MEDIATEK

JVET-O0293

CE6-related: Latency reduction for LFNST signalling

Authors: Man-Shu Chiang, Chih-Wei Hsu, Tzu-Der Chuang, Ching-Yeh Chen, Yu-Wen Huang, Shaw-Min Lei

Presenter: Man-Shu Chiang

Overall Summary

- Two methods are proposed to reduce latency for LFNST
  - Method 1: Signal LFNST index at the end of the first non-zero TB
  - Method 2: Signal LFNST index at the beginning of TU
-

#	Config.	VTM-5.0		
		Y	U	V
Proposed method 1	AI	0.07%	0.13%	0.19%
	RA	0.01%	0.12%	0.09%
Proposed method 2	AI	0.09%	-0.11%	-0.07%
	RA	0.03%	-0.06%	-0.10%

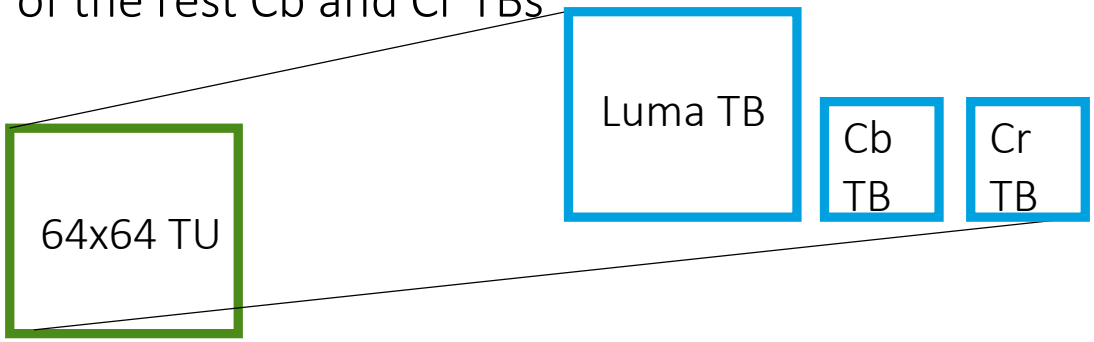
Introduction to LFNST Issues

- If the following conditions are satisfied, LFNST index can be signalled
  - numSigCoeff (# of nonzero coefficients in the current CU) is larger than a threshold (T)
  - numZeroOutSigCoeff (# of nonzero coefficients in LFNST zero-out region for the current CU) is equal to 0
- Counting numSigCoeff and numZeroOutSigCoeff for LFNST causes pipeline issue and latency issue
- The pipeline issue can be solved by any one or more proposed methods in JVET-O0291
- This contribution proposes to further reduce latency for LFNST

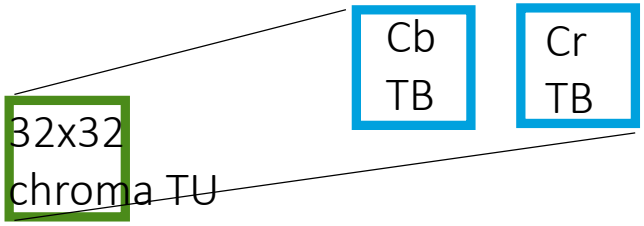
MEDIATEK

Introduction to LFNST Latency Issue

- latency issue (Here take 64x64 TU as an example)
  - When shared tree is applied, to count numSigCoeff and numZeroOutSigCoeff, luma TB needs to wait for the coefficient parsing of the rest Cb and Cr TBs



  - When chroma separate tree (CST) is applied, to count numSigCoeff and numZeroOutSigCoeff, Cb TB needs to wait for the coefficient parsing of the Cr TBs



Proposed Method 1

- For CST, LFNST index is signalled at the first TB with nonzero coefficient and is shared with the TB in the same TU
  - numSigCoeff: # of nonzero coefficients within the the first nonzero TB
    - T = 1 for luma
    - T = ((cb_cbf && cr_cbf) ? 0 : 1) for chroma
  - numZeroOutSigCoeff: # of nonzero coefficients within the zero-out region of the first nonzero TB
- For shared tree, the LFNST index is signalled at luma TB and is shared with the chroma TBs in the same TU
  - numSigCoeff: # of nonzero coefficients within luma TB
    - T = 1
  - numZeroOutSigCoeff: # of nonzero coefficients within the zero-out region of luma TB

Proposed Method 1

- After parsing the LFNST index, the coefficients of the first nonzero TB can be processed by inverse LFNST without waiting for the coefficient parsing of the rest TBs

	All Intra Main10 Over VTM-5.0				
	Y	U	V	EncT	DecT
Class A1	0.08%	0.20%	0.32%	99%	100%
Class A2	0.07%	0.25%	0.37%	101%	101%
Class B	0.06%	0.18%	0.17%	100%	100%
Class C	0.06%	-0.02%	0.09%	100%	99%
Class E	0.06%	0.08%	0.04%	100%	105%
Overall	0.07%	0.13%	0.19%	100%	101%
Class D	0.07%	0.25%	0.16%	101%	99%
Class F	0.07%	-0.05%	0.17%	101%	100%

	Random access Main10 Over VTM-5.0				
	Y	U	V	EncT	DecT
Class A1	0.03%	0.21%	0.02%	99%	98%
Class A2	0.02%	0.13%	0.16%	101%	101%
Class B	0.01%	0.31%	0.11%	101%	97%
Class C	-0.01%	-0.18%	0.06%	100%	98%
Class E					
Overall	0.01%	0.12%	0.09%	100%	98%
Class D	0.03%	0.24%	0.40%	101%	103%
Class F	0.01%	0.05%	0.00%	101%	97%

Proposed Method 2

- The positions of last significant coefficients of Cb and Cr are signalled before the luma coefficients
- The LFNST index is signalled at the beginning of a TU (before signalling the coefficients of the first TB) if (1) and (2) are satisfied
- (1) estNumSigCoeff is larger than 1, where estNumSigCoeff is derived as follows
  - Initialized as 0 for a TU
  - Increase by 1 if last significant coefficient is at 0 (DC) position
  - Increase by 2 if last significant coefficient is at > 0 position
- (2) Last significant coefficient for each TB in the TU is not located within the LFNST zero-out region

MEDIATEK

Proposed Method 2

- The flag for joint coding of chrominance residuals is also moved before the positions of last significant coefficients for Cb and Cr
- Also, if the LFNST is enabled, the coefficient parsing within the zero-out region is skipped and inferred as 0
  - For example,
    - If the zero-out region includes scan indices ranging from 8 to 15, the significance flags for the coefficients with the coefficient scan index within 8 to 15 are inferred as 0
    - If the zero-out region includes scan indices ranging from 16 to 47. the coefficient group (CG) significance flags for the second and the third CG are inferred as 0
-

Proposed Method 2

- With the proposed method 2, the inverse LFNST can be processed after the first 8 or 16 coefficients are parsed without waiting for the coefficient parsing of the first TB and the rest TBs

	All Intra Main10 Over VTM-5.0				
	Y	U	V	EncT	DecT
Class A1	0.13%	-0.14%	-0.08%	101%	101%
Class A2	0.09%	-0.12%	-0.10%	101%	100%
Class B	0.10%	-0.11%	-0.06%	101%	100%
Class C	0.06%	-0.11%	-0.03%	101%	98%
Class E	0.08%	-0.04%	-0.09%	101%	99%
Overall	0.09%	-0.11%	-0.07%	101%	100%
Class D	0.06%	0.07%	0.01%	100%	100%
Class F	0.03%	-0.02%	0.06%	101%	100%

	Random access Main10 Over VTM-5.0				
	Y	U	V	EncT	DecT
Class A1	0.01%	-0.21%	-0.06%	101%	97%
Class A2	0.03%	0.04%	-0.01%	101%	98%
Class B	0.06%	-0.03%	-0.29%	101%	97%
Class C	0.02%	-0.06%	0.03%	100%	97%
Class E					
Overall	0.03%	-0.06%	-0.10%	100%	97%
Class D	0.02%	0.09%	-0.11%	100%	100%
Class F	-0.02%	-0.08%	0.00%	101%	97%

Conclusions

- This contribution proposes to further reduce latency for LFNST after decoder pipeline issue is solved
- With Method 1, LFNST index can be parsed at TB level without waiting for the rest TBs in the TU
- With method 2, LFNST index can be parsed without waiting for the coefficient parsing of any TB in the TU
- It is suggested to consider method 1 and 2 in CE