Versatile Video Coding – Algorithms and Specification

IEEE ICME, 10.07.2020

Mathias Wien, Lehrstuhl für Bildverarbeitung, RWTH Aachen University
Benjamin Bross, Heinrich Hertz Institute, Fraunhofer Gesellschaft

Heinrich Hertz Institute

Outline I

Outline II

Outline III

- Intra Sub-Partitions
- Matrix weighted Intra Prediction

8. Inter Prediction
   - Extended motion vector prediction
   - Symmetric motion vector difference coding
   - Extended merge mode
   - Merge with motion vector difference
   - History-based Motion Vector Prediction
   - Affine motion compensated prediction
   - Subblock-based temporal motion vector prediction
   - Adaptive motion vector resolution
   - Motion field storage
   - Bi-prediction with CU-level weights
   - Bi-directional optical flow
   - Decoder side motion vector refinement

Outline IV

Outline V

Versatile Video Coding – Algorithms and Specification

Part 1 | Introduction

IEEE ICME, 10.07.2020

Mathias Wien, Lehrstuhl für Bildverarbeitung, RWTH Aachen University
Benjamin Bross, Heinrich Hertz Institute, Fraunhofer Gesellschaft

Fraunhofer Heinrich Hertz Institute | RWTH AACHEN UNIVERSITY

Video coding standardization organisations

- **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
  (Continuosly active, currently being restructured inside SC29)
- **ITU-T VCEG = “Video Coding Experts Group”**
  ITU-T SG16/Q6 = International Telecommunications Union – Telecommunications Standardization Sector, Study Group 16, Working Party 3, Question 6
- **JVT = “Joint Video Team”** collaborative team of MPEG & VCEG, responsible for developing AVC (discontinued in 2009)
- **JCT-VC = “Joint Collaborative Team on Video Coding”** team of MPEG & VCEG, responsible for developing HEVC (established January 2010)
- **JVET = “Joint Video Experts Team”** exploring potential for new technology beyond HEVC (established Oct. 2015 as Joint Video Exploration Team, renamed Apr. 2018)

History of Video Coding Standards

- ITU-T H.120 [1] (not used much)
  - "Codecs for video-conferencing using primary digital group transmission"
  - First video conferencing specification
  - DPCM structure
  - Conditional replenishment

History of Video Coding Standards

- ITU-T H.261 [2]
  - “Video codec for audiovisual services at $p \times 64$ kbit/s”
  - Video conferencing (broad application)
  - First-time hybrid coding scheme
    Motion compensated prediction, transform
  - CIF resolution (Common Intermediate Format, 352×288 in Europe)

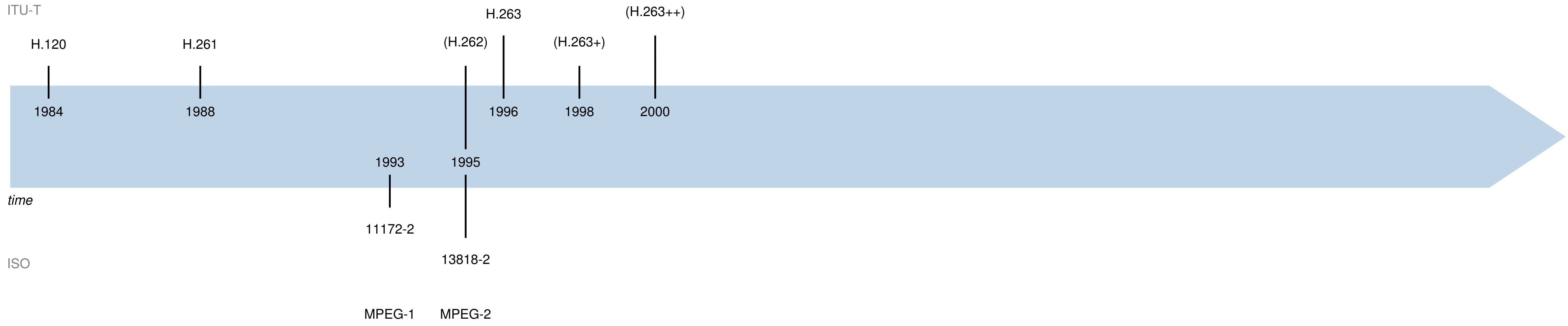
History of Video Coding Standards

- ISO 11172-2 MPEG-1 [3]
  - “Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s – Part 2: Video"
  - Video CD – first distribution of digital video format
  - Hybrid coding scheme with bi-prediction, “D-frames” for search
  - Audio layer-3: “mp3” format

History of Video Coding Standards

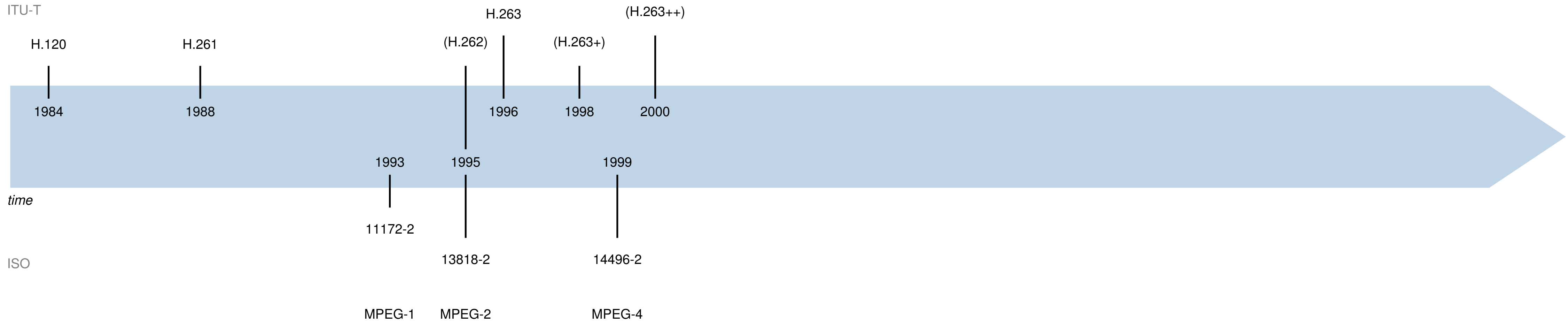
- ISO 13818-2 MPEG-2 [4]
  - “Information technology – Generic coding of moving pictures and associated audio information – Part 2: Video”
  - Digital TV, DVD, DVB-x – Widespread usage of digital video format
  - Support for interlaced video format (full SD resolution)
  - First standard supporting scalability features

History of Video Coding Standards



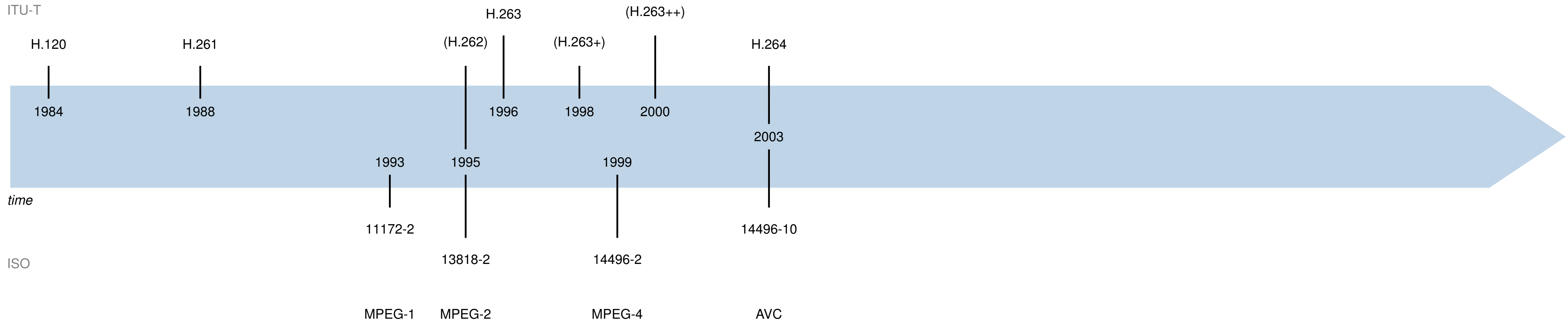
- ITU-T H.263 [5]
  - "Video coding for low bit rate communication"
  - Video communication applications
  - Improved compression performance
  - Large set of extensions (18 Annexes)
    Organization into profiles

History of Video Coding Standards



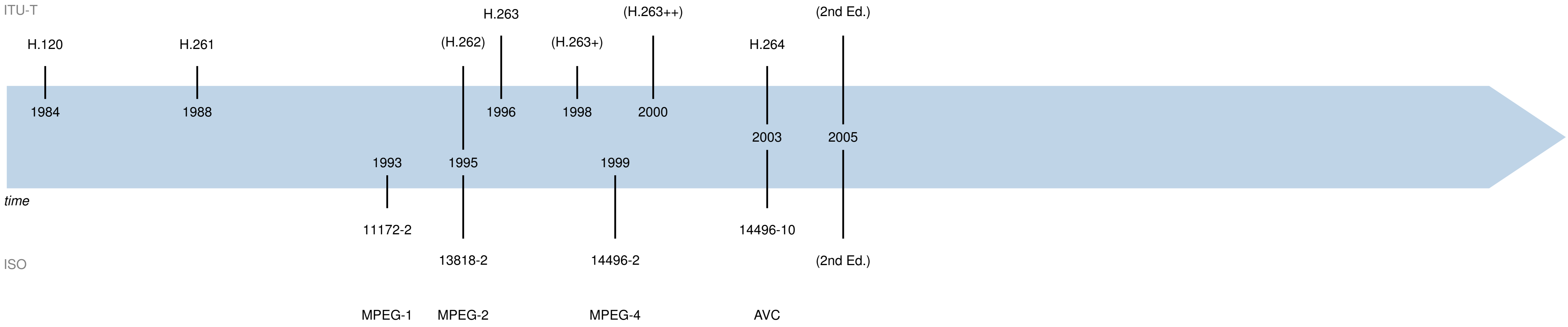
- ISO 14496-2 MPEG-4 [6]
  - “Information technology – Coding of audio-visual objects – Part 2: Visual”
  - Multimedia object representation
  - Basically H.263 coding tools (plus quarter sample, global motion)
  - Arbitrarily shaped objects
    mostly used profile: Advanced simple profile

History of Video Coding Standards



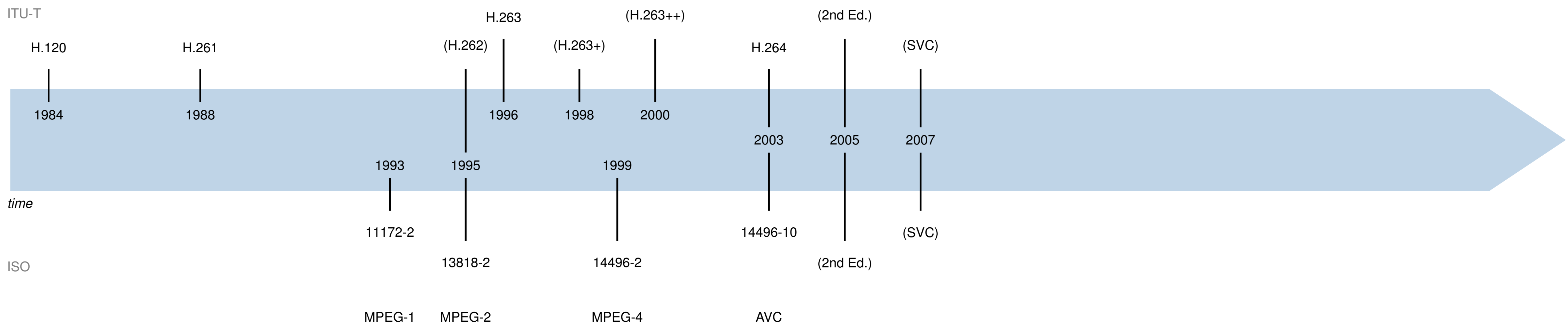
- AVC (ITU-T H.264, ISO 14496-10) [7, 8]
  - Advanced Video Coding
  - HDTV distribution, YouTube, cameras (AVCHD), conferencing, . . .
  - Bit-exact specification, 4×4 integer transform
    in-loop filter, low-complex arithmetic coding
  - High compression performance

History of Video Coding Standards



- AVC (ITU-T H.264, ISO 14496-10) [7, 8]
  - Advanced Video Coding (High Profiles)
  - HDTV distribution, YouTube, cameras (AVCHD), conferencing, ...
  - Improvements due to competitors (e. g. VC-1)
  - Additional 8×8 transform
  - Additional color spaces

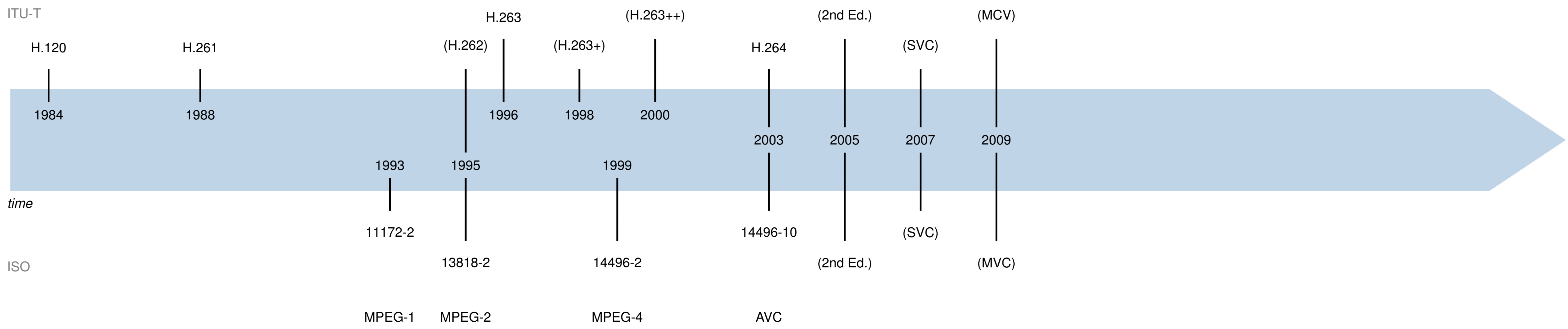
History of Video Coding Standards



- AVC (ITU-T H.264, ISO 14496-10) [7, 8]
  - Scalable Video Coding (extension)
  - General purpose, used in video conferencing applications (Vidyo)
  - Temporal, quality, spatial scalability
  - Single-loop decodability

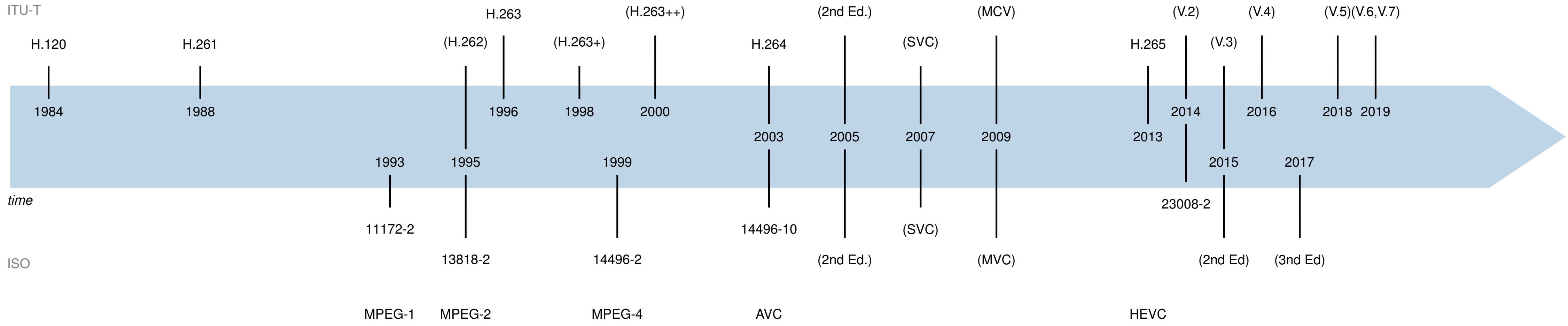
Fraunhofer Heinrich Hertz Institute
RWTH AACHEN UNIVERSITY

History of Video Coding Standards



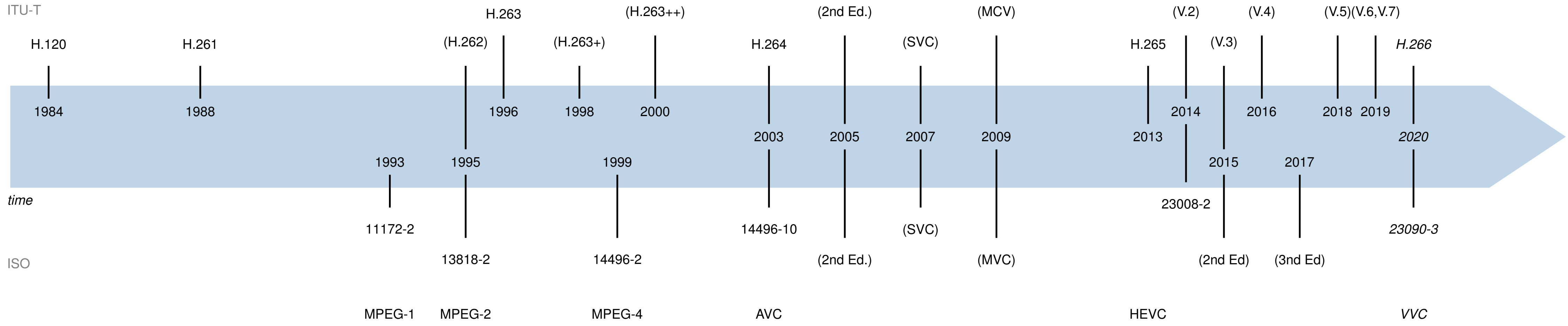
- AVC (ITU-T H.264, ISO 14496-10) [7, 8]
  - Multiview video coding (extension)
  - Stereo and 3D video applications (Blu-Ray)
  - Two or more parallel views of the scene
  - No modifications on tool level

History of Video Coding Standards



- HEVC (ITU-T H.265, ISO 23008-2) [9, 10]
  - High efficiency video coding, general purpose, HD to UHD resolutions
  - Improved compression performance for application space
  - Edition 2 (10/2014): Range extensions, multiview, scalability; Edition 3 (04/2015): 3D video coding; Edition 4 (12/2016): Screen Content Coding; Edition 5 (02/2018): Omnidirectional Video SEI; Edition 6 (06/2019): SEI manifest; Edition 7 (11/2019): Fisheye and Annotated Regions SEIs

History of Video Coding Standards



- VVC (ITU-T H.266, ISO 23090-3) [11]
  - Versatile video coding, general purpose, UHD and larger resolutions
  - Extended application space ($360^\circ, \ldots$)
  - Improved compression performance for application space
  - Based on the hybrid coding scheme

Motivation for improved video compression

- **Video is continually increasing by resolution**
  - HD existing, UHD (4K$\times$2K, 8K$\times$4K) appearing
  - Mobile services going towards HD/UHD
  - Stereo, multi-view, 360° video
- **Devices available to record and display ultra-high resolutions**
  - Becoming affordable for home and mobile consumers
- **Video has multiple dimensions to grow the data rate**
  - Frame resolution, Temporal resolution
  - Color resolution, bit depth
  - Multi-view
  - Visible distortion still an issue with existing networks
- **Necessary video data rate grows faster than feasible network transport capacities**
  - Better video compression (than current HEVC) needed in next decade, even after availability of 5G

Wide Color Gamut / High Dynamic Range Coding

- Color space
  - Standard Dynamic Range (SDR) video
    - Contrast approx. $1\,000 : 0$
    - ITU-R BT.709 colour space [12]
  - High Dynamic Range (HDR) video
    - Contrast approx. $1\,000\,000 : 0$
    - ITU-R BT.2020 colour space [13]
    - ITU-R BT.2100 colour space [14]

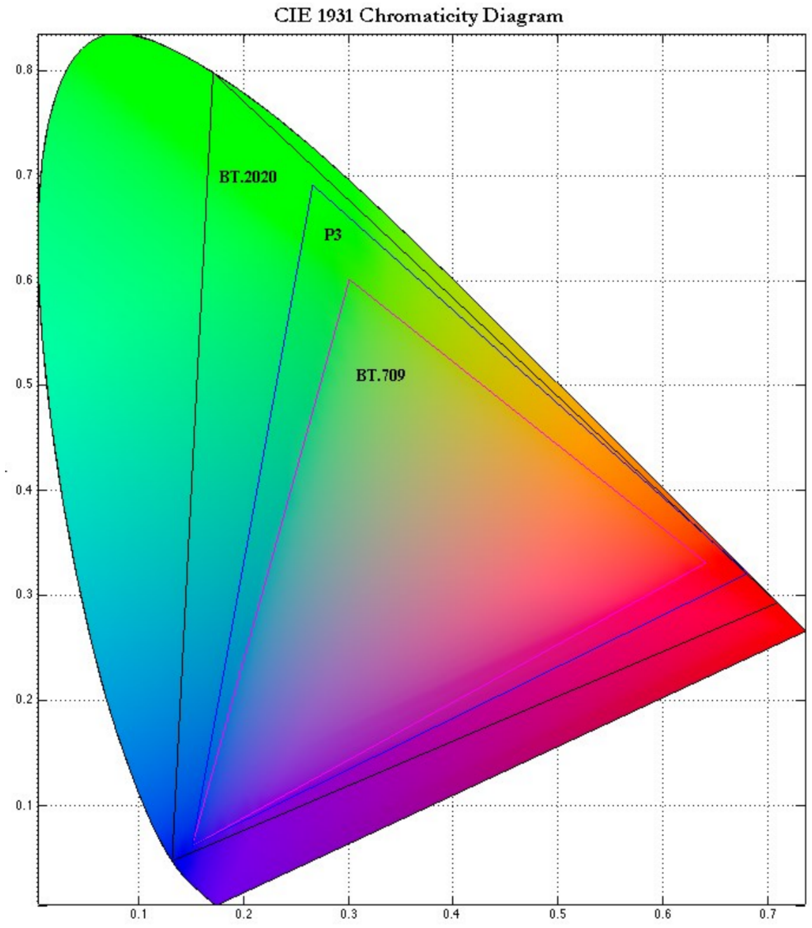


Figure from N15084 [15]

Wide Color Gamut / High Dynamic Range Coding

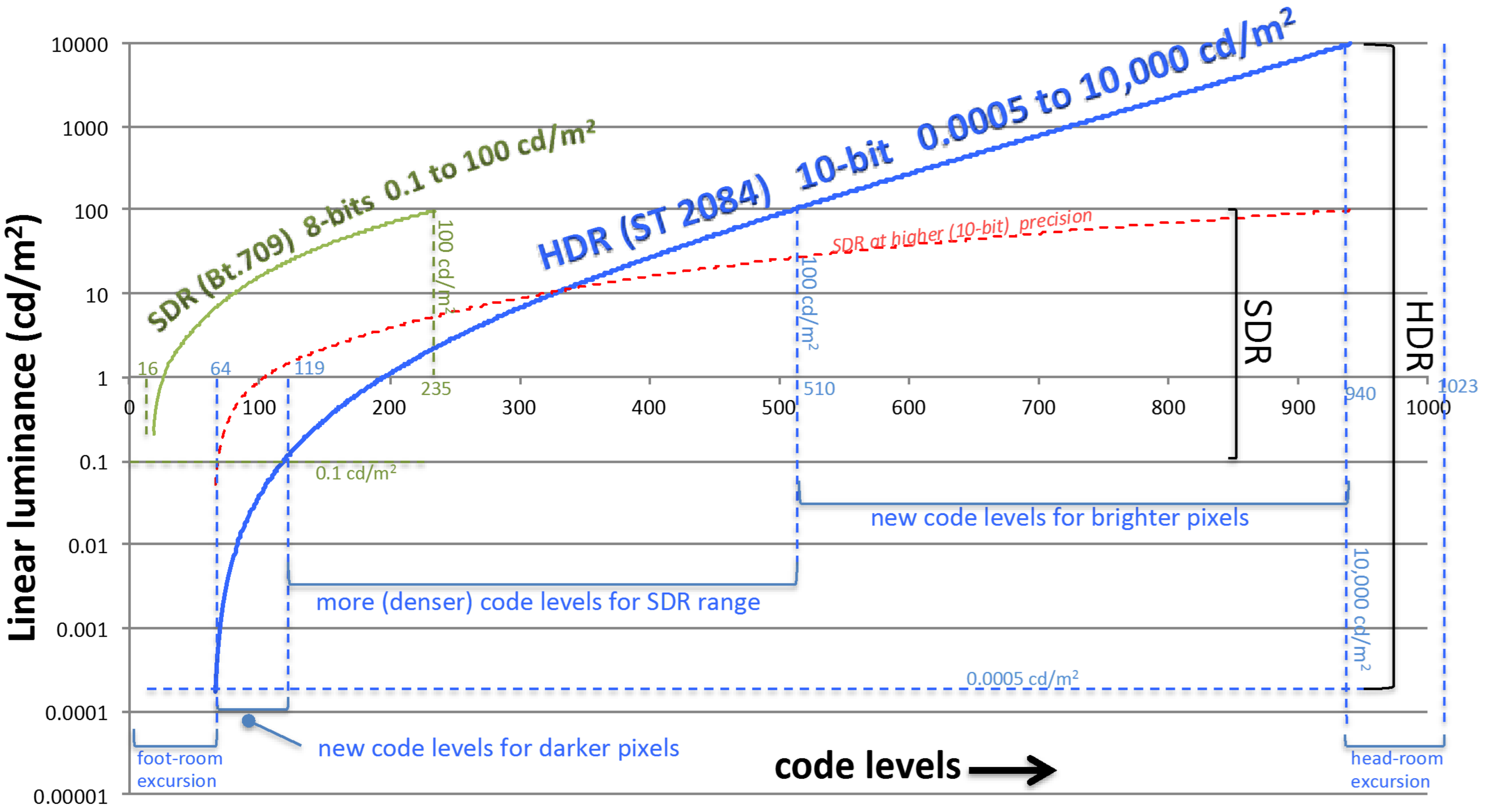


Figure from N15084 [15]

Fraunhofer Heinrich Hertz Institute
RWTH AACHEN UNIVERSITY

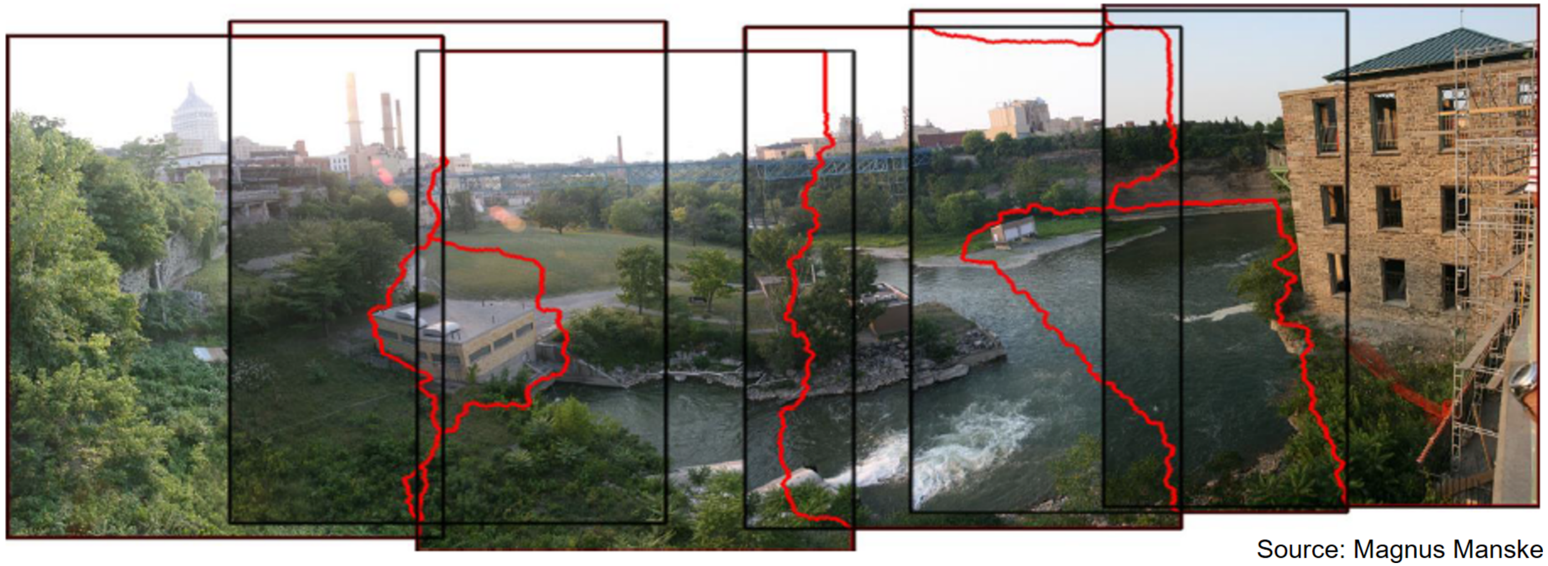
VR / 360° Video Acquisition

Omnidirectional Capture → Panorama Stitching → Packing → Encoding → Decoding → Viewport Rendering

- New omnidirectional cameras allow acquiring panoramic video (by mosaic stitching)
- Appropriate rendering to a head mounted display allows adapting the viewpoint according to head movements in real-time
- With appropriate projection, the panorama can be packed into a 2D frame
- Large resolution $\geq 4\mathrm{K} \ldots 8\mathrm{K}$ required for appropriate resolution of viewport

VR / 360°: Stitching

- Stitching requires registration
  - Identification of matching key points
  - Geometric warping of pictures
- Optimum stitching path can be based on
  - Minimum sample difference
  - Depth cues (for appropriate occlusion handling)
- Some blending / filtering / hole filling may be necessary to mask artifacts
- In video: Avoid temporal variation of stitching path



Source: Magnus Manske

360° Projection Formats: Cube Map Projection (CMP)

- Examples of projection formats: Cubemap with 3x2 packing
- 6 Faces can be treated as rectangular video

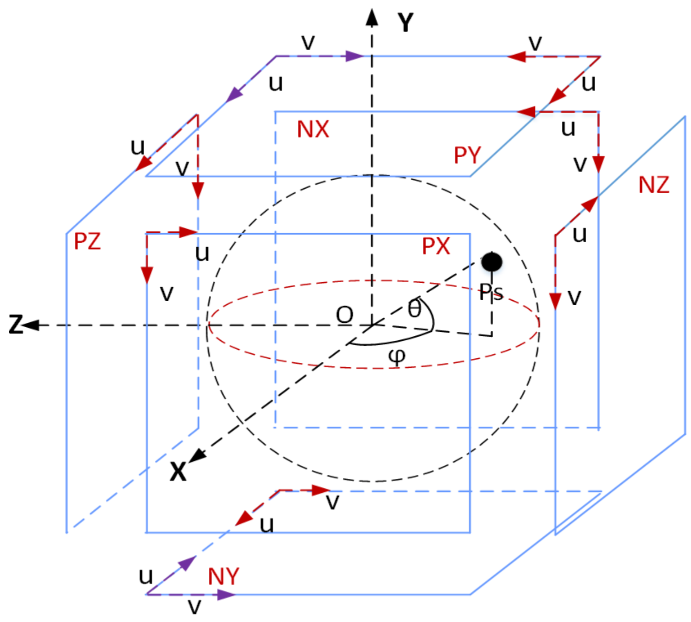


Figure from JVET-G1003 [16]

360° Projection Formats: Equirectangular Projection (ERP)

- Examples of projection formats: Equirectangular
- The whole sphere is projected into a rectangular picture
- Extreme geometric distortions, in particular at the poles
- Non-uniform sampling inherent

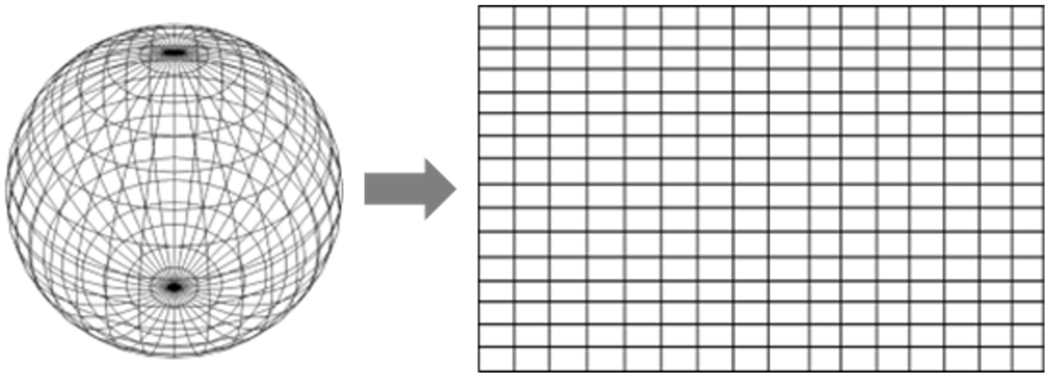


Figure from JVET-C0050 [17]

Contents

Steps towards next generation standard – Versatile Video Coding (VVC)

- **Experimental software "Joint Exploration Model" (JEM) developed by JVET**
  – Intended to investigate potential for better compression beyond HEVC
  – Was initially started extending HEVC software by additional compression tools, or replace existing tools
  – First JEM version Oct. 2015, JVET-A1001 [18]

- **Substantial benefit was shown over HEVC, both in subjective quality and objective metrics**
  – Proven in "Call for Evidence" (July 2017) JVET-F1002, JVET-G1004 [19, 20]
  – JEM was however not designed for becoming a standard (regarding all design tradeoffs)
  – Call for Proposals was issued by MPEG and VCEG (October 2017), JVET-H1002 [21]

- **Call for Proposals very successful (responses received by April 2018)**
  – 32 companies in 21 proponent groups responded
  – 46 category-specific submissions: 22 in SDR, 12 each in HDR and 360° video
  – All responses clearly better than HEVC, some evidently better than JEM
  – This marked the starting point for VVC development

Fraunhofer Heinrich Hertz Institute

Steps towards next generation standard – Versatile Video Coding (VVC)

- **What does "Versatile" stand for?**
- **VVC should be usable for many types of data**
  - SDR and HDR up to extreme high resolutions
  - All kind of camera generated content
  - Computer generated content
  - Non-camera video modalities e. g. medical data
  - 360°, lightfield, depth, and volumetric video
- **VVC should support flexible random and localized access**
- **VVC should be easily configurable for various application needs**
- **The core of VVC should consist of minimum amount of necessary and well-understood building blocks**

Joint Call for Proposals (CfP) on Video Compression with Capability beyond HEVC

- Document JVET-H1002 [21]

- **Test categories**
  - Standard dynamic range (SDR): 5 UHD and 5 HD sequences
  - High dynamic range (HDR): 3 HLG and 5 PQ sequences
  - 360° video (360): 5 sequences in ERP format
- **Constraint sets**
  - Constraint set 1 (C1): Random access configuration
    - Max 1.1s random access intervals, structural delay max 16 pictures
  - Constraint set 2 (C2): Low delay configuration only evaluated for SDR HD sequences
    - No picture reordering between input and output
- **Encoding constraints**
  - No pre-processing, post-processing only within the coding loop
  - Static quantizer setting with one-time change to meet target bitrate
  - Relevant optimization methods to be reported

VVC CfP Test Sequences

- SDR-A: 3840 × 2160
- SDR-B: 1920 × 1080
- HDR (PQ HD, HLG 4K)
- 360 Video (8K, 6K)


FoodMarket4 60p


CatRobot1 60p


DaylightRoad2 60p


ParkRunning3 50p


Campfire 30p


BasketballDrive 50p


Cactus 50p


BQTerrace 60p


RitualDance 60p


MarketPlace 60p


Market3 HD50p


Hurdles HD50p


Starting HD50p


ShowGirls2 HD25p


Cosmos1 HD24p


DayStreet 60p


PeopleInShop...


SunsetBeach 60p


ChairliftRide 30p


KiteFlite 30p


Harbor 30p


Trolley 30p


Balboa 60p

CfP Performance Analysis

- **Submissions had to provide coded/decoded sequences**
  - 4 rate points each, two constraint conditions "low delay" (LD) and "random access" (RA)
  - SDR: 5$\times$ HD (both LD and RA), 5$\times$ UHD-4K (only RA)
  - HDR: 5$\times$ HD (PQ grading), 3x UHD-4K (HLG grading)
  - 360°: 5 sequences 6K/8K for the full panorama
- **Double stimulus test with two hidden anchors HEVC-HM and JEM**
  - Rate points defined with lowest rate was typically less than "fair" quality for HEVC, but still possible to code
  - Quality was judged to be distinguishable when confidence intervals were non-overlapping
- **Evaluation:** Three ways of judging benefit:
  - Mean MOS over all test cases (28×4 test points: 23×4 C1, 5×4 C2)
  - Count cases where a proposal was visually better/worse than JEM
  - Count cases where a proposal was visually better than HEVC (HEVC at higher rate point)
- **Reports:** Input subjective test JVET-J0080 [45], output CfP results JVET-J1003 [46]

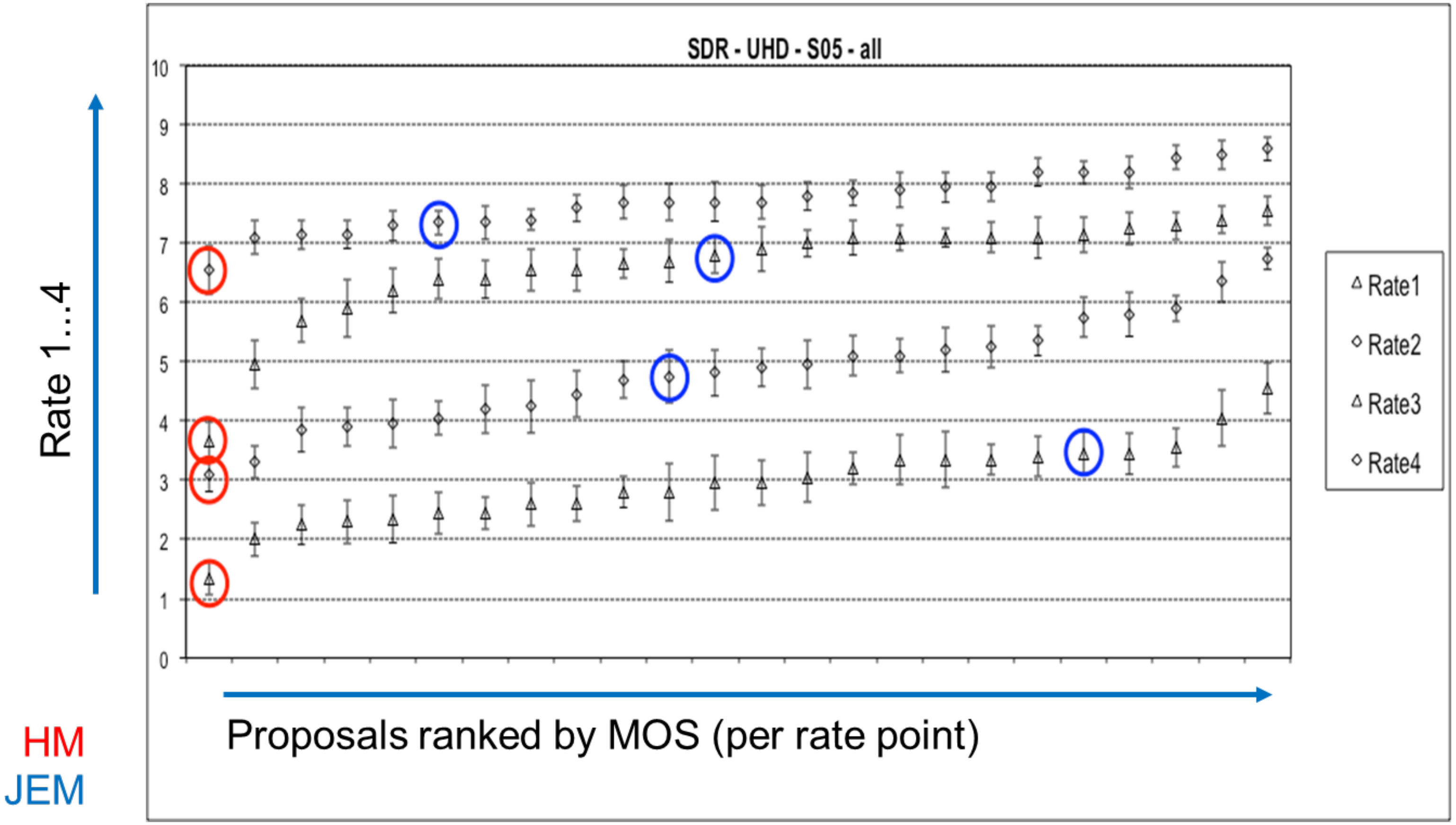
CfP Performance Analysis

- **Measured by objective performance (PSNR)**, best performers report $> 40\%$ bit rate reduction compared to HEVC, $> 10\%$ compared to JEM (for SDR case)
  - Similar ranges for HDR and 360°
  - Obviously, proposals with more elements show better performance
  - Some proposals showed similar performance as JEM with significant complexity/run time reduction
  - 2 proposals used some degree of subjective optimization, not measurable by PSNR

  Results of subjective tests generally show similar (or even better) tendency
  - Benefit over HEVC very clear
  - Benefit over JEM visible at various points
  - Proposals with subjective optimization also showing benefit in some cases

CfP Performance Analysis

- JVET-J1003 [46]: Report of subjective evaluation contains 28 plots as shown, one per sequence
- Count significant cases of positive/negative benefit with non-overlapping confidence interval against JEM



CfP Performance Analysis

- "Mean" and "significance-count" method suggested at least 7 proposals that were obviously better than JEM

Mean MOS

Pxx	6,53	Pnn	6,04
Pxx	6,46	Pnn	6,04
Pxx	6,41	Pnn	6,03
Pxx	6,37	Pnn	6,03
Pxx	6,33	Pnn	6,01
Pxx	6,33	JEM	6,01
Pxx	6,26	Pnn	6,00
Pnn	6,23	Pnn	5,96
Pnn	6,17	Pnn	5,94
Pnn	6,15	Pnn	5,88
Pnn	6,13	Pnn	5,86
Pnn	6,11	HM	4,57

–60 ... +60

Significance vs. JEM

Pxx	10	Pnn	1
Pxx	8	JEM	0
Pxx	8	Pnn	0
Pxx	6	Pnn	-1
Pxx	6	Pnn	-1
Pxx	6	Pnn	-1
Pxx	6	Pnn	-2
Pnn	3	Pnn	-2
Pnn	3	Pnn	-2
Pnn	2	Pnn	-3
Pnn	2	Pnn	-4
Pnn	1	HM	-36

Fraunhofer Heinrich Hertz Institute

CfP Performance Analysis

- Similar tendency in HDR and 360° categories
- Mostly same coding tools as in SDR provide good benefit

HDR **–32 ... +32**

Mean MOS		Signif. vs. JEM	
Pxx	6,04	Pxx	7
Pxx	6,00	Pxx	3
Pxx	5,94	Pxx	2
Pxx	5,93	Pxx	2
Pxx	5,86	Pxx	2
Pnn	5,85	Pnn	1
Pnn	5,80	Pnn	1
Pnn	5,67	JEM	0
JEM	5,62	Pnn	0
Pnn	5,60	Pnn	0
Pnn	5,59	Pnn	-1
Pnn	5,45	Pnn	-1
Pnn	5,11	Pnn	-6
HM	4,14	HM	-20

360° **–20 ... +20**

Mean MOS		Signif. vs. JEM	
Pxx	6,20	Pxx	9
Pxx	6,19	Pxx	9
Pxx	6,06	Pxx	8
Pxx	6,03	Pnn	7
Pxx	5,99	Pxx	7
Pxx	5,96	Pxx	6
Pxx	5,86	Pxx	5
Pnn	5,69	Pxx	4
Pnn	5,67	Pnn	2
Pnn	5,51	Pnn	1
Pnn	5,45	Pnn	1
JEM	5,11	JEM	0
HM	3,79	HM	-9
Pnn	3,45	Pnn	-12

CfP Performance Analysis: Comparison to HEVC

- The subjective quality of best performing proposals is always equal or sometimes better ($1/3$ of cases) than HEVC at next higher rate point, over all categories (with approx. 40% less rate)
- The subjective quality of best performing proposals is always equal or sometimes better ($1/5$ of cases) than HEVC at 2nd higher rate point, in SDR-UHD category (with approx. 65% less rate)
- Though it is not always the same proposal that performs best at a given rate point, it can be anticipated that merits of different proposals can be combined
- 50% (or more) bit rate reduction with same quality will probably be achievable by the new standard

Versatile Video Coding – Algorithms and Specification

Part 2 | Coding Structures and High Level Syntax

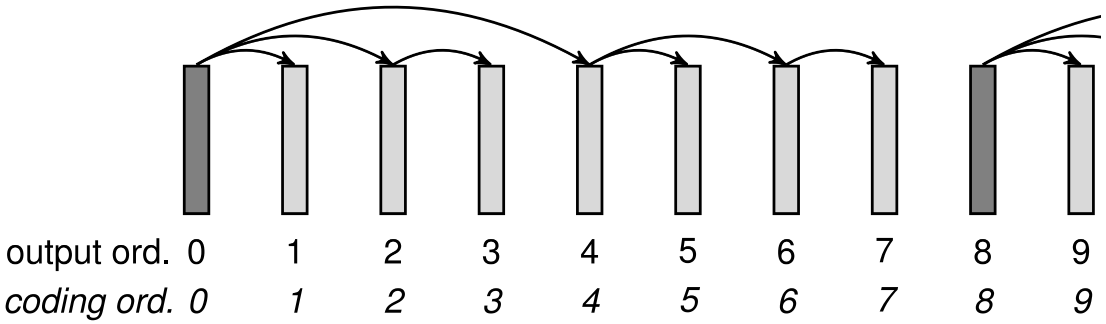
IEEE ICME, 10.07.2020

Mathias Wien, Lehrstuhl für Bildverarbeitung, RWTH Aachen University
Benjamin Bross, Heinrich Hertz Institute, Fraunhofer Gesellschaft

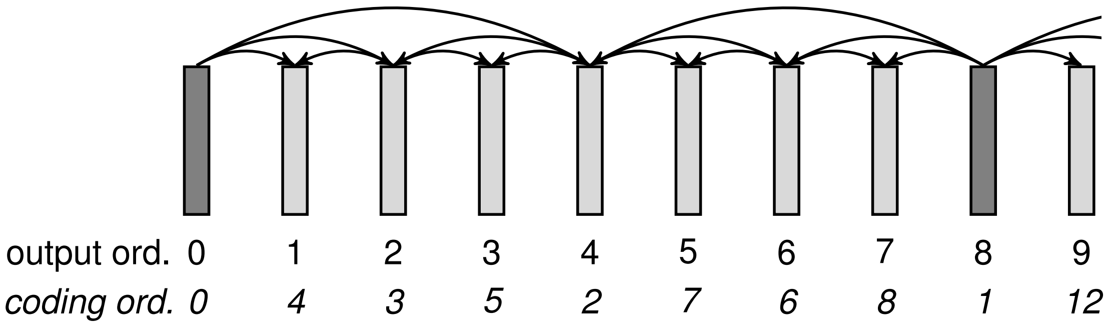
Contents

Temporal Coding Structures

- Coding order and output order may differ (affects delay)
  - (a) Low delay for conversational and interactive applications
  - (b) Reordering for broadcast and streaming applications
- Group of Pictures (GOP)
- Intra period for random access



(a) Hierarchical-P coding structure (coding order = output order)



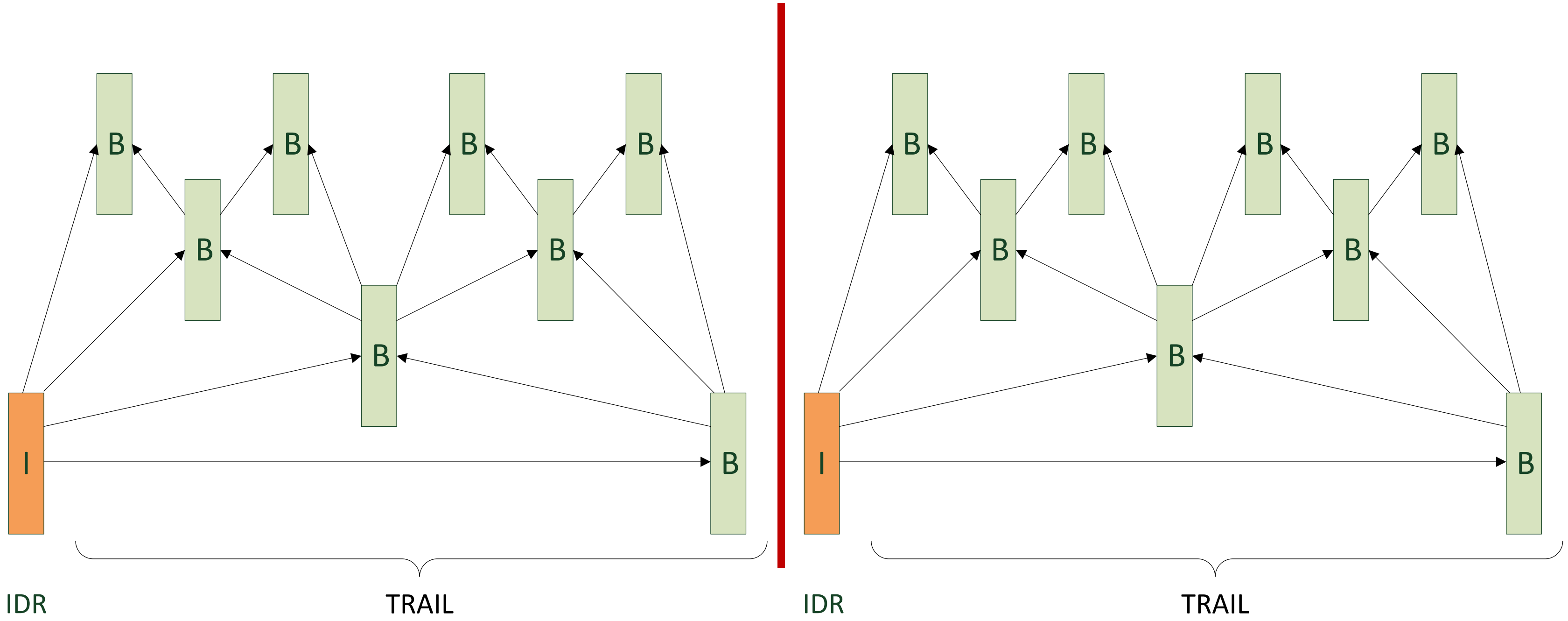
(b) Hierarchical-B coding structure (coding order ≠ output order)

Picture Types

- **Intra Random Access Point (IRAP)**
  – Instantaneous Decoding Refresh (IDR): Reset of decoder, start of new coded video sequence
  – Clean Random Access (CRA): Keep buffers intact if decoded within a coded video sequence, can be start point for decoding a coded video sequence
- **Gradual Decoding Refresh (GDR)**
  – Decoding can start at inter picture with only an intra region being decoded correctly
  – Intra region moves over time so the entire picture will be decoded correctly at some point.
  – New picture type in VVC replacing GDR by recovery point SEI in H.265|HEVC and H.264|AVC
- **Leading Pictures**
  – Follow the associated IRAP picture in coding order
  – Precede the associated IRAP picture in output order
  – Random Access Decodable Leading pictures (RADL): always decodable
  – Random Access Skipped Leading pictures (RASL): skip if decoding starts at associated CRA
- **Trailing Pictures**
  – Follow the associated IRAP picture in coding and output order
- **Stepwise Temporal Layer Access STSA**
  – Allow to switch temporal layer to the temporal layer of STSA picture

Closed GOP I

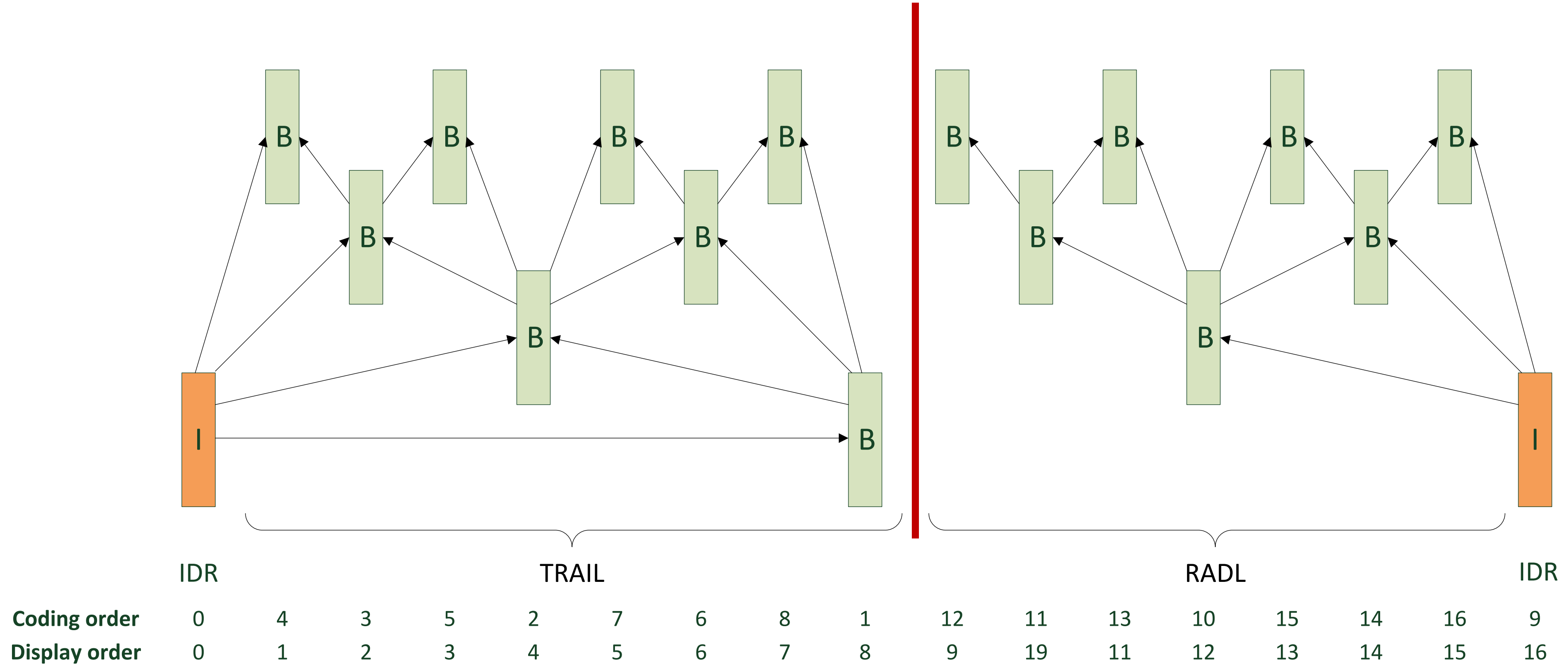
- Closed GOPs are independent from each other (H.264|AVC and older)
- IRAP with IDR picture



Coding order	0	4	3	5	2	7	6	8	1	9	13	12	14	11	16	15	17	10
Display order	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

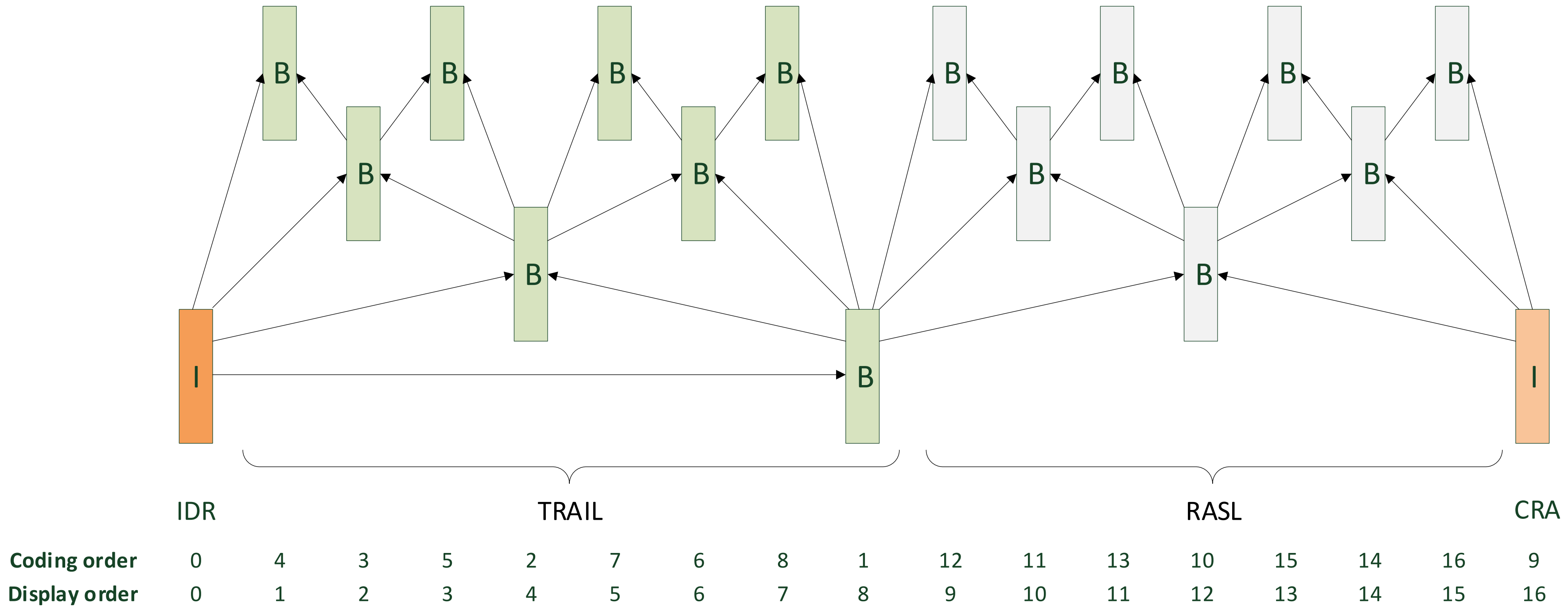
Closed GOP II

- IRAP with IDR picture and leading RADL pictures (introduced in H.265|HEVC)
- Prevents two consecutive key pictures at segment boundary



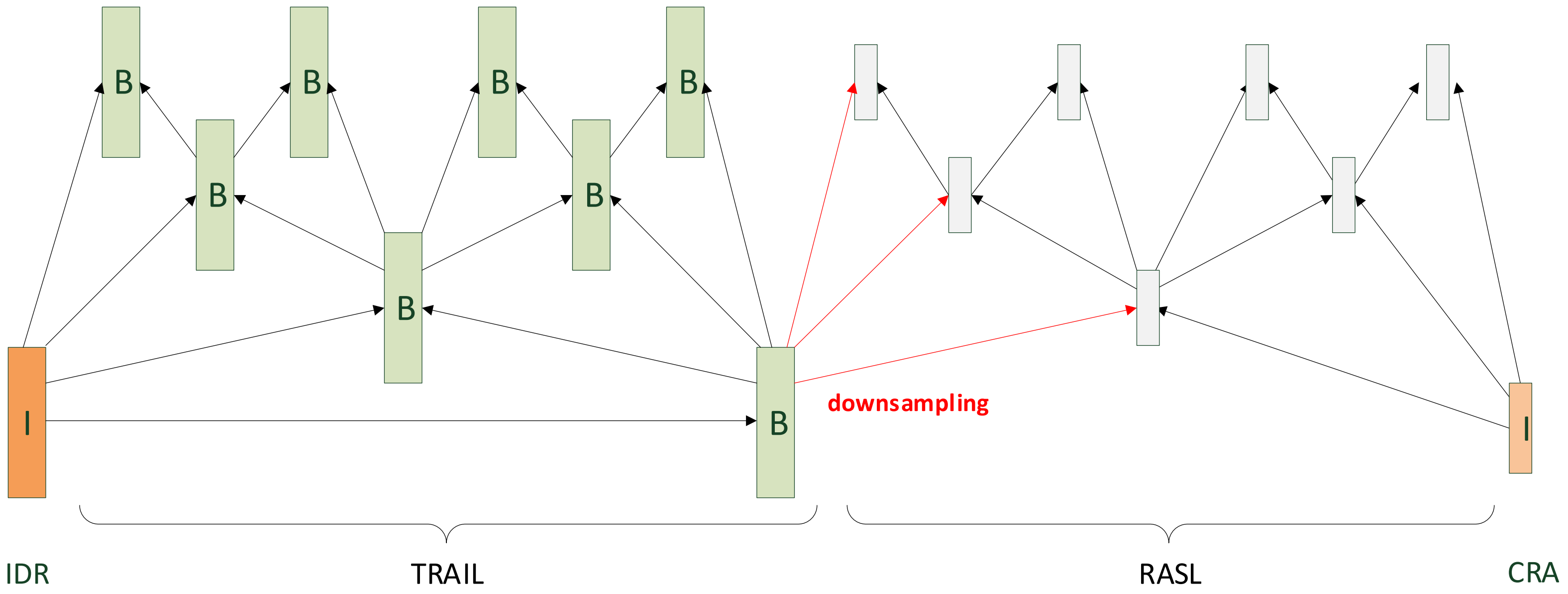
Open GOP

- Neighboring GOPs share at least one reference picture (introduced in H.265|HEVC)
- IRAP with CRA and leading RASL pictures



Reference Picture Resampling

- VVC introduces up- and downsampling of reference pictures
- Enables open GOP in adaptive streaming with resolution change (coding efficiency)



Coding order	0	4	3	5	2	7	6	8	1	12	11	13	10	15	14	16	9
Display order	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

Reference Picture Resampling

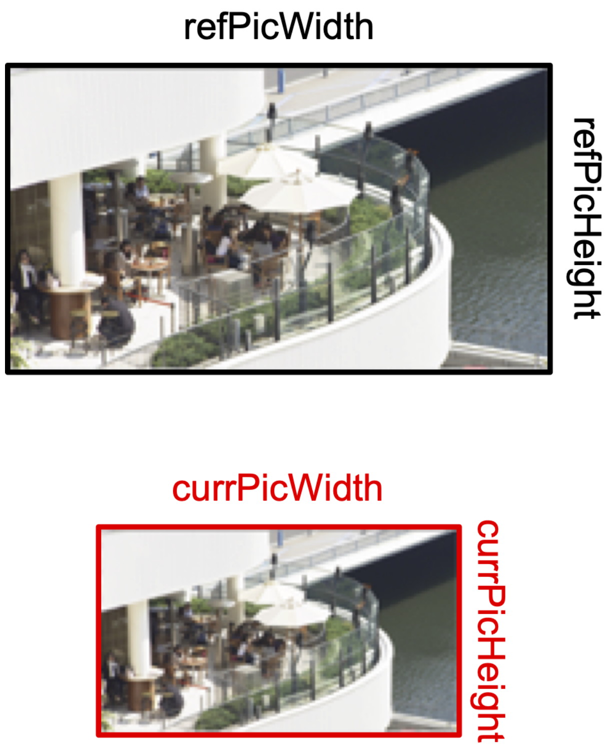
- Horizontal and vertical scaling limited:
  from $2\times$ downsampling to $8\times$ upsampling
- Three sets of resampling filters with different frequency cut-off (16 phases for luma and 32 phases for chroma):

$$\begin{aligned} 2 &\geq \text{scalingRatio} > 1.75 \\ 1.75 &\geq \text{scalingRatio} > 1.25 \\ 1.25 &\geq \text{scalingRatio} > 0.128 \text{ (same as for motion compensation)} \end{aligned}$$

- Different set of filters for translational and affine motion compensation
- Scaling ratios derived based on picture width and height of reference and current picture

$$\text{scalingRatioX} = \frac{\text{refPicWidth}}{\text{currPicWidth}}$$

$$\text{scalingRatioY} = \frac{\text{refPicHeight}}{\text{currPicHeight}}$$



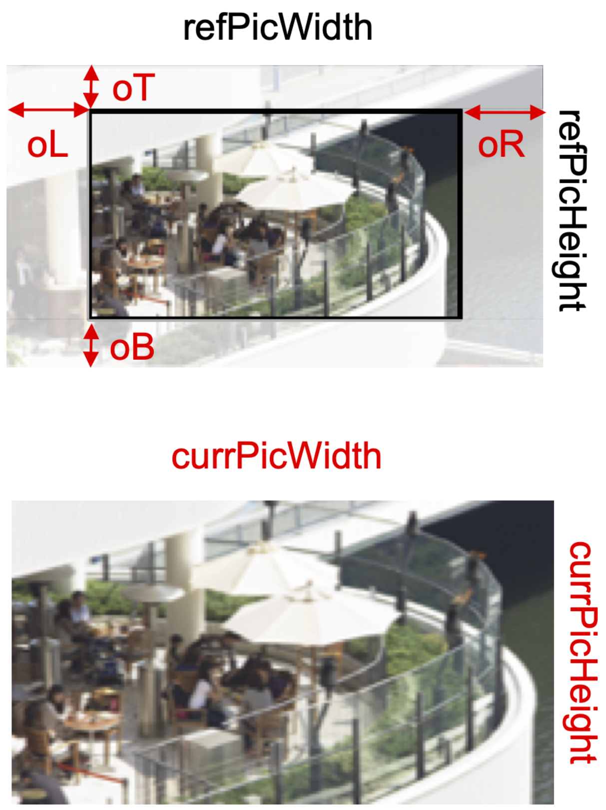
Reference Picture Resampling

- **Scaling window** can be signaled with left (oL), right (oR), top (oT) and bottom (oB) scaling offsets

$$\text{scalingRatioX} = \frac{\text{refPicWidth}}{\text{currPicWidth} - \text{oL} - \text{oR}}$$

$$\text{scalingRatioY} = \frac{\text{refPicHeight}}{\text{currPicHeight} - \text{oT} - \text{oB}}$$

- Provides additional functionality beyond resolution change, e.g. zooming



Reference Picture Management

- Decoded Picture Buffer (DPB) stores pictures for future use as reference
- Two reference picture lists (RPLs) L0 and L1 containing pictures from DPB used for current picture
- H.265|HEVC:
  - Reference picture sets (RPS) to explicitly signal the current DPB state
  - RPLs are either implicitly constructed or explicitly signaled
- VVC explicitly signals RPLs
  - Active reference pictures (used as reference)
  - Inactive reference pictures (remain in DPB)

Contents

Fraunhofer
Heinrich Hertz Institute

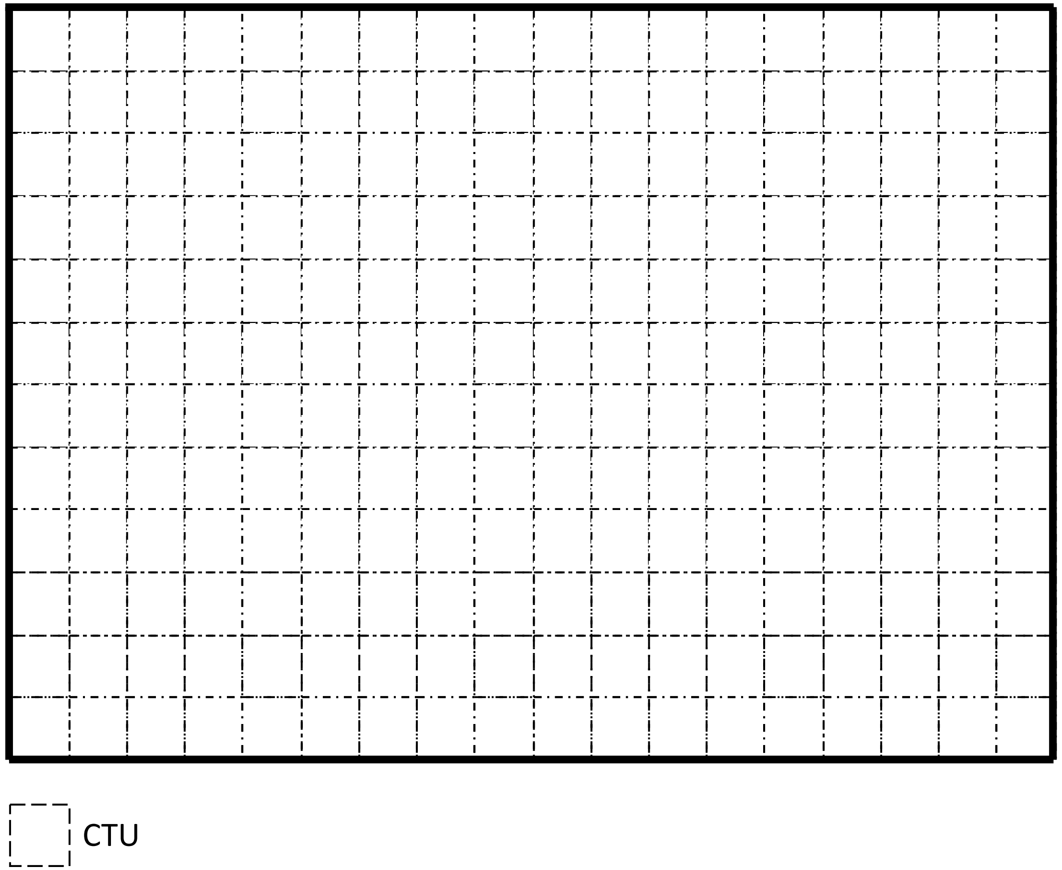
VVC Coding Tree Units

Coding Tree Unit (CTU)

- Basic processing unit for coding blocks of samples as in H.265|HEVC
- Contains blocks of samples and syntax to code these
- Fixed size for whole video sequence
- Max. size 128×128 (64×64 in H.265|HEVC)
- Min. size 32×32 (16×16 in H.265|HEVC)
- Root of the multi-type tree partitioning [link]

Coding Tree Block (CTB)

- Block of samples for one color component
- CTU can contain up to three CTBs (e.g. Y, Cb, Cr)



VVC Slices and Tiles

Extended set of elements

- **Slice**: An integer number of complete tiles or an integer number of consecutive complete CTU rows within a tile of a picture that are exclusively contained in a single NAL unit.
- **Tile**: A rectangular region of CTUs within a particular tile column and a particular tile row in a picture.

Example

- Raster-scan slice partitioning of a picture, where the picture is divided into 12 tiles (3 tile columns and 4 tile rows) and 3 raster-scan slices

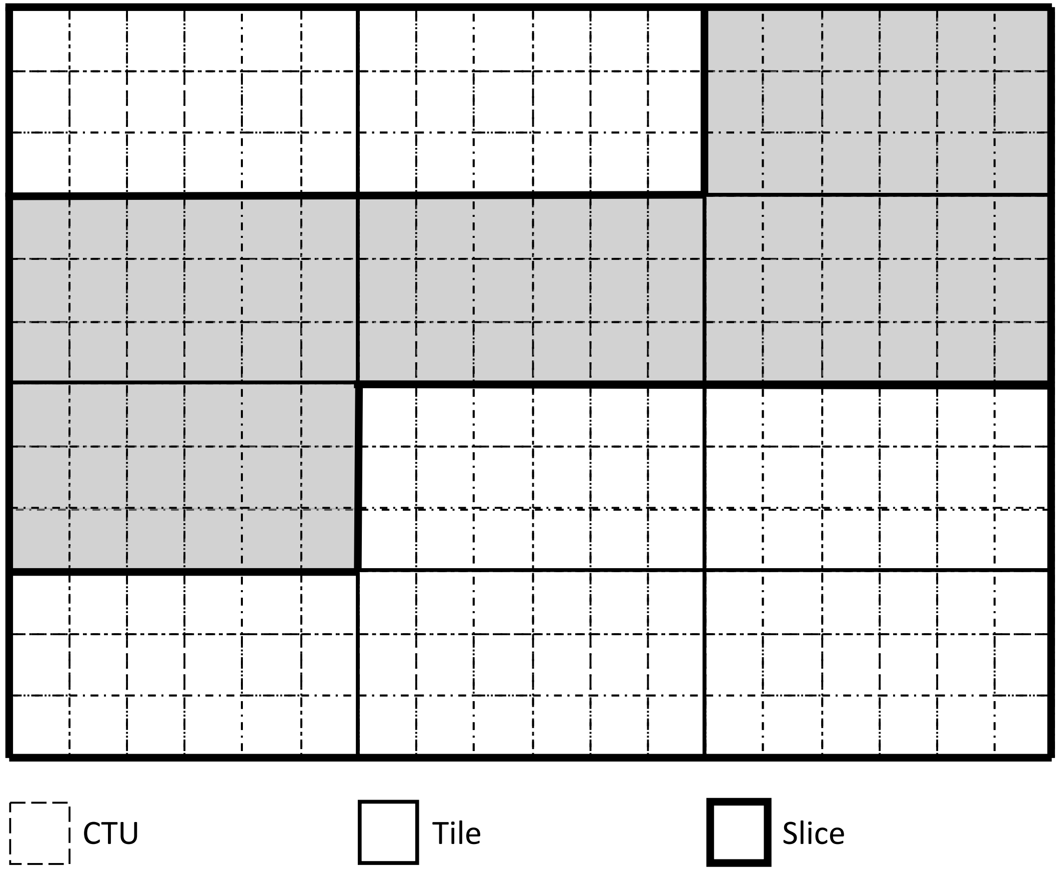


Figure from JVET-S2001 [47]

VVC Slices and Tiles

Extended set of elements

- **Slice**: An integer number of complete tiles or an integer number of consecutive complete CTU rows within a tile of a picture that are exclusively contained in a single NAL unit.
- **Tile**: A rectangular region of CTUs within a particular tile column and a particular tile row in a picture.

Example

- Rectangular slice partitioning of a picture, where the picture is divided into 24 tiles (6 tile columns and 4 tile rows) and 9 rectangular slices

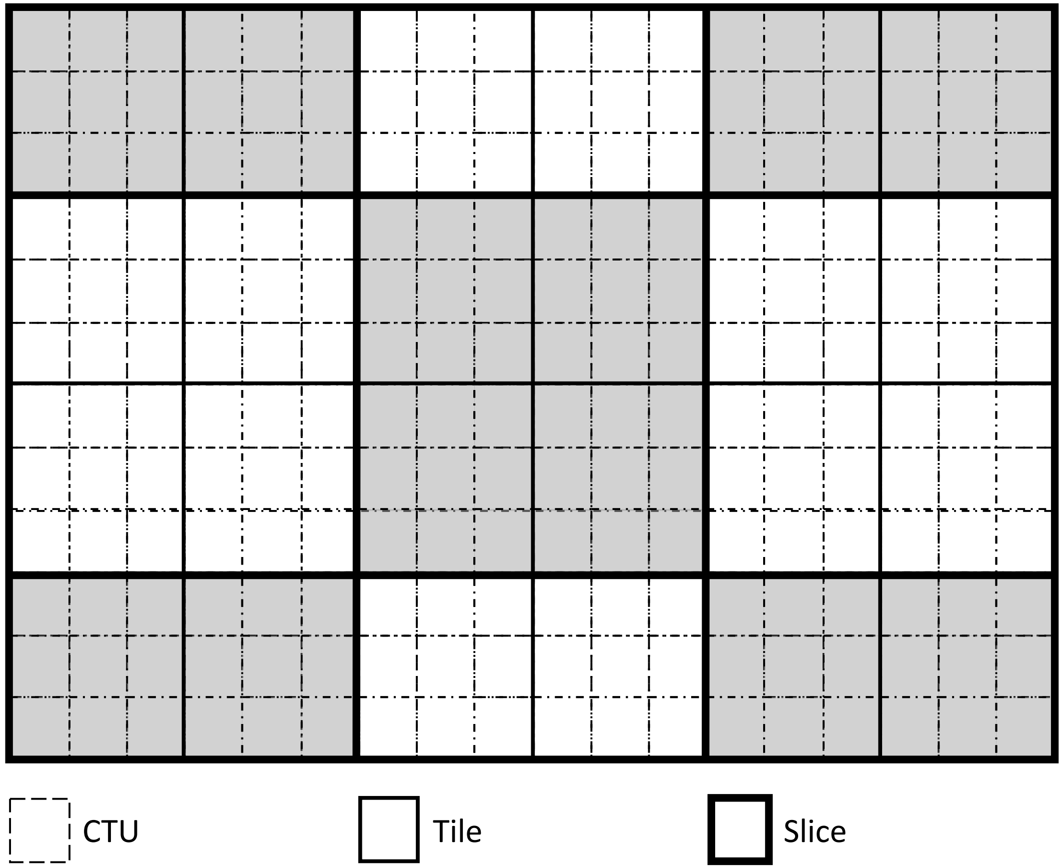


Figure from JVET-S2001 [47]

VVC Slices and Tiles

Extended set of elements

- **Slice**: An integer number of complete tiles or an integer number of consecutive complete CTU rows within a tile of a picture that are exclusively contained in a single NAL unit.
- **Tile**: A rectangular region of CTUs within a particular tile column and a particular tile row in a picture.

Example

- A picture that is partitioned into 4 tiles and 4 rectangular slices

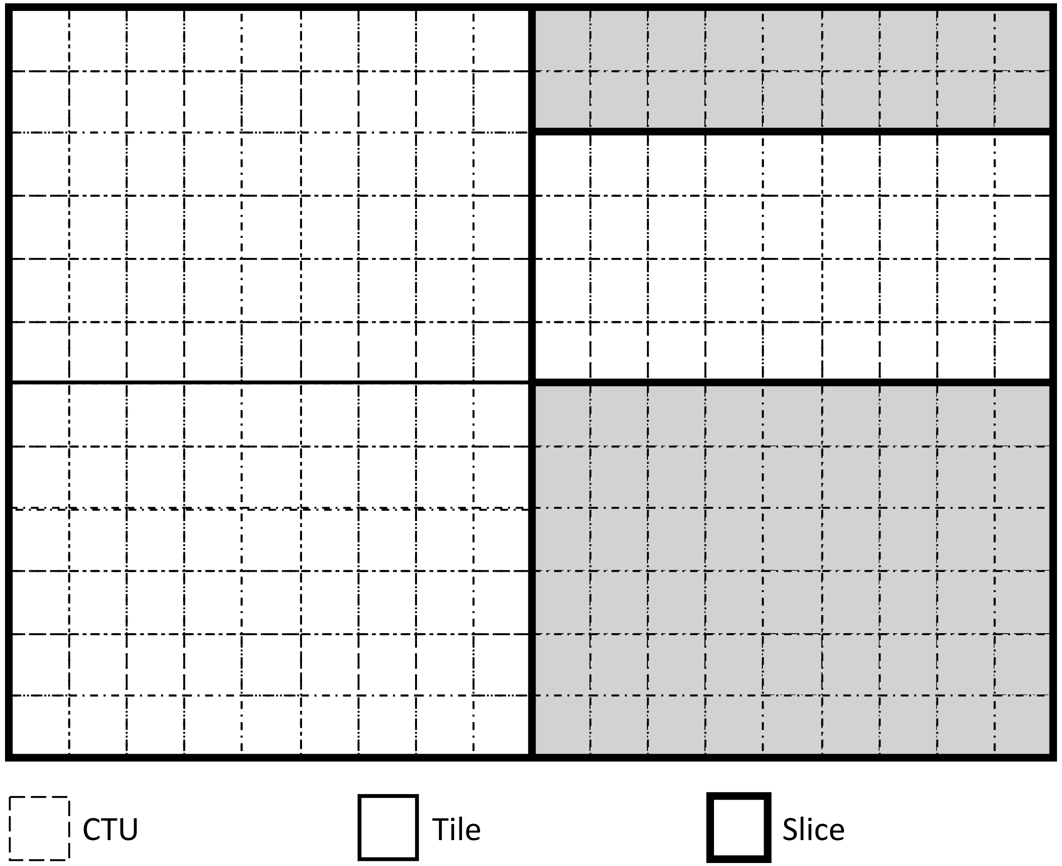


Figure from JVET-S2001 [47]

VVC Subpictures

- Realization of local random access
- Subpicture boundaries treated as picture boundaries
  - Padding for motion vectors pointing outside of region
- Subpictures independently decodable
- Re-composition of subpictures for various applications possible without header re-writing

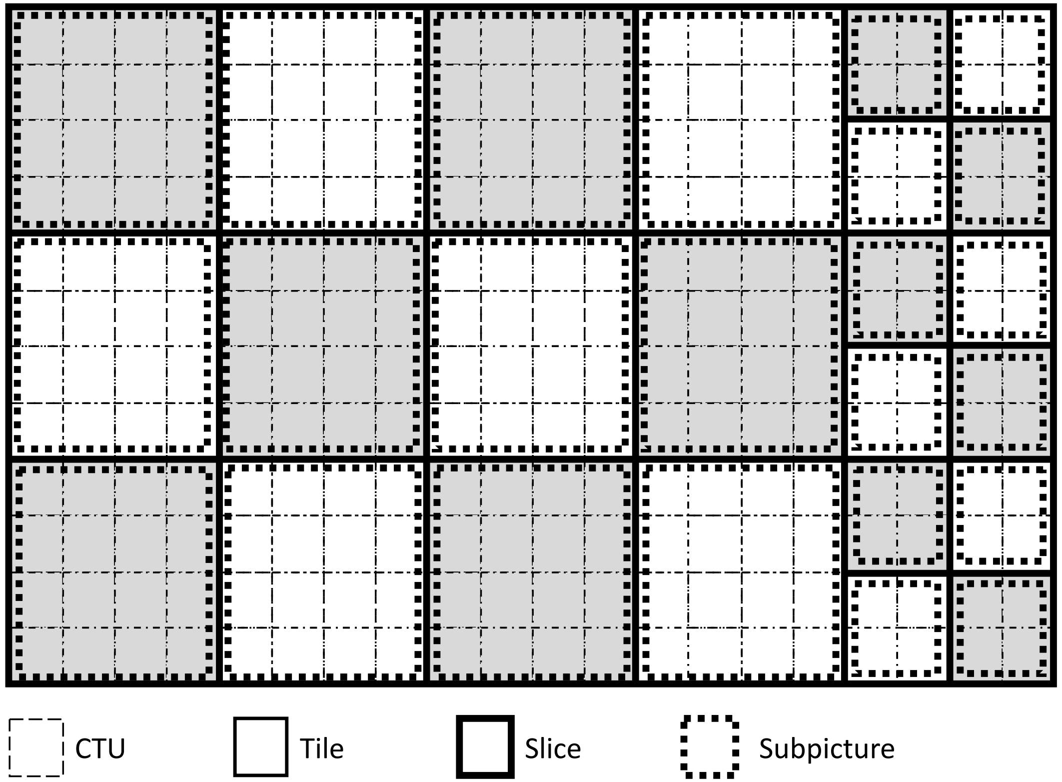


Figure from JVET-O2001 [48]

VVC Wavefront Parallel Processing

- CABAC contexts are inherited from row above
- Allows to run several processing threads in a slice over rows of CTUs in parallel including CABAC entropy decoding
- Parallel decoding within one slice requires signaling of entry point offsets for each CTU row
- Same concept as in H.265|HEVC but with 1 CTU delay instead of 2 CTUs to accommodate higher resolutions

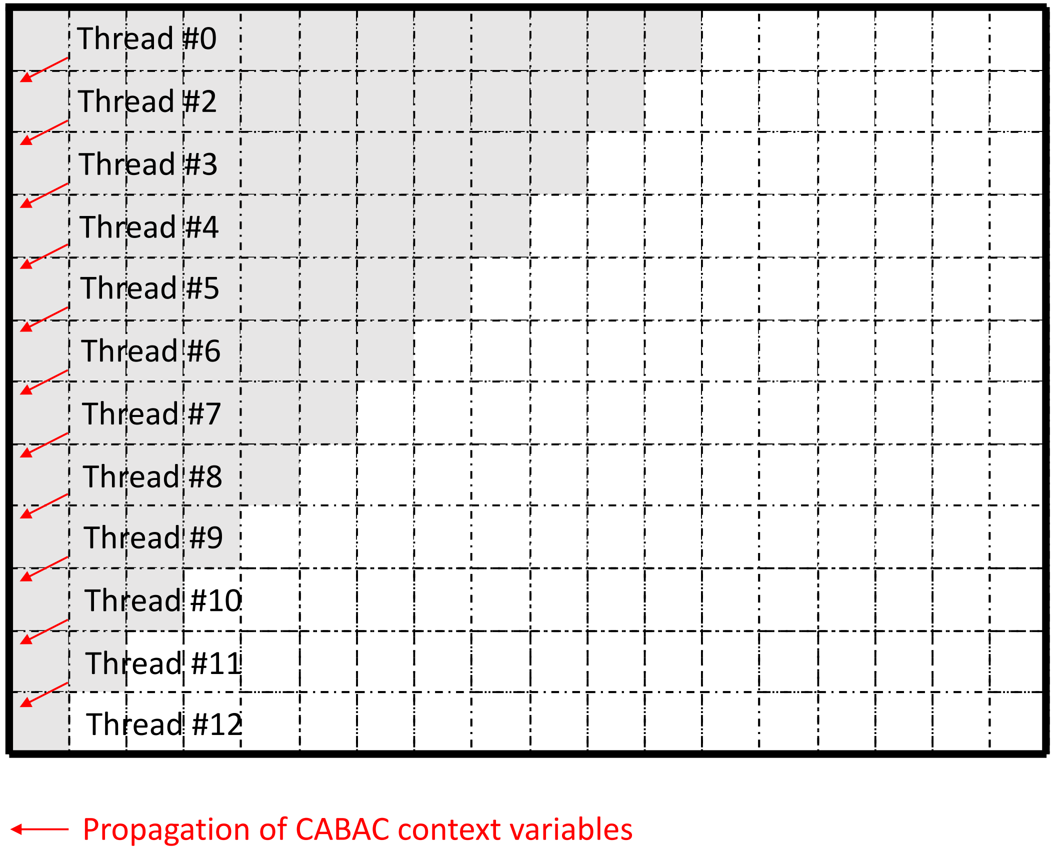


Figure from JVET-O2001 [48]

Contents

High-Level Design

Network abstraction layer (NAL)

Systems and transport interfaces

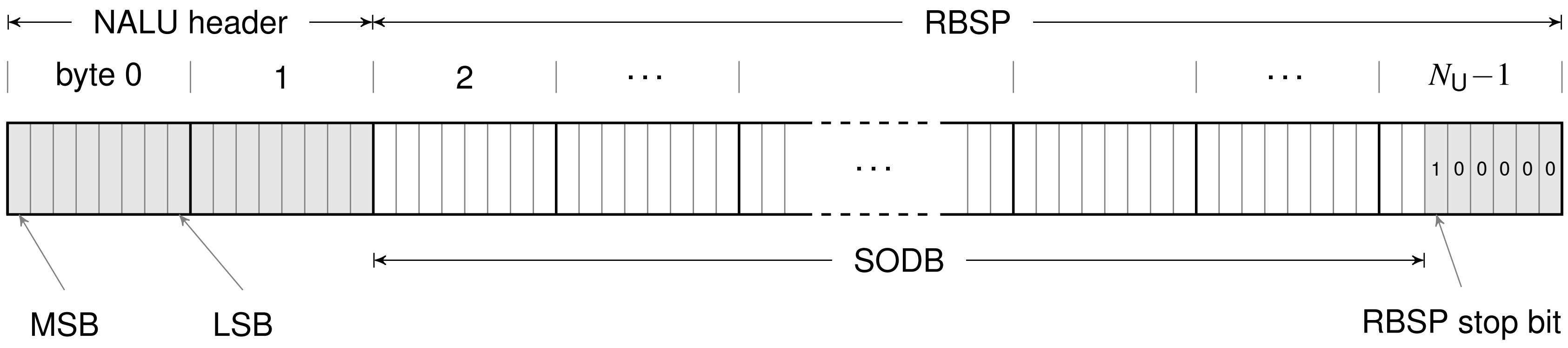
- Temporal structures
  - Random access
  - Reference picture management
- Spatial structures, picture partitioning into slices, tiles and subpictures
- Parameter sets
- NAL units

Video coding layer (VCL)

Coding efficiency

- Coded representation of picture samples
- Decoding involves (hybrid video coding):
  - Block Partitioning
  - Intra Prediction
  - Inter Prediction
  - Transforms Coding
  - Entropy Coding
  - Loop Filtering

NAL Unit Structure

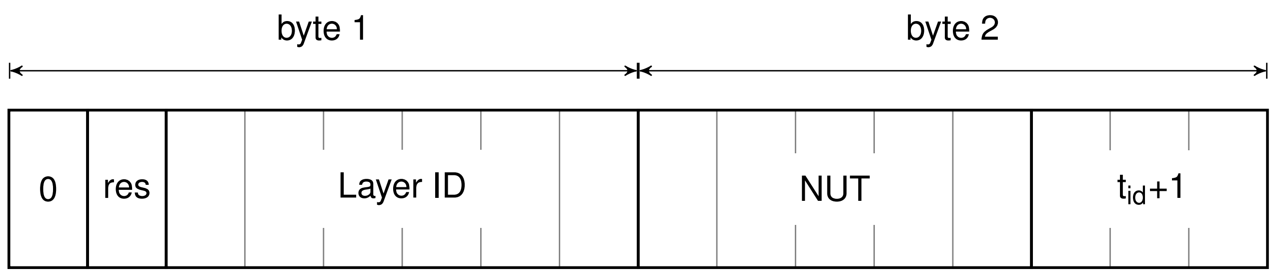


- RBSP: Raw byte sequence payload
  - Sequence of bytes comprising the coded NAL unit payload
  - RBSP stop bit (='1') plus zero bytes for byte alignment
- SODB: String of data bits
  - Concatenation of bits in the RBSP bytes from MSB to LSB
  - All bits needed for the decoding process
  - Only the bits needed for the decoding process

MSB: Most significant bit LSB: Least significant bit

Fraunhofer Heinrich Hertz Institute

VVC NAL Unit Header



- forbidden zero bit, reserved zero bit
- NAL unit header layer id: l_{id}
- nal unit type (NUT)
- nuh temporal id plus1: $t_{id} =$nuh temporal id plus1-1. Shall not be zero.

NAL Unit Types (NUTs)

- 32 different NAL unit types
  - 0-12: Coded slices for different picture types
  - 13-19: Parameter sets and picture header
  - 20-31: non-VCL NAL units
- Reserved NUTs for potential future use of ITU and ISO/IEC
- Unspecified NUTs for use specified outside of video standardization
- Reduced number of NAL unit types compared to HEVC

Figure from JVET-S2001 [47]

nal_unit_type	Name of nal_unit_type	Content of NAL unit and RBSP syntax structure	NAL unit type class
0	TRAIL_NUT	Coded slice of a trailing picture or subpicture* slice_layer_rbsp( )	VCL
1	STSA_NUT	Coded slice of an STSA picture or subpicture* slice_layer_rbsp( )	VCL
2	RADL_NUT	Coded slice of a RADL picture or subpicture* slice_layer_rbsp( )	VCL
3	RASL_NUT	Coded slice of a RASL picture or subpicture* slice_layer_rbsp( )	VCL
4..6	RSV_VCL_4.. RSV_VCL_6	Reserved non-IRAP VCL NAL unit types	VCL
7 8	IDR_W_RADL IDR_N_LP	Coded slice of an IDR picture or subpicture* slice_layer_rbsp( )	VCL
9	CRA_NUT	Coded slice of a CRA picture or subpicture* silce_layer_rbsp( )	VCL
10	GDR_NUT	Coded slice of a GDR picture or subpicture* slice_layer_rbsp( )	VCL
11 12	RSV_IRAP_11 RSV_IRAP_12	Reserved IRAP VCL NAL unit types	VCL
13	DCI_NUT	Decoding capability information decoding_capability_information_rbsp( )	non-VCL
14	VPS_NUT	Video parameter set video_parameter_set_rbsp( )	non-VCL
15	SPS_NUT	Sequence parameter set seq_parameter_set_rbsp( )	non-VCL
16	PPS_NUT	Picture parameter set pic_parameter_set_rbsp( )	non-VCL
17 18	PREFIX_APS_NUT SUFFIX_APS_NUT	Adaptation parameter set adaptation_parameter_set_rbsp( )	non-VCL
19	PH_NUT	Picture header picture_header_rbsp( )	non-VCL
20	AUD_NUT	AU delimiter access_unit_delimiter_rbsp( )	non-VCL
21	EOS_NUT	End of sequence end_of_seq_rbsp( )	non-VCL
22	EOB_NUT	End of bitstream end_of_bitstream_rbsp( )	non-VCL
23 24	PREFIX_SEI_NUT SUFFIX_SEI_NUT	Supplemental enhancement information sei_rbsp( )	non-VCL
25	FD_NUT	Filler data filler_data_rbsp( )	non-VCL

Access Units and Picture Units

Definitions JVET-S2001 [47]

- **Access Unit**
  - A set of picture units that belong to different layers and contain coded pictures associated with the same time for output from the decoded picture buffer (DPB).
- **Picture Unit**
  - A set of NAL units that are associated with each other according to a specified classification rule, are consecutive in decoding order, and contain exactly one coded picture.

Parameter Sets

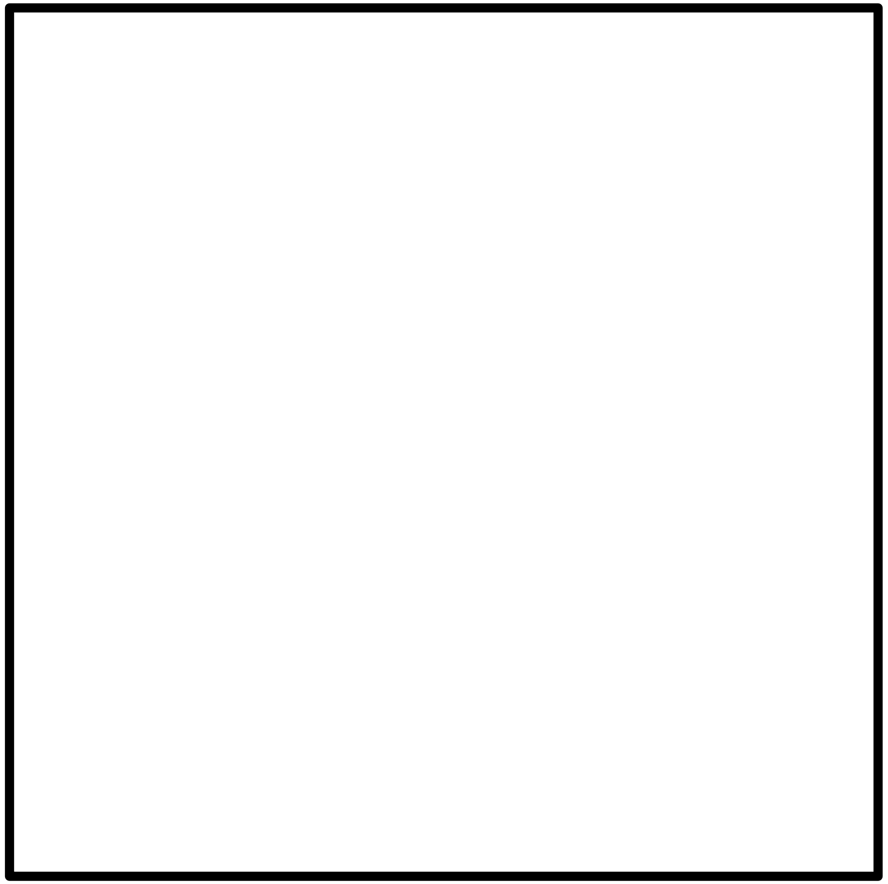
- **Decoding capability information (DCI)** (new)
  - Specifies decodable coded video sequences as subsets of the bitstream
- **Video parameter set (VPS)**
  - Specifies the layer structure of a layered Coded Video Sequence (e.g. scalability, multiview)
- **Sequence parameter set (SPS)**
  - Specifies general parameters which apply to all pictures of the a layer of a coded video sequence
- **Picture parameter set (PPS)**
  - Specifies tool usage and initial tool parameter settings (multiple PPS in one Coded Video Sequence possible)
- **Adaptation parameter set (APS)** (new)
  - Carries persisting parameters, e. g. for the Adaptive Loop Filter

- **Picture header** (new)
  - Specifies parameter settings for all slices in the coded picture, includes sub-picture IDs (if present)
- **Slice header**
  - Slice-specific settings, sub-picture ID

Contents

VVC Multi-type Tree Block Partitioning

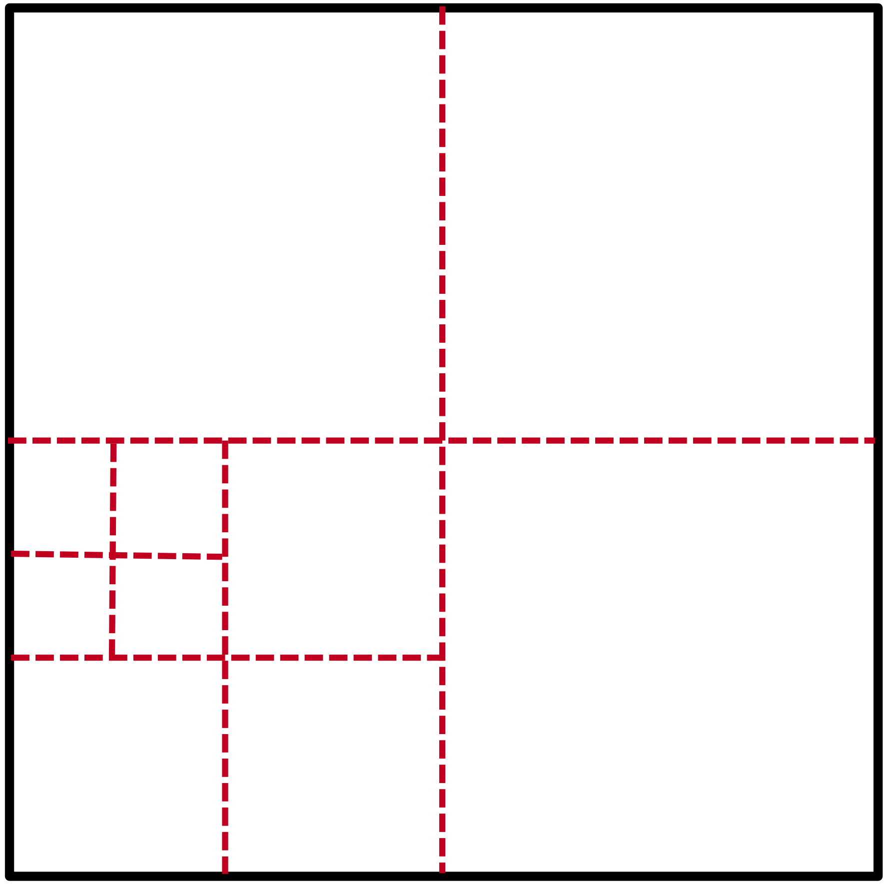
New block partitioning of a CTU using...



VVC Multi-type Tree Block Partitioning

New block partitioning of a CTU using...

- Recursive **quadtree (QT)** split

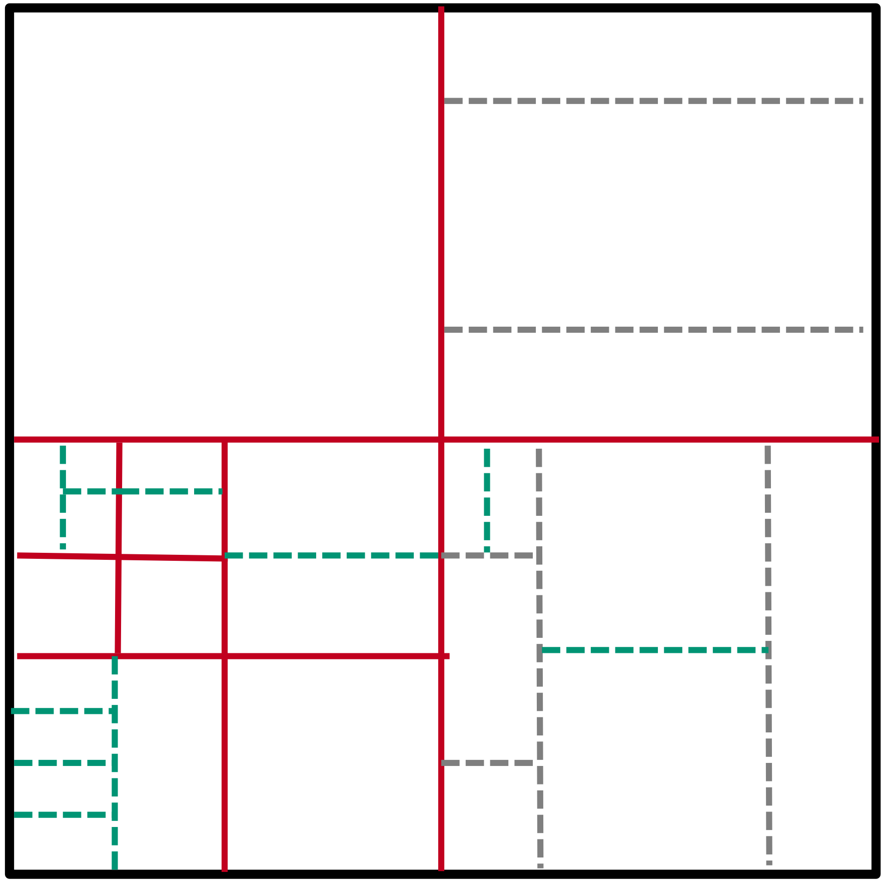


VVC Multi-type Tree Block Partitioning

New block partitioning of a CTU using...

- Recursive **quadtree (QT)** split
- Nested recursive **multi-type tree (MTT)** splits with

binary split or ternary split

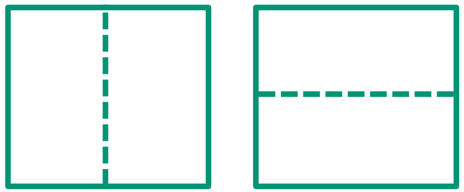


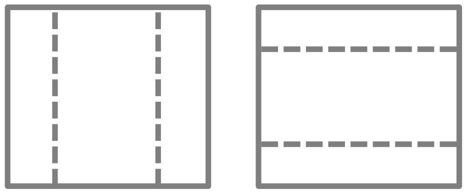
VVC Multi-type Tree Block Partitioning

New block partitioning of a CTU using...

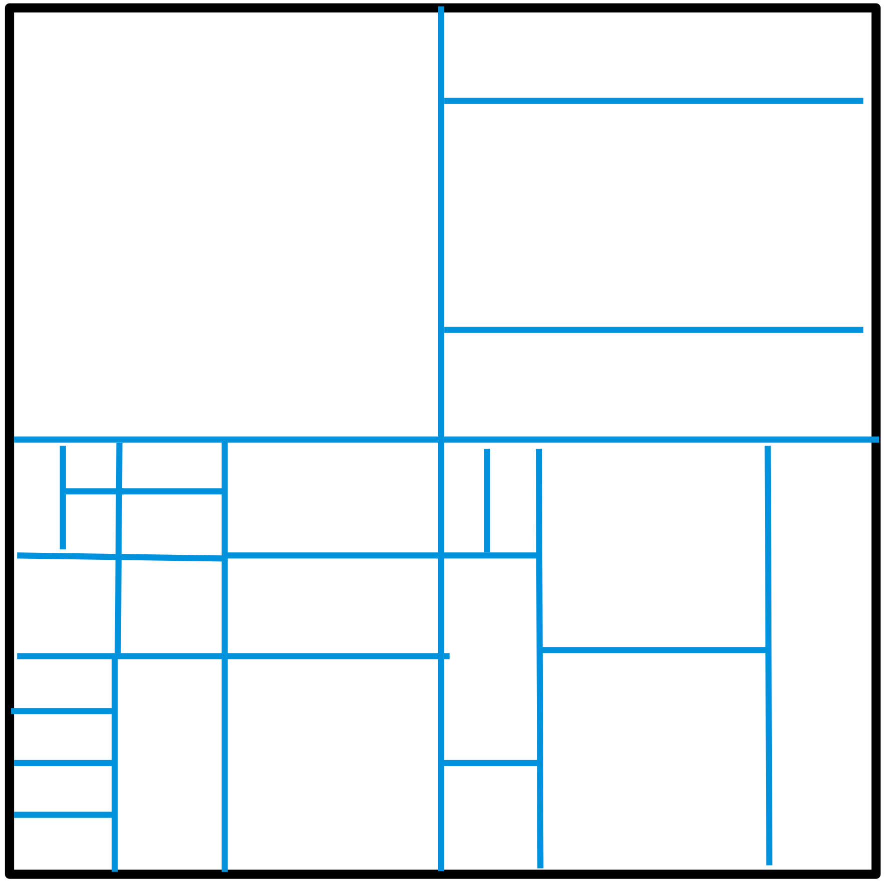
- Recursive **quadtree (QT)** partitioning
- Nested recursive **multi-type tree (MTT)** partitioning with

binary split or ternary split





- Variable size **Coding Units (CU)**
  - CUs can range from CTU size (max. 128×128) to 4×4 luma samples
  - MTT restrictions to prevent redundant signaling of partitions



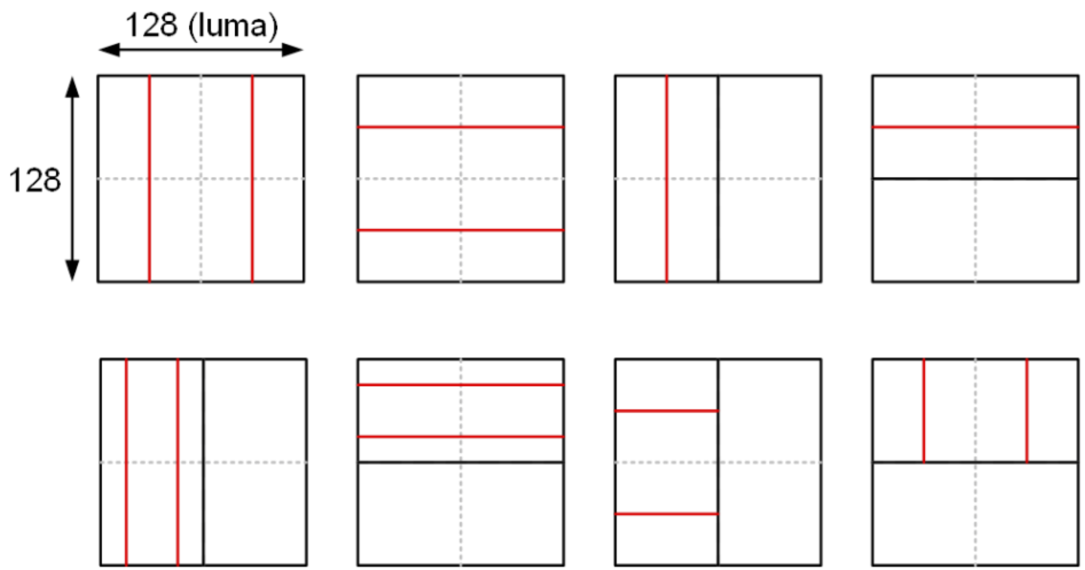
VVC Separate Chroma Partitioning

- **Dual tree (chroma separate tree)**
  - Separate MTT coding tree for chroma
  - Typically higher tree depth needed for luma (details/structure)
  - Partitioning can stop early for chroma (faster encoding)
  - Luma and chroma blocks not aligned anymore

- **Local Dual tree**
  - No further splitting of $4 \times N$ and $N \times 4$ chroma intra blocks (including IBC and palette modes)
  - Improves hardware decoding throughput by allowing better pipelining

VVC Virtual Pipeline Data Units (VPDUs)

- Increased throughput for pipeline processing in hardware decoders
- Critial for intra blocks that depend on reconstructed neighboring samples
- Set to the maximum transform / intra prediction size, i.e. 64×64
- Requires MTT split restrictions, e.g. 128×128 intra CU is implicitly split into 64×64 CUs.



Versatile Video Coding – Algorithms and Specification

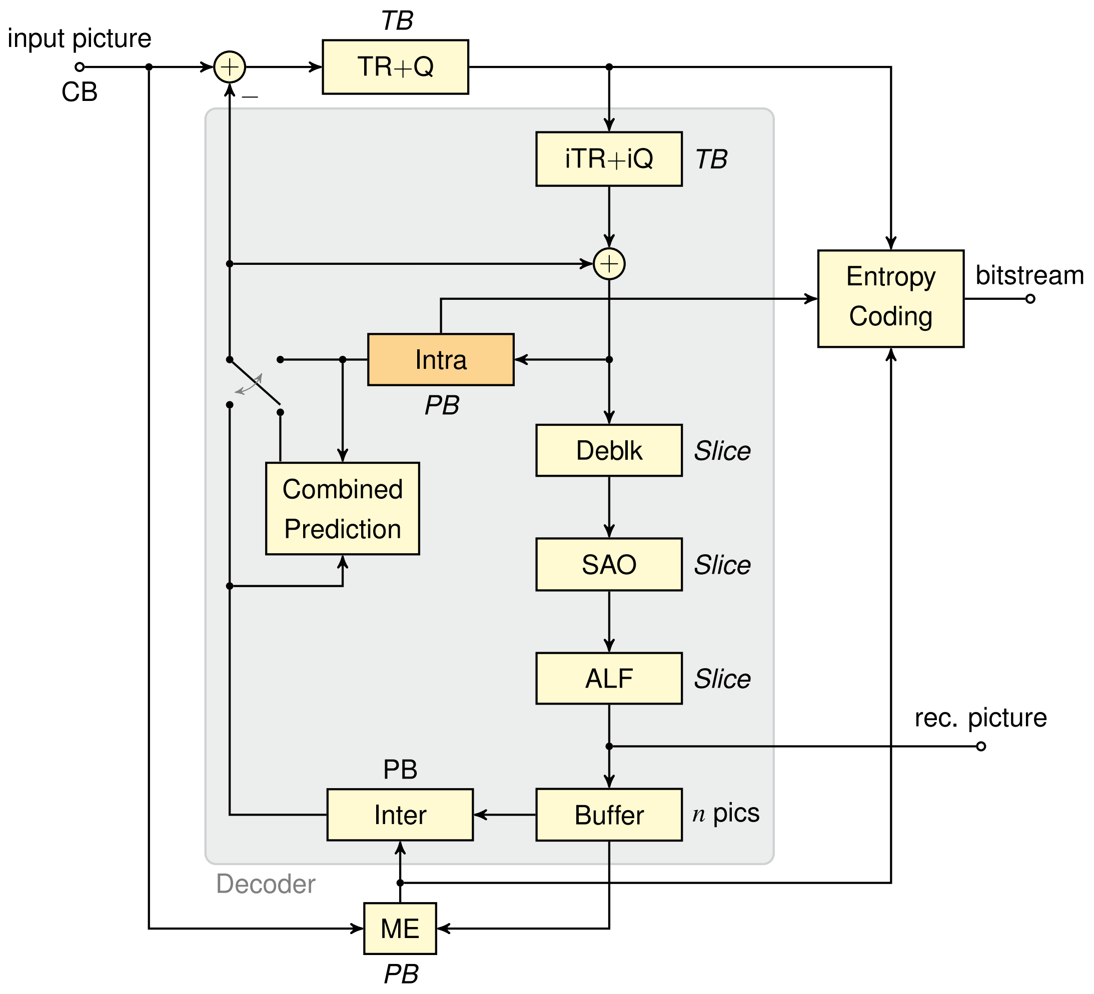
Part 3 | Coding Tools I

IEEE ICME, 10.07.2020

Mathias Wien, Lehrstuhl für Bildverarbeitung, RWTH Aachen University
Benjamin Bross, Heinrich Hertz Institute, Fraunhofer Gesellschaft

Contents

Intra Prediction



CB – Coding Block
ME – Motion Estimation
PB – Prediction Block
Q – Quantization
TB – Transform Block
TR – Transform
SAO – Sample Adaptive Offset
ALF – Adaptive Loop Filter

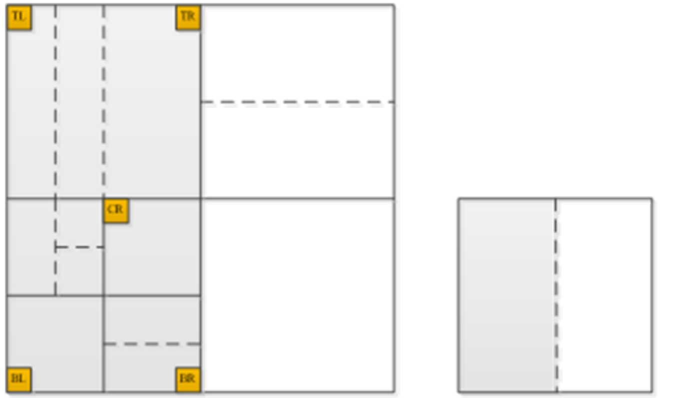
Directional intra prediction modes

Concept of HEVC as basis

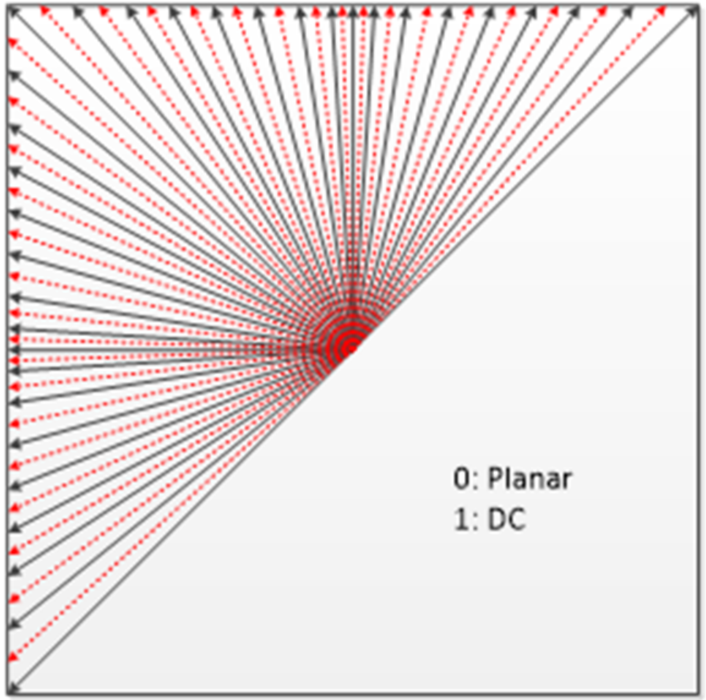
- Higher number of prediction modes (double resolution)
- Larger maximum block size

Chroma

- Prediction modes from neighbors
- Derived modes from collocated luma



(a) Luma QTBT structure (b) Chroma QTBT structure



Figures from JVET-G1001 [49]

Wide angular modes

- For rectangular blocks, prediction directions with angles beyond 45/135 degrees are reasonable
- This can be implemented by adding modes at both ends
- 85 directional intra modes available (plus DC and planar), width/height ratio dependent mapping for signaling

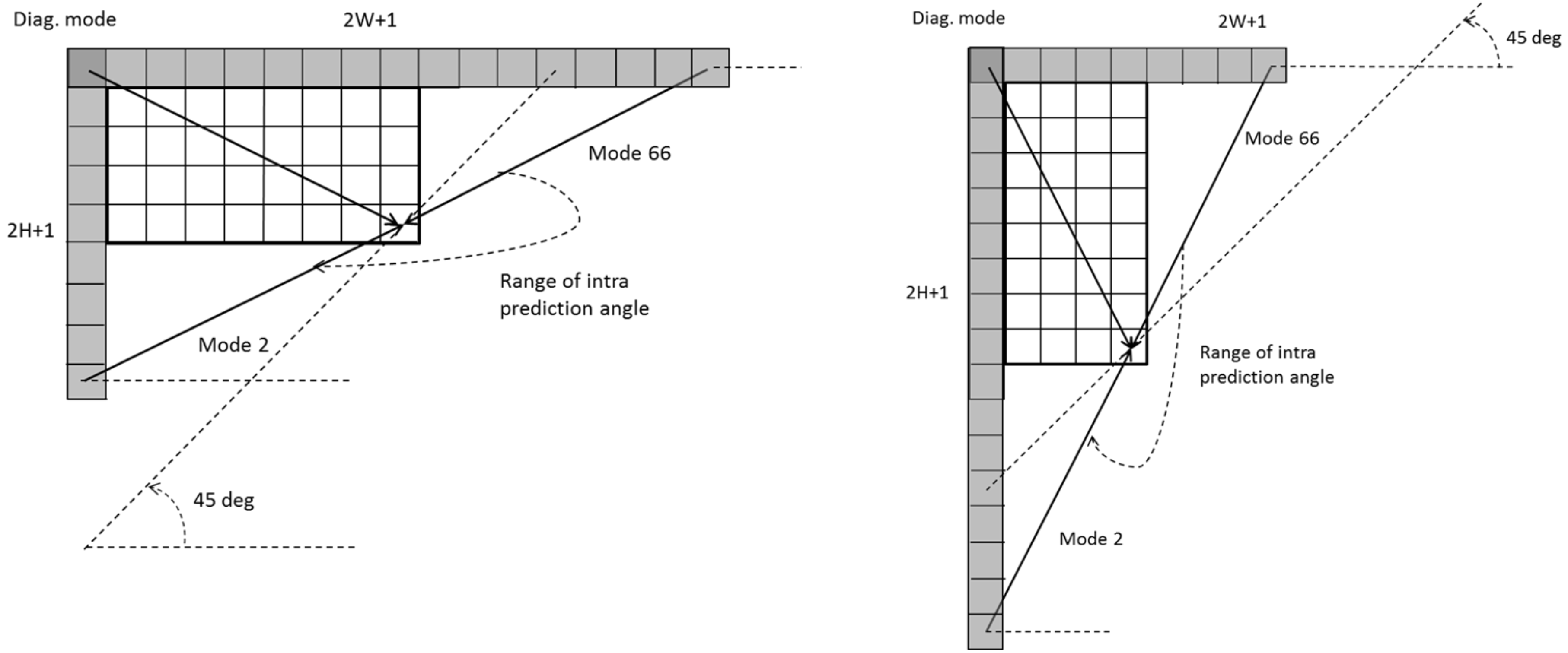


Figure from JVET-K0500 [50]

Cross-Component Linear Model Prediction (CCLM)

- Chroma samples predicted using corresponding reconstructed luma samples

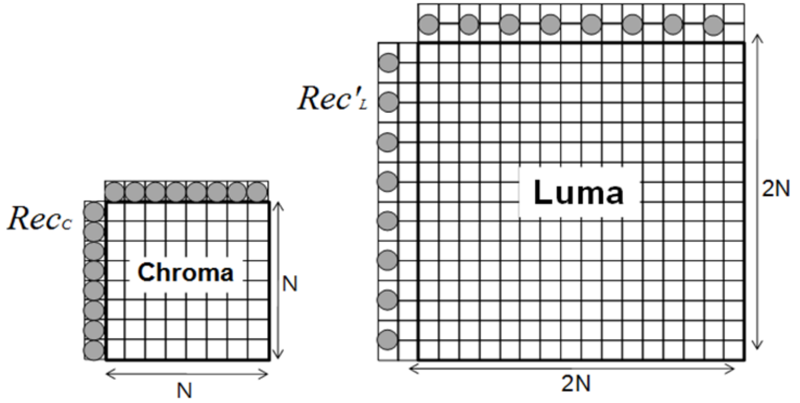
$$\mathrm{pred}_{\mathrm{C}}(i,j) = \alpha \cdot \mathrm{rec}'_{\mathrm{L}}(i,j) + \beta$$

- Parameters α and β: minimize regression error between neighboring reconstructed luma and chroma samples around current block

- Selection of left/top neighbors via 3 modes INTRA_LT_CCLM, INTRA_L_CCLM, and INTRA_T_CCLM

- Further prediction between chroma components with updated parameters

$$\mathrm{pred}^{*}_{\mathrm{Cr}}(i,j) = \mathrm{pred}_{\mathrm{Cr}}(i,j) + \alpha \cdot \mathrm{resi}'_{\mathrm{Cb}}(i,j)$$



Position Dependent Intra Prediction Combination for Planar Mode (PDPC)

- Combination of the un-filtered boundary reference samples and HEVC-style intra prediction with filtered boundary reference samples
- Position-dependent weighting of filtered and unfiltered reference, configurable by four weighing parameters (hor/ver + corner)
- Filtered reference: linear comination of un-filtered reference and lowpass, configurable weight
- Three predefined lowpass filters selectable (3-tap, 5-tap, 7-tap)

- Prediction parameters stored per block size

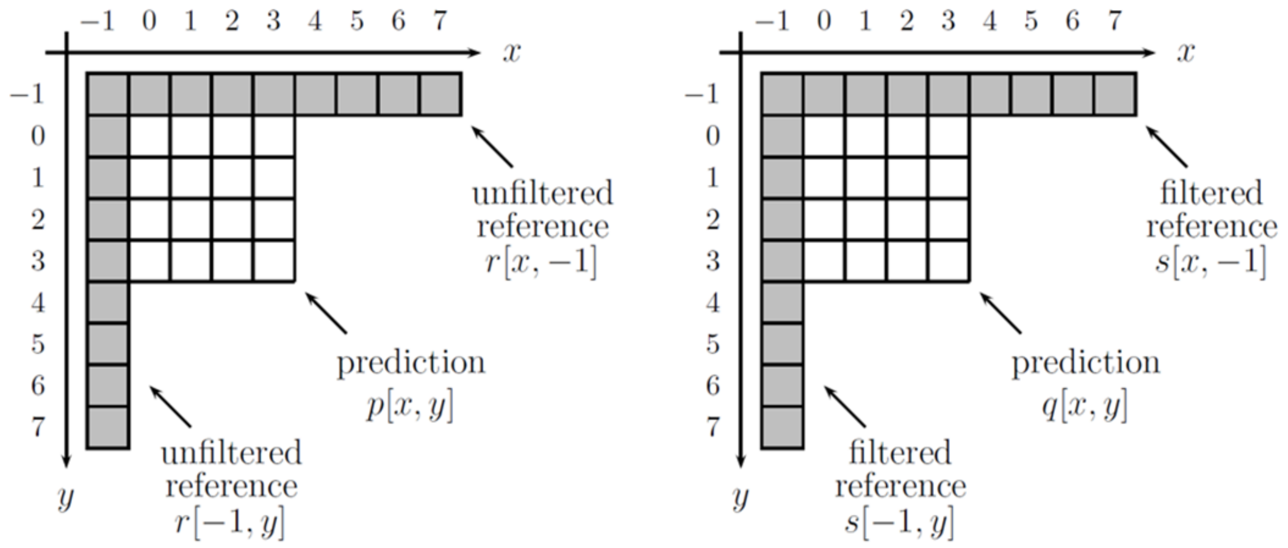


Figure from JVET-G1001 [49]

Multiple Reference Line Intra Prediction

- Two additional reference lines, luma only
- Activation of lines by syntax element `mrl_idx`
- Applied for DC, directional prediction
- `mrl_idx` > 0:
  - No PDPC
  - No boundary filter
  - no edge filter

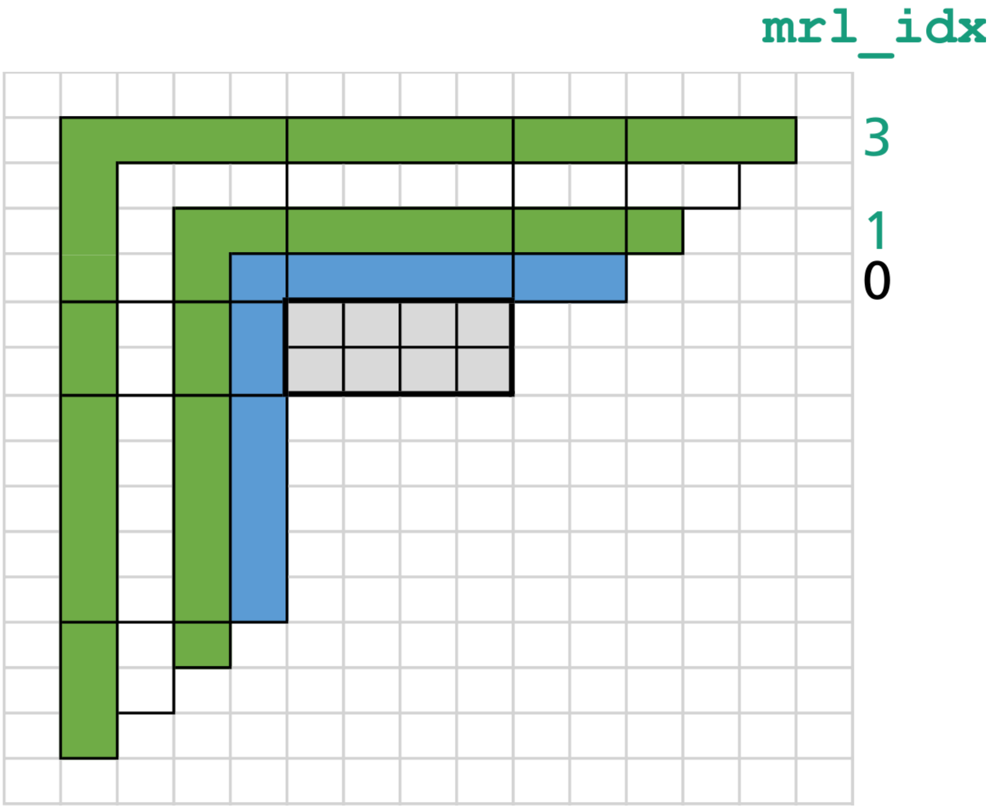
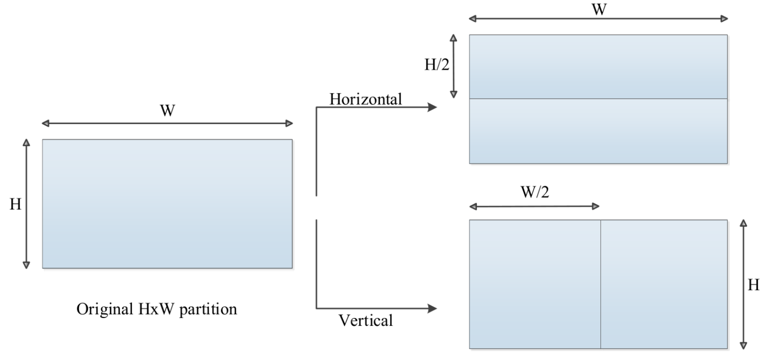


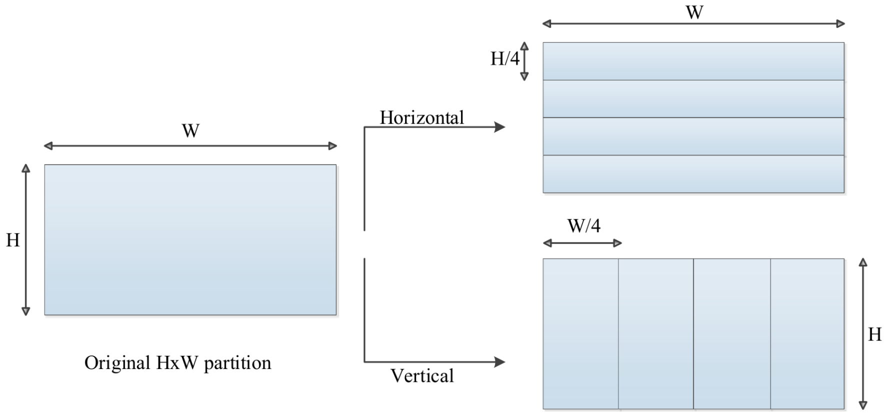
Figure from JVET-J0014 [25]

Intra Sub-Partitions (ISP)

- CB is further partitioned into sub-partitions for intra prediction / transform
- Reduce distance of predicted samples from reference to increase spatial correlation
- Enables 1D transform blocks , e.g. 32×1 resulting from splitting a 32×4 CB
- Special case: prediction on CB for sub-partitions with W or H less than 4
- Variation of concept in HEVC (wich used the residual quadtree)



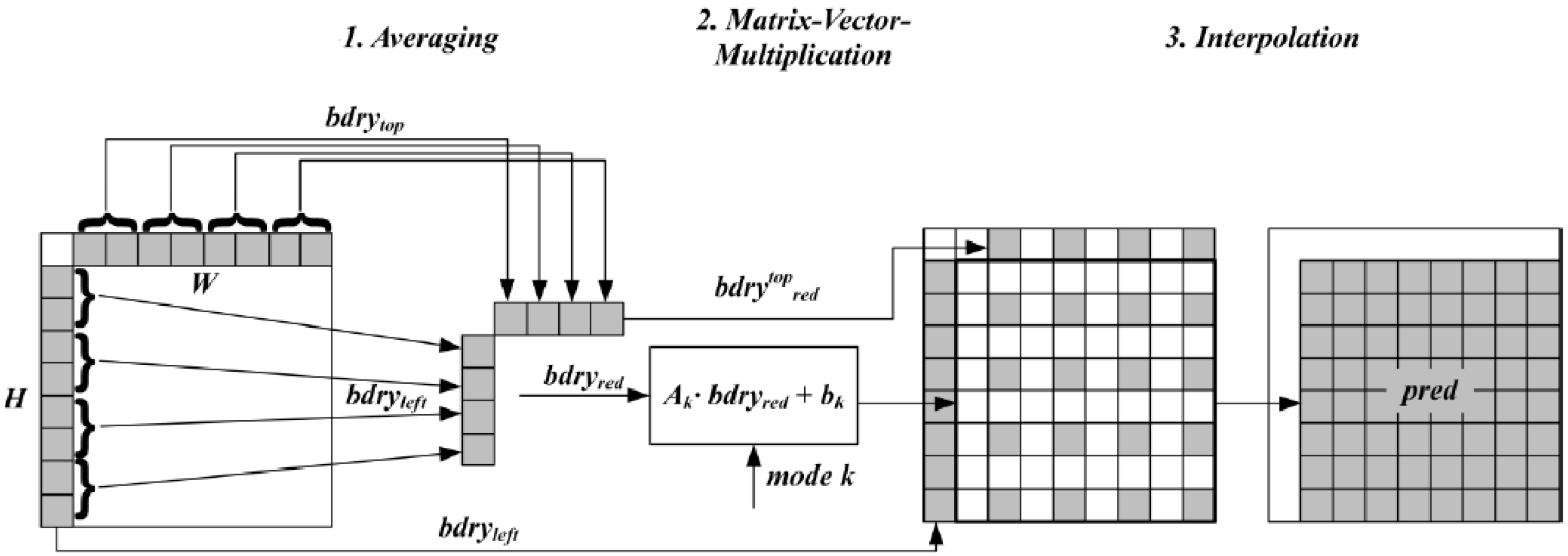
a) Examples of sub-partitions for 4x8 and 8x4 CUs



b) Examples of sub-partitions for CUs other than 4x8, 8x4 and 4x4

Figure from JVET-Q2002 [51]

Matrix weighted Intra Prediction (MIP)

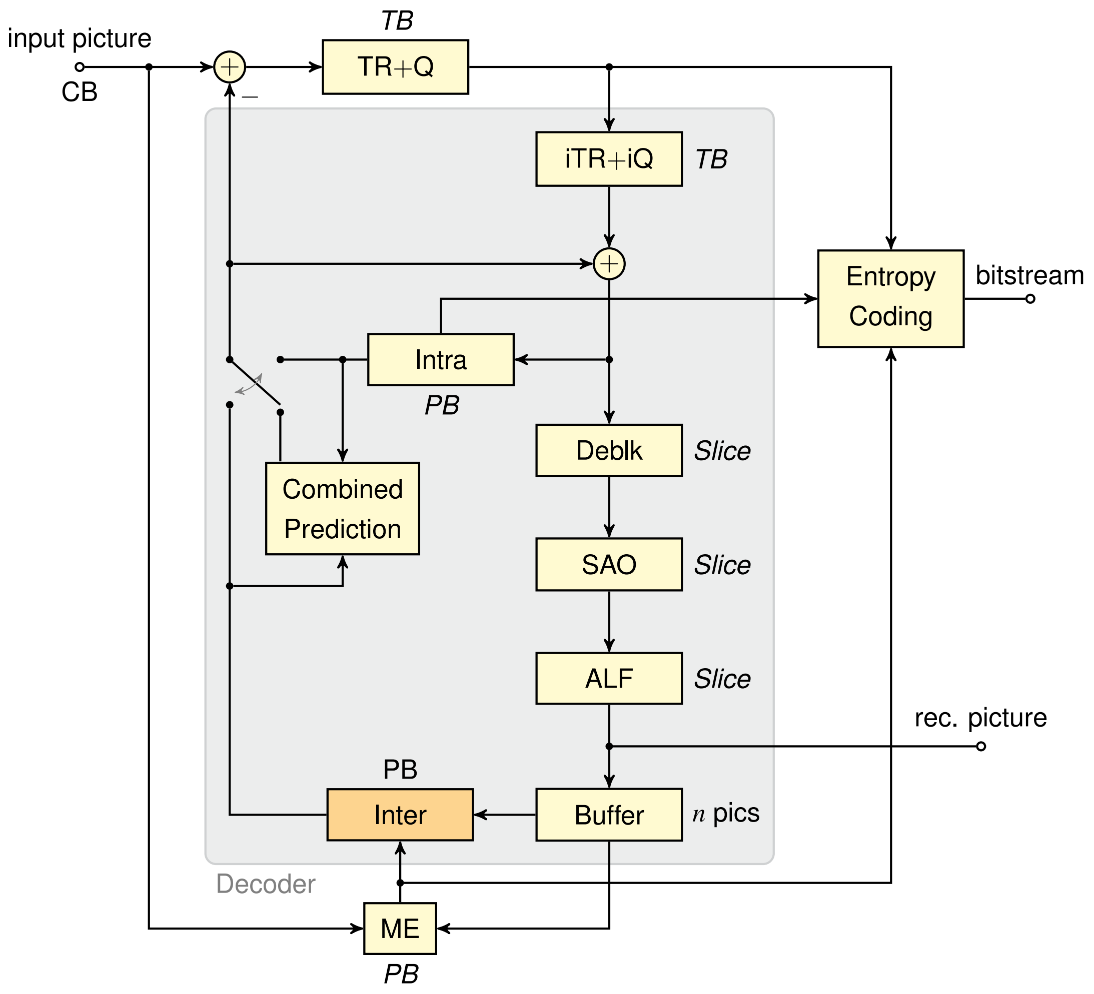


- First (originally) learned tool in a video coding standard
- Up to 16 weight tables depending on block size

Figure from JVET-Q2002 [51]

Contents

Inter Prediction

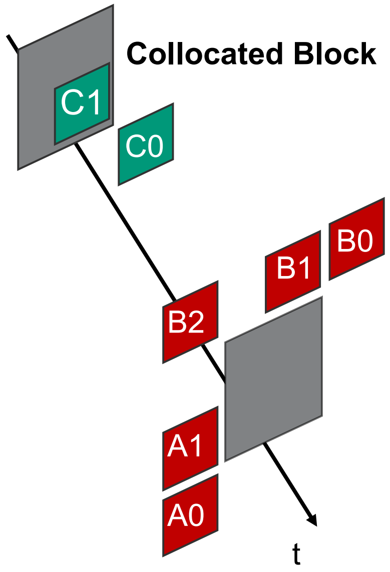


CB – Coding Block
ME – Motion Estimation
PB – Prediction Block
Q – Quantization
TB – Transform Block
TR – Transform
SAO – Sample Adaptive Offset
ALF – Adaptive Loop Filter

Extended Motion Vector Prediction

MV = MVP + MVD

- MVD and reference picture list (RPL) index explicitly signaled
- MVP derived from list of 2 candidates
  - Spatial (A,B) and temporal (C) neighbors (same as HEVC)
  - History-based candidates
  - Zero MV (same as HEVC)



Symmetric MVD coding (SMVD)

- Signal only motion MVP, RPL index and MVD for List0
- Apply inverse MVD on opposite reference picture from List1

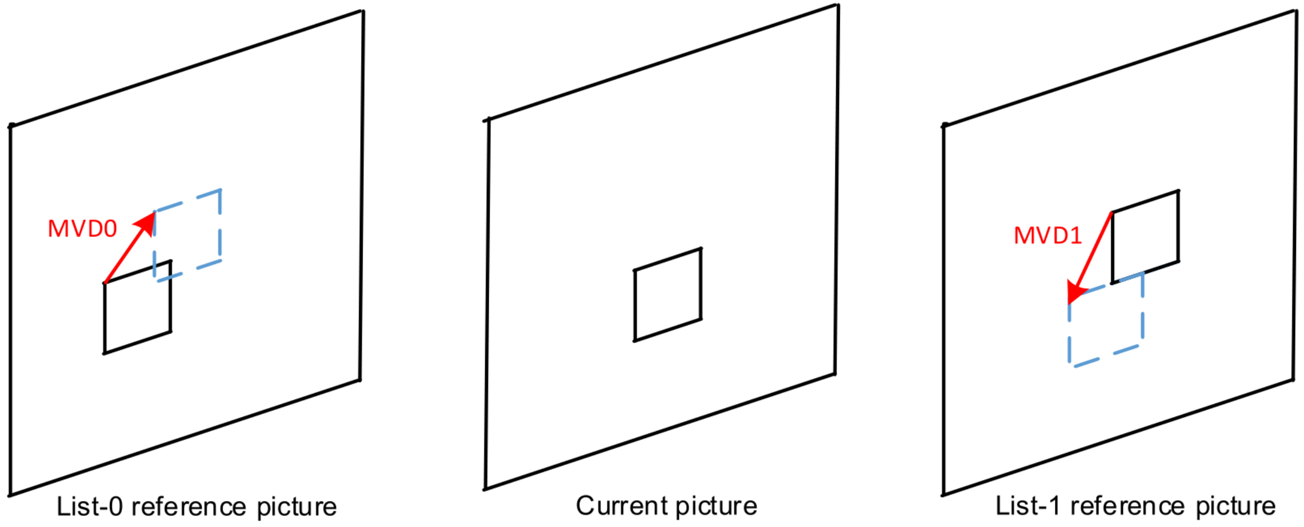
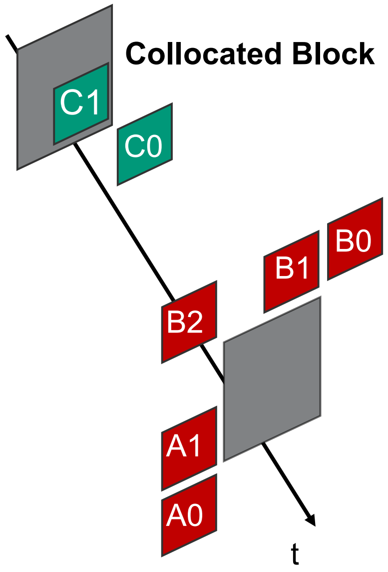
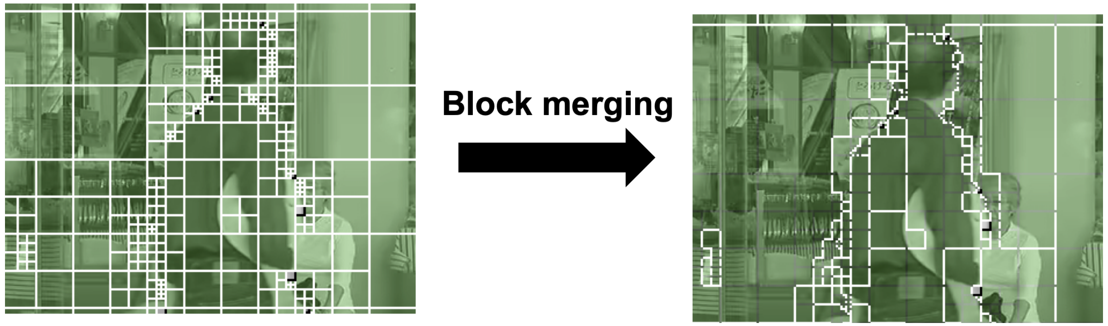


Figure from JVET-Q2002 [51]

Extended merge prediction

- Create regions of equal motion parameters
- Motion vector (MV) and reference picture index derived from list of 6 candidates
  - Spatial (A,B) and temporal (C) neighbors (same as HEVC)
  - History-based candidates
  - Pair-wise average candidates (pre-defined combination of previous candidates)
  - Zero MV (same as HEVC)
- Only merge index to candidate list is signaled



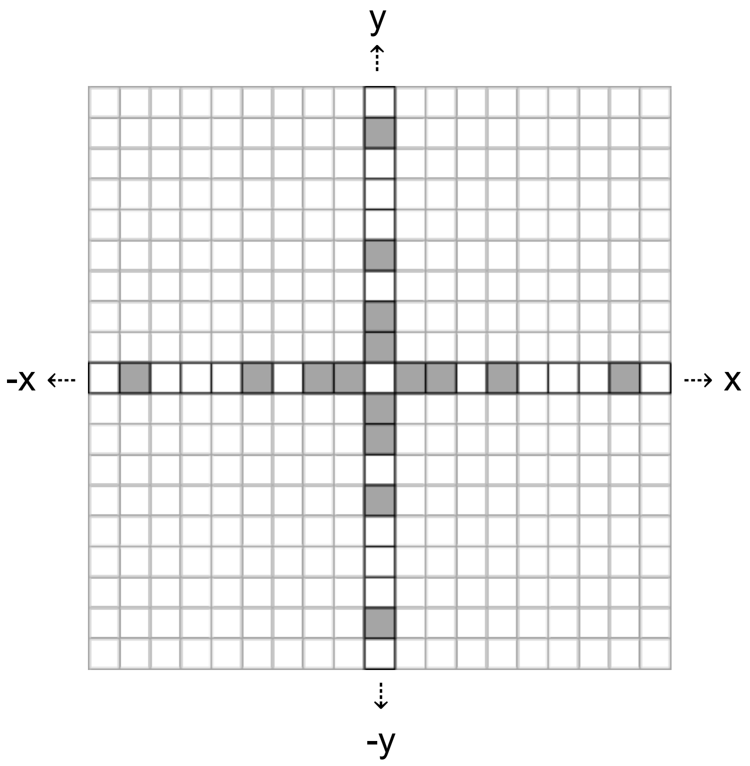


Merge with motion vector difference (MMVD)

- Motion parameters from 1st or 2nd merge list entry
- Additional MVD to merge MV signaled using index to discrete set of distances and directions:

  Distances=$\{\frac{1}{4}, \frac{1}{2}, 1, 2, 4, 8, 16, 32\}$-samples

  Directions={x,-x,y,-y}



History-based Motion Vector Prediction (HMVP)

- Additional method of MV prediction
- HMVP candidates added to merge/MVP list after the spatial MVP and TMVP
- MVs of up to 5 non-subblock inter-coded CUs stored in table
- FIFO buffer with redundancy check
- Reset for new CTU row, and at tile / slice boundaries
- Number of HMVP candidates in merge/MVP list construction depends on number of previous candidates

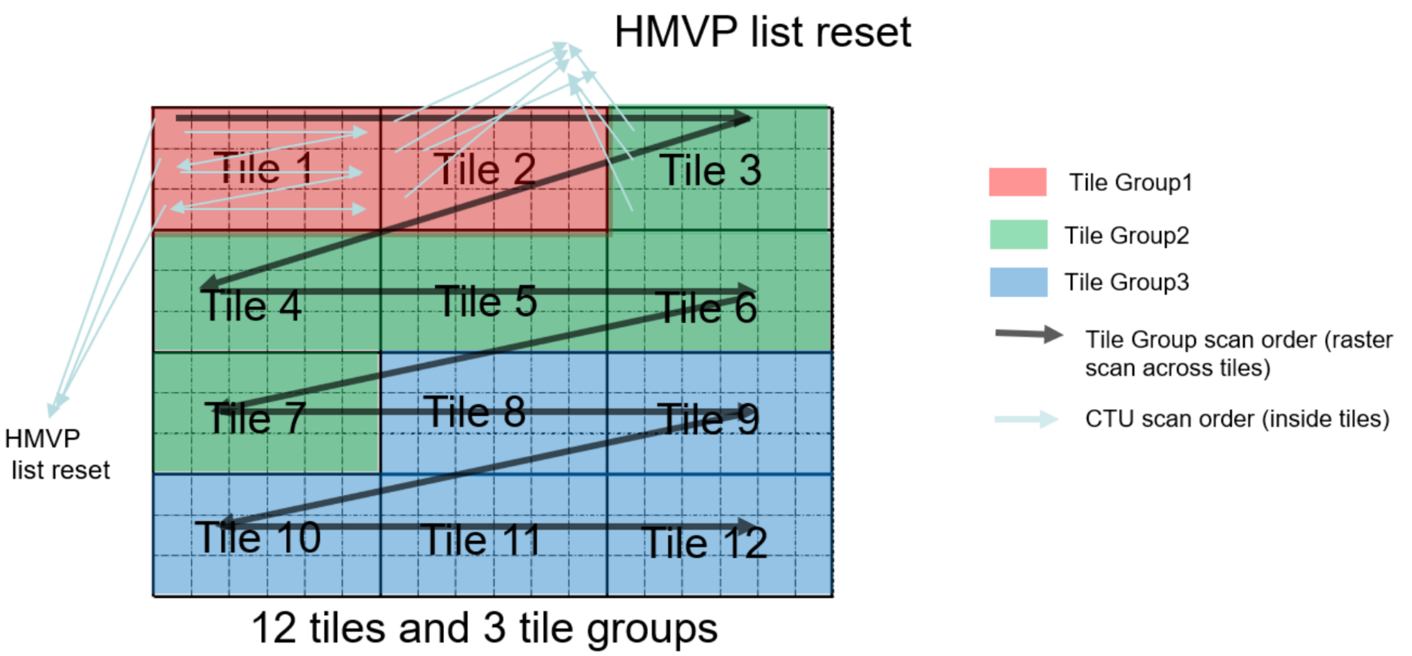


Figure from JVET-M0300 [52]

Fraunhofer Heinrich Hertz Institute
RWTH AACHEN UNIVERSITY

Affine Motion Vector Derivation for MC

- Motion vector field (MVF) for CU, applicable MV derived for each 4×4 block at $1/16$-sample resolution
- 4-parameter or 6-parameter models available
  - Control point motion vector (CPMV) for the {4|6}-parameter case

$$v_x = \frac{v_{1x} - v_{0x}}{w} \cdot x - \frac{v_{\{1|2\}y} - v_{0y}}{w} \cdot y + v_{0x}$$
$$v_y = \frac{v_{1y} - v_{0y}}{w} \cdot x + \frac{v_{\{1|2\}x} - v_{0x}}{w} \cdot y + v_{0y}$$

- Additional luma prediction refinement with optical flow (PROF) which is aligned with BDOF
- Dedicated 6-tap interpolation filters
- CPMV signaling
  - as CP-MVP+CP-MVD
  - using affine subblock merge mode (extension of SbTMVP)

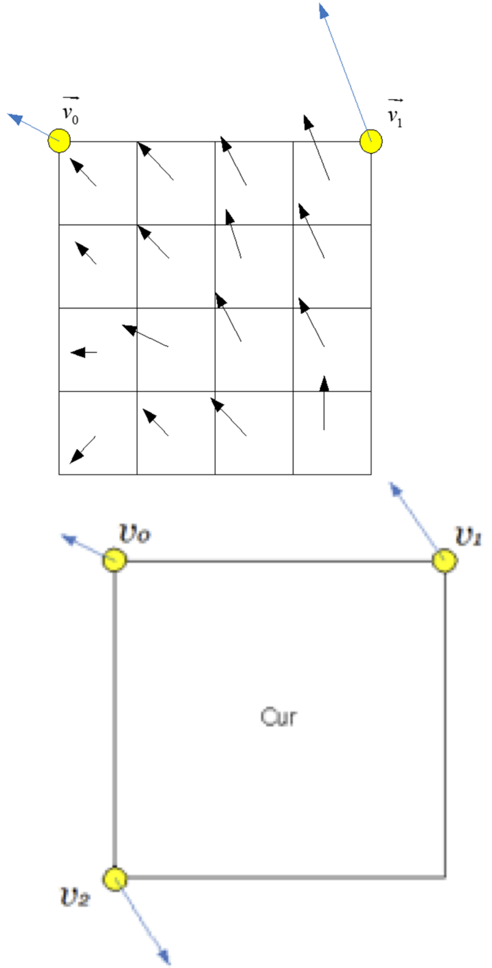


Figure from JVET-R2002 [53]

Fraunhofer Heinrich Hertz Institute
RWTH AACHEN UNIVERSITY

Subblock-based temporal motion vector prediction (SbTMVP)

- Derive local motion vector field from reference picture at 4×4 grid
- Step 1: Derive offset from local neighbor MVs (similar to MVP candidates)
- Step 2: MVs from corresponding region scaled accoring to temporal relation

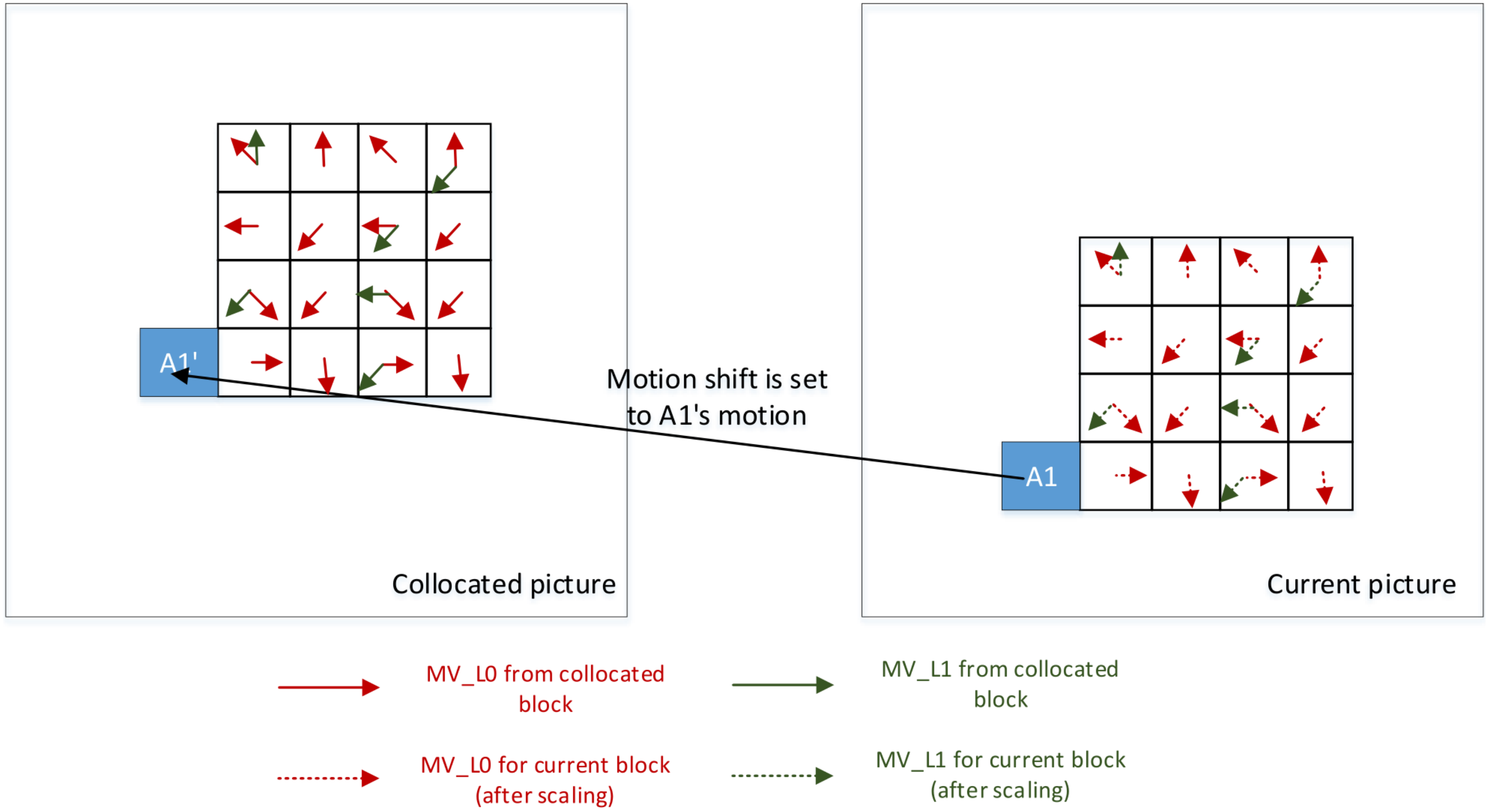
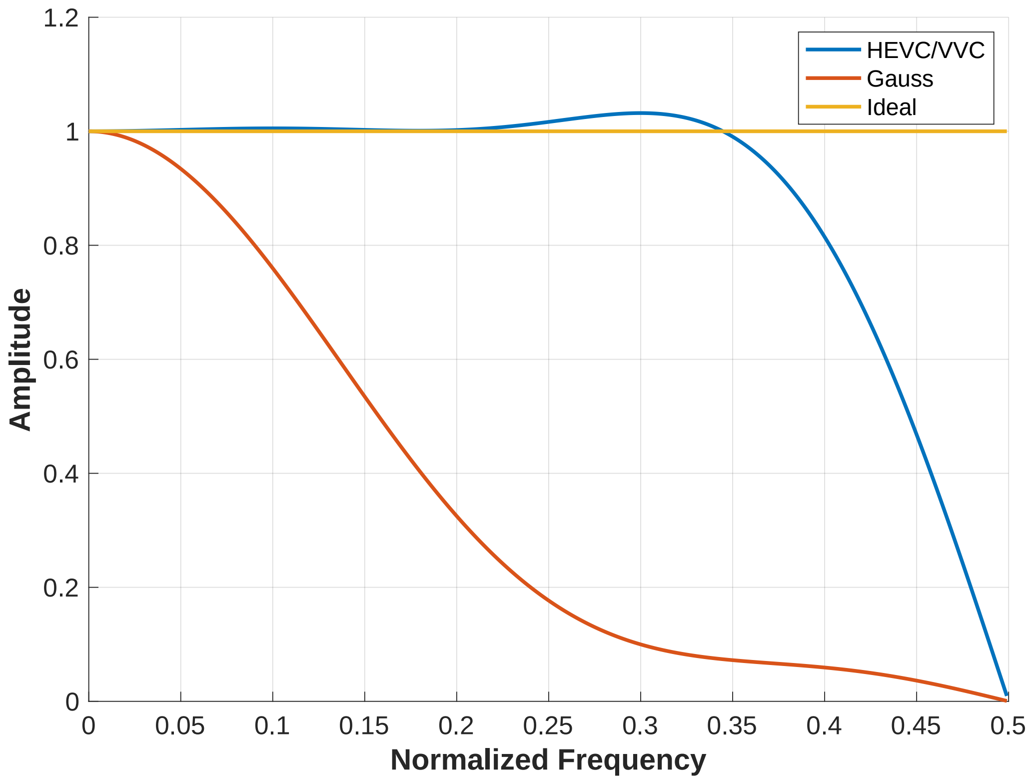


Figure from JVET-Q2002 [51]

Adaptive Motion Vector Resolution (AMVR)

- CU-level AMVR index signals resolution of MVDs for
  - translational motion compensation: $\{\frac{1}{4}, \frac{1}{2}, 1, 4\}$-samples
  - affine motion compensation: $\{\frac{1}{4}, \frac{1}{16}, 1\}$-samples
- Only signaled if all MVDs not equal to 0
- MVP rounded to signaled resolution
- Allows to trade off prediction accuracy and MV signaling overhead

- **Alternative Interpolation filter for HPEL resolution**:
  - Default HEVC/VVC filter replaced by smoothing based on Gaussian kernel
  - Symmetrical 6-tap filter: $\frac{1}{64}\{3, 9, 20, 20, 9, 3\}$
  - Merge block: HPEL IF flag inherited from neighbor

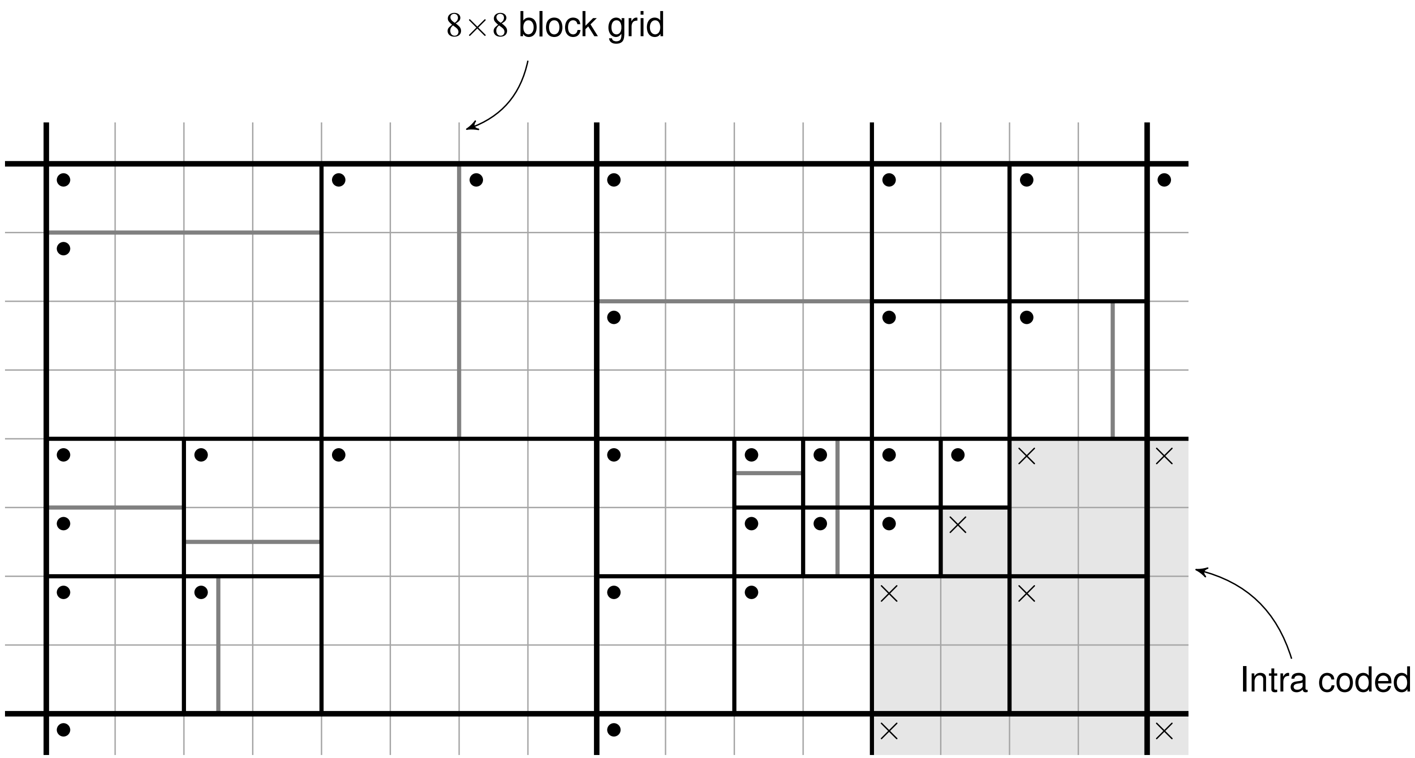


Fraunhofer Heinrich Hertz Institute

RWTH AACHEN UNIVERSITY

Motion Field Storage

- Sub-sampling of motion vector information in reference pictures
- 8×8 block grid
- Higher accuracy compared to H.265|HEVC (16×16) beneficial for SbTMVP
- Motion vectors stored at 16th-pel precision
- Motion vector at top-left block corner stored



Bi-prediction with CU-level weights (BCW)

- Weighting of the two prediction blocks in bi-prediction case (using both, List0 and List1)

$$P_{\text{bi-pred}} = \left[(8-w)\cdot P_0 + (w\cdot P_1 + 4) \gg 3\right]$$

- Candidate weights: $w \in \{-2, 3, 4, 5, 10\}$
- Minimum block size: $W \cdot H \geq 256$
- Non-Merge blocks: $w \in \{3, 4, 5\}$
  index signalled
- Merge blocks: index inherited from neighbor

Bi-directional optical flow (BDOF)

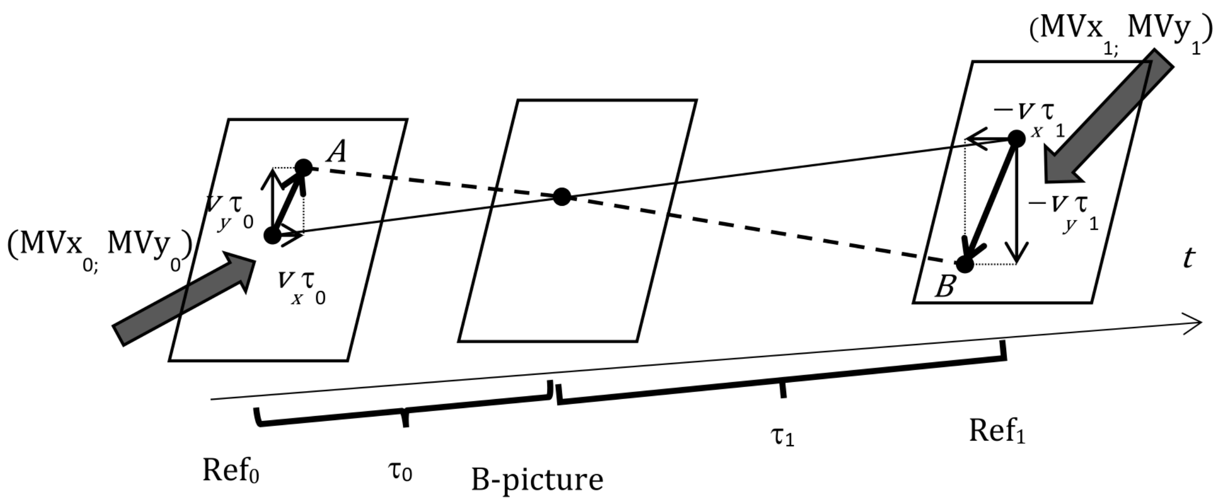


Figure from JVET-G1001 [49]

- Refine the bi-prediction signal of a CU at the 4×4 subblock level
- Conditions
  - CU equal or larger than 8×8, coded using “true” bi-prediction mode, i.e. using past and future reference pictures
  - POC difference of the reference pictures is the same, both short-term reference pictures
  - BCW indicates equal weight
  - No Affine, SbTMVP, Weighted Prediction (WP), CIIP

Decoder side motion vector refinement (DMVR)

- Applicable for merge blocks
- Conditions (similar to BDOF)
  - CU equal or larger than 8×8, coded using “true” bi-prediction mode, i.e. using past and future reference pictures
  - POC difference of the reference pictures is the same, both short-term reference pictures
  - BCW indicates equal weight
  - No WP, CIIP
- Original MV used for spatial MV prediction and deblocking
- Derived MVs used for temproal motion vector prediction

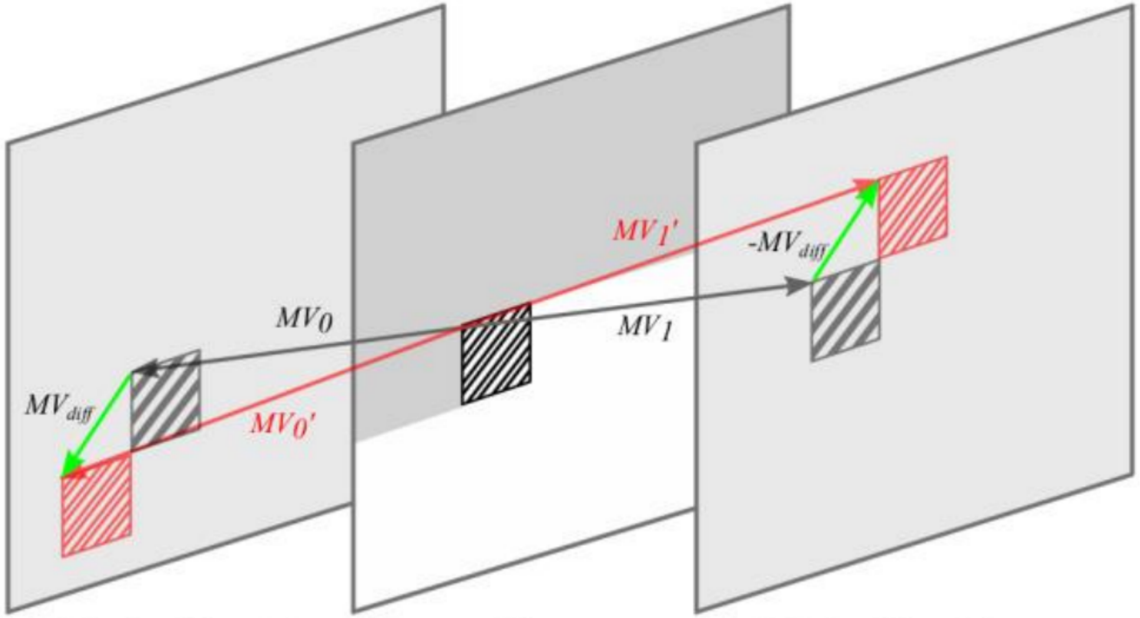


Figure from JVET-Q2002 [51]

Geometric Partitioning (GEO)

- Applied on top of merge mode
- CU split into two partitions with one merge index each
- Only uni-prediction is allowed for each partition
- Splitting line location derived from angle and offset parameters
- 64 different partition layouts supported
  (CU sizes from 8×8 to 64×64, excluding 8×64 and 64×8)
- Samples along the geometric partition edge are combined using blending

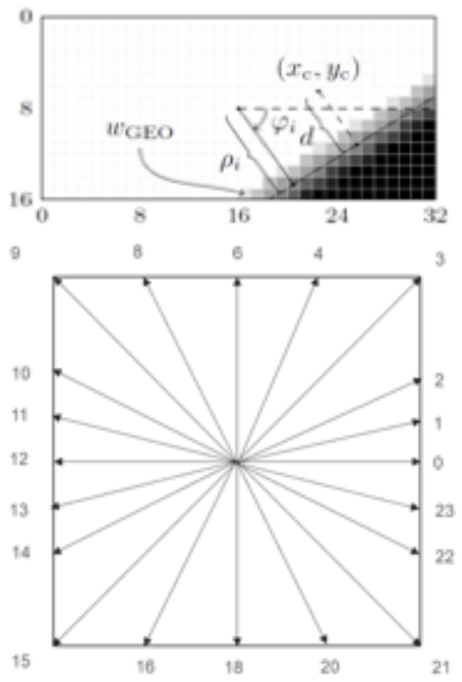
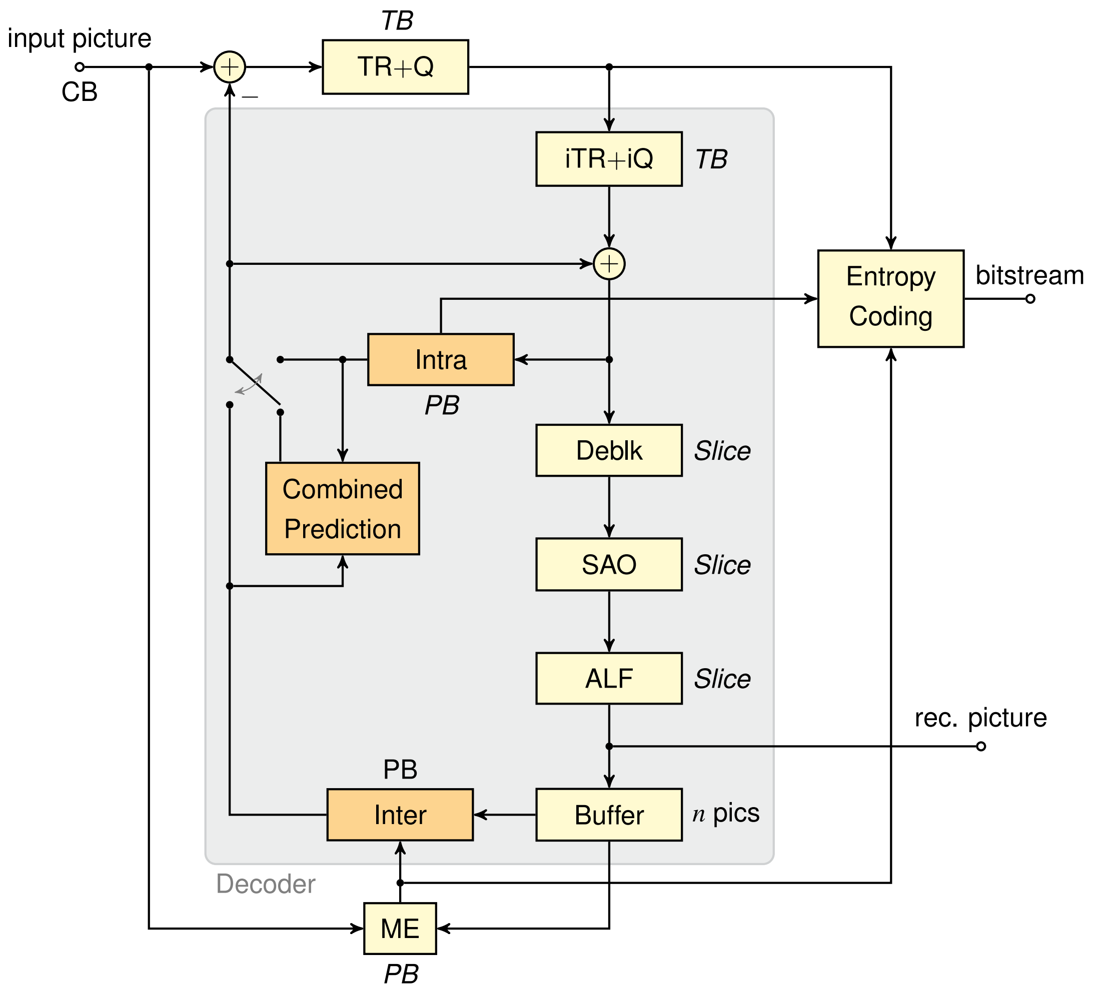


Figure from JVET-Q0024 [54]

Combined Inter and Intra Prediction



CB – Coding Block
ME – Motion Estimation
PB – Prediction Block
Q – Quantization
TB – Transform Block
TR – Transform
SAO – Sample Adaptive Offset
ALF – Adaptive Loop Filter

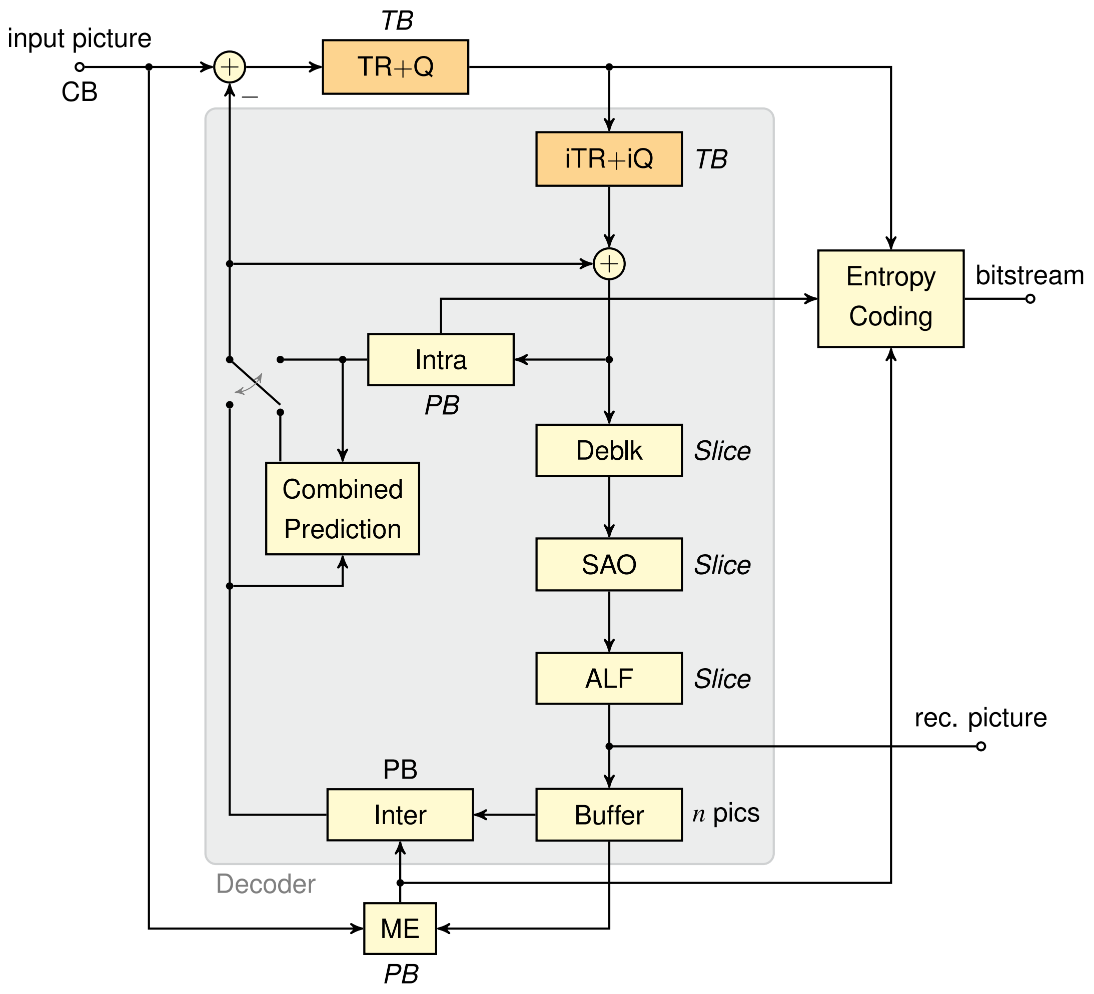
Combined inter and intra prediction (CIIP)

- Condition: CU contains at least 64 samples, width and height < 128
- Syntax element signals combined prediction
- Inter prediction: using process of merge mode
- Intra prediction: using planar mode
- Weight between inter and intra depends on number of neighboring intra blocks, $1 \leq w_{\text{CIIP}} \leq 3$

$$P_{\text{CIIP}} = \left[(4 - w_{\text{CIIP}}) \cdot P_{\text{inter}} + w_{\text{CIIP}} \cdot P_{\text{intra}} + 2\right] \gg 2$$

Contents

Transforms and Residual Coding



CB – Coding Block
ME – Motion Estimation
PB – Prediction Block
Q – Quantization
TB – Transform Block
TR – Transform
SAO – Sample Adaptive Offset
ALF – Adaptive Loop Filter

VVC Transform Matrices

- Transforms of the DCT/DST family used in VVC as primary transforms
- Combination of different transforms in horizontal and vertical direction possible
- Maximum luma transform block size is 64×64, minimum 4×4
- Secondary non-separable transform can be applied to residuals of intra predicted blocks.

- Existing HEVC transform cores reused
  - 4×4 DCT-2 and DST-7
  - 8×8, 16×16, 32×32 DCT-2
- New matrices
  - 64×64 DCT-2
  - 4×4 DCT-8
  - 8×8, 16×16, 32×32 DST-7 and DCT-8

Discrete Cosine Transform (DCT)

Family of transforms based on the $\cos$ function (even bases)

$$t_n^{\mathrm{I}}(m) = \sqrt{\frac{2}{N}} \cdot a(n) \cdot \cos\left(\frac{\pi m n}{N}\right), \quad m,n = 0,\dots,N-1,$$

$$t_n^{\mathrm{II}}(m) = \sqrt{\frac{2}{N}} \cdot a(n) \cdot \cos\left(\frac{\pi(2m+1)n}{2N}\right), \quad m,n = 0,\dots,N-1,$$

$$t_n^{\mathrm{III}}(m) = \sqrt{\frac{2}{N}} \cdot a(m) \cdot \cos\left(\frac{\pi m(2n+1)}{2N}\right), \quad m,n = 0,\dots,N-1,$$

$$t_n^{\mathrm{IV}}(m) = \sqrt{\frac{2}{N}} \cdot a(m) \cdot \cos\left(\frac{\pi(2m+1)(2n+1)}{4N}\right), \quad m,n = 0,\dots,N-1$$

$$a(n) = \begin{cases} 1/\sqrt{2} & : \quad n = 0, \\ 1 & : \quad n = 1,\dots,N-1. \end{cases}$$

Transform Matrices

- Block transform $\Rightarrow$ matrix notation

$$\mathbf{T}_{\mathsf{DCT},N} = \begin{bmatrix} \mathbf{t}_0 \\ \mathbf{t}_1 \\ \vdots \\ \mathbf{t}_{N-1} \end{bmatrix}$$

- The transform matrices are orthonormal:

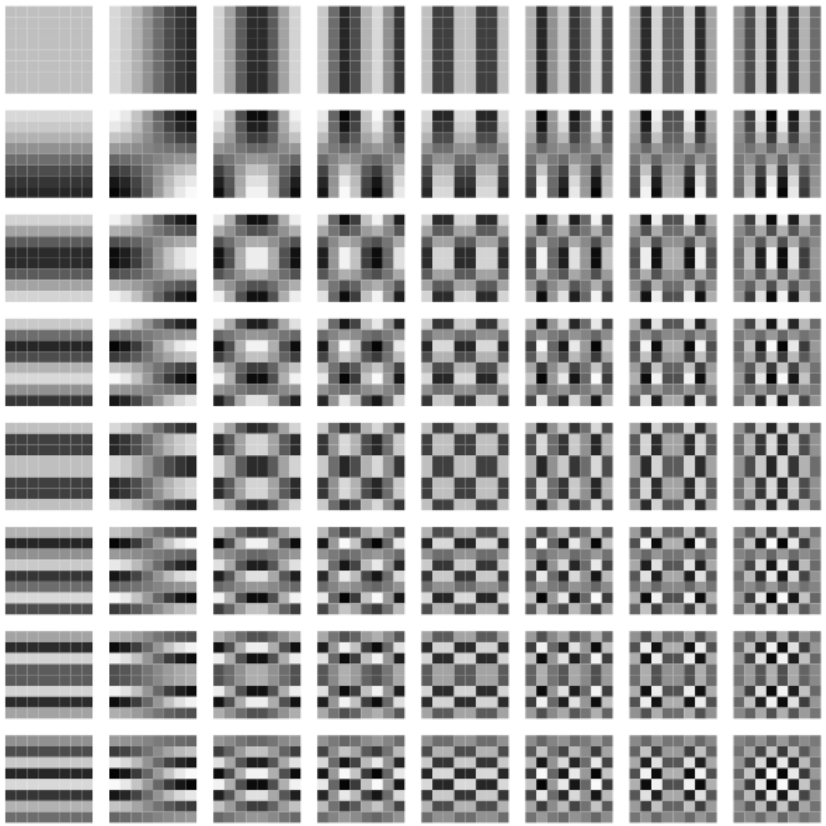
$$\mathbf{T}_{\mathsf{DCT},N} \cdot \mathbf{T}^{-1}_{\mathsf{DCT},N} = \mathbf{I}_N$$

- The transform matrices are unitary:

$$\mathbf{T}^{-1}_{\mathsf{DCT-II},N} = \mathbf{T}^{\mathsf{T}}_{\mathsf{DCT-II},N}$$

- Relation between DCT-II and DCT-III:

$$\mathbf{T}^{\mathsf{T}}_{\mathsf{DCT-II},N} = \mathbf{T}_{\mathsf{DCT-III},N}$$



Fraunhofer Heinrich Hertz Institute | RWTH AACHEN UNIVERSITY

DCT Matrix Features

- Symmetry of the DCT coefficients

$$t_n^{\mathrm{II}}(k-m) = \begin{cases} t_n^{\mathrm{II}}(m) & : \quad m = 0,\ldots,k/2-1, \quad n \text{ even},\\ -t_n^{\mathrm{II}}(m) & : \quad m = 0,\ldots,k/2-1, \quad n \text{ odd}. \end{cases}$$

- Recursive construction of transform bases, $K = 2N$

$$\mathbf{t}^{\mathrm{II}}_{2n,K} = \frac{1}{\sqrt{2}} \cdot \left[\mathbf{t}_{n,N}, \quad \mathbf{J}_N \cdot \mathbf{t}^{\mathrm{II}}_{n,N}\right],$$

  with $\mathbf{J}_N$ opposite identity or reversal matrix of size $N \times N$

- Overall $N-1$ different coefficient values in a $N \times N$ transform matrix

Discrete Sine transform (DST)

Family of transforms based on the $\sin$ function (odd bases)

$$t_n^{\mathsf{V}}(m) = \frac{2}{\sqrt{2N-1}} \cdot \sin\left(\frac{2\pi mn}{2N-1}\right), \quad m,n = 1,\dots,N-1$$

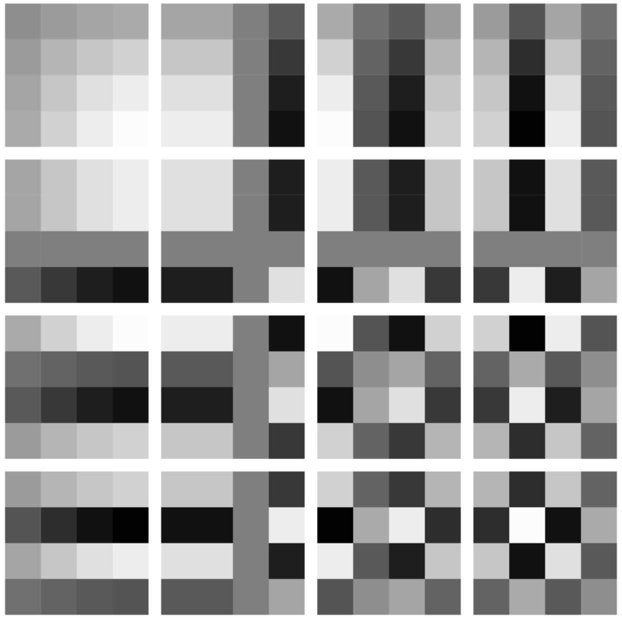
$$t_n^{\mathsf{VI}}(m) = \frac{2}{\sqrt{2N-1}} \cdot \sin\left(\frac{\pi(2m-1)n}{2N-1}\right), \quad m,n = 1,\dots,N-1$$

$$t_n^{\mathsf{VII}}(m) = \frac{2}{\sqrt{2N-1}} \cdot \sin\left(\frac{\pi m(2n-1)}{2N-1}\right), \quad m,n = 1,\dots,N-1$$

$$t_n^{\mathsf{VIII}}(m) = \frac{2}{\sqrt{2N-1}} \cdot \sin\left(\frac{\pi(2m-1)(2n-1)}{2(2N-1)}\right), \quad m,n = 1,\dots,N-1$$

HaSaRo10 [55], JCTVC-B024 [56]

Example: 4×4 DST base pictures



Integer Transform Implementation

- Normalization for integer transform required (single-norm example)

$$\mathbf{T} \cdot \mathbf{T}^{\mathsf{T}} = \|\mathbf{T}\|^2 \cdot \mathbf{I}$$

- Implementation into quantization process and reconstruction, quantizer step size Δ_{q}

$$\mathbf{n}_{\mathsf{q}} = \text{round}\left\{ \frac{1}{\|\mathbf{T}\|^2 \cdot \Delta_{\mathsf{q}}} \cdot \mathbf{T} \cdot \mathbf{X} \cdot \mathbf{T}^{\mathsf{T}} \right\}$$

$$\mathbf{X}_{\mathsf{r}} = \text{round}\left\{ \frac{\Delta_{\mathsf{q}}}{\|\mathbf{T}\|^2} \cdot \mathbf{T}^{\mathsf{T}} \cdot \mathbf{n}_{\mathsf{q}} \cdot \mathbf{T} \right\}$$

- Integer approximation

$$\frac{f_{\mathsf{q}}}{2^{N_{\mathsf{e}}}} \approx \frac{1}{\|\mathbf{T}\|^2 \cdot \Delta_{\mathsf{q}}} \quad \text{and} \quad \frac{g_{\mathsf{q}}}{2^{N_{\mathsf{d}}}} \approx \frac{\Delta_{\mathsf{q}}}{\|\mathbf{T}\|^2}$$

VVC DST7/DCT8 Design

Features of the N-point DST-7/DCT-8 transform core

- N distinct coefficient values with varying sign, or additional zeros
- Some base vectors: sum of several (2 or 3) coefficients equals to the sum of another several (1 or 2)
- Some replicated patterns with symmetric or anti-symmetric characteristics
- Some base vector(s) with very few (1 or 2) distinct number(s) without considering the sign changes.

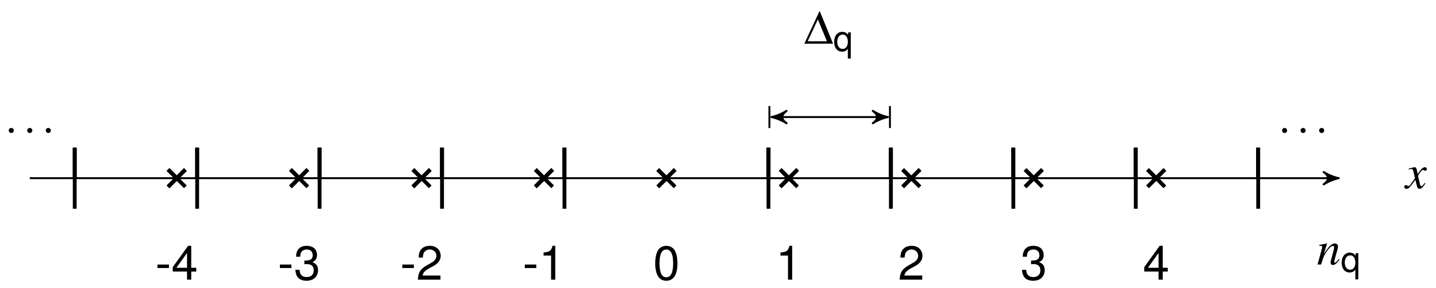
Approach

- Tune integer coefficient values to fulfill same properties
- Implementation exploiting features
  $\Rightarrow$ up to 50% reduction of multiplications

```
{ a  b  c  d  e  f  g  h  i  j  k  l  m  n  o  p }
{ c  f  i  l  o  o  l  i  f  c  0 -c -f -i -l -o }
{ e  j  o  m  h  c -b -g -l -p -k -f -a  d  i  n }
{ g  n  l  e -b -i -p -j -c  d  k  o  h  a -f -m }
{ i  o  f -c -l -l -c  f  o  i  0 -i -o -f  c  l }
{ k  k  0 -k -k  0  k  k  0 -k -k  0  k  k  0 -k }
{ m  g -f -n -a  l  h -e -o -b  k  i -d -p -c  j }
{ o  c -l -f  i  i -f -l  c  o  0 -o -c  l  f -i }
{ p -a -o  b  n -c -m  d  l -e -k  f  j -g -i  h }
{ n -e -i  j  d -o  a  m -f -h  k  c -p  b  l -g }
{ l -i -c  o -f -f  o -c -i  l  0 -l  i  c -o  f }
{ j -m  c  g -p  f  d -n  i  a -k  l -b -h  o -e }
{ h -p  i -a -g  o -j  b  f -n  k -c -e  m -l  d }
{ f -l  o -i  c  c -i  o -l  f  0 -f  l -o  i -c }
{ d -h  l -p  m -i  e -a -c  g -k  o -n  j -f  b }
{ b -d  f -h  j -l  n -p  o -m  k -i  g -e  c -a }
```

Figure from JVET-K0291 [57]

Basic Quantizer Design



- Quantizer step size Δ_q derived from quantization parameter QP
- Logarithmic relation of quantizer step sizes
- Double step size every 6 QP

$$\Delta_q(\mathrm{QP}+1) = \sqrt[6]{2} \cdot \Delta_q(\mathrm{QP})$$

- Definition: $\Delta_q = 1$ for $\mathrm{QP} = 4$, thereby

$$\Delta_{q,0} \in \left\{ 2^{-\frac{4}{6}}, 2^{-\frac{3}{6}}, 2^{-\frac{2}{6}}, 2^{-\frac{1}{6}}, 1, 2^{\frac{1}{6}} \right\}$$

- Quantizer step sizes for $\mathrm{QP} > 5$

$$\Delta_q(\mathrm{QP}) = \Delta_{q,0}(\mathrm{QP} \bmod 6) \cdot 2^{\left\lfloor \frac{\mathrm{QP}}{6} \right\rfloor}$$

Quantizer Step Size

- Quantizer range for 8 bit video

$$\mathsf{QP} = 0, \ldots, 62$$

- Resulting quantizer step sizes

$$0.630 \le \Delta_\mathsf{q} \le 812.7$$

- Covering an extended value range (HEVC: $\mathsf{QP}_\mathsf{max} = 51$)

- Higher input bit depth:
  - Extension towards finer quantization
  - Range extended by 6 QP steps per additional bit

Integer Quantizer Implementation

- Integer approximation (see basics lecture)

$$\frac{f_\mathsf{q}}{2^{N_\mathsf{e}}} \approx \frac{1}{\|\mathbf{T}^2\| \cdot \Delta_\mathsf{q}} \quad \text{and} \quad \frac{g_\mathsf{q}}{2^{N_\mathsf{d}}} \approx \frac{\Delta_\mathsf{q}}{\|\mathbf{T}\|^2}$$

- Here: scaled integer quantizer step sizes, $g_\mathsf{q} = \mathrm{round}\left\{2^6 \cdot \Delta_\mathsf{q}\right\}$

$$g_{\mathsf{q},0} \in \{40, 45, 51, 57, 64, 72\}$$

- Encoder side division by quantizer step size: scale and shift with $f_\mathsf{q} = \mathrm{round}\left\{\frac{2^{14}}{\Delta_\mathsf{q}}\right\}$

$$f_{\mathsf{q},0} \in \{26214, 23302, 20560, 18396, 16384, 14564\}$$

- Resulting $f_{\mathsf{q},0} \cdot g_{\mathsf{q},0} \approx 2^{20}$ at high precision

Multiple Transform Selection (MTS)

- No 64-length DST7 and DCT8
  - No MTS syntax sent when either dimension is larger than 32
- Only DCT2, DST7 and DCT8
- Applied only for luma

Table 8-14 – Specification of trTypeHor and trTypeVer depending on tu_mts_idx[x][y]

tu_mts_idx[x0][y0]	**0**	**1**	**2**	**3**	**4**
trTypeHor	0	1	2	1	2
trTypeVer	0	1	1	2	2

Table 8-15 – Specification of trTypeHor and trTypeVer depending on cu_sbt_horizontal_flag and cu_sbt_pos_flag

cu_sbt_horizontal_flag	**cu_sbt_pos_flag**	trTypeHor	trTypeVer
0	**0**	2	1
0	**1**	1	1
1	**0**	1	2
1	**1**	1	1

Figure from JVET-O2001 [48]

Large Transform Coefficient Scan

- High frequency coefficients of large transform blocks forced to zero outside of top-left 32×32 area (complexity consideration), JVET-M0297 [58]
- Scanning of transform coefficients adapted, JVET-M0257 [59]

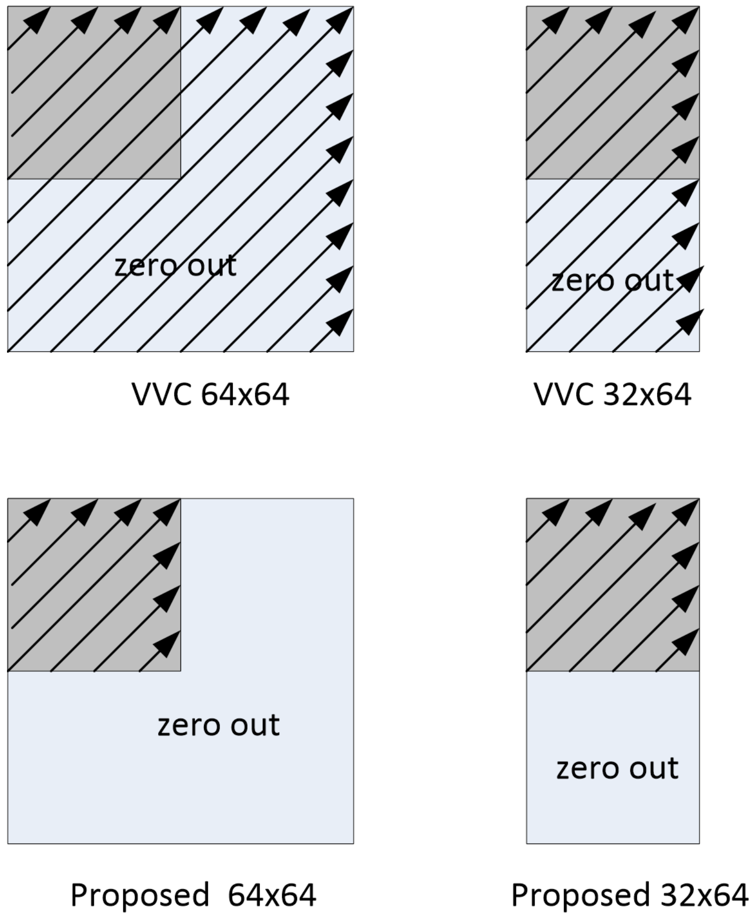


Figure from JVET-M0257 [59]

Subblock Transforms (SBT) for Inter CUs

- Inter CU with CBF=1 split into
  - two subblocks (horizontally or vertically)
  - four subblocks (quad split)
- Only one subblock contains coefficients
- Implicit MTS mapping

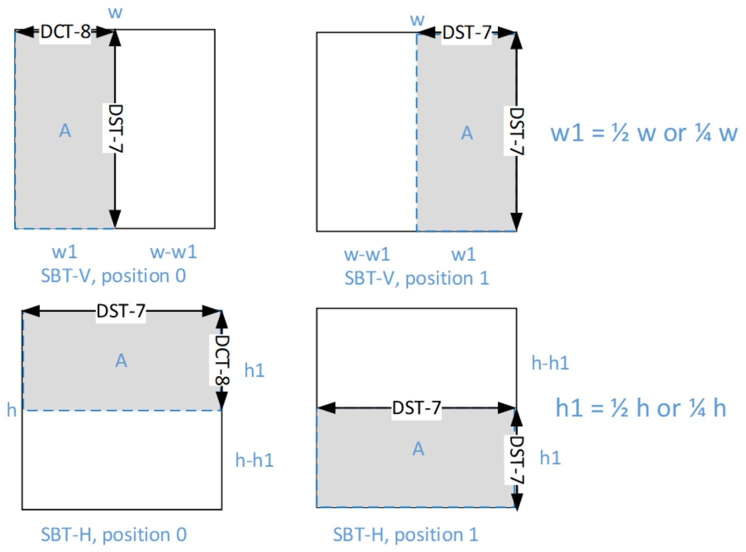


Figure 1 Illustration of sub-block transform modes SBT-V and SBT-H

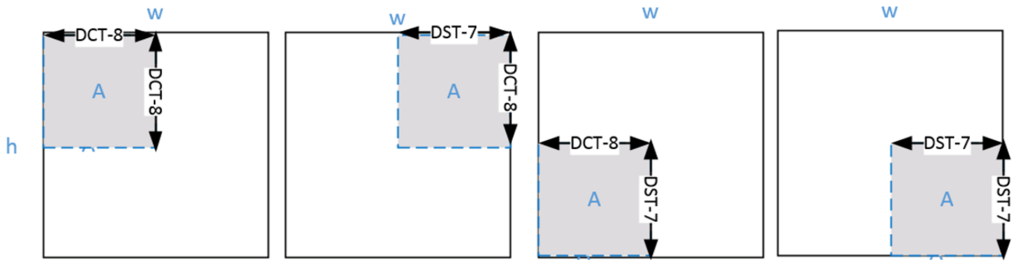


Figure 2 Illustration of sub-block transform modes SBT-Q used in CE6.4.1-e

Figure from JVET-M0140 [60]

Low frequency non-separable transform (LFNST)

- Strive for maximum decorrelation of residual signal!
- LFNST can be applied to residuals of intra-coded blocks
- 4×4 or 8×8 non-separable transform to low-frequency coefficients (determined by minimum block width / height)
- Input to transform:
  - 16 primary transform coefficients in case of 4×4 LFNST
  - 48 primary transform coefficients in case of 8×8 LFNST
- Output of transform
  - 8 coefficients in case of 4×4 LFNST
  - 16 coefficients in case of 8×8 LFNST

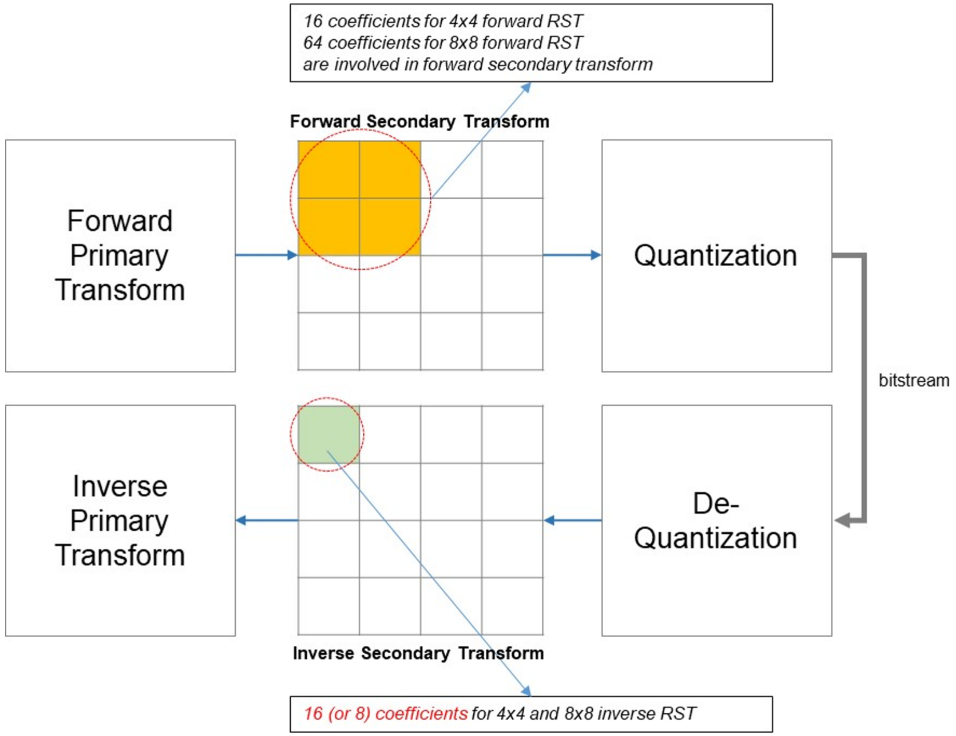
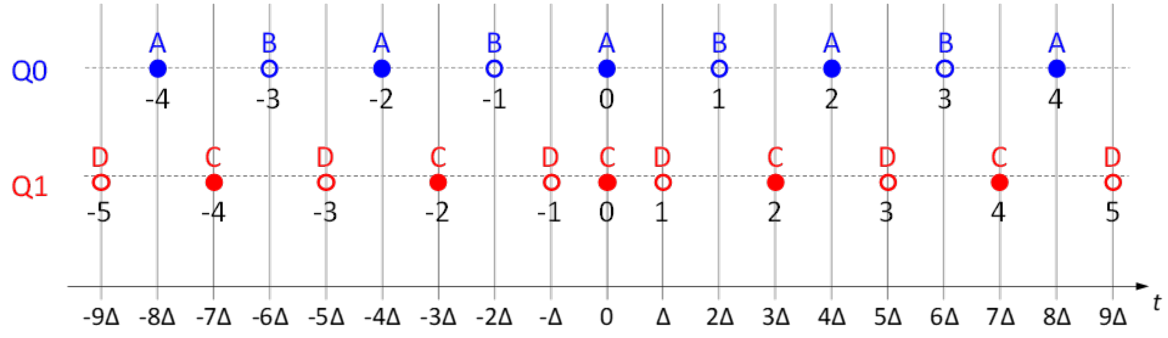
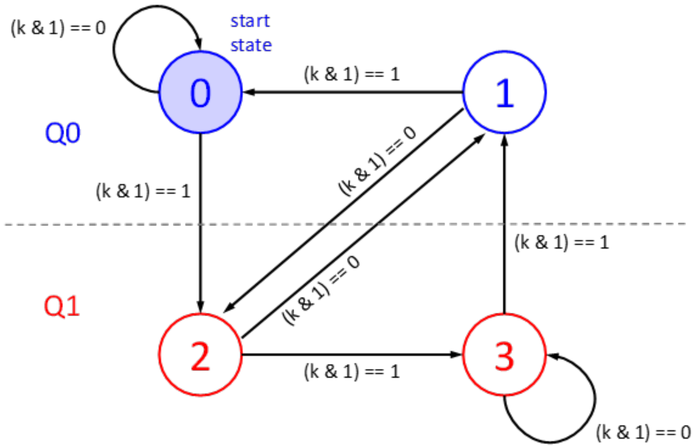


Figure from JVET-N0193 [61]

Dependent quantization

- Alternating between two quantizers based on state transition rule allows to select an optimum sequence of reconstruction values (e.g. by trellis-like search)
- Decoder needs to implement the sequential state transition rule
- CABAC contexts needs to be modified as well for this case (greater than 0/1/2/... would have different meaning depending on Q0/Q1)





current state	next state for ...	
	(k & 1) == 0	(k & 1) == 1
0	0	2
1	2	0
2	1	3
3	3	1

Figures from JVET-K0071 [62]

Joint coding of chroma residuals (JCCR)

tu_cbf_cb	tu_cbf_cr	reconstruction of Cb and Cr residuals	mode
1	0	resCb[x][y] = resJointC[x][y] resCr[x][y] = (CSign * resJointC[x][y]) >> 1	1
1	1	resCb[x][y] = resJointC[x][y] resCr[x][y] = CSign * resJointC[x][y]	2
0	1	resCb[x][y] = (CSign * resJointC[x][y]) >> 1 resCr[x][y] = resJointC[x][y]	3

Figure from JVET-Q2002 [51]

- Three modes available for intra blocks, only mode 2 for inter blocks
- Usage generally indicated by a flag on PPS and slice header level
- Implicit usage if CBF for both chroma planes indicate coefficients

Versatile Video Coding – Algorithms and Specification

Part 4 | Coding Tools II and Performance

IEEE ICME, 10.07.2020

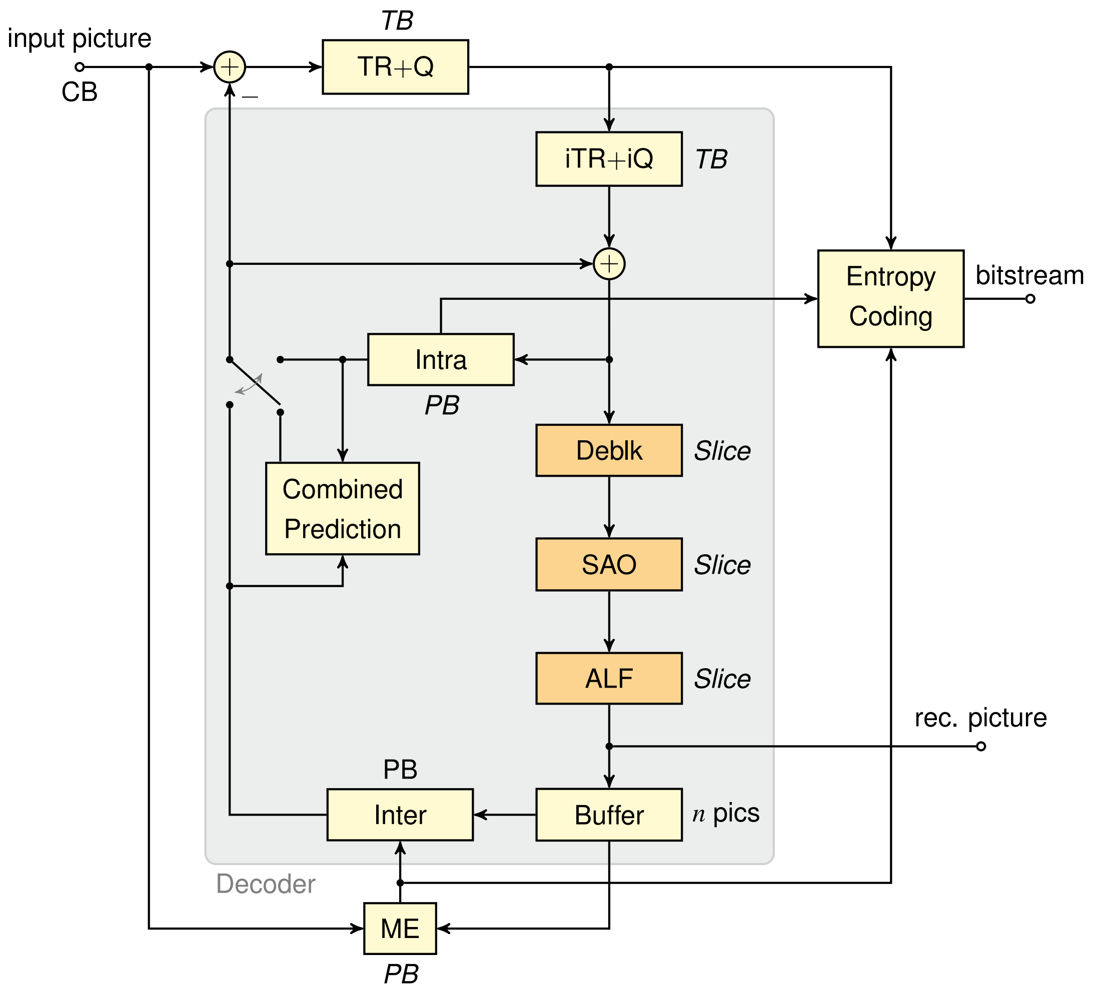
Mathias Wien, Lehrstuhl für Bildverarbeitung, RWTH Aachen University
Benjamin Bross, Heinrich Hertz Institute, Fraunhofer Gesellschaft

Fraunhofer Heinrich Hertz Institute

RWTH AACHEN UNIVERSITY

Contents

Loop Filtering



CB – Coding Block
ME – Motion Estimation
PB – Prediction Block
Q – Quantization
TB – Transform Block
TR – Transform
SAO – Sample Adaptive Offset
ALF – Adaptive Loop Filter

Reshaper – Luma Mapping with Chroma Scaling (LMCS)

LMCS operation

- Scaling of input signal across dynamic range
- In-loop mapping of luma based on adaptive piecewise linear models
  - Mapping function with 16 segments
  - Signaled in the APS (up to 4 APSs per sequence)
- Luma-dependent chroma residual scaling for chroma
  - Compensate for interaction between luma and corresponding chroma
  - Activation by flag

- Originally proposed for HDR Video, applied before regular loop filters

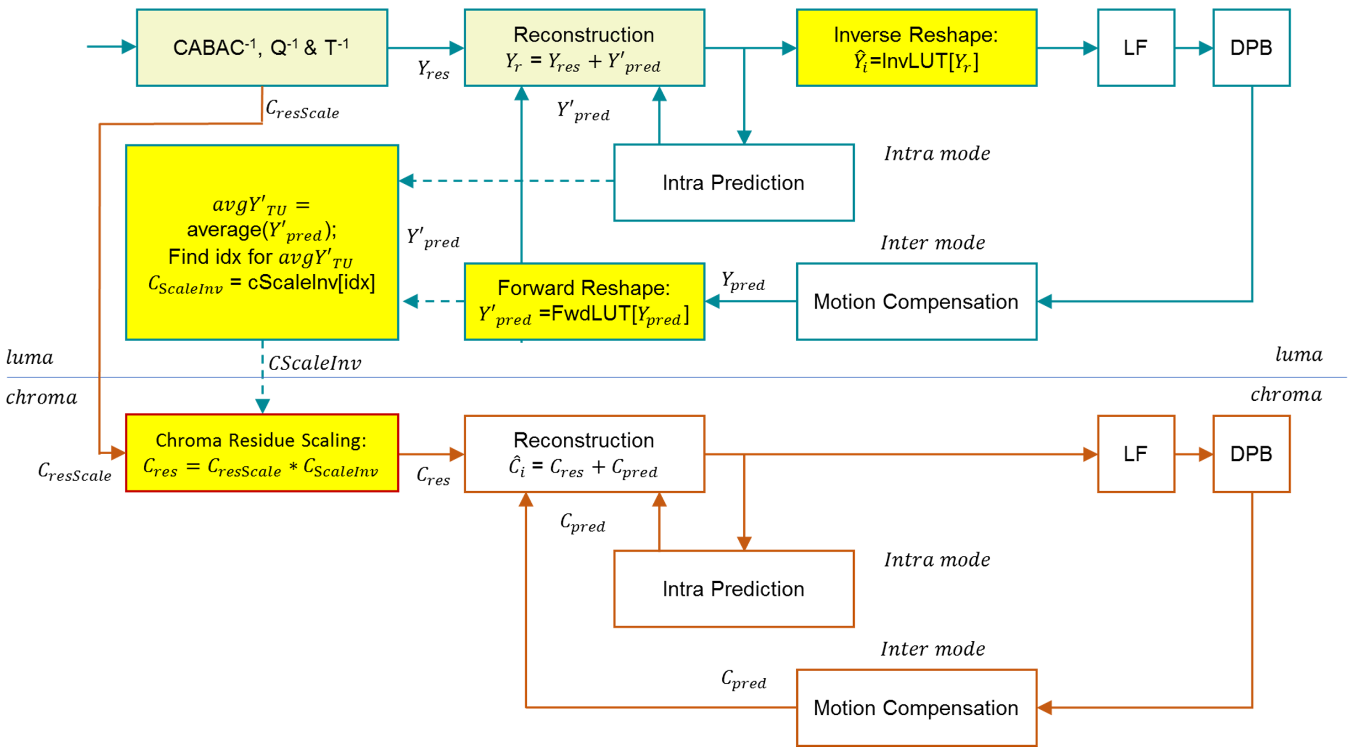
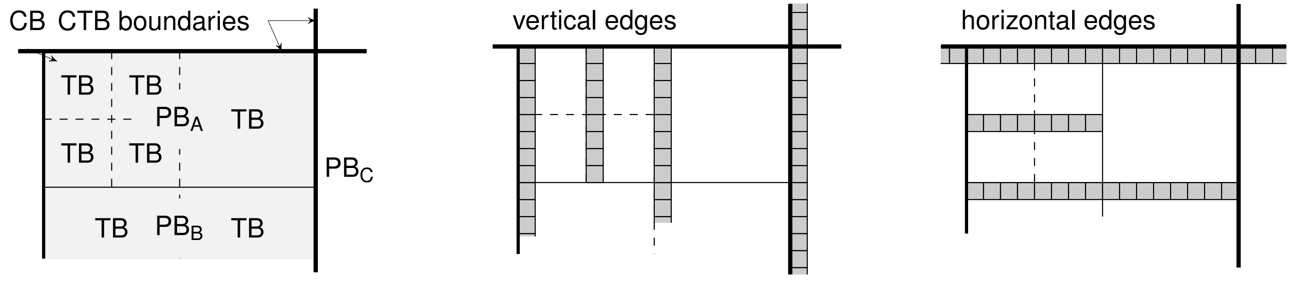


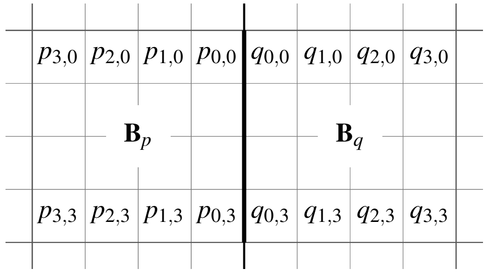
Figure from JVET-M0427 [63]

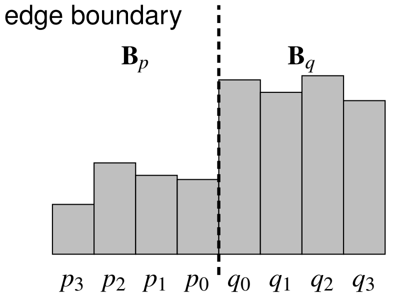
Deblocking Filter

- **Deblocking filter application**
  - Filtering at CU and transform subblock boundaries on 4×4 grid
    - Subblock Transform (SBT), Intra Sub-Partitioning (ISP)
  - Filtering at prediction subblock boundaries on 8×8 grid
    - Subblock-based Temporal Motion Vector Prediction (SbTMVP), affine modes
  - Vertical edges 1st, horizontal edges 2nd
  - Parallel processing of edge types enabled, also CTU-based processing possible



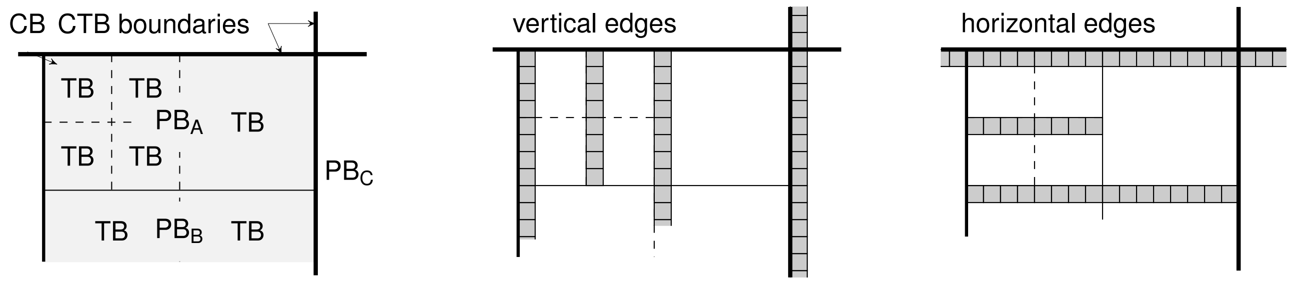
(a) (b) (c)



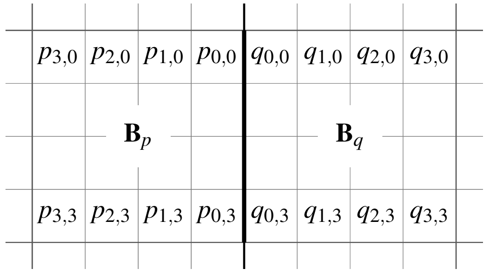


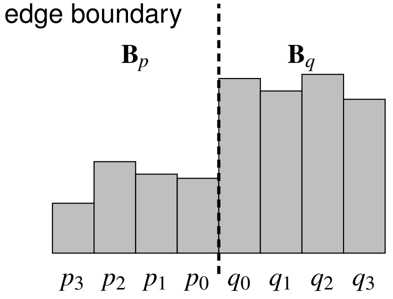
Deblocking Filter

- **Deblocking filter operation** similar to HEVC
  - Boundary processed in 4-sample sections (edges)
  - Boundary strength, threshold parameters, depending on QP
  - *New*: Long strong filtering for luma (block ≥ 32, filtering 7 samples)
  - *New*: Strong filtering for chroma (block ≥ 8, filtering 3 samples)
- **Luma-Adaptive Deblocking** (for HDR content)
  - Additional QP offset based on average luma level, signaled in SPS
  - Application-based derivation, e. g., based on EOTF and OOTF of video contents



(a) (b) (c)





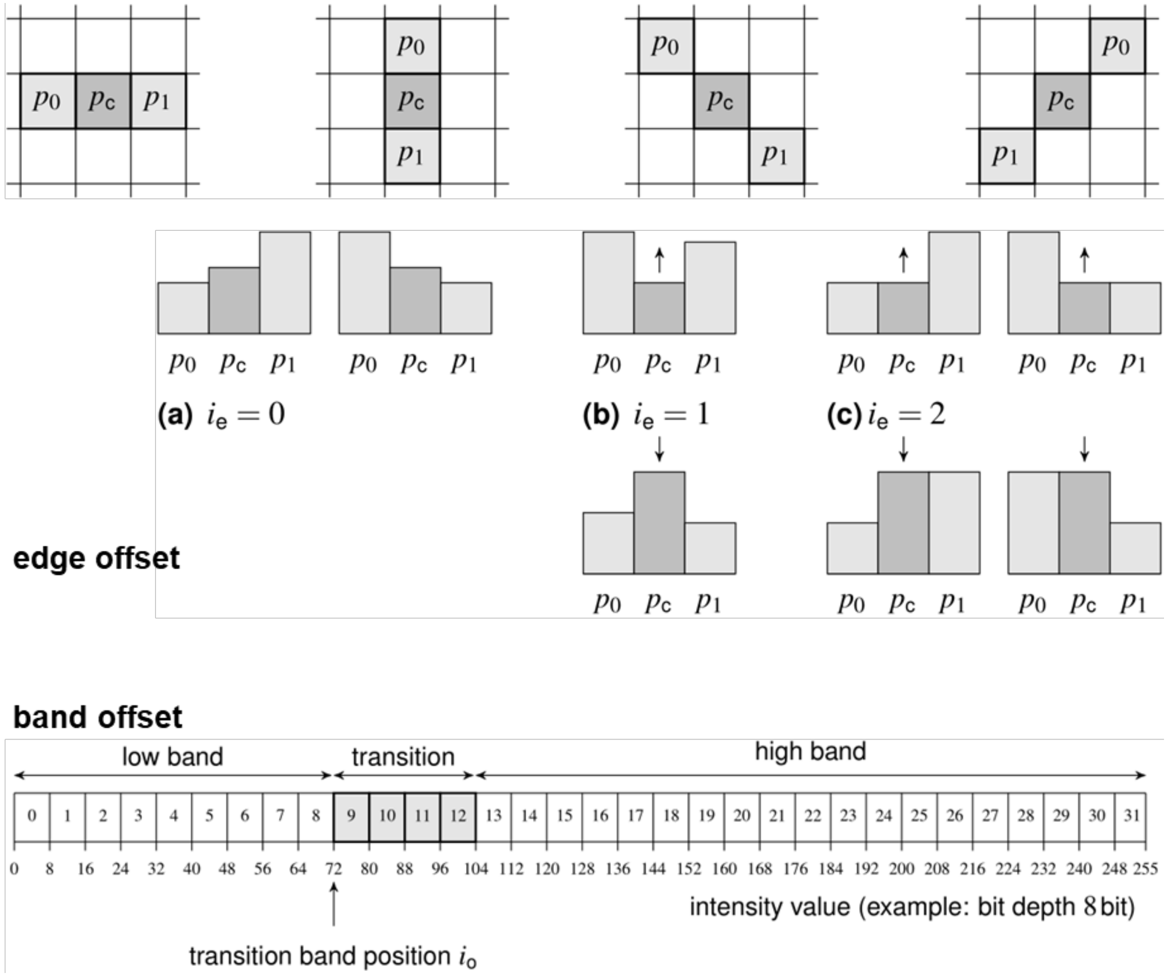
EOTF = Electro-Optical Transfer Function, OOTF = Opto-Optical Transfer Function, ITURBT2390 [64]

Sample Adaptive Offset Filter (SAO)

- **HEVC SAO filtering also used in VVC**

- **Local processing of samples**
  - Depending on local neighborhood (**edge offset**)
    - Direction signaled, smoothing only
  - Depending on sample value (**band offset**)
    - Configurable correction of sample intensity values for four transition bands

- Operation independent of processed samples $\rightarrow$ parallel processing
- Local filter parameter adaptation
- Four different offset values available (plus SAO off)
- Dedicated SAO parameters for Y, Cb, Cr
  - Common SAO mode for chroma components



Adaptive loop filter (ALF)

- **Luma component**
  - 25 filters available for each 4×4 block, based on direction and activity of local gradients
  - Diamond filter shapes (5×5, 7×7)
  - Classification into 25 classes, based on
    - Activity index, directionality index

  Chroma components
  - Diamond filter shape 5×5
  - No classification
  - Single set of filter coefficients
- Geometric transformations based on data from classification
  - Transpose, vertical flip, rotation
- Clipping to minimize difference of filtered samples and neighbor samples $\rightarrow$ non-linear operation
- Signaling of filter coefficients in APS

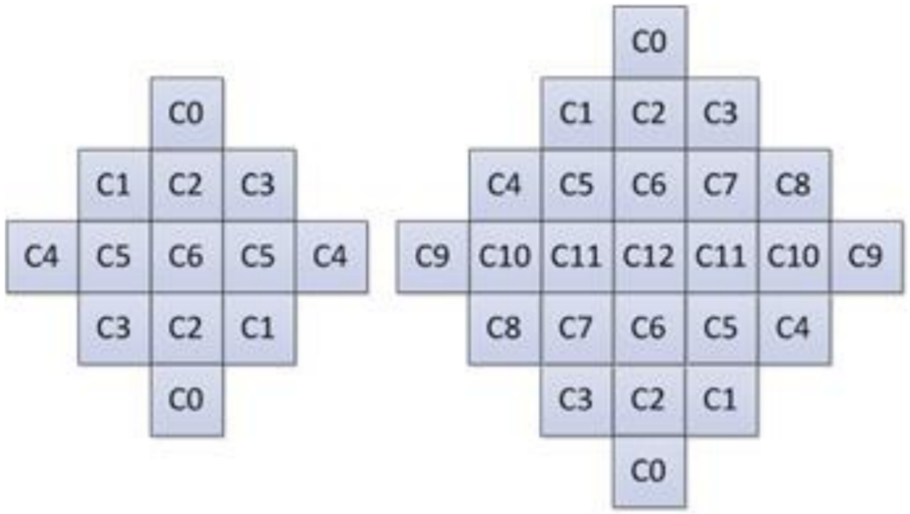


Figure from JVET-Q2002 [51]

ALF – Virtual Boundary Processing

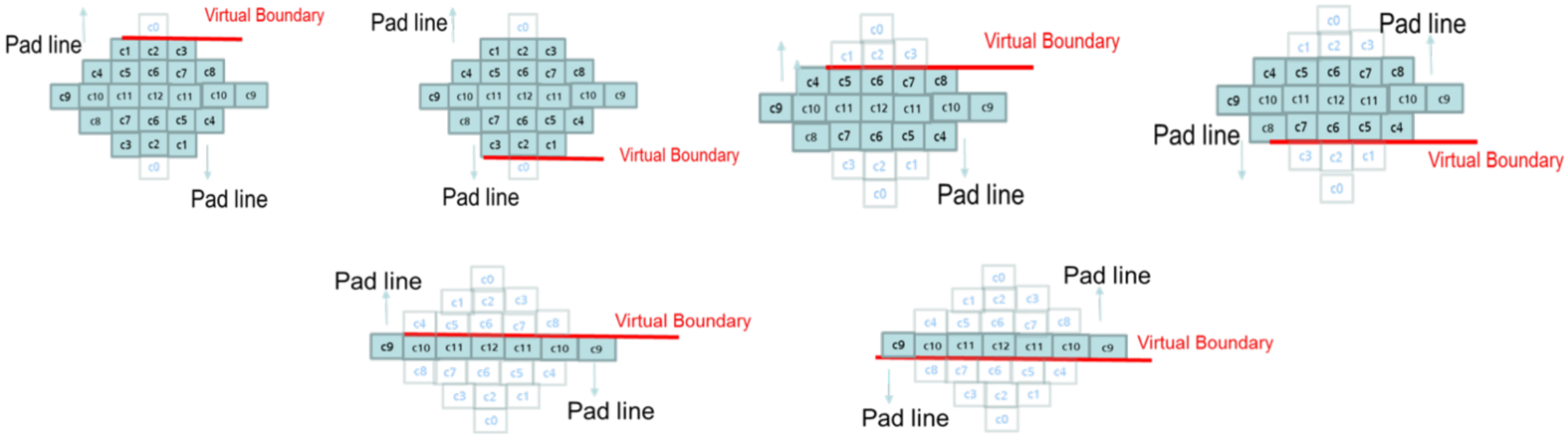
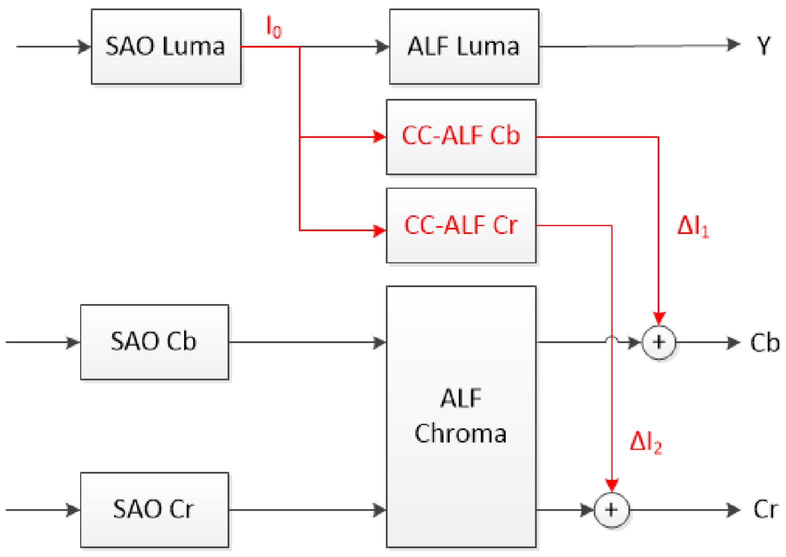


Figure from JVET-Q2002 [51]

- For reduced line buffer requirement of ALF, modified block classification and filtering are employed for the samples near horizontal CTU boundaries
- Padding applied at slice, tile, subpicture, and picture boundaries when filter across the boundaries is disabled

Cross-Component Adaptive Loop Filter

- Operates in parallel to luma ALF
- Diamond-shaped high-pass linear filter on luma
- Output of this filtering operation used for chroma refinement



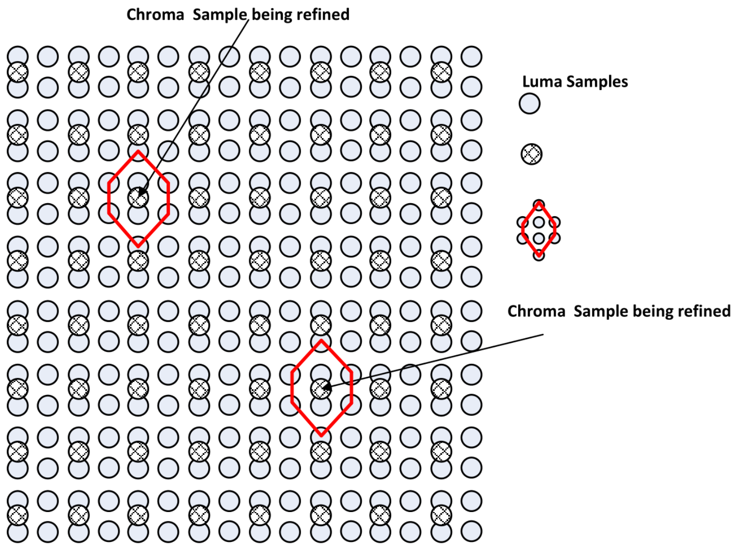
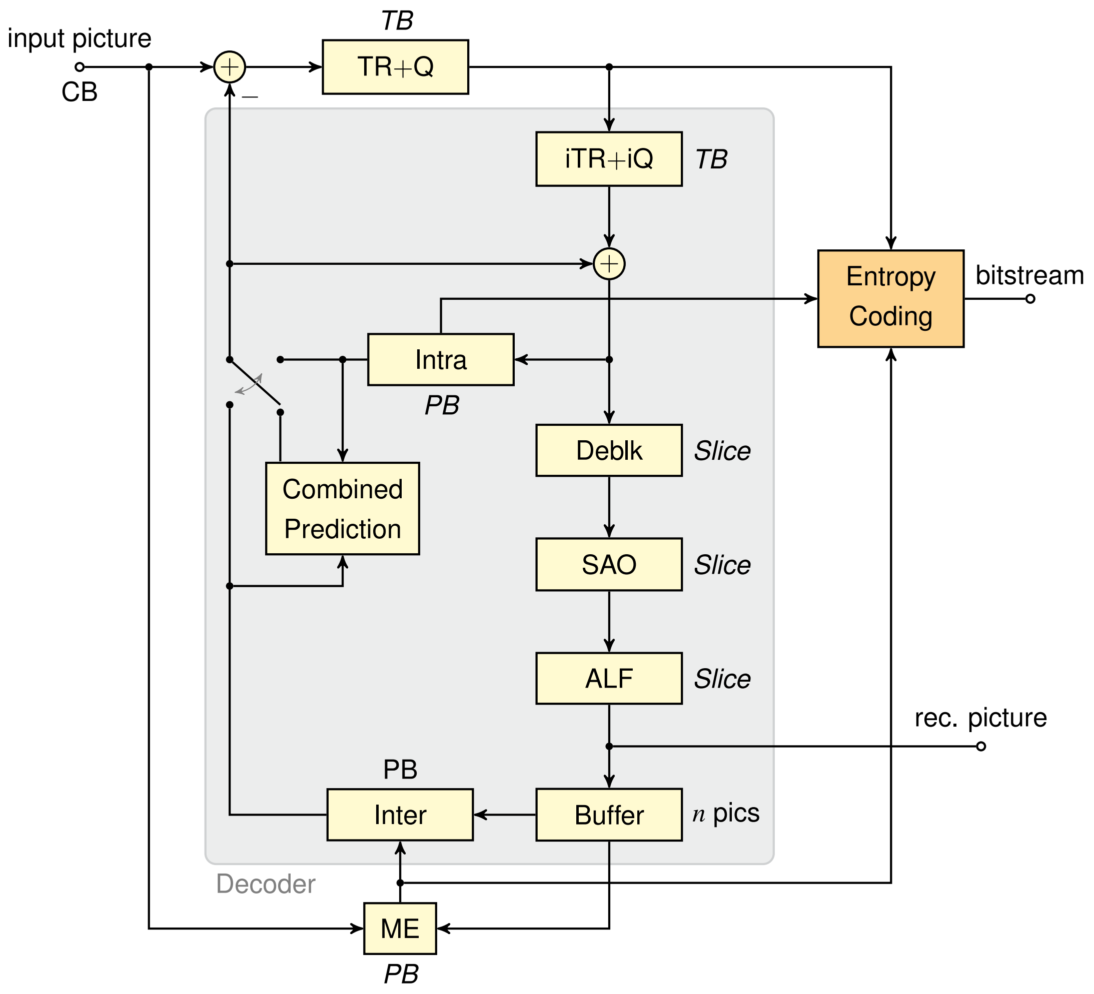


Figure from JVET-R2002 [53]

Contents

Entropy Coding



CB – Coding Block
ME – Motion Estimation
PB – Prediction Block
Q – Quantization
TB – Transform Block
TR – Transform
SAO – Sample Adaptive Offset
ALF – Adaptive Loop Filter

Entropy Coding

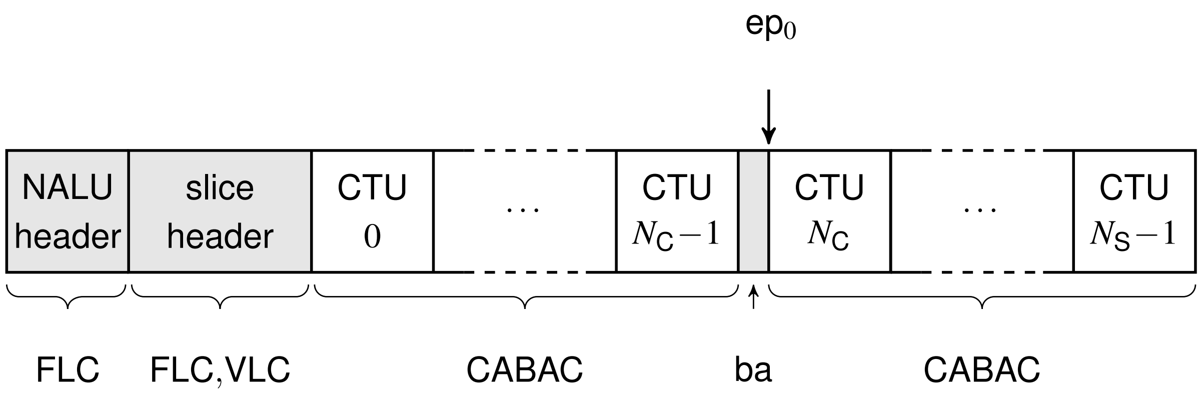
Fixed length and variable length codes (FLC, VLC)

- High-level syntax
- Parameter sets, slice segment header
- SEI messages
- Fixed-length codes, Exp-Golomb codes

Arithmetic coding

- Slice level, CTUs
- Context-based adaptive coding
- Bypass coding (complexity, throughput)

Variable Length and Arithmetic Coding



VCL NAL Unit

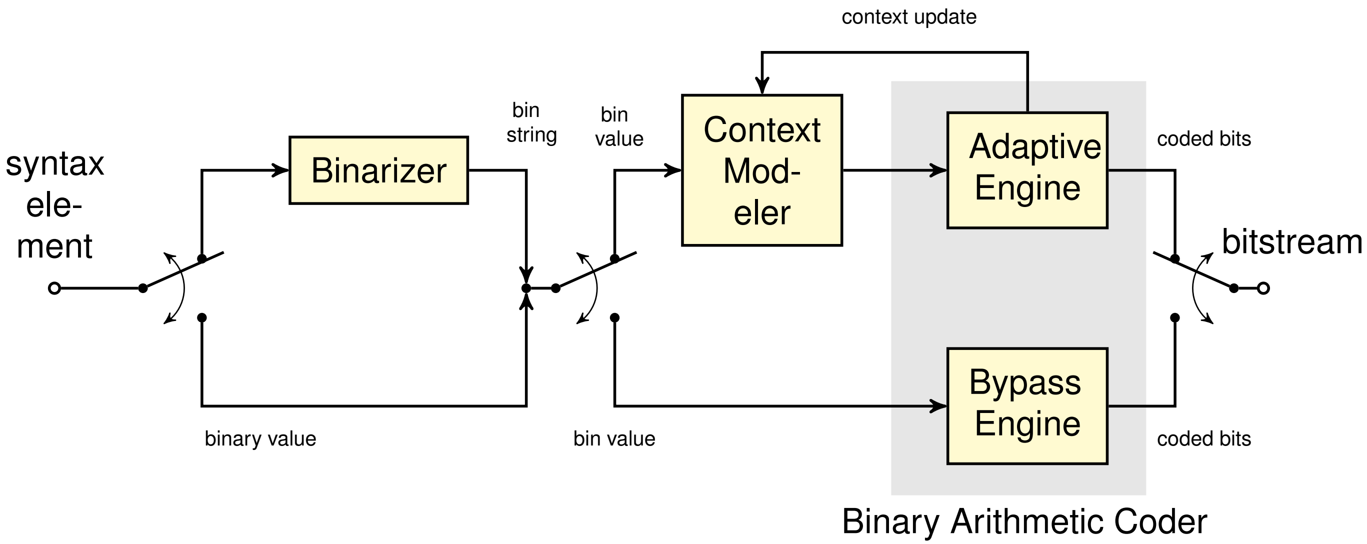
- FLC, VLC for header information
- CABAC for CTUs
- Byte alignment in case of multiple tiles, or with wavefront parallel processing

CABAC = Context-based Adaptive Binary Arithmetic Coding

ba = byte alignment

Context-Based Adaptive Binary Arithmetic Coding – CABAC

Process Overview



- Binarization
- Context model selection
- Binary arithmetic coding

MaScWi03 [65]

Binary Arithmetic Coding

Representation of engine state

- Coding of ’1’s and ’0’s
- Most probable symbol (MPS) / least probable symbol (LPS)
- Representation
  - Probability of the LPS: p_{LPS}
  - Value of the MPS: $v_{\mathrm{MPS}} \in \{0, 1\}$

***New*: Multi-hypothesis probability update model**

- Probability estimates P_0 and P_1 associated with each context model
- Updated independently with different (pre-trained) adaptation rates for each bin
- Average of hypotheses as probability estimate P used for interval subdivision in binary arithmetic coder

Contents

Intra Block Copy (IBC)

- Simplified compared to IBC in HEVC v4 (SCC) [9]
- Ref buffer restricted to current CTU and CTU to the left, 3 VPDUs max
- IBC merge mode for block vector coding
- Adaptive block vector resolution (1-sample or 4-sample precision)

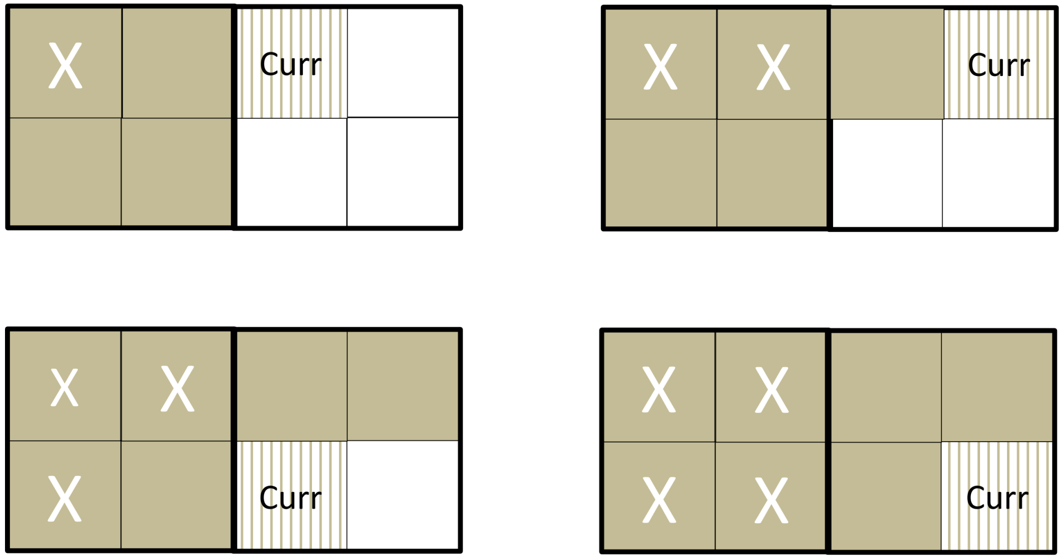


Figure from JVET-R2002 [53]

Fraunhofer Heinrich Hertz Institute
RWTH AACHEN UNIVERSITY

SCC Residual Coding

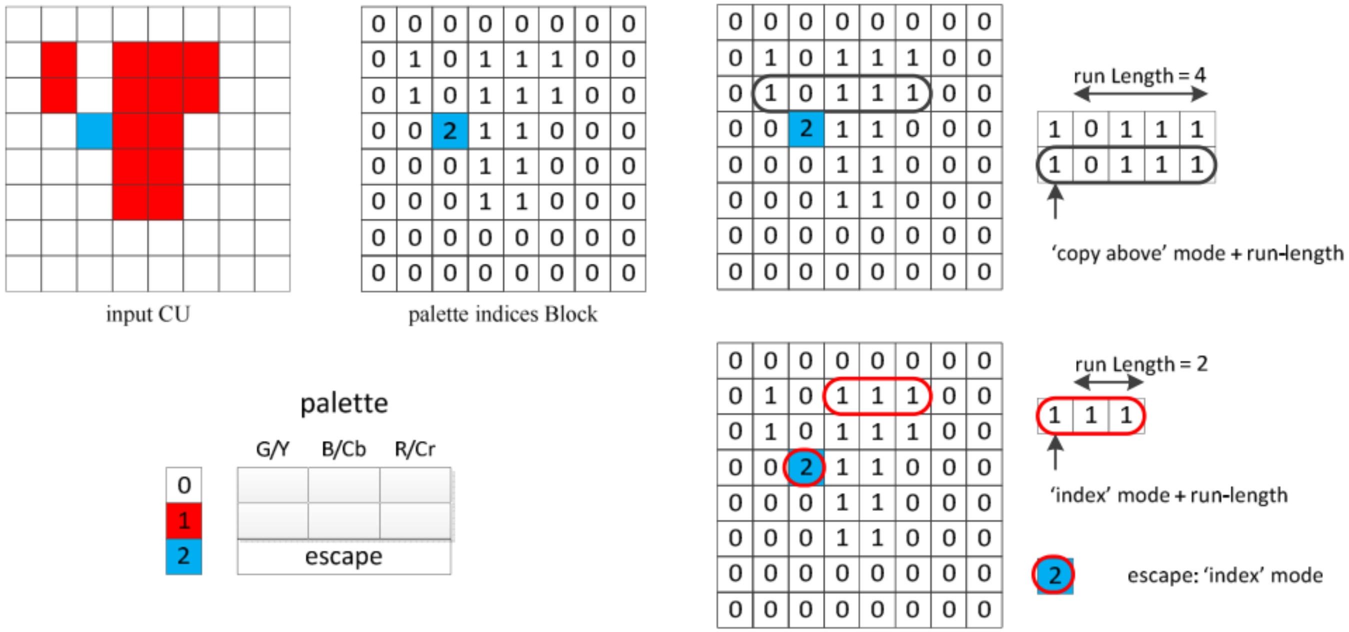
Block differential pulse coded modulation (BDPCM)

- Avaiable for intra coded CUs
- Application of DPCM instead of residual transform (good for SCC type of content)
- Indication of horizontal or vertical mode, connection to intra prediction mode

Transform skip residual coding (TSRC)

- Direct coding of the residual (also available in HEVC)
- *New*: Adaptation of entropy coding stage
  - No first non-zero coefficient
  - Context dependent sign coding
  - Modified absolute value coding: Higher cut-off between unary and Rice binarization
    $\rightarrow$ more "greater than X" flags
- Max block size 32×32
- Handling on 4×4 basis

Palette mode



Pixel value prediction instead of ’regular’ prediction and transform

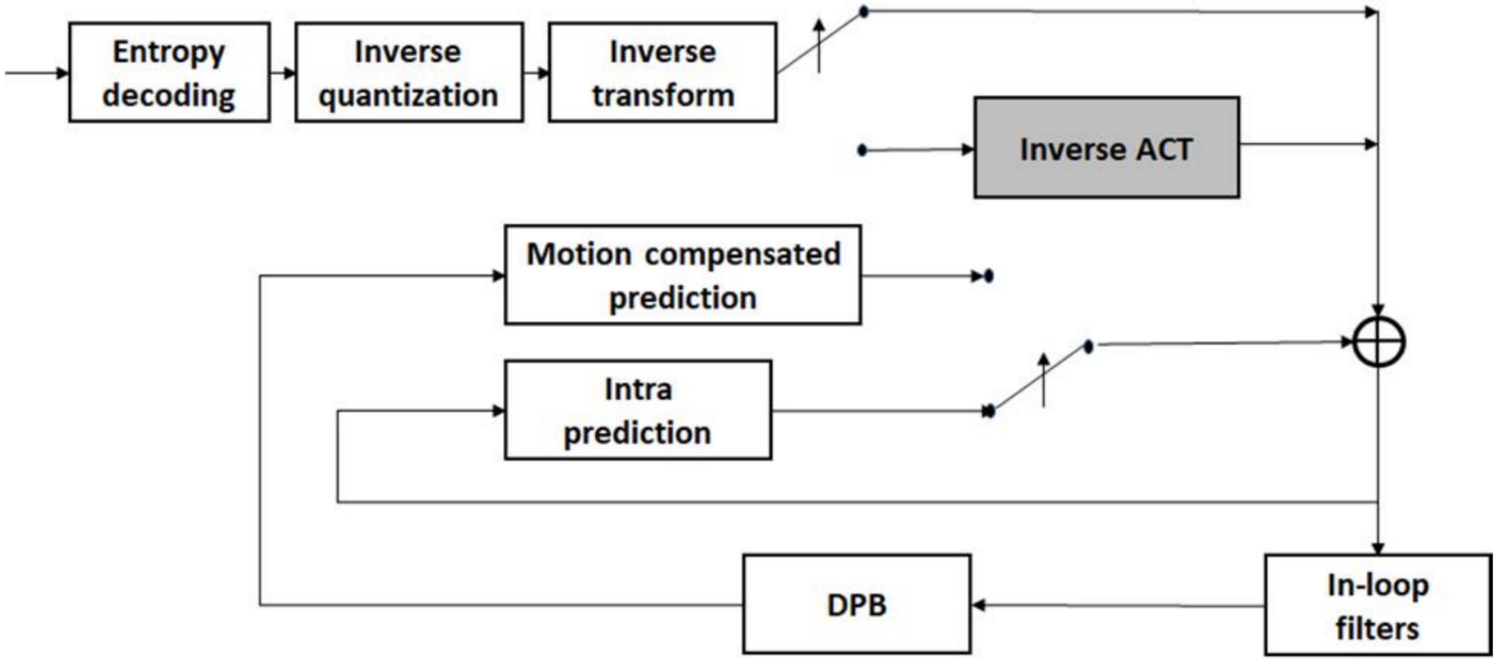
- Concept proposed for HEVC Screen Content Coding
- Palette predictor, size configurable in SPS
- Differential coding of new palette entries

Figure from JCTVC-S1014 [66]

Adaptive colour transform

- SCC content in 4:4:4 color format – redundancy among components
- Cross-component transform for decorrelation selectable on a block basis (YCgCo, or original color space, e. g. GBR)

$$
\begin{aligned}
\text{tmp} &= Y - C_\text{g}/2 \\
G &= C_\text{g} + \text{tmp} \\
B &= \text{tmp} - C_\text{o}/2 \\
R &= B + C_\text{o}
\end{aligned}
$$



Figures from JVET-Q2002 [51]

SCC Performance Comparison HEVC SCC vs. VVC SCC

	All Intra Main10									
	YUV 4:4:4					RGB 4:4:4				
	Y	U	V	EncT	DecT	Y	U	V	EncT	DecT
TGM 1080p	-12.24%	-7.09%	-7.28%	1790%	190%	-12.06%	-8.67%	-7.46%	1985%	181%
TGM 720p	-16.89%	-11.99%	-11.51%	1757%	185%	-18.23%	-6.12%	-16.69%	1919%	174%
Animation	-22.33%	-21.18%	-21.72%	1912%	196%	-21.85%	-17.21%	-15.90%	2321%	188%
Mixed content	-17.33%	-12.49%	-13.14%	1871%	188%	-21.27%	-8.42%	-14.60%	1918%	178%
Overall	-16.82%	-12.67%	-12.84%	1823%	189%	-17.89%	-9.72%	-13.43%	2018%	180%

	Random access Main10									
	YUV 4:4:4					RGB 4:4:4				
	Y	U	V	EncT	DecT	Y	U	V	EncT	DecT
TGM 1080p	-22.48%	-18.31%	-18.45%	774%	155%	-21.29%	-21.56%	-18.62%	812%	159%
TGM 720p	-24.76%	-22.13%	-21.90%	798%	169%	-25.53%	-9.21%	-24.34%	876%	175%
Animation	-29.70%	-29.96%	-31.11%	634%	166%	-31.67%	-24.05%	-20.03%	756%	180%
Mixed content	-25.94%	-24.17%	-24.51%	841%	149%	-31.71%	-13.43%	-25.00%	879%	159%
Overall	-25.42%	-23.15%	-23.45%	762%	160%	-26.96%	-16.82%	-21.92%	831%	168%

	Low delay B Main10									
	YUV 4:4:4					RGB 4:4:4				
	Y	U	V	EncT	DecT	Y	U	V	EncT	DecT
TGM 1080p	-27.38%	-24.29%	-24.49%	449%	136%	-24.73%	-25.95%	-23.23%	465%	134%
TGM 720p	-31.41%	-30.89%	-31.60%	423%	116%	-30.39%	-16.12%	-29.12%	458%	119%
Animation	-30.67%	-30.23%	-31.22%	397%	137%	-30.07%	-24.89%	-19.84%	457%	149%
Mixed content	-37.13%	-36.93%	-37.87%	448%	122%	-41.16%	-26.13%	-33.73%	481%	124%
Overall	-31.32%	-30.16%	-30.83%	430%	127%	-31.02%	-22.95%	-26.44%	465%	130%

Figure from JVET-S0264 [67]

Horizontal wrap around motion compensation

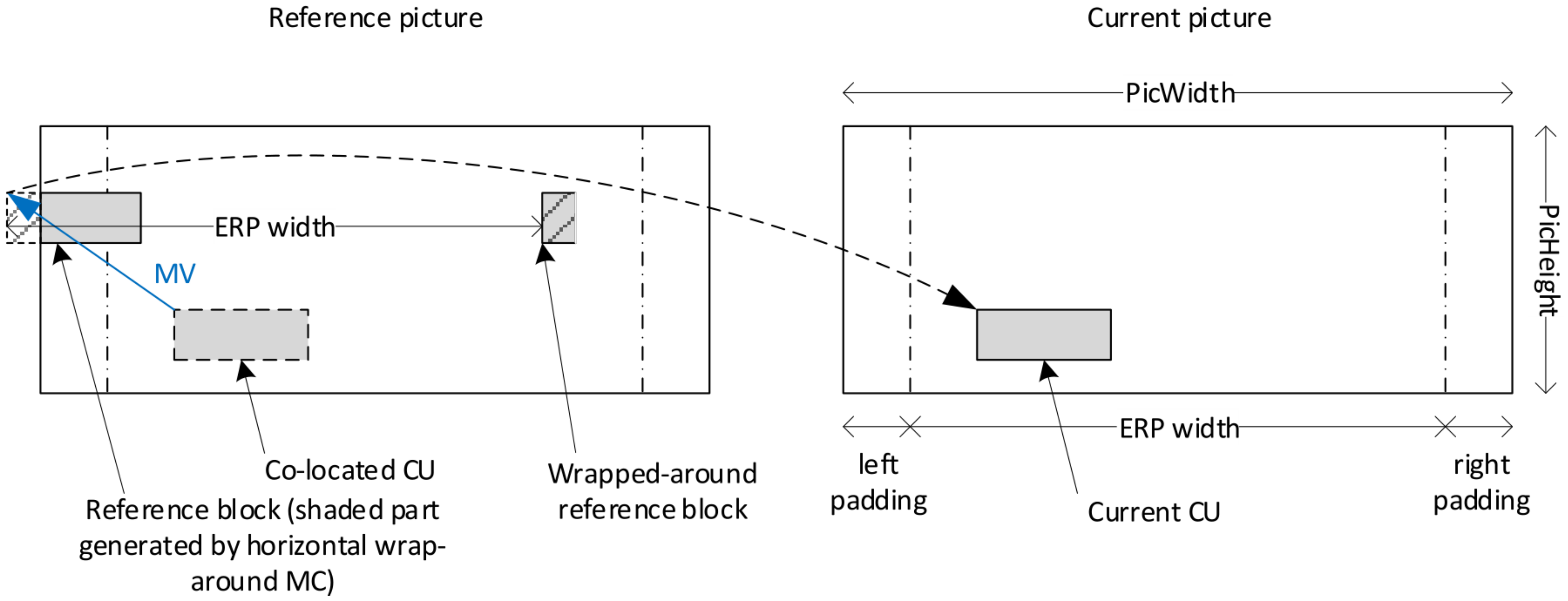


Figure from JVET-Q2002 [51]

- Apply 'wrap-around' replacement of pixels in case of blocks pointing outside of coded area for ERP pictures
- Specification takes into account potential padding of picture

Loopfilter handling at virtual boundaries

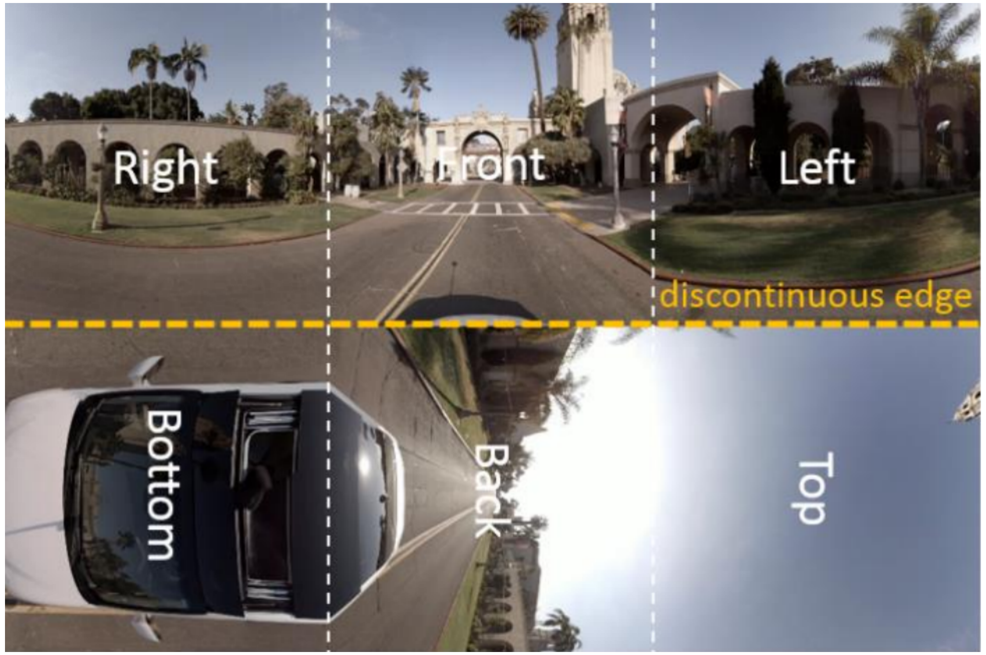


Figure from JVET-Q2002 [51]

- No loop filtering across virtual boundaries
- Virtual boundaries defined in SPS or PH, e. g. at boundaries in cube map projections

VVC Layered Coding

Requirements on VVC, MPEGN17074 [68], Sec. 5.14/15:

- Scalability modalities (such as temporal, spatial, and SNR scalability) shall be supported.
- The standard shall support the coding of stereo and multiview content.

Approach, JVET-O1159 [69]

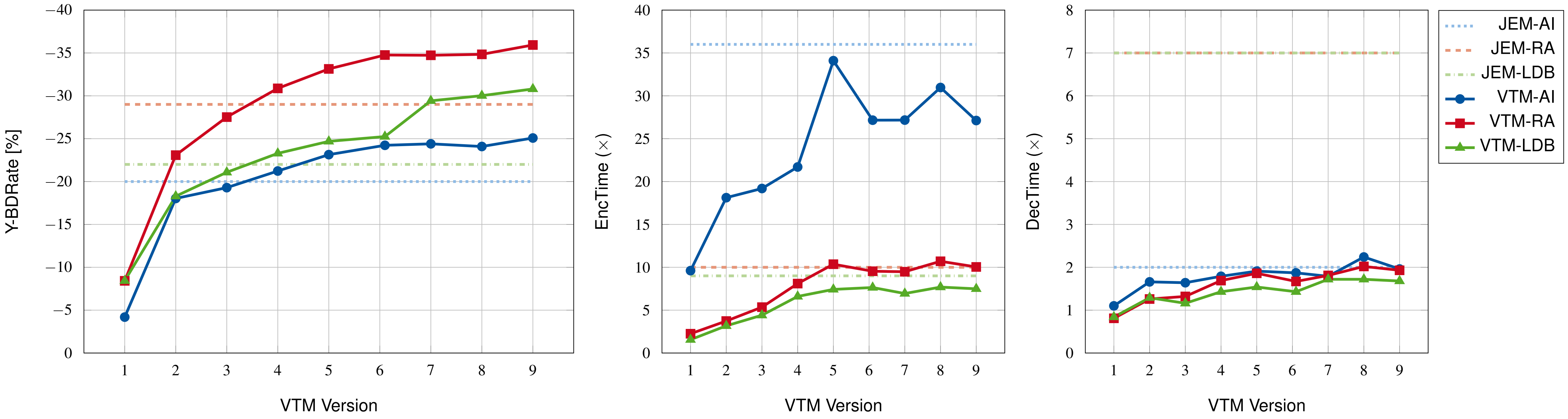
- Scalability (beyond temporal) enabled already in version 1 of VVC
- Spatial scalability enabled on top of temporal scalability by basic design of RPS+RPR
  - Picture marking only within each layer
  - All inter-layer referencing is within the same AU
  - No "diagonal" referencing would be supported, similar to SHVC
- Mode type indication in DCI
  - Only highest layer is output, or
  - All layers are output, or
  - Explicit indication of which layers to output
- No inter-layer motion vector prediction foreseen (meeting report JVET-O2000 [70]: only 0.7% gain)

Separate profiles for layered coding: Multilayer Main 10, Multilayer Main 10 4:4:4 profile

Contents

VTM Performance Evolvement

- Evolvement of VTM compression improvement and encoder / decoder run times
- Comparison to JEM 7.0 for reference, data from JVET-H0001 [71]



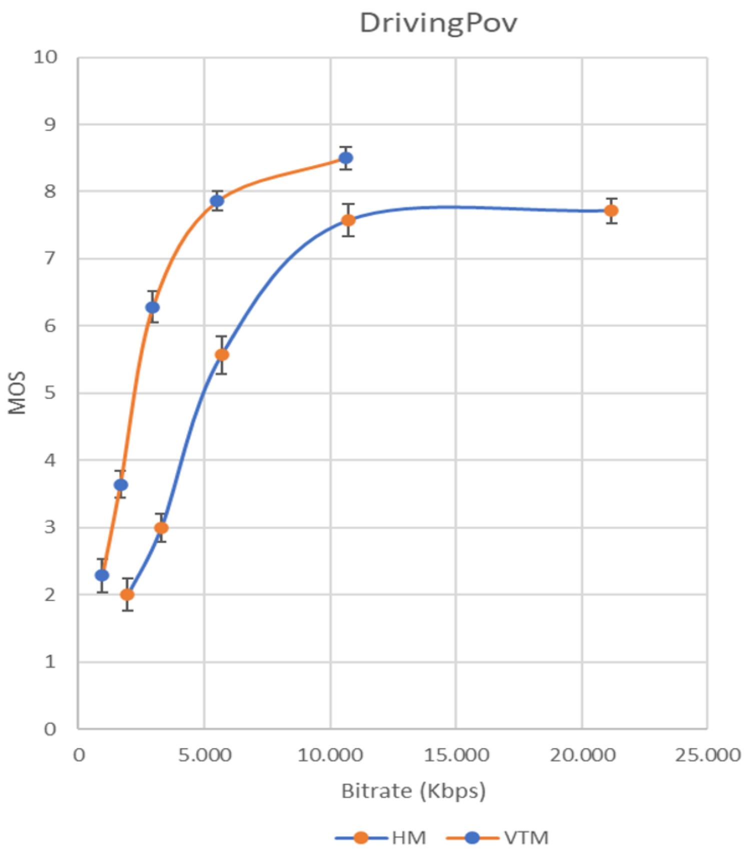
VTM results from JVET AHG3 reports, JVET-S0003 [72]

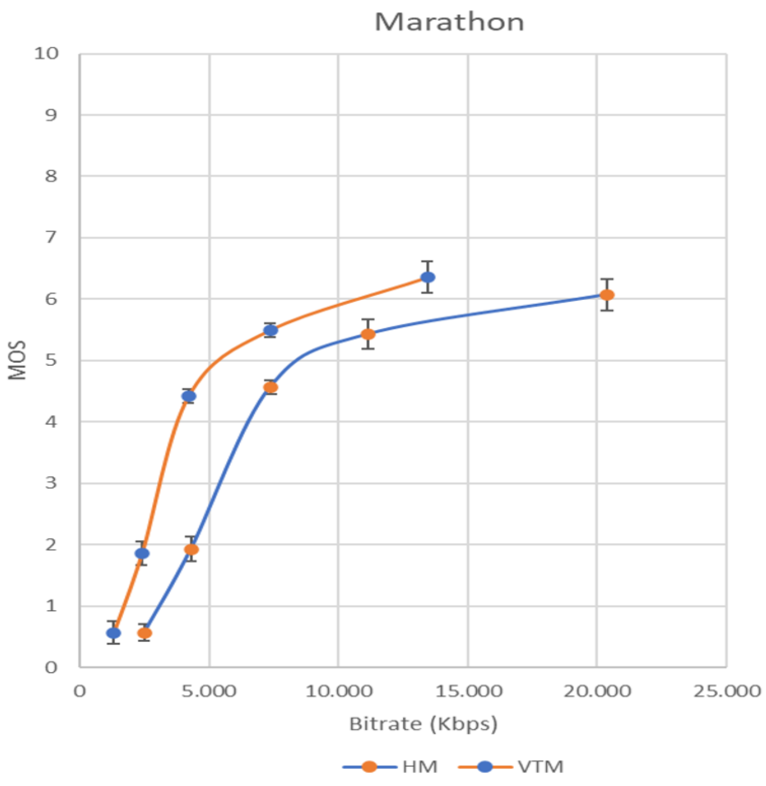
VTM Performance Assessment

- **VVC verification tests being prepared** JVET-S2009 [73]
- Goal: Assert achieved performance improvement relative to reference
- **Test categories, 1st round**
  - SDR UHD content, Random Access configuration
  - SDR HD content, Low Delay configuration
  - 360$\circ$ video content, Random Access configuration, assessed by pre-defined dynamic viewports (HD resolution)
  - HDR UHD PQ content, Random Access configuration
  - HDR UHD HLG content, Random Access configuration
- **Test categories, 2nd round**
  - Screen content
  - Scalability
  - 4:4:4 content

VTM Performance Assessment

- Preliminary results for SDR UHD content (preliminary dry run) JVET-S0246 [74]

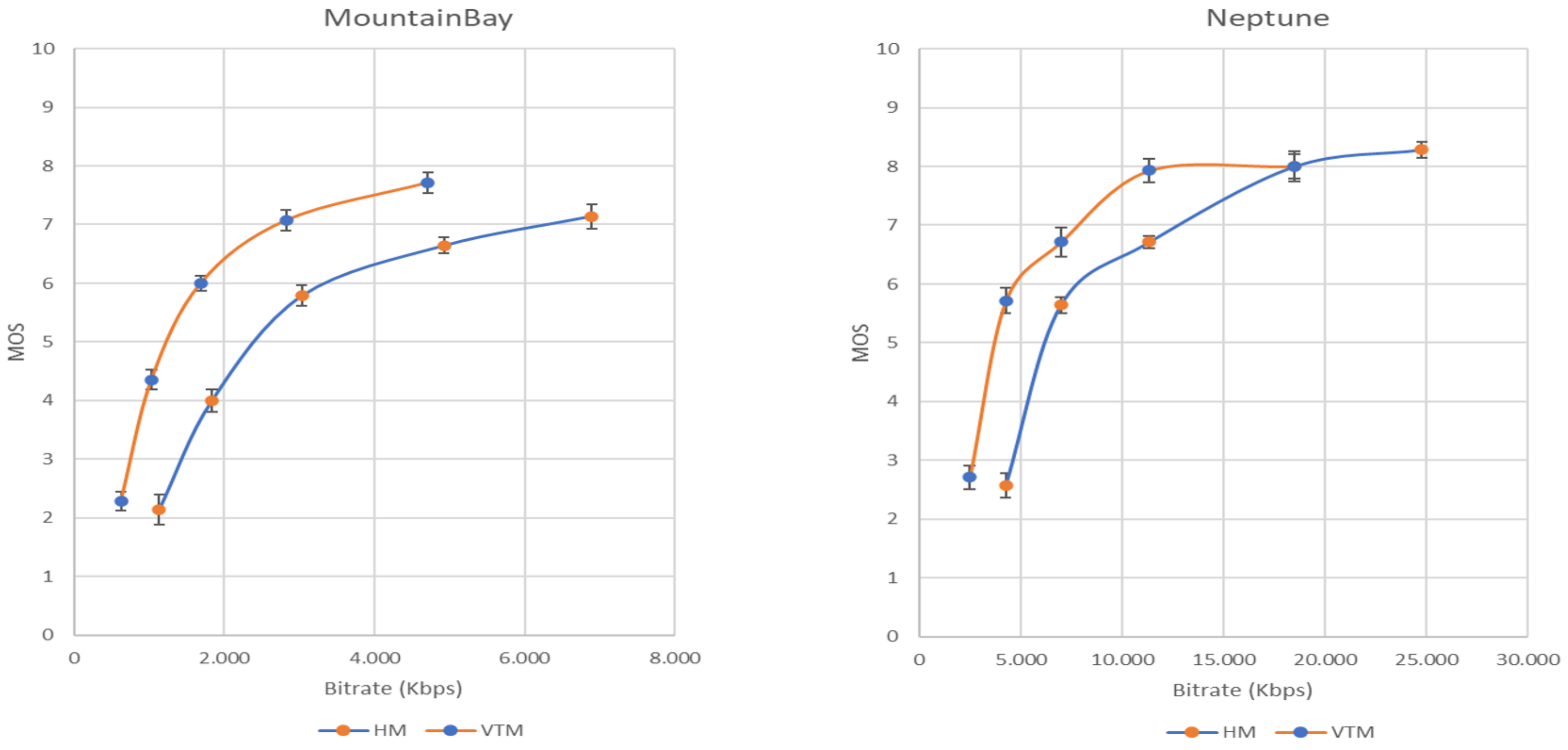




Figures from JVET-S0246 [74]

VTM Performance Assessment

- Preliminary results for SDR UHD content (preliminary dry run) JVET-S0246 [74]

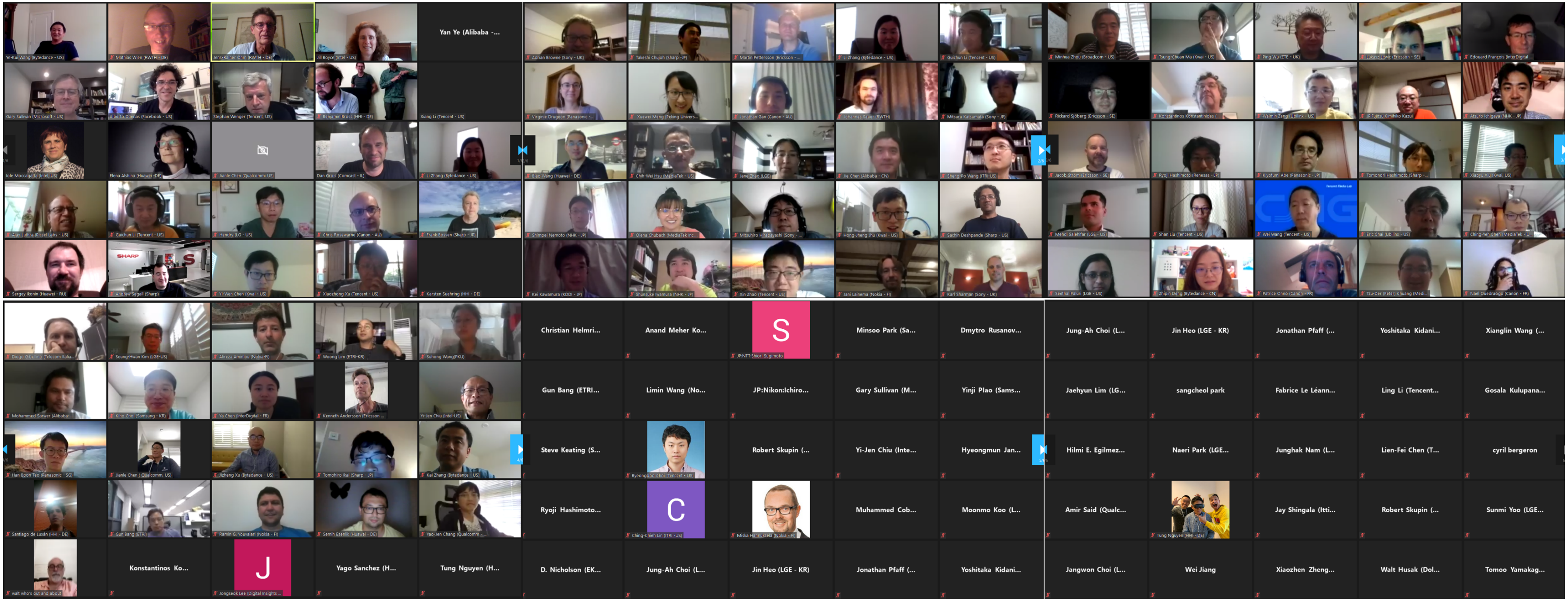


Figures from JVET-S0246 [74]

Conclusions

- Versatile Video Coding has been finalized on July 1st 2020!
- Significant number of tools adopted into the specification
- Significant increase in subjective and objective compression performance
- Significant effort to balance compression performance and computational complexity
- Subjective performance verification in preparation

The Joint Video Experts Team



The JVET after finalization of VVC, 19th meeting (virtual), July 1st, 2020, 21:00h UTC

Thank you for your attention

Any questions?

Contents

Links and Further Information

- Document archives (publicly accessible)
  - JVET / VVC:
    - http://phenix.it-sudparis.eu/jvet
    - http://ftp3.itu.ch/av-arch/jvet-site
  - JCT-VC / HEVC:
    - http://phenix.it-sudparis.eu/jct
    - http://ftp3.itu.ch/av-arch/jctvc-site
- VTM Software: https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/
- VTM Trac: https://jvet.hhi.fraunhofer.de/trac/vvc/

Contents

Acronyms I

ABT – Asymmetric binary tree
AF INTER – Affine motion vector derivation inter mode
AF MERGE – Affine motion vector derivation merge mode
AFP – Adaptive frame packing
AHG – Ad-hoc group
AI – All intra (CTC)
AIF* – Adaptive interpolation filter
ALF – Adaptive loop filter
ALTR – Adaptive long-term reference
AMT – Adaptive multiple core transform (also EMT, MTS)
AMVP – Advanced motion vector prediction
AMVR – Adaptive motion vector resolution
AQS – Adaptive quantization step size scaling
ARC – Adaptive resolution change (cp. RPR)
ATMVP – Alternative temporal motion vector prediction
BARC – Block adaptive resolution coding
BCBR – Block-composed background reference
BCW – Bi-prediction with CU based weighting
BDIP – Bi-directional intra prediction
BCPCM – Block differential pulse coded modulation
BDSNR – Bjøntegaard Delta PSNR
BD-rate – Bjøntegaard Delta rate
BIO – Bi-directional optical flow (BDOF)
BDOF – Bi-directional optical flow (formerly known as BIO)
BLA – Broken link access

Acronyms II

CABAC – Context adaptive binary arithmetic coding
CAVLC – Context adaptive variable length coding
CB – Coding block
CCIP – Cross-component intra prediction
CCLM – Cross-component linear model prediction
CF – Combined filter (intra prediction)
CfE – Call for evidence
CfP – Call for proposals
CIIP – Combined inter and intra prediction
CMP – Cube map projection
CNNLF – Convolution neural network loop filter
CNNSR – CNN for super-resolution
CPMV – Control point motion vector
CPR – Current picture referencing
CR-CNN – CNN for compact-resolution
CRA – Clean random access
CS – Constraint set (in CfP)
CTC – Common testing conditions
CTU – Coding tree unit
CTX – CABAC context
CU – Coding unit
CVS – Coded Video Sequence
CVSG – Coded Video Sequence Group
DCT – Discrete cosine transform
DST – Discrete sine transform

Acronyms III

DPB – Decoded picture buffer
DCT – Discrete cosine transform
DCTIF – DCT interpolation filter
DMVD – Decoder side motion vector derivation
DMVR – Decoder-side motion vector refinement
DPB – Decoded picture buffer
DRA – Dynamic range adaptation
DST – Discrete sine transform
EAC – Enhanced angular cubemap
EDSR – Enhanced deep residual network for super-resolution
EMT – Explicit multiple core transforms (also AMT)
EOTF – Electro-optical transfer function
ERP – Equirectangular projection
FRUC – Frame rate up conversion
GALF – Geometry transformation-based adaptive loop filter
GEO – Geometric partitioning
GRL – Givens rotation layer
HAC – Hybrid angular cubemap
HEVC – High efficiency video coding
HM – HEVC test model
HMVP – History based motion vector prediction
HWT – Hadamart-Walsh transform
HyGT – Hypercube-Givens transform
IBC – Intra block copy
IBDI* – Internal bit-depth increase

Acronyms IV

IDCT – Inverse discrete cosine transform
IDR – Instantaneous decoder refresh
IPR – Inter prediction refinement
ISP – Intra subblock partitioning
JCCR – Joint coding of chroma residuals
JCT – Joint collaborative team (of ISO and ITU)
JCT-VC – Joint collaborative team on video coding
JCT-3V – Joint collaborative team on 3D video coding extension development
JEM – Joint exploration model (JVET test model)
JTC – Joint technical committee
JVET – Joint video exploration team
KLT – Karhunen-Loève transform
KTA – Key Technical Areas (H.264 based exploration software of VCEG)
LAMVR – Locally adaptive motion vector resolution
– Low-frequency non-separable transform
LGT – Layered Givens transform
LIC – Local illumination compensation
LIP – Linear intra prediction
LM – Linear model prediction
LMCS – Luma mapping with chroma scaling
LPS – Least probably symbol
MAP – Merge assistant prediction
MBF – Multiple boundary filtering
MCP – Modified cubemap projection
MDCS – Mode-dependent coefficient scanning

Heinrich Hertz Institute

Acronyms V

MDIS – Mode-dependent intra reference sample smoothing
MDNSST – Mode-dependent non-separable secondary transforms
Merge – Merge Mode (MV prediction)
MFLM – Multiple filter linear model prediction
MIP – Multi-line intra prediction
MIP – Multi-combined intra prediction
MMLM – Multi-Model Cross-component linear model prediction
MMVD – Merge with motion vector difference
MNLM – Multiple neighbor linear model
MPCR – Motion predictor candidate refinement
MPEG – Moving picture experts group
MPM – Most probable mode
MPS – Most probably symbol
MRIP – Multi-reference intra prediction
MRM – Motion refinement mode
MSB – Most Significant Bit
MSE – Mean squared error
MTS – Multiple transform selection
MTT – Multi-type tree
MV – Motion vector
MVD – Motion vector difference
NAL – Network abstraction layer
MB – Macroblock (H.264 | AVC)
NAL – Network abstraction layer
NALU – NAL unit

Acronyms VI

NLMLF – Non-local mean loop filter
NLSF – Non-local structure-based filter
NSF – Noise suppression filter
NSST – Non-separable secondary transforms
NUH – NAL unit header
NUT – NAL unit type
OBMC – Overlapped block motion compensation
OETF – Opto-electrical transfer function
PAU – Parallel-to-axis uniform cubemap projection format
PB – Prediction block
PDPC – Position dependent intra prediction combination for planar mode
PMMVD – Pattern matched motion vector derivation
PMVD – Pattern matched motion vector derivation
PMVR – Pattern-matched motion vector refinement
POC – Picture order count
PPS – Picture parameter set
PROF – Prediction refinement with optical flow
PSNR – Peak signal to noise ratio
PU – Prediction unit
QP – Quantization parameter
QT – Quad-tree
QTBTT – Quad-tree plus binary tree and ternary tree (also QTBTTT)
RA – Random access (CTC)
RADL – Random access decodable leading picture
RAP – Random access point

Acronyms VII

RASL – Random access skipped leading picture
RBSP – Raw byte sequence payload
RC-ALF – Reduced-complexity adaptive loop filter
RD – Rate-distortion
RDO – Rate-distortion optimization
RDOQ – Rate-distortion optimized quantization
RPL – Reference picture list
RPR – Reference picture resampling (cp. ARC)
RPS – Reference picture set
RQT – Residual quad-tree
RSP – Rotated sphere projection
RST – Reduced secondary transform
SAO – Sample adaptive offset
SATD – Sum of absolute transformed differences
SBT – Subblock transform
SDP – Signal dependent transform
SEI – Supplemental enhancement information
SODB – String of data bits
SHVC – Scalable high efficiency video coding
SRCC – Scan region-based coefficient coding
STMVP – Spatial-temporal motion vector prediction
STSA – Stepwise temporal sub-layer access
SUCO – Split unit coding order
SVD – Singular value decomposition
SVT – Spatial varying transform

Acronyms VIII

TB – Transform block
TD-RSP – Residual signs prediction in transform domain
TMM – Template matchted merge mode
TMVP – Temporal motion vector predictor
TSA – Temporal sub-layer access
TSB – Transform sub-block
TSR – Transform syntax reorder
TSRC – Transform skip residual coding
TU – Transform unit
UHD – Ultra High Definition
UMVE – Ultimate motion vector expression
UWP – Unequal weight planar prediction
VCEG – Video coding experts group
VCL – Video coding layer
VLC – Variable length code
VPDU – Virtual pipeline data units
VPS – Video parameter set
VUI – Video usability information
WCG – Wide colour gamut
WPP – Wavefront parallel processing
XGA – Extended Graphics Array 1024×768
XYZ – XYZ color space, also color format
YCbCr – Color format with luma and two chroma components
YUV – Color format with luma and two chroma components

Literature I

[1] *Codecs for Videoconferencing using primary digital group transmission*. Rec. ITU-T, Mar. 1993. URL: http://www.itu.int/rec/T-REC-H.120/en.

[2] *Video codec for audiovisual services at p×64 kbit/s*. Rec. ITU-T, Mar. 1993. URL: http://www.itu.int/rec/T-REC-H.261/en.

[3] *Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s – Part 2: Video*. ISO/IEC, June 2006. URL: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=22411.

[4] *Information technology – Generic coding of moving pictures and associated audio information – Part 2: Video*. ISO/IEC, Sept. 2013. URL: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=61152.

[5] *Video coding for low bit rate communication*. Rec. ITU-T, Jan. 2005. URL: http://www.itu.int/rec/T-REC-H.263/en.

[6] *Information technology – Coding of audio-visual objects – Part 2: Visual*. ISO/IEC, Sept. 2009. URL: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=39259.

[7] *Advanced video coding for generic audiovisual services*. Rec. ITU-T, Feb. 2014. URL: http://www.itu.int/rec/T-REC-H.264/en.

[8] *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*. ISO/IEC, Apr. 2012. URL: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=61490.

[9] *High efficiency video coding (7th Ed.)* Version 7. ITU-T, Nov. 29, 2019. URL: https://www.itu.int/rec/T-REC-H.265.

[10] *Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High efficiency video coding*. ISO/IEC, Nov. 2013. URL: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=35424.

[11] Benjamin Bross, Jianle Chen, and Shan Liu. *Versatile Video Coding (Draft 9)*. Doc. JVET-R2001. Online, 18th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 24, 2020. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/18_Alpbach/wg11/JVET-R2001-v10.zip.

[12] *Parameter values for the HDTV standards for production and international programme exchange*. Rec. ITU-R, Apr. 2002. URL: http://www.itu.int/rec/R-REC-BT.709/en.

[13] *Parameter values for ultra-high definition television systems for production and international programme exchange*. Rec. ITU-R, Aug. 2012. URL: http://www.itu.int/rec/R-REC-BT.2020/en.

[14] *BT.2100: Image parameter values for high dynamic range television for use in production and international programme exchange*. ITU-R, June 2017. URL: http://www.itu.int/rec/R-REC-BT.2100-1-201706-I/en.

Literature II

[15] Ajay Luthra, Edouard Francois, and Walt (Eds.) Husak. *Requirements and Use Cases for HDR and WCG Content Coding*. Doc. N15084. Geneva, CH, 111th meeting: MPEG, Feb. 2015. URL: http://phenix.it-sudparis.eu/mpeg/doc_end_user/documents/111_Geneva/wg11/w15084.zip.

[16] Yan Ye, Elena Alshina, and Jill Boyce. *Algorithm descriptions of projection format conversion and video quality metrics in 360Lib Version 4*. Tech. rep. JVET-G1003. Torino, IT, 7th meeting: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 21, 2017. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/7_Torino/wg11/JVET-G1003-v2.zip.

[17] Kwang Pyo Choi et al. *Test sequence formats for virtual reality video coding*. Doc. JVET-C0050. 3rd meeting, Geneva, CH: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, June 1, 2016. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/3_Geneva/wg11/JVET-C0050-v3.zip.

[18] Jianle Chen et al. *Algorithm Description of Joint Exploration Test Model 1*. Doc. JVET-A1001. Geneva, CH, 1st meeting: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Oct. 2015. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/1_Geneva/wg11/JVET-A1001-v1.zip.

[19] Mathias Wien et al. *Joint Call for Evidence on Video Compression with Capability beyond HEVC*. Tech. rep. JVET-F1002. Hobart, AU, 6th meeting: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Oct. 18, 2017. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/6_Hobart/wg11/JVET-F1002-v1.zip.

[20] Vittorio Baroncini et al. *Results of the Joint Call for Evidence on Video Compression with Capability beyond HEVC*. Tech. rep. JVET-G1004. Torino, IT, 7th meeting: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 21, 2017. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/7_Torino/wg11/JVET-G1004-v2.zip.

[21] Andrew Segall et al. *Joint Call for Proposals on Video Compression with Capability beyond HEVC*. Tech. rep. JVET-H1002. Macao, CN, 8th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Oct. 18, 2017. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/8_Macau/wg11/JVET-H1002-v6.zip.

[22] Zhao Wang et al. *Description of SDR video coding technology proposal by DJI and Peking University*. Doc. JVET-J0011. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0012-v2.zip.

Literature III

[23] Rickard Sjöberg et al. *Description of SDR and HDR video coding technology proposal by Ericsson and Nokia*. Doc. JVET-J0012. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0012-v2.zip.

[24] Jung Won Kang et al. *Description of SDR video coding technology proposal by ETRI and Sejong University*. Doc. JVET-J0013. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0013-v3.zip.

[25] M. Albrecht et al. *Description of SDR, HDR, and 360° video coding technology proposal by Fraunhofer HHI*. Doc. JVET-J0014. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0014-v4.zip.

[26] Xiaoyu Xiu et al. *Description of SDR, HDR, and 360° video coding technology proposal by InterDigital Communications and Dolby Laboratories*. Doc. JVET-J0015. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0015-v3.zip.

[27] Kei Kawamura, Yoshitaka Kidani, and Sei Naito. *Description of SDR video coding technology proposal by KDDI*. Doc. JVET-J0016. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0016-v4.zip.

[28] Moonmo Koo et al. *Description of SDR video coding technology proposal by LG Electronics*. Doc. JVET-J0017. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0017-v2.zip.

[29] Chih-Wei Hsu et al. *Description of SDR video coding technology proposal by MediaTek*. Doc. JVET-J0018. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0018-v2.zip.

[30] Jian-Liang Lin et al. *Description of 360° video coding technology proposal by MediaTek*. Doc. JVET-J0019. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0019-v3.zip.

Literature IV

[31] Tadamasa Toma et al. *Description of SDR video coding technology proposal by Panasonic*. Doc. JVET-J0020. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0020-v2.zip.

[32] Y. Chen et al. *Description of SDR, HDR and 360° video coding technology proposal by Qualcomm and Technicolor – low and high complexity versions*. Doc. JVET-J0021. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0021-v5.zip.

[33] Philippe Bordes et al. *Description of SDR, HDR and 360° video coding technology proposal by Qualcomm and Technicolor – medium complexity version*. Doc. JVET-J0022. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0022-v3.zip.

[34] Max Blaeser, Johannes Sauer, and Mathias Wien. *Description of SDR and 360° video coding technology proposal by RWTH Aachen University*. Doc. JVET-J0023. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0023-v3.zip.

[35] Sri Nitchith Akula et al. *Description of SDR, HDR and 360° video coding technology proposal considering mobile application scenario by Samsung, Huawei, GoPro, and HiSilicon*. Doc. JVET-J0024. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0024-v5.zip.

[36] Huanbang Chen et al. *Description of SDR, HDR and 360° video coding technology proposal considering mobile application scenario by Huawei, GoPro, HiSilicon, and Samsung*. Doc. JVET-J0025. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0025-v4.zip.

[37] Kiran Misra et al. *Description of SDR and HDR video coding technology proposal by Sharp and Foxconn*. Doc. JVET-J0026. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0026-v11.zip.

[38] Shunsuke Iwamura et al. *Description of SDR and HDR video coding technology proposal by NHK and Sharp*. Doc. JVET-J0027. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0027-v5.zip.

Literature V

[39] Teruhiko Suzuki, Masaru Ikeda, and Karl Sharman. *Description of SDR and HDR video coding technology proposal by Sony*. Doc. JVET-J0028. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0028-v2.zip.

[40] Xiang Li et al. *Description of SDR video coding technology proposal by Tencent*. Doc. JVET-J0029. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0029-v4.zip.

[41] Alexandre Gabriel and Emmanuel Thomas. *Description of 360° video coding technology proposal by TNO*. Doc. JVET-J0030. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0030-v2.zip.

[42] David Bull, Fan Zhang, and Mariana Afonso. *Description of SDR video coding technology proposal by University of Bristol*. Doc. JVET-J0031. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0031-v3.zip.

[43] Feng Wu et al. *Description of SDR video coding technology proposal by University of Science and Technology of China, Peking University, Harbin Institute of Technology, and Wuhan University (IEEE 1857.10 Study Group)*. Doc. JVET-J0032. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0032-v4.zip.

[44] Yule Sun et al. *Description of 360° video coding technology proposal by Zhejiang University*. Doc. JVET-J0033. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0033-v3.zip.

[45] Vittorio Baroncini. *Results of Subjective Testing of Responses to the Joint CfP on Video Compression Technology with Capability beyond HEVC*. Doc. JVET-J0080. 10th meeting, San Diego, CA, USA: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/10_San%20Diego/wg11/JVET-J0080-v2.zip.

[46] Vittorio Baronchini, Jens-Rainer Ohm, and Gary Sullivan. *Report of results from the Call for Proposals on Video Compression with Capability beyond HEVC*. Doc. JVET-J1003. San Diego, CA, USA, 10th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 10, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=3540.

Literature VI

[47] Benjamin Bross, Jianle Chen, and Shan Liu. *Versatile Video Coding (Draft 10)*. Doc. JVET-S2001. Teleconference, 19th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, June 28, 2020. URL: http://phenix.it-sudparis.eu/jvet/doc_end_user/documents/19_Teleconference/wg11/JVET-S2001-v4.zip.

[48] Benjamin Bross, Jianle Chen, and Shan Liu. *Versatile Video Coding (Draft 6)*. Doc. JVET-O2001. Gothenburg, SE, 15th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 12, 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/15_Gothenburg/wg11/JVET-O2001-v14.zip.

[49] Jianle Chen et al. *Algorithm Description of Joint Exploration Test Model 7 (JEM 7)*. Doc. JVET-G1001. Torino, IT, 7th meeting: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 21, 2017. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/7_Torino/wg11/JVET-G1001-v1.zip.

[50] Fabien Racapé et al. *CE3-related: Wide-angle intra prediction for non-square blocks*. Doc. JVET-K0500. Ljubljana, SI, 11th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 18, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/11_Ljubljana/wg11/JVET-K0500-v6.zip.

[51] Jianle Chen, Yan Ye, and Seung Hwan Kim. *Algorithm description for Versatile Video Coding and Test Model 8 (VTM 8)*. Doc. JVET-Q2002. Brussels,BE, 17th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Jan. 17, 2020.

[52] A. M. Kotra et al. *CE4-related: HMVP and parallel processing with tiles and tile groups*. Doc. JVET-M0300. Marrakech, MA, 13th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Jan. 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/13_Marrakech/wg11/JVET-M0300-v3.zip.

[53] Jianle Chen, Yan Ye, and Seung Hwan Kim. *Algorithm description for Versatile Video Coding and Test Model 9 (VTM 9)*. Doc. JVET-R2002. Online, 18th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Apr. 24, 2020. URL: http://phenix.int-evry.fr/jvet/doc_end_user/current_document.php?id=10156.

[54] Chun-Chi Chen et al. *CE4: Summary report on inter prediction with geometric partitioning*. Doc. JVET-Q0024. Brussels, BE, 17th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Jan. 8, 2020. URL: http://phenix.it-sudparis.eu/jvet/doc_end_user/current_document.php?id=9353.

[55] Jingning Han, Ankur Saxena, and Kenneth Rose. "Towards jointly optimal spatial prediction and adaptive transform in video/image coding". In: *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP '10*. Mar. 2010, pp. 726–729. DOI: 10.1109/ICASSP.2010.5495043.

[56] C. Yeo et al. *Mode-dependent fast separable KLT for block-based intra coding*. Doc. JCTVC-B024. Geneva, CH, 2nd meeting: Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 2010.

Literature VII

[57] Zhaobin Zhang et al. *CE6-related: Fast DST-7/DCT-8 with dual implementation support*. Doc. JVET-K0291. Ljubljana, SI, 11th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 18, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/11_Ljubljana/wg11/JVET-K0291-v2.zip.

[58] Moonmo Koo et al. *CE6-related: 32 point MTS based on skipping high frequency coefficients*. Doc. JVET-M0297. Marrakech, MA, 13th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Jan. 9, 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/13_Marrakech/wg11/JVET-M0297-v4.zip.

[59] Muhammed Coban and Marta Karczewicz. *CE7-related: coefficient scanning and last position coding for TUs of greater than 32 width or height*. Doc. JVET-M0257. Marrakech, MA, 13th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Jan. 9, 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/13_Marrakech/wg11/JVET-M0257-v4.zip.

[60] Yin Zhao et al. *CE6: Sub-block transform for inter blocks (CE6.4.1)*. Doc. JVET-M0140. Marrakech, MA, 13th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Jan. 9, 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/13_Marrakech/wg11/JVET-M0140-v3.zip.

[61] Moonmo Koo et al. *CE6: Reduced Secondary Transform (RST) (CE6-3.1)*. Doc. JVET-N0193. Geneva, CH, 14th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Mar. 27, 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/14_Geneva/wg11/JVET-N0193-v5.zip.

[62] Heiko Schwarz et al. *CE7: Transform Coefficient Coding and Dependent Quantization (Tests 7.1.2, 7.2.1)*. Doc. JVET-K0071. Ljubljana, SI, 11th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 18, 2018. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/11_Ljubljana/wg11/JVET-K0071-v2.zip.

[63] Taoran Lu et al. *CE12: Mapping functions (test CE12-1 and CE12-2)*. Doc. JVET-M0427. Marrakech, MA, 13th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Jan. 9, 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/13_Marrakech/wg11/JVET-M0427-v3.zip.

[64] BT.2390-8. *High dynamic rangetelevision for production and international programme exchange*. Report BT.2390-8. ITU-R, Oct. 2017. URL: https://www.itu.int/pub/publications.aspx?lang=en&parent=R-REP-BT.2390-8-2020.

[65] Detlev Marpe, Heiko Schwarz, and Thomas Wiegand. "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard". In: *IEEE Transactions on Circuits and Systems for Video Technology* 13.7 (July 2003), pp. 620–637. DOI: 10.1109/TCSVT.2003.815173.

[66] Rajan Joshi et al. *Screen content coding test model 3 (SCM 3)*. Doc. JCTVC-S1014. 19th meeting, Strasbourg, F: Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Oct. 2014.

Literature VIII

[67] Xiaozhong Xu and Shan Liu. *Performance Comparison of Screen Content Coding between HEVC and VVC*. Doc. JVET-S0264. Teleconference, 19th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, June 22, 2020. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/19_Teleconference/wg11/JVET-S0264-v2.zip.

[68] MPEG Requirements. *Requirements for a Future Video Coding Standard v5*. Doc. N17074. Torino, IT, 119th meeting: ISO/IEC JTC1/SC29/WG11, July 21, 2017. URL: https://www.itu.int/en/ITU-T/studygroups/2017-2020/16/Documents/video/ToR-JVET-TD-PLEN-0155A3-A20171027-Reqs-MPEG-WG11-N17074.pdf.

[69] Ye-Kui Wang et al. *AHG8: Scalability for VVC – Text for approach 1*. Doc. JVET-O1159. Gothenburg, SE, 15th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 12, 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/15_Gothenburg/wg11/JVET-O1159-v1.zip.

[70] Gary J. Sullivan and Jens-Rainer Ohm. *Meeting report of the 15th meeting of the Joint Video Exploration Team (JVET), Gothenburg, SE, 3–12 July 2019*. Doc. JVET-O2000. Gothenburg, SE, 15th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, July 12, 2019. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/15_Gothenburg/wg11/JVET-O2000-v1.zip.

[71] Marta Karczewicz and Elena Alshina. *JVET AHG report: Tool evaluation (AHG1)*. Doc. JVET-H0001. Macau, CN, 8th meeting: Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Oct. 24, 2017. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/8_Macau/wg11/JVET-H0001-v1.zip.

[72] Frank Bossen, Xiang Li, and Karsten Suehring. *JVET AHG report: Test model software development (AHG3)*. Doc. JVET-S0003. Teleconference, 19th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, June 28, 2020. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/19_Teleconference/wg11/JVET-S0003-v2.zip.

[73] Mathias Wien et al. *Verification Test Plan (Draft 3)*. Doc. JVET-S2009. Teleconference, 19th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, June 28, 2020. URL: http://phenix.it-sudparis.eu/jvet/doc_end_user/documents/19_Teleconference/wg11/JVET-S2009-v1.zip.

[74] Vittorio Baroncini and Mathias Wien. *Results of dry run subjective assessment of SDR UHD verification test*. Doc. JVET-S0246. Teleconference, 19th meeting: Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, June 28, 2020. URL: http://phenix.int-evry.fr/jvet/doc_end_user/documents/19_Teleconference/wg11/JVET-S0246-v1.zip.