

Description of SDR, HDR, and 360° video coding technology proposal by InterDigital and Dolby

JVET-J0015

Introduction

- JEM is an experimental platform
 - Provides good coding performance
 - For CS1, JEM anchors provide 33.15% average BD rate savings over HM CfP anchors
 - But has many complexity issues, not just in terms of encoding/decoding time
- Many JEM coding tools are well understood by the experts, and some simplified flavor of them will likely be included in our next standard
- We target simplification with the following important design considerations for the future video coding standard

Decoder complexity is still one of the most important design aspects

Any increase in decoder complexity must be justified

Real-time decoder must be implementable in the near-term

The SDR codec should serve as the core coding engine for HDR and 360 video

Modifications/customizations can be allowed if justified, e.g., subjective quality improvement

Summary of J0015

- This joint response follows these design considerations

- For SDR:

- Great attention is paid to reduce decoder complexity, many tools are simplified with little performance loss
 - Further performance improvement achieved through 2 additional tools

	BD rate savings			Enc time *	Dec time
	Y	U	V		
CS1	35.72%	44.75%	44.95%	1710%	263%
CS2	27.18%	43.81%	44.51%	1827%	301%

* Encoding time increase mostly due to MTT implementation

- For HDR:

- Improve HDR coding efficiency using coding tools that reside inside the core codec

- For 360 video:

- Main goal is to solve face seam artifacts for this type of video → small modifications are proposed to some SDR tools
 - Exploration of projection format as a “coding tool” for coding performance improvement

- This presentation is organized by categories

HDR Category

HDR Summary

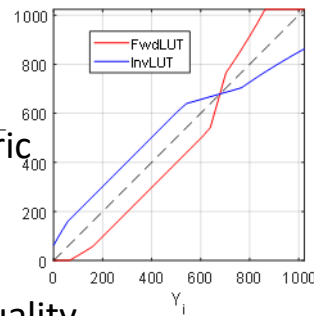
- Two major HDR specific coding tools
 - In-loop reshaping
 - Luma-based QP prediction
- Pre- and Post-processing are unnecessary
 - All proposed tools reside inside core codec
 - Keep conformance point same as SDR
- HDR CfP submitted bitstreams target better subjective performance
 - In-loop reshapers can also be configured to maximize HDR objective metrics

Why In-loop Reshaper?

	ETM Reshaper (out-loop reshaper)	In-loop Reshaper
Architecture	Pre/post processing based, completely out of coding loop; Conformance point issue - output decoded picture is not “viewable”; Extra delay due to post processing.	Completely inside coding loop; transparent to users; no conformance point shift; no extra delay
Codec input	Pre-processed video and corresponding reshaping info metadata	Original video
Codec output	Decoded non-original video which needs post-processing	Decoded video (no need of post processing)
Block level adaptation	Not implemented: boundary issue	ROI reshaping
Slice adaptation	Scene based: when ref and cur slice have different reshaper curve, it might hurt coding efficiency	Yes
Additional picture buffer	Need to store two copies of each picture for decoding/display purpose	No need

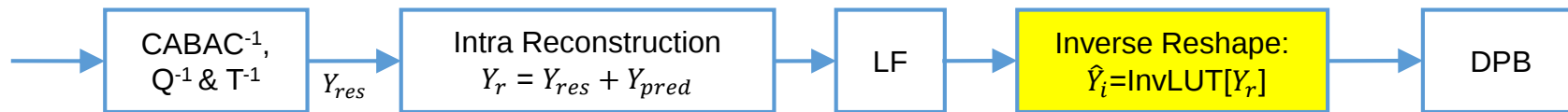
In-loop Reshaper: Luma Only

- Reshaper: A pair of one dimensional, 10-bit, 1024-entry mapping tables (1D-LUT): FwdLUT and InvLUT
 - Remaps Luma code values inside the coding loop for better HDR coding efficiency
 - Reconstructed pictures in DPB is in the same form as original source picture
- Design philosophy :
 - Default mode:
 - Uses a pre-defined Luma mapping curve derived from weights from the wPSNR metric
 - Helps validate the in-loop reshaping architecture
 - Adaptive mode:
 - Allows adaptive adjustment to the Luma mapping curve for better HDR subjective quality
 - Enables future commercial optimization using better HVS models



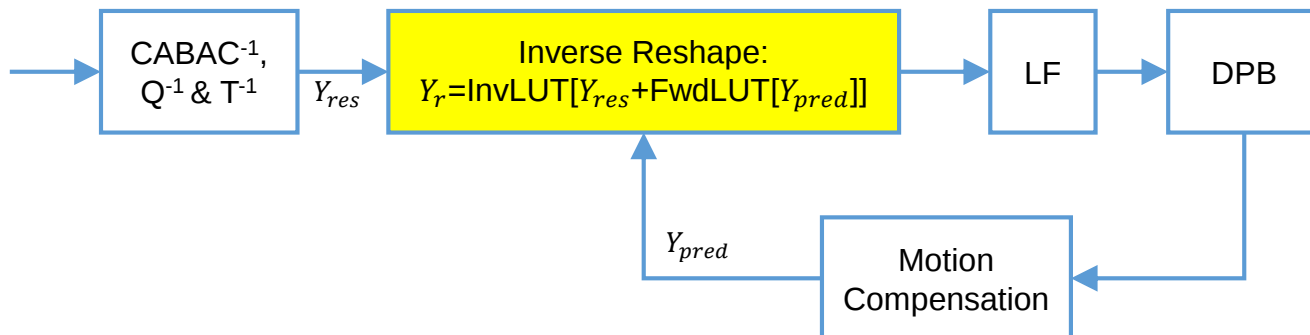
In-loop Reshaper: Intra Slice

- Coding process is performed in the reshaped domain, followed by inverse reshaping before storing in DPB
- Encoder:
 - Forward reshaping, intra coding in reshaped picture, inverse reshaping reconstructed picture before stored in DPB buffer.
- Decoder:
 - Intra decoding, inverse reshaping reconstructed picture before stored in DPB buffer



In-loop Reshaper: Inter Slice

- Reshaping is performed within the prediction loop
- Encoder:
 - Modify how residue is generated: $Y_{res} = FwdLUT[Y_i] - FwdLUT[Y_{pred}]$
- Decoder:
 - Modify how reconstructed sample is generated:
 $Y_r = InvLUT[Y_{res} + FwdLUT[Y_{pred}]]$



Luma Reshaper LUT

- Default reshaper: pre-defined luma mapping tables based on wPSNR
- Adaptive reshaper: use parametric model for codeword distribution
 - Divides Luma channel into 32 bins or intervals.
 - Parametric model:
 - signals picture profiling and bin importance
 - pre-defines how codewords are allocated for Luma intervals based on profile and bin importance.
 - Typical size of reshaper syntax is around 30 bits.
- In CfP submission, adaptive reshaper is applied and syntax is only sent in intra slices.

In-loop Reshaper: Additional Configurable Coding Tool

- Region-of-Interest reshaping:
 - Reshaping is only enabled in ROI
 - ROI is defined as a rectangular area in multiples of CTU size
 - For intra slice, to avoid boundary artifacts, inverse reshaping is applied to ROI before loop filter
 - Used on HLG content in the submission
- Luma-based Chroma QP Offset
 - To compensate for Luma signal interaction with Chroma signal, an additional Chroma QP Offset is used to adjust Chroma QPs
- Loop Filter Optimization
 - For inter slice, control if loop filter is applied in reshaped domain or not

Luma-based QP Prediction

- Target: reduce bitrate overhead of the deltaQP syntax when spatial Luma-dependent QP adaptation (LumaDQP) is used.
- Assumption: CU average predicted luma value is very close to that of original

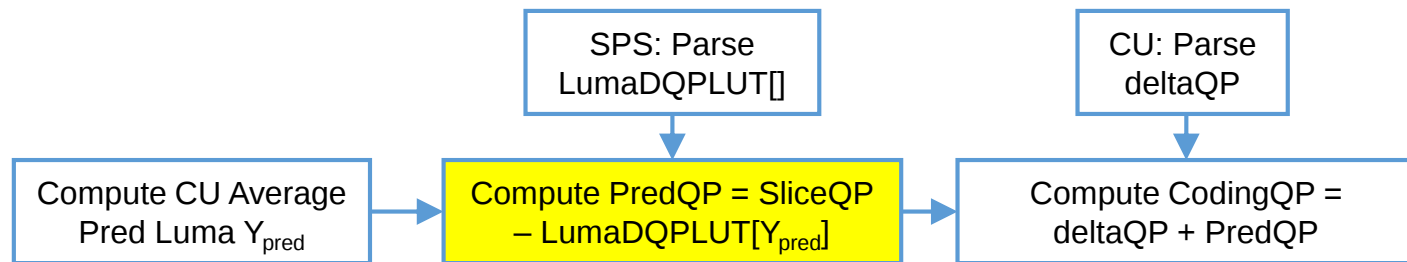
- Encoder:

$$\text{CodingQP} = \text{SliceQP} - \text{LumaDQPLUT}[\text{Yorg}]$$

$$\text{PredQP} = \text{SliceQP} - \text{LumaDQPLUT}[\text{Ypred}]$$

$$\text{deltaQP} = \text{CodingQP} - \text{PredQP} = \text{LumaDQPLUT}[\text{Ypred}] - \text{LumaDQPLUT}[\text{Yorg}]$$

- Decoder:



Additional Encoder Operation

- Joint optimization of reshaper and luma QP adaptation
- RDO in Mode Decision:
 - intra: SSE
 - inter: wSSE
- Chroma QP offset adjustment
 - PQ: same as anchor
 - HLG: chromaQPoffset is derived in a similar way as that used for the PQ anchors

Compression Performance:

- Setting:
 - RA configuration similar to SDR setting
 - Both HDR specific coding tools and all SDR coding tools are applied.
- CfP submission: adaptive reshaping
- Additional results: default reshaping

CfP Submission: Compared to HM and JEM

		Over HM anchor					Over JEM-7.0 anchor				
		BD-rate					BD-rate				
		DE100	PSNRL100	wPsnrY	wPsnrU	wPsnrV	DE100	PSNRL100	wPsnrY	wPsnrU	wPsnrV
HDR-A	DayStreet	-59.0%	-34.6%	-26.8%	-58.2%	-71.5%	-2.0%	2.9%	13.5%	-7.1%	-8.6%
	PeopleInShoppingCenter	-51.7%	-27.6%	-23.2%	-62.2%	-58.3%	-3.7%	2.5%	7.1%	-12.9%	-11.0%
	SunsetBeach	-60.4%	-28.1%	-22.7%	-88.3%	-88.0%	-10.0%	-2.7%	-2.0%	-14.9%	-22.8%
HDR-B	Market	-29.7%	-28.4%	-24.3%	-31.1%	-24.2%	15.7%	-5.3%	1.0%	27.2%	36.8%
	ShowGirl2	-25.1%	-23.7%	-16.8%	-50.1%	-22.8%	4.0%	3.8%	15.0%	27.5%	15.4%
	Hurdles	-42.3%	-39.8%	-33.1%	-54.9%	-47.4%	13.9%	-8.1%	0.7%	17.4%	18.9%
	Starting	-38.7%	-30.3%	-24.6%	-26.7%	-30.7%	-7.7%	0.8%	8.6%	12.4%	-0.7%
	Cosmos1	-32.8%	-30.0%	-17.7%	-61.3%	-76.6%	-3.5%	-7.8%	8.8%	14.8%	1.2%
Average HDR-A		-57.0%	-30.1%	-24.2%	-69.6%	-72.6%	-5.2%	0.9%	6.2%	-11.7%	-14.1%
Average HDR-B		-33.7%	-30.4%	-23.3%	-44.8%	-40.4%	4.5%	-3.3%	6.8%	19.8%	14.3%
Average all		-42.5%	-30.3%	-23.7%	-54.1%	-52.4%	0.8%	-1.7%	6.6%	8.0%	3.6%
Enc Time[%]		1462%					183%				
Dec Time[%]		292%					33%				

Additional Results: Default reshaper

	Anchor HM: BD-rate							
	DE100	PSNRL100	wPsnrY	wPsnrU	wPsnrV	psnrY	psnrU	psnrV
Market	-38.8%	-28.8%	-28.8%	-41.5%	-43.2%	-25.4%	-44.7%	-43.1%
ShowGirl2	-38.0%	-29.7%	-30.6%	-69.3%	-52.7%	-27.7%	-70.6%	-56.6%
Hurdles	-47.5%	-38.9%	-37.7%	-57.8%	-53.1%	-34.0%	-54.0%	-49.4%
Starting	-36.7%	-34.4%	-33.0%	-36.0%	-34.9%	-30.3%	-37.3%	-30.0%
Cosmos1	-36.4%	-26.4%	-26.2%	-71.1%	-83.8%	-24.2%	-70.9%	-83.9%
Average HDR-B	-39.5%	-31.6%	-31.3%	-55.1%	-53.6%	-28.3%	-55.5%	-52.6%

	Anchor JEM: BD-rate							
	DE100	PSNRL100	wPsnrY	wPsnrU	wPsnrV	psnrY	psnrU	psnrV
Market	1.7%	-5.9%	-5.2%	8.9%	4.2%	-2.6%	-8.9%	-16.5%
ShowGirl2	-15.2%	-4.5%	-4.4%	-23.4%	-35.2%	-2.7%	-34.0%	-46.0%
Hurdles	4.6%	-7.0%	-6.5%	6.7%	11.3%	-4.7%	3.2%	8.1%
Starting	-3.8%	-5.3%	-4.0%	-2.5%	4.4%	-2.4%	-8.0%	-9.0%
Cosmos1	-8.7%	-3.0%	-2.8%	-13.9%	-30.7%	-1.9%	-14.4%	-29.3%
Average HDR-B	-4.3%	-5.1%	-4.6%	-4.8%	-9.2%	-2.9%	-12.4%	-18.5%

Subjective Results: Market @R2



JEM anchor



Response

Subjective Results: Starting@R1



JEM anchor



Response

In-loop Reshaper Tool-Off Test

- Default reshaper is configured to perform the test.
- The anchor of the “tool-off” test is the CfP response using default in-loop reshaper.

		BD-rate							
		DE100	PSNRL100	wPsnrY	wPsnrU	wPsnrV	psnrY	psnrU	psnrV
HDR-B	Market	7.1%	7.8%	5.1%	15.1%	12.3%	-1.2%	5.2%	4.4%
	ShowGirl2	5.6%	2.9%	1.6%	9.2%	12.7%	-1.9%	-1.0%	1.1%
	Hurdles	8.5%	6.5%	5.0%	11.8%	5.8%	2.6%	7.0%	4.5%
	Starting	18.2%	6.0%	2.3%	10.6%	14.1%	-1.0%	4.8%	5.8%
	Cosmos1	1.6%	3.4%	1.3%	-4.5%	-4.4%	-0.1%	-6.5%	-9.2%
	Average all	8.2%	5.3%	3.1%	8.4%	8.1%	-0.3%	1.9%	1.3%
Enc Time[%]		96%							
Dec Time[%]		94%							

HDR summary

- HDR tools should be part of the base codec:
 - Wide commercial success of HDR in the market place.
- Concept of the reshaper has proven to be beneficial
 - Increasing compression efficiency.
- Tools:
 - Simplified method includes the in-loop reshaper in the core codec
 - Enhanced luma-based QP prediction mechanism.
- Propose inclusion of tools in Test Model.

THANK YOU