



JVET-AD0157, EE1-related: Neural-network loop filters in EE1-1.1.2 with further complexity reduction

April 22, 2023

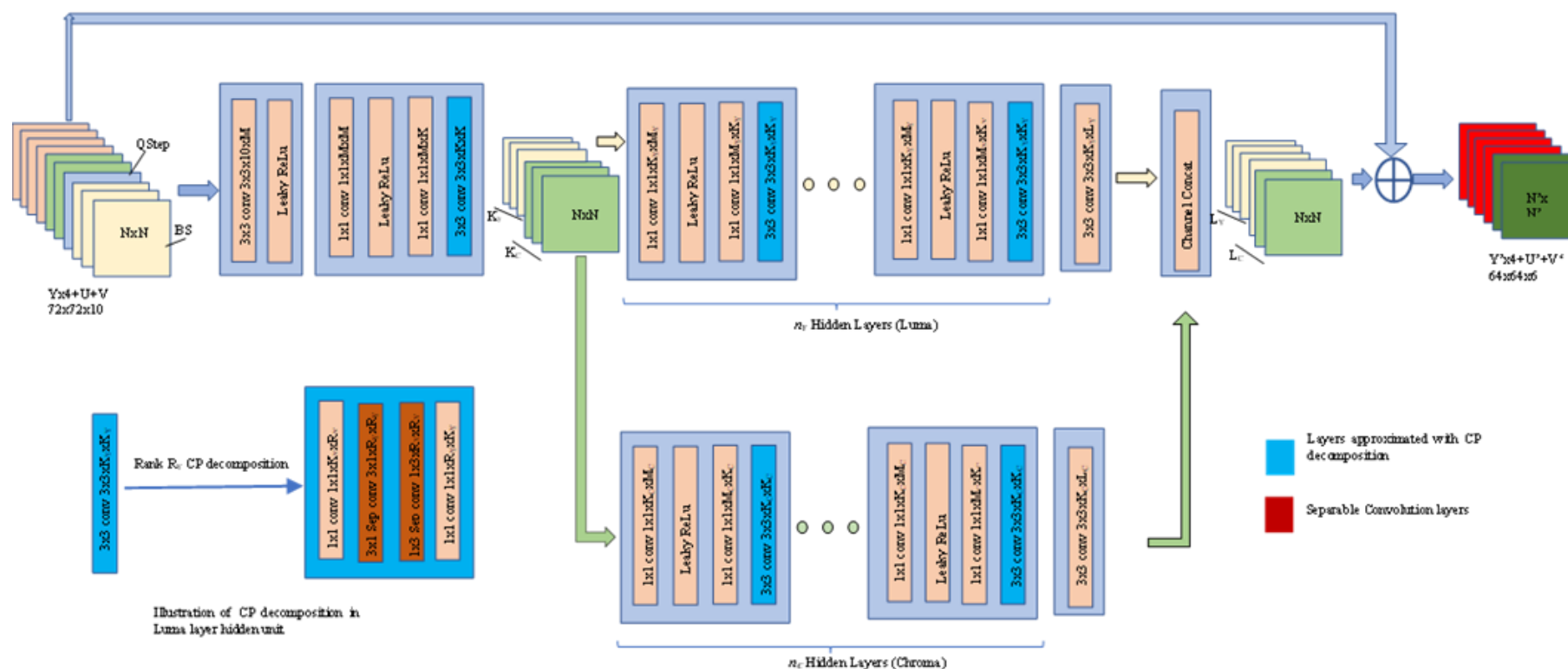
T. Shao, A. Arora, P. Yin, F. Pu, T. Lu, Sean McCarthy (Dolby), J. N. Shingala, A. Shyam, A. Suneja, S. Badya (Ittiam)

Outline

- ❑ Results of EE1-1.1.2, CP fused model with split luma chroma (24L, 8C, 10H): 17.7 KMAC/Pixel
- ❑ 3 variations of further complexity based on EE1-1.1.2
 - (16L, 8C, 10H): 9.9 KMAC/Pixel
 - (16L, 8C, 6H): 7.3 KMAC/Pixel
 - (12L, 4C, 10H): 4.9 KMAC/Pixel
- ❑ Experimental results
- ❑ Conclusions

Baseline model: EE1-1.1.2, CP fused model with split luma chroma

- (24L, 8C, 10H): 24 luma channels, 8 chroma channels, 10 hidden layers, with worst case block level complexity of 17.7KMAC/Pixel



Baseline model: EE1-1.1.2, CP fused model with split luma chroma

❑ SADL fp32, NNVC-4.0 anchor:

❑ AI {-4.99%, -7.46%, -7.03%}

❑ RA {-5.66%, -9.30%, -7.63%}

Table 1 AI performance of EE1-1.1.2 (SADL, fp32, NNVC-4.0 anchor)

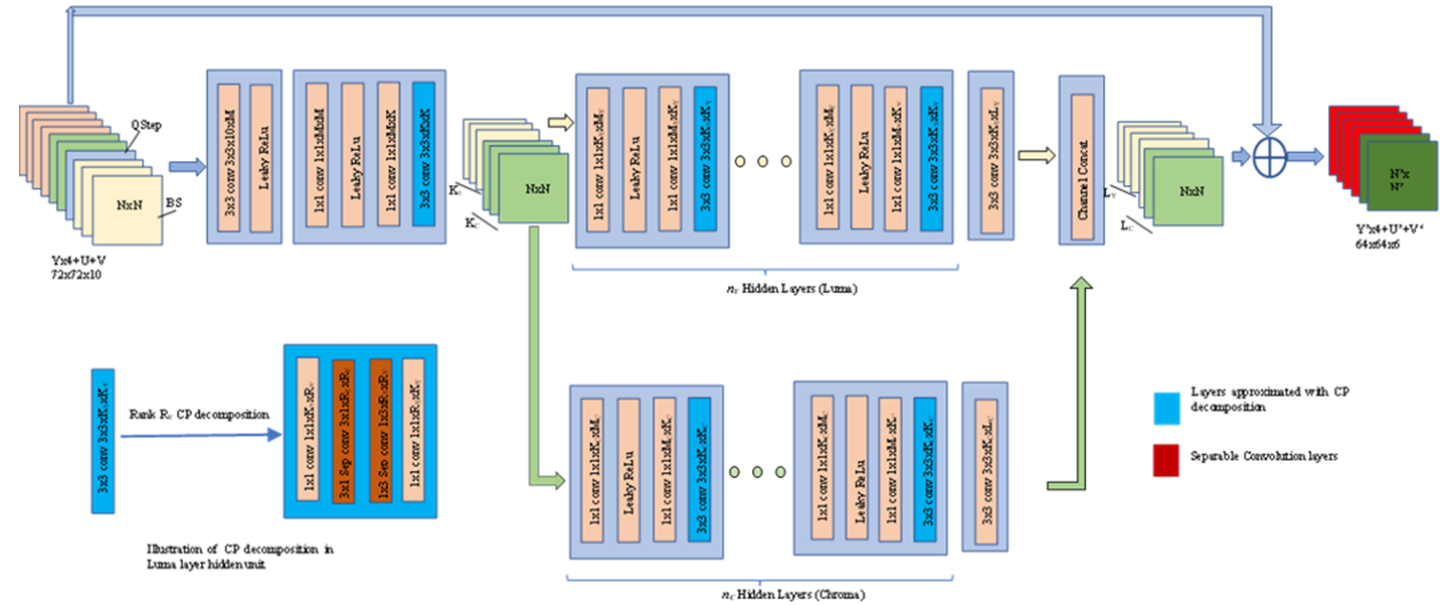
	All Intra Main10				
	BD-rate Over VTM-11.0_nnvc-4.0				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-4.62%	-5.80%	-6.37%	143%	6773%
Class A2	-4.53%	-8.11%	-5.94%	147%	5218%
Class B	-4.43%	-7.76%	-7.39%	132%	5319%
Class C	-5.03%	-8.00%	-7.70%	129%	3977%
Class E	-6.66%	-7.27%	-7.30%	144%	6342%
Overall	-4.99%	-7.46%	-7.03%	137%	5329%
Class D	-5.48%	-6.80%	-8.58%	130%	4105%

Table 2 RA performance of EE1-1.1.2 (SADL, fp32, NNVC-4.0 anchor)

	Random Access Main10				
	BD-rate Over VTM-11.0_nnvc-4.0				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-5.85%	-6.90%	-6.23%	159%	8199%
Class A2	-6.01%	-9.04%	-5.40%	151%	7302%
Class B	-5.28%	-9.85%	-8.88%	161%	8623%
Class C	-5.74%	-10.59%	-8.78%	146%	7859%
Class E					
Overall	-5.66%	-9.30%	-7.63%	155%	8056%
Class D	-6.94%	-10.04%	-10.44%	145%	7865%

Improvements for further complexity reduction based on EE1-1.1.2

- ❑ 1) reduced number of luma and/or chroma channels;
- ❑ 2) CP decomposed first and last convolutional layer;
- ❑ 3) reduced number of luma/chroma path's hidden layers.



Improvements for further complexity reduction based on EE1-1.1.2

- ❑ 1) reduced number of luma and/or chroma channels;
- ❑ 2) CP decomposed first and last convolutional layer;
- ❑ 3) reduced number of luma/chroma path's hidden layers.

Table 1 Model architectures and block-level complexity in KMac/Pixel

Models	Num. of luma channels	Num. of chroma channels	Num. of luma/chroma path hidden layers	Is first and last conv. layer CP decomposed?	Complexity in KMAC/Pixel
(24L, 8C, 10H), EE1-1.1.2 baseline	24	8	10	no	17.7
(16L, 8C, 10H)	16	8	10	yes	9.9
(16L, 8C, 6H)	16	8	6	yes	7.3
(12L, 4C, 10H)	12	4	10	yes	4.9

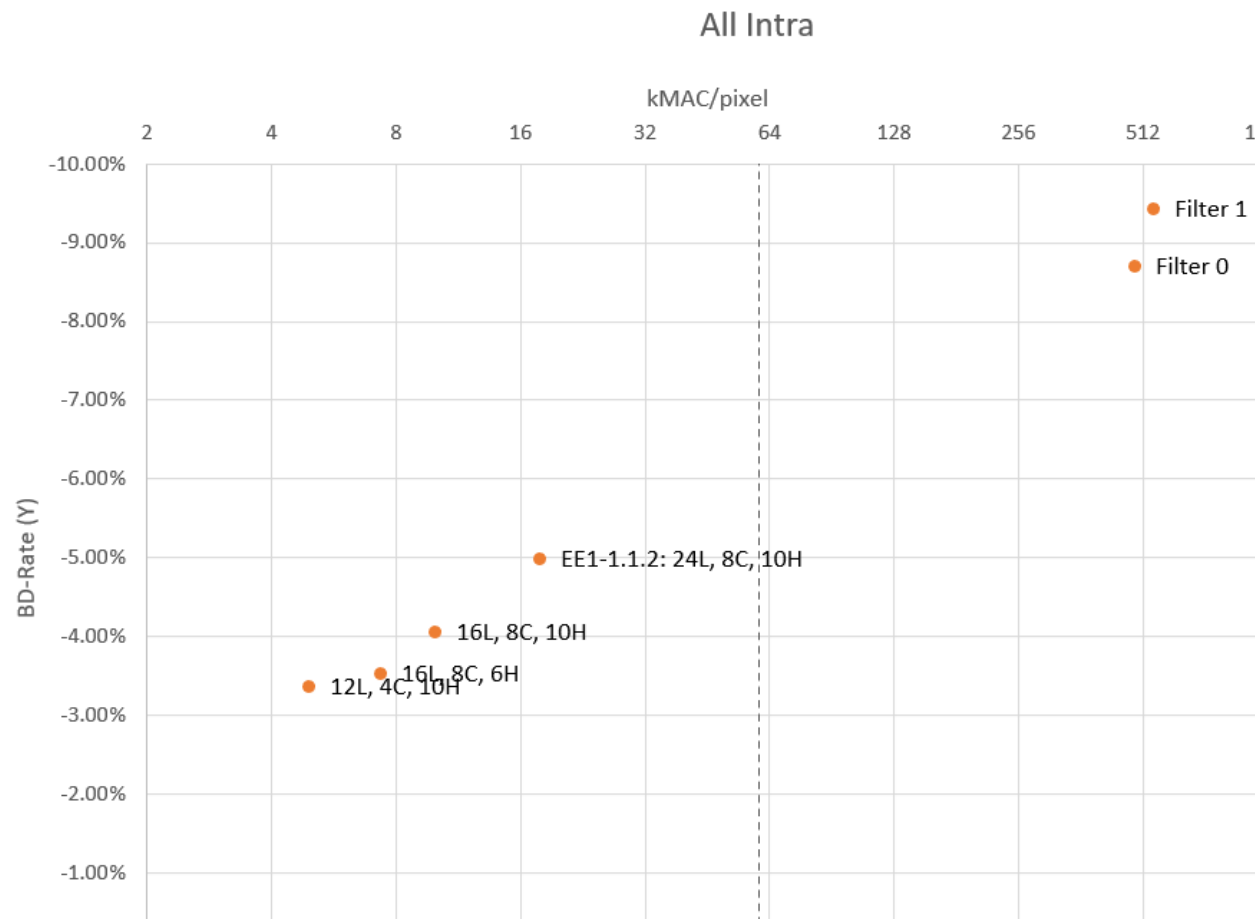
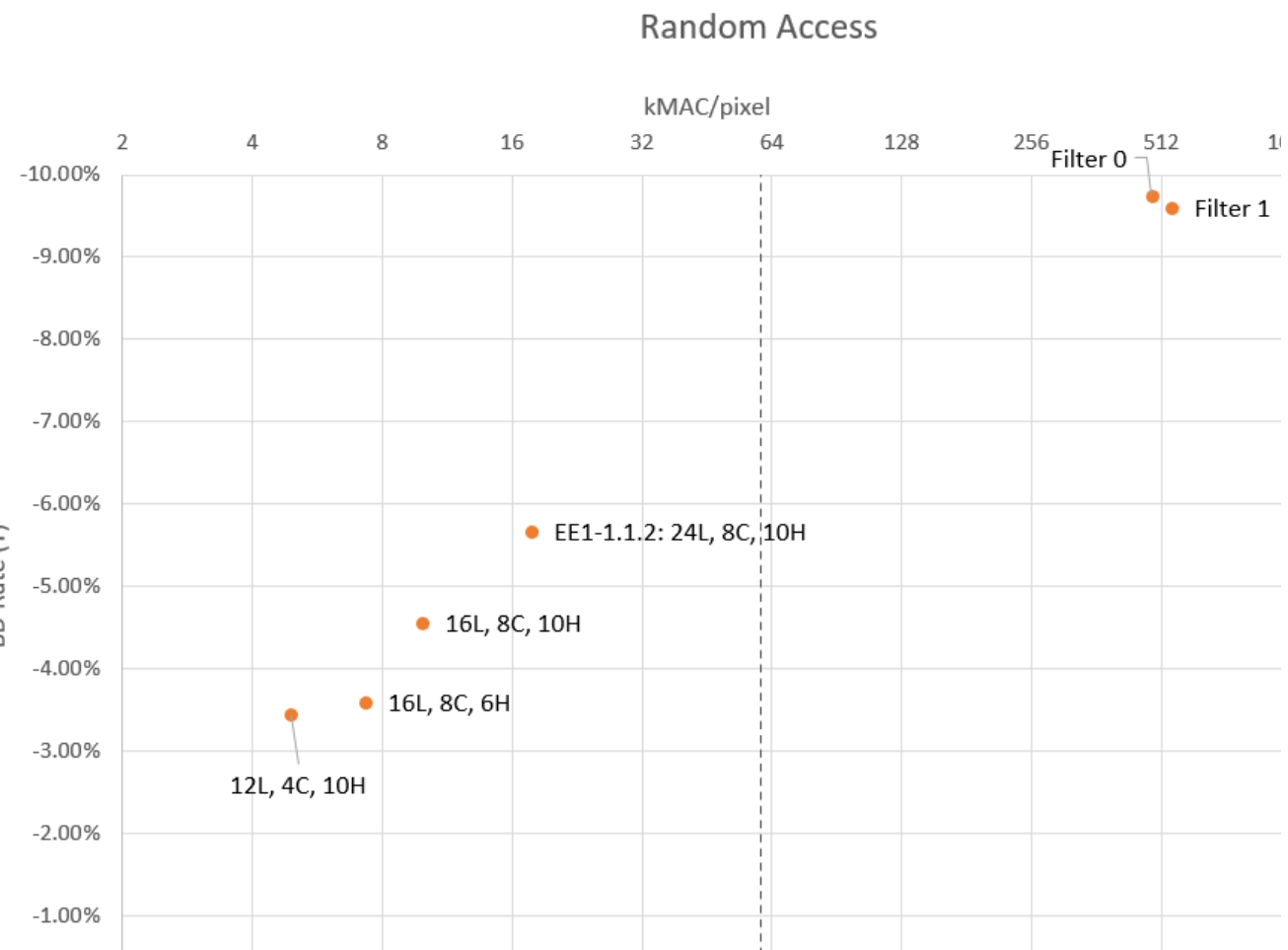
Improvements for further complexity reduction based on EE1-1.1.2

- ❑ (24L, 8C, 10H): fp32: AI {-4.99%, -7.46%, -7.03%}, RA {-5.66%, -9.30%, -7.63%}
- ❑ (16L, 8C, 10H): fp32: AI {-4.07%, -7.97%, -7.99%}, RA {-4.55%, -8.85%, -10.42%}
- ❑ (16L, 8C, 6H): fp32: AI {-3.54%, -6.07%, -5.85%}, RA {-3.60%, -6.68%, -5.72%}
- ❑ (12L, 4C, 10H): fp32: AI {-3.38%, -3.74%, -3.43%}, RA {-3.44%, -4.62%, -2.23%}

Table 1 Model architectures and block-level complexity in KMacs/Pixel

Models	Num. of luma channels	Num. of chroma channels	Num. of luma/chroma path hidden layers	Is first and last conv. layer CP decomposed?	Complexity in KMAC/Pixel
(24L, 8C), EE1-1.1.2 baseline	24	8	10	no	17.7
(16L, 8C, 10H)	16	8	10	yes	9.9
(16L, 8C, 7H)	16	8	6	yes	7.3
(12L, 4C, 10H)	12	4	10	yes	4.9

Improvements for further complexity reduction based on EE1-1.1.2



Experimental results

Table 4 AI performance of (16L, 8C, 10H) (SADL, fp32, NNVC-4.0 anchor)

	All Intra Main10				
	BD-rate Over VTM-11.0_nnvc-4.0				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-3.89%	-5.87%	-7.12%		2767%
Class A2	-3.80%	-8.38%	-6.79%		2163%
Class B	-3.68%	-8.40%	-8.68%		2207%
Class C	-3.96%	-8.75%	-8.56%		1666%
Class E	-5.30%	-7.88%	-8.13%		2600%
Overall	-4.07%	-7.97%	-7.99%		2205%
Class D	-4.43%	-8.05%	-9.49%		1726%

Table 5 RA performance of (16L, 8C, 10H) (SADL, fp32, NNVC-4.0 anchor)

	Random Access Main10				
	BD-rate Over VTM-11.0_nnvc-4.0				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-5.03%	-5.66%	-8.92%		4318%
Class A2	-4.88%	-7.81%	-7.64%		3925%
Class B	-4.17%	-9.85%	-11.71%		4563%
Class C	-4.42%	-10.78%	-12.04%		4222%
Class E					
Overall	-4.55%	-8.85%	-10.42%		4289%
Class D	-5.36%	-9.89%	-12.97%		4397%

Table 6 AI performance of (16L, 8C, 7H) (SADL, fp32, NNVC-4.0 anchor)

	All Intra Main10				
	BD-rate Over VTM-11.0_nnvc-4.0				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-3.23%	-4.73%	-5.24%		2078%
Class A2	-3.27%	-6.50%	-5.05%		1624%
Class B	-3.20%	-6.42%	-6.40%		1658%
Class C	-3.57%	-6.52%	-6.16%		1261%
Class E	-4.66%	-5.82%	-5.93%		1951%
Overall	-3.54%	-6.07%	-5.85%		1659%
Class D	-4.08%	-6.02%	-6.62%		1306%

Table 7 RA performance of (16L, 8C, 7H) (SADL, fp32, NNVC-4.0 anchor)

	Random Access Main10				
	BD-rate Over VTM-11.0_nnvc-4.0				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-4.00%	-3.93%	-4.48%		2739%
Class A2	-3.94%	-6.55%	-4.33%		2450%
Class B	-3.29%	-7.28%	-6.55%		2882%
Class C	-3.41%	-8.10%	-6.64%		2609%
Class E					
Overall	-3.60%	-6.68%	-5.72%		2689%
Class D	-4.24%	-7.97%	-7.98%		2558%

Experimental results

Table 8 AI performance of (12L, 4C, 10H) (SADL, fp32, NNVC-4.0 anchor)

	All Intra Main10				
	BD-rate Over VTM-11.0_nnvc-4.0				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-3.06%	-2.91%	-3.23%		7053%
Class A2	-3.22%	-3.90%	-2.73%		5485%
Class B	-3.09%	-3.94%	-3.77%		5579%
Class C	-3.29%	-4.39%	-3.63%		4145%
Class E	-4.44%	-3.18%	-3.52%		6637%
Overall	-3.38%	-3.74%	-3.43%		5574%
Class D	-3.67%	-3.79%	-3.74%		4283%

Table 9 RA performance of (12L, 4C, 10H) (SADL, fp32, NNVC-4.0 anchor)

	Random Access Main10				
	BD-rate Over VTM-11.0_nnvc-4.0				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-3.63%	-2.64%	-1.90%		7292%
Class A2	-3.89%	-3.83%	-1.58%		6453%
Class B	-3.22%	-5.51%	-2.88%		7610%
Class C	-3.23%	-5.59%	-2.16%		6928%
Class E					
Overall	-3.44%	-4.62%	-2.23%		7120%
Class D	-3.79%	-5.30%	-2.28%		6898%

Conclusions

- ❑ Propose to further study in the next EE.
- ❑ Two aspects to be studied
 - 1) Explore further complexity reduction based on CP Fused Split Luma Chroma model, aimed at sub 10 KMAC/Pixel
 - 2) SIMD friendly model settings

Q&A

Thanks!

