達摩院
ALIBABA DAMO ACADEMY

Enhanced Compression Model for Beyond-VVC Capability

“Future video coding – advanced signal processing, AI and standards”
ITU-T SG21 and ISO/IEC SC 29 Joint Workshop
January 17, 2025

Yan Ye

Head of DAMO Video Technology Lab, Alibaba Group U.S.

Outline

01 Introduction to ECM

02 ECM tools and performance

03 ECM tool assessment

04 Future outlook

達摩院
ALIBABA DAMO ACADEMY

01 Introduction to ECM

Enhanced Compression Model (ECM)

- In April 2021, shortly after VVC version 1 was finalized, JVET received contribution JVET-U0100 that showed 11.5% coding performance gain vs. VVC
  - Includes tools proposed to VVC but not adopted and extensions of VVC coding tools
- In July 2021, ECM software platform was established to facilitate exploration toward next-gen video codec standard
  - Performance steadily improves with more tools included into newer versions of ECM
- The latest ECM-15.0 (output from Nov. 2024 meeting) achieves **26.6%** coding performance gain vs. VVC
- Four ECM-related Ad hoc Groups (AHGs) for well-coordinated exploration

AHG12	AHG4 & AHG17	AHG6	AHG7
• ECM common test conditions • Alg. descriptions • Gen. coordination	• Subjective quality assessment • Beyond CTC testing	• ECM software implementation & maintenance	• ECM tool assessment and tool-off testing

- Based on hybrid block-based video coding framework, mainly using traditional signal processing algorithms
  - More data-driven training used by various tools

V. Seregin et al., “JVET AHG report: ECM software development (AHG6)”, JVET-AK0006, Jan. 2025
Y.-J. Change et al., “Compression efficiency methods beyond VVC”, JVET-U0100, April 2021

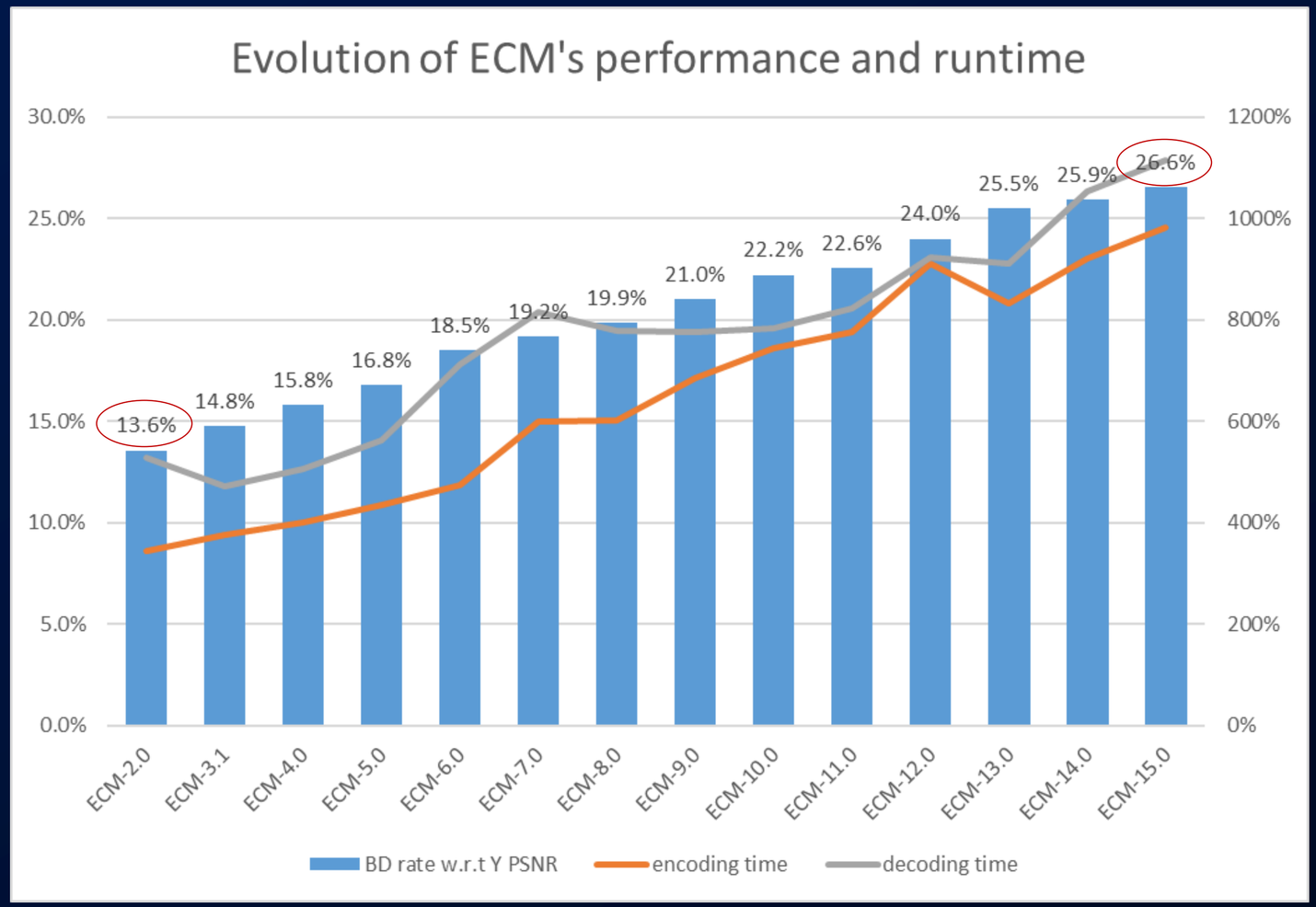
ECM evolution: performance vs. enc./dec. runtime

Random Access (RA) config. of ECM common test conditions

YUV4:2:0 10b coding, covering SD, HD, UHD resolutions

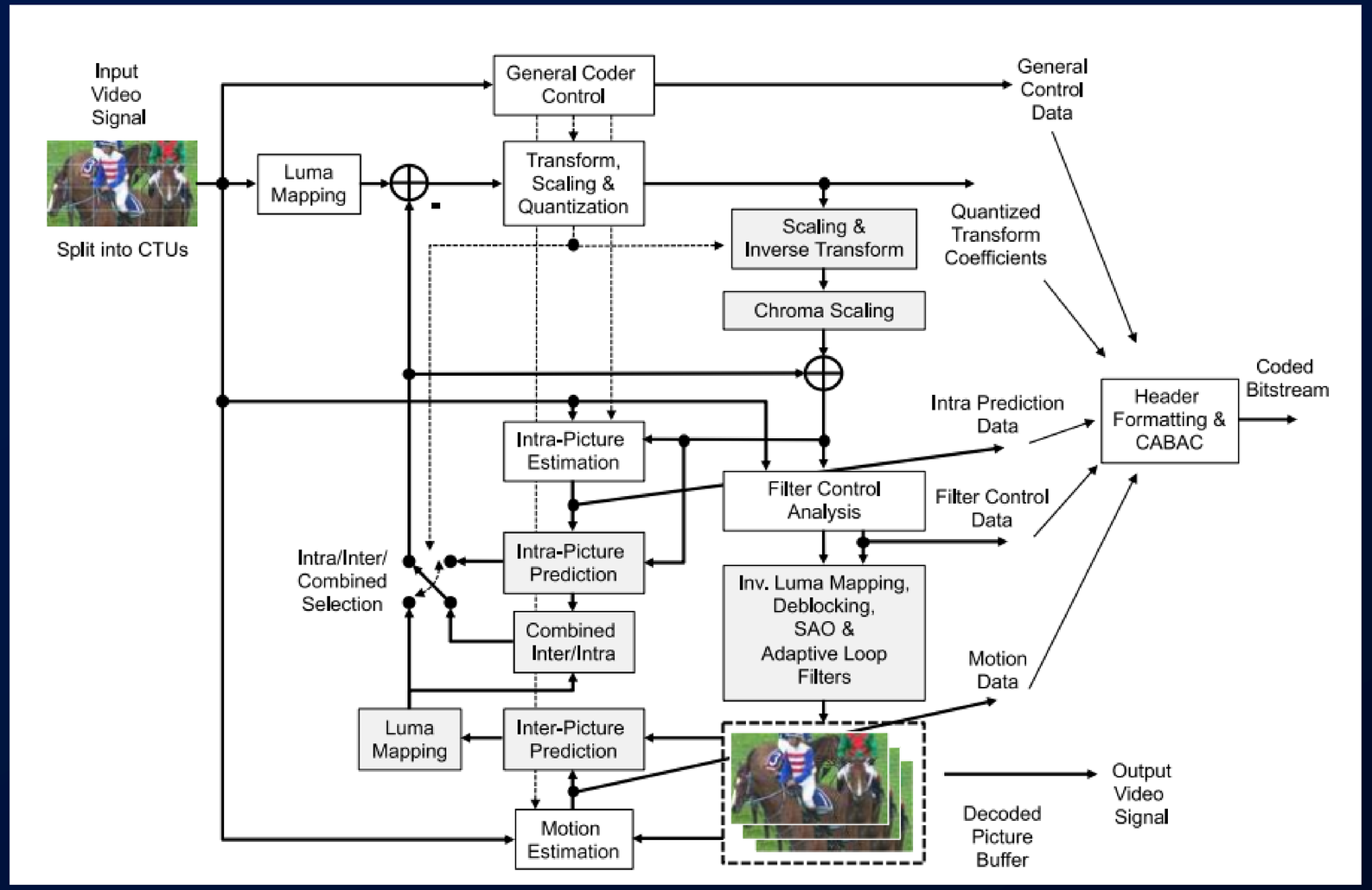
BD rate savings (%) in terms of Y PSNR

Encoding and decoding time relative to VVC ref. sw. VTM



02 ECM tools and performance

Versatile Video Coding (H.266/VVC)



Mostly inherited from H.265/HEVC

- CTU, QT block partition
- Intra pred. modes
- Advanced MV pred. and merge
- Translational motion comp. pred.
- IBC and palette (SCC ext.)
- Deblocking, SAO
- CABAC

New elements

- Larger CTU, binary/ternary tree
- Larger transform, multi. transf. set
- 65 angular pred. directions, WAIP
- History-based merge
- Geometric partitions
- Combined Intra/Inter Pred.
- Affine, PROF, BDOF, DMVR
- New loop filters: LMCS, ALF, CC-ALF
- Ref. pic. resampling
- And more…

B. Bross et al., “Overview of the Versatile Video Coding (VVC) Standard and Its Applications”, IEEE Transactions on Circuits and Systems for Video Technology, 2021

ECM coding tools

- Enhancements and combinations of various VVC coding tools
- More on-the-fly adaptation at decoder
- More tools based on data-driven training

Intra prediction (27)

- Conv. cross-comp. model (CCCM), other cross-comp. pred.
- Extrapolation-based intra prediction (EIP)
- NN-based intra prediction
- Decoder side deriv., etc.

Inter prediction (42)

- Inter template matching & reordering
- Enh. DMVR, enh. affine, enh. GPM, OBMC, etc.
- Non-adjacent candidates, chained MV, etc.
- IBC for natural content

Transform and quant. (11)

- Non-sep primary transform
- Enhanced MTS, SBT and LFNST
- 8-state depend. quant.
- Quant. center shift, etc.

Loop filters

- Various enh. to ALF, chroma ALF, and CC-ALF
- Bilateral filter for luma and chroma
- Cross-comp. SAO

Entropy coding (5)

- Extended precision
- Multi. hypo. prob. estimation
- Slice-type based window size, etc

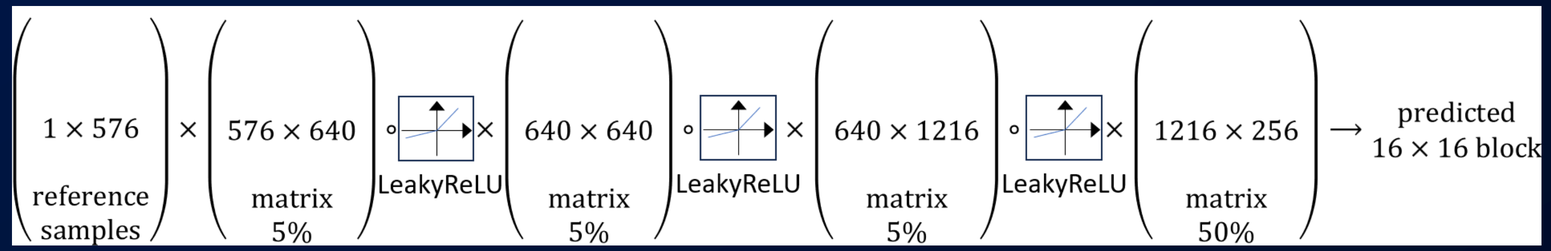
More coding tools based on data-drive training, e.g. PDP, NN intra, NSPT

M. Coban et al.,, “Algorithm description of Enhanced Compression Model 15 (ECM15)”, JVET-AJ2025, Nov. 2024

NN-based intra prediction

- NN-based intra pred. is a well-published concept, and the basis for VVC's matrix-based intra prediction (MIP)
- Incorporated into the neural network-based video coding (NNVC) exploration in JVET since Jan 2023
- Simplified NN intra pred. recently adopted into ECM-15.0
  - Training of NN models follows process defined by NNVC
  - A total of 6 NN models supporting 17 block sizes

Mode size	Block sizes supported
4×4	4×4
8×4	$8 \times 4, 4 \times 8$
16×4	$16 \times 4, 4 \times 16$ $32 \times 4, 4 \times 32$
8×8	8×8
16×8	$16 \times 8, 8 \times 16$ $32 \times 8, 8 \times 32$
16×16	$16 \times 16, 32 \times 16$ $16 \times 32, 32 \times 32$ 64×64



16x16 NN model: sequential matrix multiplications and LeakyReLUs (piecewise-linear functions)

NN-based intra prediction (cont.)

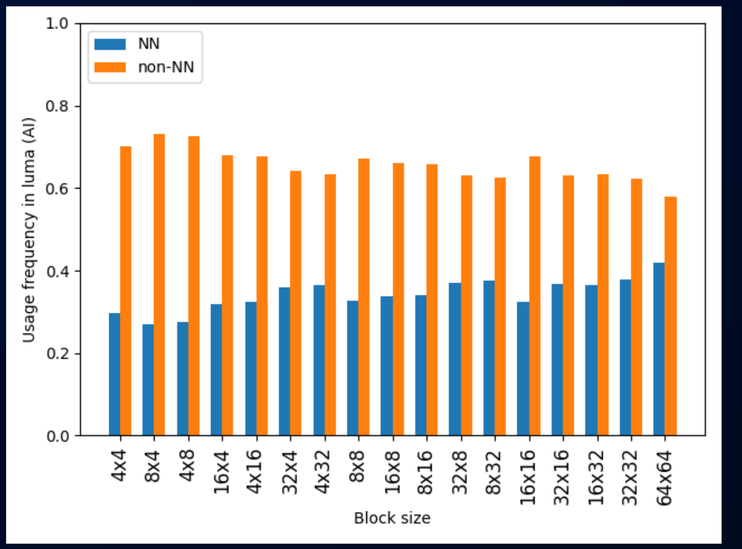
達摩院 ALIBABA DAMO ACADEMY

Perf. vs. complexity evolution of NN intra tool in the context of ECM (All Intra config)

2024.04	2024.07	2024.11
JVET-AH0156	JVET-AI0201 JVET-AI0225*	JVET-AJ0249
0.93% Y-BD rate 0.75% U-BD rate 0.79% V-BD rate	0.74% Y-BD rate 0.57% U-BD rate 0.57% V-BD rate	0.58% Y-BD rate 0.40% U-BD rate 0.40% V-BD rate
109.7% EncT 141.2% DecT	105.6% EncT 118.3% DecT	101.6% EncT 109.8% DecT

* JVET-AI0225 is a joint contribution from two companies, and used to provide performance and runtime data

Model	# non-zero params	MACs/pixel	Memory (MB)
4x4	50823	3176	0.195
8x4	61886	1934	0.224
16x4	123420	1928	0.450
8x8	132260	2066	0.472
16x8	178754	1396	0.590
16x16	215376	841	0.596



F. Urban. et al., “AHG12 : neural network-based intra prediction”, JVET-AH0156, April 2024
F. Urban. et al., “AHG12 : neural network-based intra prediction”, JVET-AI0201, July 2024
S. Eadie, et al, “AhG12: Neural network-based intra prediction with DIMD mode derivation”, JVET-AI0225, July 2024
T. Dumas, et al, “EE2-2.20_2.21: Neural network-based intra prediction with DIMD mode derivation”, JVET-AJ0249, Nov. 2024

ECM-15.0 performance

All Intra config

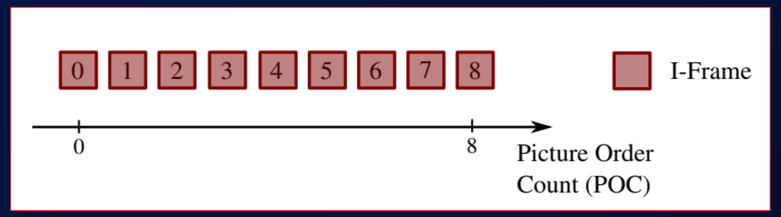
	All Intra Main 10 Over VTM-11.0ecm15.0				
	Y	U	V	EncT	DecT
Class A1	-14.3%	-15.2%	-26.5%	1136%	528%
Class A2	-20.8%	-23.5%	-28.0%	1131%	569%
Class B	-14.4%	-21.9%	-19.7%	1052%	601%
Class C	-14.3%	-11.2%	-12.3%	1008%	551%
Class E	-18.6%	-21.8%	-20.1%	996%	623%
Overall	**-16.1%**	**-18.6%**	**-20.7%**	**1059%**	**575%**
Class D	-12.2%	-8.2%	-8.9%	990%	585%
Class F	-29.9%	-33.5%	-33.7%	744%	673%
Class TGM	-43.1%	-48.8%	-48.0%	576%	704%

Natural content

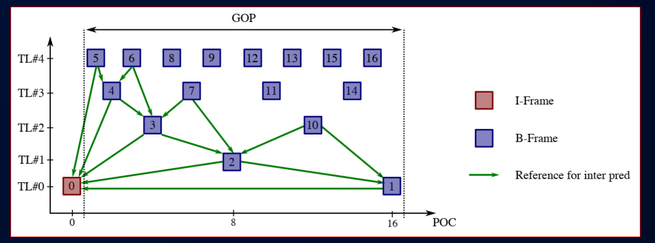
Screen content

Random Access config

	Random Access Main 10 Over VTM-11.0ecm15.0				
	Y	U	V	EncT	DecT
Class A1	-26.9%	-23.4%	-35.5%	1148%	1018%
Class A2	-30.1%	-33.3%	-38.7%	1103%	1216%
Class B	-24.5%	-31.5%	-28.8%	934%	1077%
Class C	-26.3%	-21.7%	-22.4%	1003%	1180%
Class E					
Overall	**-26.6%**	**-27.6%**	**-30.4%**	**1026%**	**1118%**
Class D	-27.1%	-22.2%	-23.4%	949%	1295%
Class F	-32.7%	-34.9%	-35.6%	870%	827%
Class TGM	-42.3%	-47.8%	-47.6%	736%	656%



No temporal prediction, still picture coding



Hierarchical-B prediction, with picture reordering

ECM-15.0 performance (cont.)

Low Delay B config

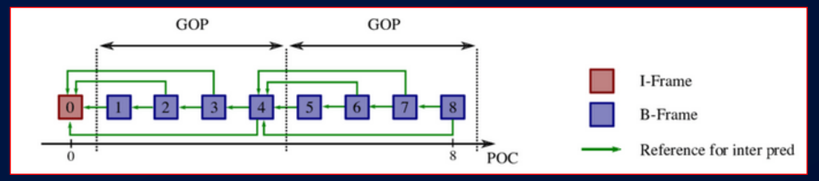
	Low delay B Main 10 Over VTM-11.0ecm15.0				
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B	-21.7%	-35.5%	-32.0%	965%	897%
Class C	-24.1%	-24.5%	-26.3%	917%	969%
Class E	-21.4%	-25.6%	-24.7%	861%	593%
Overall	**-22.4%**	**-29.4%**	**-28.3%**	**922%**	**830%**
Class D	-25.5%	-25.4%	-26.3%	934%	1111%
Class F	-30.3%	-38.2%	-37.9%	824%	729%
Class TGM	-40.6%	-50.2%	-50.0%	740%	622%

Low Delay P config

	Low delay P Main 10 Over VTM-11.0ecm15.0				
	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B	-19.5%	-44.6%	-41.8%	836%	876%
Class C	-22.1%	-34.0%	-34.6%	774%	896%
Class E	-19.9%	-35.5%	-35.9%	779%	628%
Overall	**-20.5%**	**-38.8%**	**-37.9%**	**801%**	**812%**
Class D	-24.8%	-36.5%	-37.2%	784%	1014%
Class F	-28.4%	-43.6%	-44.1%	815%	743%
Class TGM	-38.7%	-52.6%	-52.1%	790%	613%

Natural content

Screen content



Temporal bi-prediction without picture reordering

達摩院
ALIBABA DAMO ACADEMY

03 ECM tool assessment

ECM tool assessment

- Ad hoc Group on ECM tool assessment (AHG7) established since the Jan 2023 JVET meeting
  - Groupings of ECM tools considering potential implementation issues
  - Ensure proper tool controls within ECM software in coordination with software AHG
  - Collect and report tool-off/tool-on results

Group 1
- Tools that interleave MV derivation with reconstruction
- Hardware pipeline issues
- Ex: inter template matching

Group 2
- Tools that interleave candidate list derivation with reconstruction
- Latency and pipeline issues
- Ex: local illumination compensation

Group 3
- Intra tools requiring decoder search
- Latency and/or hardware cost
- Ex: intra template matching

Group 4
- Tools needing more processing on neighboring reconstructed samples
- Latency and/or hardware cost
- Ex: conv. cross-component model

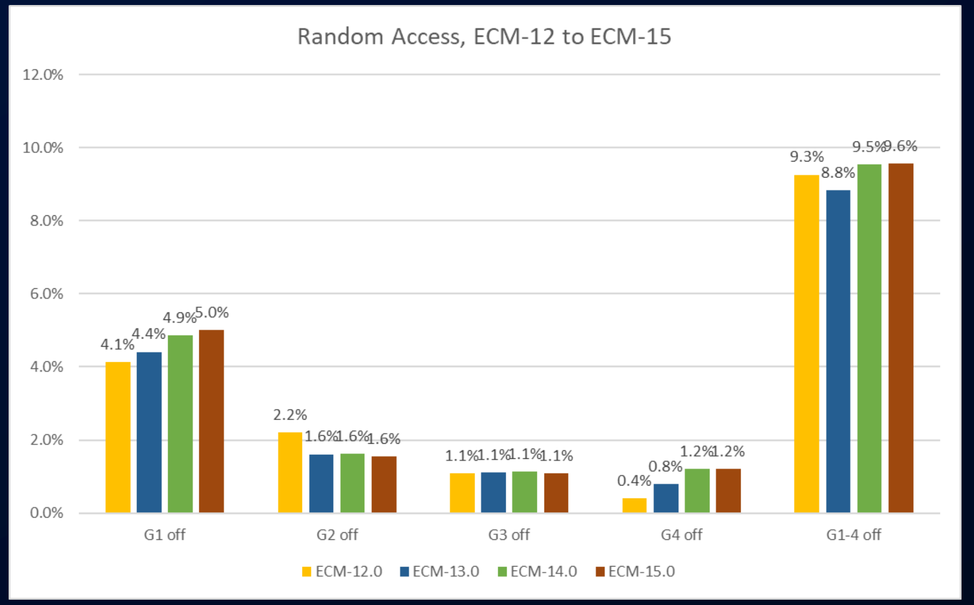
Group 5 (new)
- Tools needing large memory footprint
- Hardware cost
- Ex: non-separable primary transform

X. Li. et al., “AHG report: ECM tool assessment (AHG7)”, JVET-AK0007, Jan 2025

Tool assessment of ECM-15.0

	Anchor ECM-15.0					Anchor VTM11ECM15		
	Y	U	V	EncT	DecT	Y	U	V
G1 off	5.0%	4.9%	5.1%	82%	67%	-22.9%	-24.1%	-26.8%
G2 off	1.6%	1.1%	1.4%	87%	98%	-25.4%	-26.8%	-29.5%
G3 off	1.1%	3.1%	4.1%	92%	97%	-25.8%	-25.5%	-27.8%
G4 off	1.2%	4.0%	4.3%	92%	97%	-25.7%	-24.9%	-27.6%
G1-4 off	9.6%	14.1%	16.0%	59%	63%	-19.5%	-17.7%	-19.8%

- Random Access config., groups 1-4, group 5 to be tested in the future
- Tool off performance shows still significant gains over VVC (VTM with encoder-only optimizations)
- Relatively stable performance over time
- Study of the tools, not representing any final conclusions on their implementation feasibility
- Closer examination of implementation issues will be conducted during standardization



04 Concluding remarks

Conclusion and future outlook

- ECM demonstrates beyond-VVC compression capability
  - Rate reduction of **16.1% (AI), 26.6% (RA), and 22.4% (LDB)** at the same quality (luma PSNR)
  - Visual assessments recently performed in coordination with AG 5, showing some subjective benefits (JVET-AH0344)
- At recent meetings, discussions started regarding next gen. video codec standard
  - Requirements and use cases being collected
- New AHG on beyond-CTC testing, many companies volunteering in the effort
  - Could potentially lead to Call for Evidence in the near future
- Next video codec standard expected to be another successful joint standard from ITU-T SG21 and ISO/IEC JTC 1/SC 29

達摩院
ALIBABA DAMO ACADEMY

Thanks