JCTVC-G530

Layered quantization matrices compression

Yunfei Wang, Tsinghua
Jianhua Zheng, HiSilicon
Xiaozhen Zheng, HiSilicon
Yun He, Tsinghua

Outline

- Summary of the contribution
- Layered Quantization Matrix coding
- Experimental Results
- Conclusions

Summary of the contribution

- **Layered quantization matrix representation and coding structure.**
- **Matrix is represented as coded layers and predicted layers, matrix coefficients are divided into multiple layers.**
- **By changing the number of coded layers, it easily shift from matrix lossless coding to matrix lossy coding with high compression ratio.**
- **Support a large range variation of compression ratio.**
- **Reduce matrix bits to 50% in average compared to F475 with similar error.**
- **140x~153x compression ratio improvements against AVC method in lossy coding and 4~6.8x improvements against AVC method in lossless coding.**

Crosscheck by JCTVC-G730(SONY) and JCTVC-G992(MediaTek)

Layered Quantization Matrix coding

- matrix coefficients are divided into multiple layers.
- Two types layers: Coding matrix from lower layers to higher layers
- coded layer – residue of the predicted value and the original element is coded.
- predicted layer – residue in higher layer is not coded to reduce the generated bits.

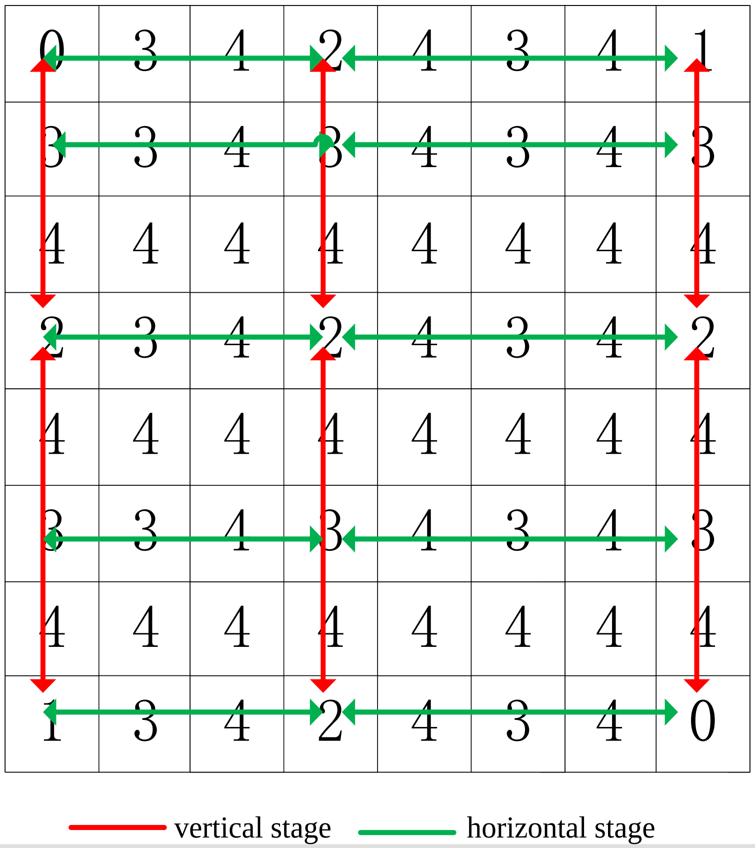
0	3	4	2	4	3	4	1
3	3	4	3	4	3	4	3
4	4	4	4	4	4	4	4
2	3	4	2	4	3	4	2
4	4	4	4	4	4	4	4
3	3	4	3	4	3	4	3
4	4	4	4	4	4	4	4
1	3	4	2	4	3	4	0

layers for quantization matrix

4x4 matrix, layer: 0~3
8x8 matrix, layer: 0~4
16x16 matrix, layer: 0~5
32x32 matrix, layer: 0~6
NxN matrix, layer: 0~$\log_2 N+1$

Layered Quantization Matrix coding

- **Prediction:**
- **Stage 1: vertical stage.** if it can be vertical interpolated, then element is vertical interpolated by reference elements R1 and R2, (R1 and R2 are the nearest coded elements in the same column).
- **Stage 2: horizontal stage.** The rest elements are predicted by horizontal interpolation. (R1 and R2 are the nearest coded elements in the same row)



Interpolation (take layer 3 as an example):

- (R1 + R2 + 1) >>1

- (4R1 + 2R2 + 3) / 6
  =R1 + (2(R2 – R1) + 3) / 6

interpolation within the stage can be paralleled.

Layered Quantization Matrix coding

- ***Copy mode***
  - Matrix is derived directly from its reference matrix One bit is signaled
  - Down sampling in copy mode: the elements with even indexes are extracted directly from the reference matrix.

- ***Symmetry mode:*** half of the elements are coded. One bit is signaled

- ***Differential coding:***
  differential coding between two adjacent matrix sets in temporal axis is employed, the values in the differential matrix are coded.

- ***Residue Coding***
  Layer >=4: coded by run-length code.
  Layer   <4: coded by signed exp-golomb code

Experimental results

- ***TEST1:*** only the coding performance of matrix itself is measured
- AVC

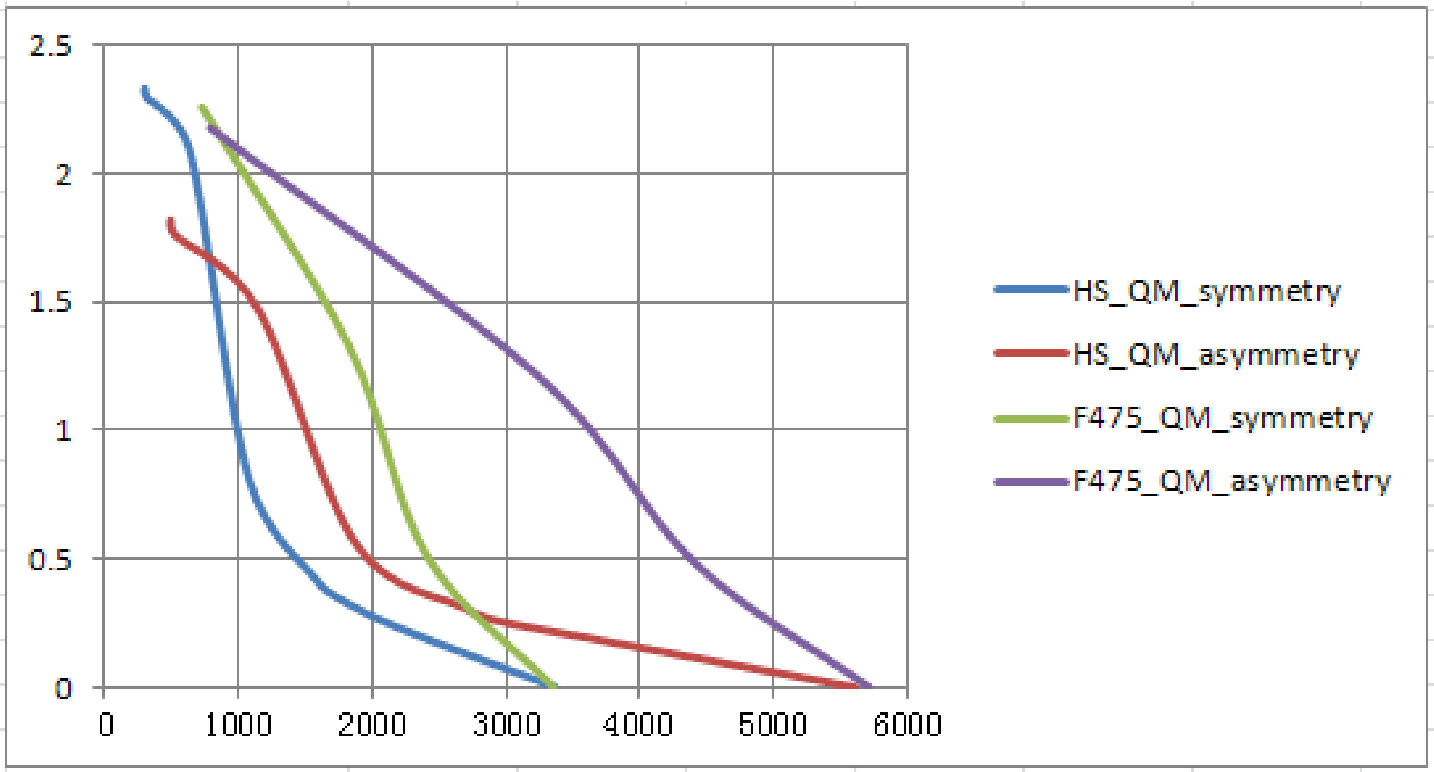
Symmetry bits		Asymmetry bits
26240		28232

- Proposed

QMSizexCodeLevels	Symmetry bits	AvgError		Asymmetry bits	AvgError
3,4,5,6	3372	0		5614	0
3,4,5,5	2187	0.234191		3661	0.188235
3,4,4,5	1740	0.349632		2984	0.253676
3,3,4,5	1580	0.419853		2690	0.308456
3,3,4,4	1076	0.814338		1876	0.559926
3,3,3,3	632	2.095221		1144	1.474265
1,2,2,3	314	2.297426		522	1.7625
1,1,2,3	297	2.332353		493	1.819853

Experimental results

- ***TEST1:*** only the coding performance of matrix itself is measured



It is observed that in the typical average coding error range (0, 2], the proposed layered quantization matrix compression method can further reduce the coded matrix bits to nearly 50% in average.

Experimental results

- ***TEST2:***

	All Intra HE			All Intra LC		
	Y	U	V	Y	U	V
Class A	-1.7%	-3.1%	-2.9%	-3.7%	-8.0%	-7.4%
Class B	-6.3%	-6.7%	-6.6%	-7.3%	-10.0%	-9.9%
Class C	-15.5%	-15.9%	-16.3%	-15.1%	-16.3%	-16.9%
Class D	-34.6%	-34.1%	-34.4%	-33.4%	-33.6%	-33.8%
Class E	-19.0%	-19.2%	-18.0%	-18.7%	-21.8%	-21.2%
Class F						
Overall	-14.8%	-15.2%	-15.1%	-15.1%	-17.3%	-17.3%
	-14.8%	-15.2%	-15.2%	-15.1%	-17.4%	-17.4%
Enc Time[%]	100%			99%		
Dec Time[%]	97%			100%		

	Random Access HE			Random Access LC		
	Y	U	V	Y	U	V
Class A	-19.4%	-19.0%	-18.6%	-19.0%	-17.7%	-16.3%
Class B	-37.6%	-35.0%	-33.9%	-36.1%	-33.4%	-32.4%
Class C	-56.8%	-55.1%	-55.3%	-55.8%	-54.2%	-54.4%
Class D	-79.2%	-77.8%	-78.0%	-78.6%	-77.5%	-77.7%
Class E						
Class F						
Overall	-47.6%	-46.0%	-45.7%	-46.7%	-45.0%	-44.4%
	-47.8%	-46.1%	-45.9%	-46.9%	-45.0%	-44.5%
Enc Time[%]	100%			99%		
Dec Time[%]	99%			100%		

	Low delay B HE			Low delay B LC		
	Y	U	V	Y	U	V
Class A						
Class B	-39.9%	-36.6%	-36.0%	-38.7%	-35.4%	-35.1%
Class C	-57.3%	-54.8%	-55.2%	-56.5%	-54.3%	-55.0%
Class D	-79.0%	-77.3%	-77.4%	-78.6%	-77.0%	-77.5%
Class E	-82.4%	-80.0%	-80.6%	-81.1%	-78.4%	-79.3%
Class F						
Overall	-62.0%	-59.5%	-59.5%	-61.1%	-58.6%	-59.0%
	-62.0%	-59.6%	-59.6%	-61.1%	-58.6%	-59.0%
Enc Time[%]	100%			100%		
Dec Time[%]	99%			97%		

	Low delay P HE			Low delay P LC		
	Y	U	V	Y	U	V
Class A						
Class B	-38.8%	-35.9%	-35.0%	-37.4%	-34.4%	-33.9%
Class C	-56.6%	-54.1%	-54.3%	-55.7%	-53.6%	-54.3%
Class D	-78.5%	-76.5%	-77.0%	-78.2%	-76.3%	-76.9%
Class E	-81.8%	-79.3%	-80.0%	-80.3%	-77.8%	-78.6%
Class F						
Overall	-61.2%	-58.7%	-58.7%	-60.2%	-57.8%	-58.1%
	-61.2%	-58.8%	-58.9%	-60.2%	-57.9%	-58.2%
Enc Time[%]	100%			100%		
Dec Time[%]	104%			102%		

Experimental results

- ***TEST2:***

QMSizexCodeLevels =2,2,2,3. symmetric matrix

All Seqs	AVC Method					proposed method				
	QP=22	QP=27	QP=32	QP=37	average	QP=22	QP=27	QP=32	QP=37	average
Intra_LC	60.1%	91.0%	135.3%	196.1%	120.7%	0.5%	0.8%	1.3%	2.3%	1.2%
Intra_HE	63.2%	96.5%	145.0%	213.4%	129.5%	0.5%	0.8%	1.4%	2.5%	1.3%
RA_LC	208.7%	313.2%	420.7%	520.3%	365.7%	2.7%	5.6%	11.3%	21.1%	10.2%
RA_HE	214.6%	321.5%	431.1%	533.8%	375.3%	2.8%	5.8%	11.7%	22.1%	10.6%
LB_LC	296.2%	434.9%	551.9%	642.3%	481.3%	3.9%	9.1%	18.7%	35.5%	16.8%
LB_HE	305.5%	444.9%	561.4%	651.1%	490.7%	4.1%	9.5%	19.6%	37.6%	17.7%
LP_LC	280.2%	424.7%	546.5%	639.7%	472.8%	3.5%	8.5%	18.1%	35.1%	16.3%
LP_HE	290.2%	435.4%	557.1%	649.6%	483.1%	3.7%	8.9%	19.1%	37.3%	17.2%

Comparing to AVC method, the matrix bits ratio has been decreased from 1-5x to 1.2-18.5%.

Experimental results

- ***TEST2:***

QMSizexCodeLevels =2,2,2,3. symmetric matrix

HDSeqs	AVC Method					proposed method				
ClassB	QP=22	QP=27	QP=32	QP=37	average	QP=22	QP=27	QP=32	QP=37	average
Intra_LC	18.0%	32.0%	53.8%	89.0%	48.2%	0.1%	0.2%	0.4%	0.7%	0.4%
Intra_HE	19.1%	34.6%	60.2%	102.9%	54.2%	0.1%	0.2%	0.4%	0.8%	0.4%
RA_LC	94.2%	196.2%	322.4%	450.6%	265.9%	0.7%	1.7%	3.7%	7.1%	3.3%
RA_HE	99.7%	206.9%	339.2%	474.7%	280.1%	0.8%	1.9%	3.9%	7.9%	3.6%
LB_LC	94.0%	203.0%	343.0%	483.4%	280.9%	0.7%	1.8%	4.1%	8.7%	3.8%
LB_HE	99.7%	214.4%	359.1%	501.5%	293.7%	0.8%	2.0%	4.5%	9.5%	4.2%
LP_LC	85.8%	190.9%	332.1%	476.7%	271.4%	0.6%	1.7%	3.9%	8.3%	3.6%
LP_HE	92.9%	204.5%	351.0%	497.6%	286.5%	0.7%	1.8%	4.2%	9.2%	4.0%

For the HD(Class B) sequences, the generated matrix bits ratio of has been reduced to 4% around.

Experimental results

- ***TEST2:***

  QMSizexCodeLevels =2,2,2,3.

	Average compression ratio against AVC	
	symmetry	asymmetry
Intra_LC	152.6	139.3
Intra_HE	152.6	139.3
RA_LC	152.3	138.9
RA_HE	152.3	138.9
LB_LC	153.5	140.7
LB_HE	153.5	140.7
LP_LC	153.5	140.7
LP_HE	153.5	140.7

Conclusions

- **The proposed method support a large range variation of compression ratio, easily shift from lossless coding to lossy coding.**
- **The proposed method achieve 140x~153x compression ratio against AVC method in lossy coding and 4~6.8x improvements against AVC method in lossless coding while the encoding and decoding time are similar.**

Recommendation:

- **Support both lossless and lossy method in quantization matrix coding in HEVC.**
- **Continue investigate improving the compression ratio of quantization matrix and support layered quantization matrix representation and coding in HM5.**

Thank you!

www.hisilicon.com