on par with H.264/AVC in compression

Gary J. Sullivan 2016-03-30

Other recent standards activity

- Reference software, conformance, and verification testing
- Reconfigurable video coding (RVC)
- Explorations on free-viewpoint / light fields / VR / AR / point clouds / plenoptic
  - Video for virtual reality and augmented reality is becoming a reality
  - Microsoft Hololens, Oculus Rift, Samsung Gear VR, Google Cardboard
- Compact descriptors for visual search
  - MPEG-7 Part 13
  - Recently completed
- Compact descriptors for video analysis
  - Call for Proposals responses evaluated Feb 2016
  - Objectives, Applications and Use Cases
  - Requirements
  - Vision document
  - Results of responses to call reviewed and summarized
  - First experimentation model
  - Core experiments started
- Very low complexity & delay (VESA DisplayStream, JPEG XS)

Gary J. Sullivan 2016-03-30

Summary and outlook

- Very active work (many documents & people per meeting)
- Multiple versions and extensions (RExt, 3D/MVC, SHVC, SCC, etc.)
- Very diverse company & university participation
- Major technical advances
- Computational/implementation complexity is reasonable
  - For decoders, that is
  - Encoders can choose their degree of complexity
- Parallelism is an increased theme
- Deliverables:
  - Video coding specification
  - Reference software
  - Conformance testing specification
  - Verification testing
- Systems support for MPEG-2 TS, ISO BMFF, DASH, etc.
- Contact: JVT, JCT-VC, JCT-3V, VCEG, MPEG video chairs:
  - Gary J. Sullivan (garysull@microsoft.com)
  - Jens-Rainer Ohm (ohm@ient.rwth-aachen.de)

Gary J. Sullivan 2016-03-30

For further info - links

- Document archives and software are publicly accessible
  - http://hevc.info
  - http://jct-vc.org (http://www.itu.int/ITU-T/studygroups/com16/jct-vc/index.html)
  - http://jct-3v.org (http://www.itu.int/en/ITU-T/studygroups/com16/video/Pages/jct3v.aspx)
  - http://phenix.it-sudparis.eu/jct
  - http://phenix.it-sudparis.eu/jct3v
  - http://ftp3.itu.ch/av-arch/jctvc-site
  - http://ftp3.itu.ch/av-arch/jct3v-site

Gary J. Sullivan 2016-03-30

For further info: some publications

- Publications
  - HEVC books by Wien and by Sze, Budagavi & Sullivan
  - "Special Issue on Emerging Research and Standards in Next Generation Video Coding (HEVC)", *IEEE T-CSVT*, Dec. 2012 (includes technical overview paper, compression capability analysis paper, complexity analysis paper, and others)
  - Nutshell article in *IEEE Commun. Magazine,* Jan. 2013.
  - "Standardized Extensions of High Efficiency Video Coding", *IEEE Journal on Selected Topics in Signal Processing*, Vol. 7, no. 6, pp. 1001–1016, Dec. 2013
  - "Performance analysis of HEVC and its format range and screen content coding extensions", *SPIE Applications of Digital Image Proc. XXXVII,* Aug. 2015
  - January 2016 special issue of *IEEE T-CSVT*, including papers on
    - Overview of HEVC Format Range Extensions
    - Overview of Scalable HEVC (SHVC)
    - Overview of Multiview and 3D HEVC
    - Overview of HEVC Screen Content Coding
    - Status of work on HDR for HEVC
    - Subjective verification test results (incl. 4K)

Gary J. Sullivan 2016-03-30

Why Royalty-Free Video Coding

- Licensing terms of recent standards are unpredictable.
- Several patent holders stay outside patent pools.
- Large volume distribution of soft clients is desired.

Arild Fuldseth
Principal Engineer, Video Coding
Cisco Systems

Royalty-Free Initiatives

- MPEG (WVC, VCB, IVC)
- IETF/netvc
  - Daala (Mozilla)
  - Thor (Cisco)
- Alliance for Open Media (AOM)
  - Daala (Mozilla)
  - Thor (Cisco)
  - VP9/VP10 (Google)

Considerations when Designing a Royalty-Free Video Codec

- Patents usually expire after 20 years.
- Obvious extensions to existing video codec technology can be used, for instance:
  - Larger block sizes
  - Better interpolations filters
- Most overly general claims overlap with prior art.
- Some patents essential to major video standards are easy to avoid without significant impact on performance.

Designing Around Patents

Next Generation Video Codec

- Joint Video Exploration Team (JVET) established October 2015 – JEM SW.
  - Initial SW (HMKTA) had 10% improvement over HEVC (February 2015).

Sequence Class	All Intra			Random Access		
	Y	U	V	Y	U	V
4kx2k	-23.6%	-21.1%	-19.9%	-36.6%	-37.6%	-39.4%
1080p	-22.7%	-22.1%	-21.7%	-39.8%	-37.5%	-37.0%
WVGA (832 x 480)	-19.7%	-20.9%	-21.1%	-30.3%	-32.1%	-32.1%
WQVGA (416x240)	-16.4%	-17.0%	-17.7%	-28.0%	-31.2%	-33.3%
720p	-28.8%	-27.1%	-27.1%			
Average	**-21.9%**	**-21.4%**	**-21.2%**	**-34.0%**	**-34.8%**	**-35.6%**

Rate reduction of HM vs. JM (HEVC vs. AVC)
- February 2013.

Sequence Class	All Intra			Random Access		
	Y	U	V	Y	U	V
4kx2k	-15.4%	-23.5%	-20.1%	-20.8%	-29.9%	-23.8%
1080p	-13.8%	-8.8%	-6.2%	-21.3%	-13.2%	-9.2%
WVGA (832 x 480)	-14.8%	-11.8%	-14.9%	-20.6%	-14.8%	-18.2%
WQVGA (416x240)	-11.8%	-7.9%	-9.4%	-20.5%	-9.8%	-12.1%
720p	-15.7%	-12.0%	-14.4%			
Average	**-14.2%**	**-12.6%**	**-12.6%**	**-20.8%**	**-16.7%**	**-15.4%**

Rate reduction of JEM vs. HM
- February 2016.

Marta Karczewicz, VP, Technology, Video R&D and Standards – Qualcomm

Tools Performance – "Tool On" Results

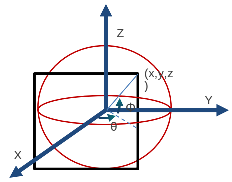
Coding Tools	Rate Reduction (over HEVC) All Intra / Random Access	
Large CTU up to 256x256	-0.2% / -0.7%	
Adaptive loop filter	-2.8% / -4.6%	
Cross component linear model prediction	-2.6% / -1.4%	
Sub-PU level motion merge	-1.7%	
Locally adaptive motion vector resolution	-0.8%	
Overlapped motion vector compensation	-1.9%	
Frame rate up-conversion mode	-4.5%	-12.5%
Bidirectional optical flow	-2.4%	
Affine motion vector prediction	-0.9%	
Local illumination compensation	-0.3%	
64x64 transform	-0.3% / -0.4%	
Adaptive multiple core transform (AMT)	-2.8% / -2.4%	-4.8%
Non-separable secondary transform (NSST)	-3.3% / -1.7%	

Tools Performance – "Tool On" Results

Coding Tools	Rate Reduction (over HEVC) All Intra / Random Access	
Position dependent intra prediction	-1.5% / -0.8%	
Reference sample adaptive filtering	-1.0% / -0.4%	
Additional boundary filter for intra prediction	-0.2% / -0.1%	
4-tap interpolation fro intra prediction	-0.4% / -0.2%	-1.7%
65 angular intra prediction directions	-0.7% / -0.2%	
Neighborhood transform coeff. context modelling	-0.9% / -0.6%	
Temporal CABAC initialization	NA / -0.2%	
CABAC update with adaptive window sizes	-0.6% / -0.3%	-1.5%
CABAC probability estimation with 2 windows	-0.7% / -0.4%	

- Sum of rate reduction for all tools off is 27.2% (JEM rate reduction 20.8%).
- Tools and improvements in EEs promising further improvements.
  - Quadtree Plus Binary Tree structure (QTBT): ~5%.
  - ALF improvements: ~1%.

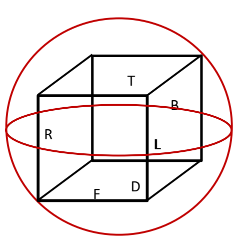
360 Video

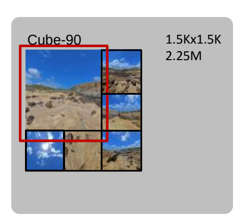


- Head Mounted Displays:
  - 2.0-2.7” on display, 490-577 ppi.
  - Assuming ~90^0 Field-of-View and equirectangular projection: ~3840x1920 resolution for 360 view.
- HVS can distinguish up to 60 pixels per degree of FOV (5400x5400 resolution per eye) **- VR-Retina would require 21600 x 10800 video**!
- Frame rate - 60fps and higher.
- High bandwidth requirement.
  - 50-60Mbps for Blu-ray 4K content.

360 Video

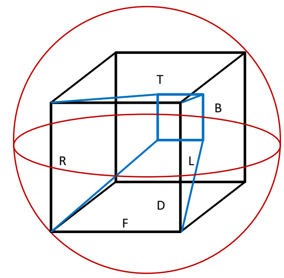
- Directional geometry/coding schemes.
  - Field-of-view coded at high resolution/quality.
  - Rest of the 360 view represented with lower resolution/quality.
  - Lower bandwidth.
  - Reduced pixel count (reduced decoding complexity).

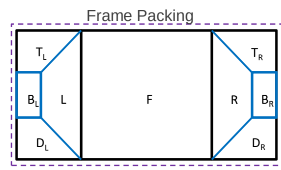


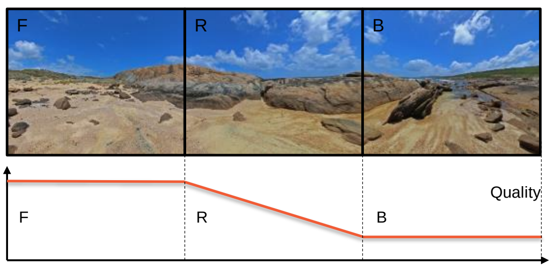


360 Video

- Sphere is projected on truncated square pyramid.
  - Pyramid faces are warped to fit into packed frame structure.
  - Pixel count reduction by 75% compared to equirectangular projection.
  - Smoothly transitions from front face (high quality) towards back face (lowest quality).





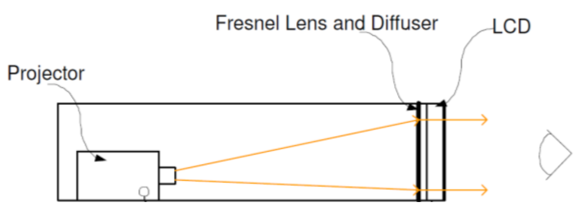


Using HEVC for HDR

Jacob Ström, Principal Researcher
Ericsson Research

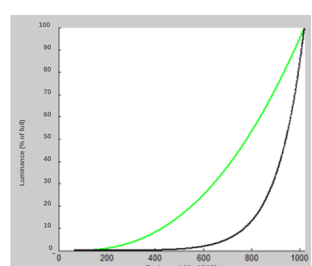
Personal reflection

- Emerging Technologies at SIGGRAPH 2003: High-Dynamic-Range Display System
- First reflections: I had thought HDR was for creating beautiful images you could shown on regular displays, not a capability of a display
- Second reflection: I want that TV in my home.



Differences when coding HDR data

- Gamma function more non-linear
- Several knock-on effects:
  - Subsampling can give luminance artifacts
    - Can be avoided by new subsampling procedure (Luma Adjustment)
  - Chroma values cluster around 0 more than for SDR data
    - Can be counteracted by encoder optimization (Chroma QP offset)
  - More bits spent in dark regions of image
    - Can be counteracted by encoder optimization (Luma Delta QP)



Subsampling artifacts

original 4:4:4

subsampled to 4:2:0
(no compression!)

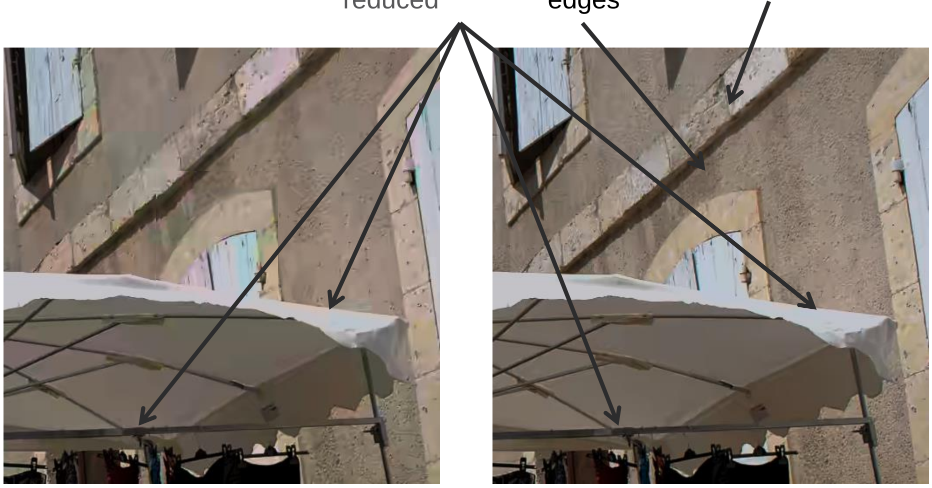
Sequence courtesy of Technicolor and the NevEx project

Examples

color artifacts reduced

sharper edges

more texture detail



Before improvement

After improvement

Sequence courtesy of Technicolor and the NevEx project

Conclusion

- HDR looks really great and I want it on my TV yesterday
- In order to get good quality out of HEVC for HDR data you need a slightly different encoder configuration than for SDR, but the decoder doesn't need to change.

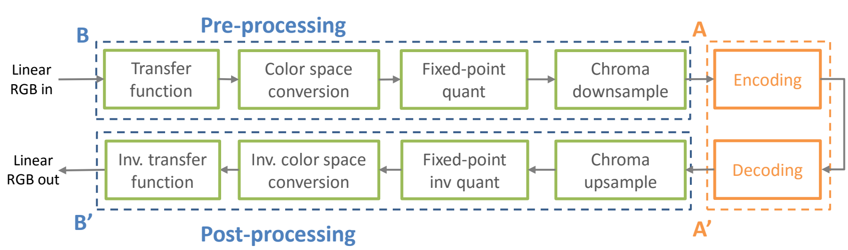
Trends for future video codec

Yan Ye, Senior Manager
InterDigital Communications

Outline

- Topic #1: High Dynamic Range and Wide Color Gamut
- Topic #2: 360 video for Virtual Reality
- Discussion

High Dynamic Range and Wide Color Gamut Video

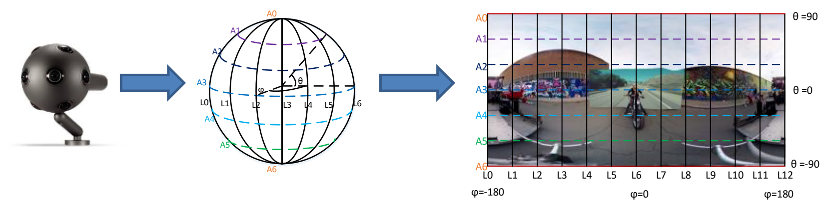


- Conventional codec boundary is between A and A'
- Input and output are fixed-point non-linear signal, often in 4:2:0 Y'Cb'Cr' format

- HDR/WCG work considered the end-to-end pipeline, including pre- and post-processing
- Input and output are floating-point linear RGB signal

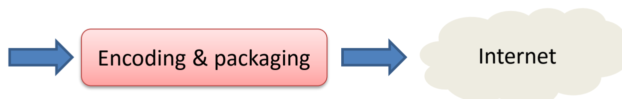
46

360 Video for Virtual Reality: pre-processing and encoding



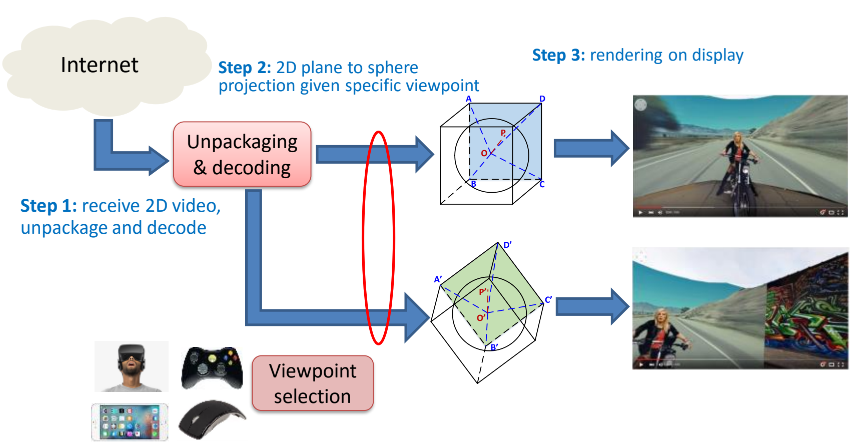
Step 1: multi-camera array captures video, then image stitching is applied to obtain spherical video

Step 2: spherical video is “unfolded” to 2D plane, e.g. using the equirectangular projection



Step 3: 2D video coding, packaging and delivery

360 Video for Virtual Reality: decoding, post-processing and rendering



Discussion

- More pre- and post-processing are being applied in the end-to-end system (camera to display)
- Such processing has become more closely related with the codec, esp. with VR video
- Support for more end-user control is being demanded

Pre- and post-processing have traditionally been left out of scope for video codec standards. Should that still be the case for the future?

In addition to high compression efficiency, the future video codec also needs to provide more functionality support, e.g., ease of perspective extraction without loss of coding efficiency.

49