

EE1-4.1: Neural-Network loop filter with further complexity reduction

JVET-AE0067



Overview of EE1-4.1 reduced complexity NN loop filter

- **New features over NNVC-5.0 LOP loop filter**
 - Split luma-chroma architecture proposed in JVET-AD0157
 - Each channel depth is multiple of 16 for SIMD and HW friendly NN operations
 - Reduced complexity of 9.87 KMAC/pixel
- **Reduced complexity filter tested in NNVC-5.1rc2 in two methods**
 - Test 1 : Proposed loop filter enabled without other NN tools
 - Anchor is NNVC VTM (NN IntraPred and>NNLF LOP disabled)
 - Test 2 : Proposed loop filter enabled with NN IntraPred tool
 - Anchor is NNVC-5.0 (NN IntraPred and>NNLF LOP enabled)
- **Inference crosscheck**
 - Crosscheck report: JVET-AE0183 (Ericsson)

Test 1 BD-Rate over NNVC-VTM Anchor			
Configuration	Y	Cb	Cr
All Intra	-4.14%	-7.90%	-7.64%
Random Access	-4.58%	-10.14%	-8.37%

Test 2 BD-Rate over NNVC-5.0 Anchor			
Configuration	Y	Cb	Cr
All Intra	0.35%	-2.59%	-2.09%
Random Access	0.48%	-3.29%	-2.14%

EE1-4.1 Split Luma-Chroma Architecture

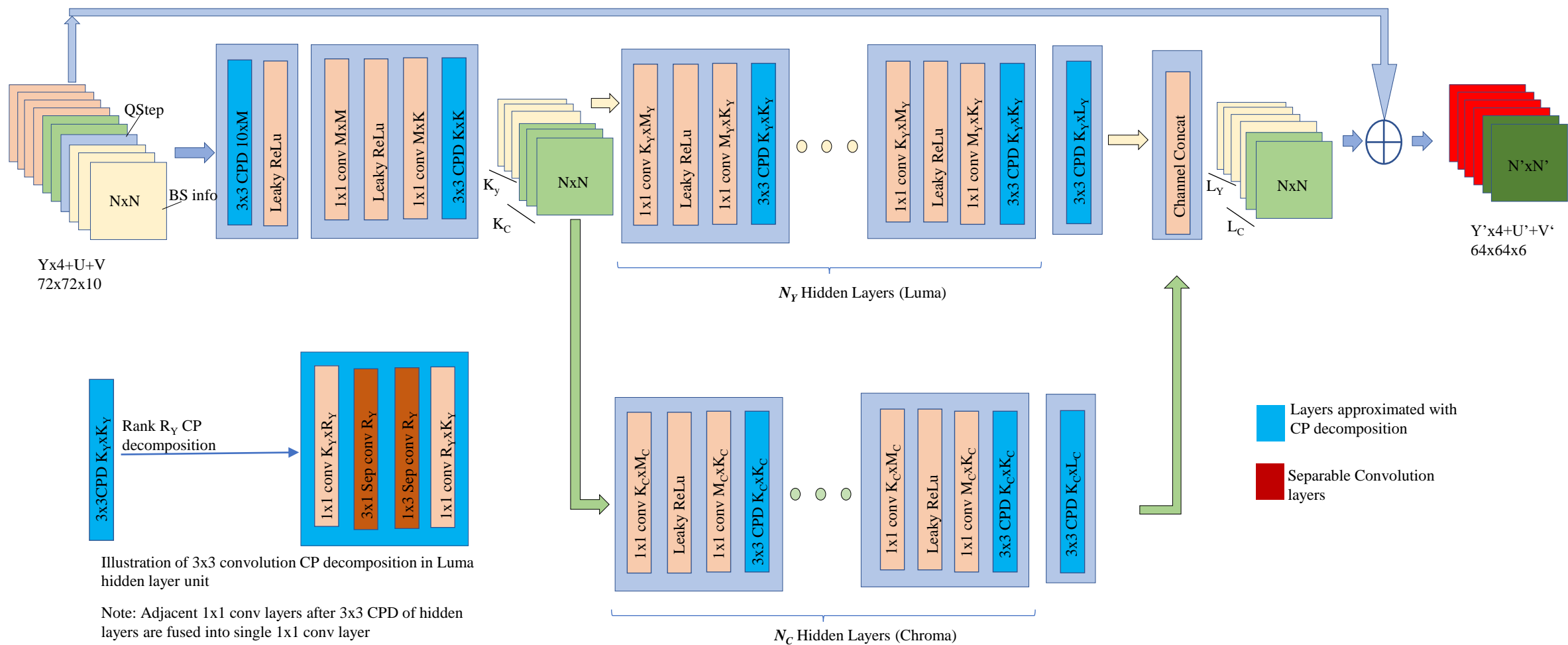


Figure 1 - Split Luma-Chroma reduced complexity loop filter with CPD decomposition and fusion

Model parameters	K	K_Y	K_C	M	M_Y	M_C	R_Y	R_C	N_Y	N_C	Kmac	#Params
Split luma chroma 16Y, 16C	32	16	16	64	48	16	16	16	10	6	9.87	32262

Test 1 Results

Test 1 :

- Loop filter enabled without other NN tools (SADL fp32)
- Anchor is NNVC-5.0 VTM (NN IntraPred and NNLF LOP disabled)
- Encoder run times are not reliable, decoder run times are reliable

Table 1 AI performance of EE1-4.1 Test 1 (SADL, fp32)

	All Intra				
	BD-rate Over VTM-11_nnvc-5.0 (NN tool off)				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-3.81%	-6.09%	-6.68%	153%	5225%
Class A2	-3.84%	-8.17%	-6.17%	147%	4942%
Class B	-3.77%	-8.09%	-8.17%	155%	5143%
Class C	-4.06%	-9.06%	-8.70%	139%	5025%
Class E	-5.49%	-7.55%	-7.74%		
Overall	-4.14%	-7.90%	-7.64%	149%	5087%
Class D	-4.57%	-8.09%	-9.01%	138%	3907%

Table 2 RA performance of EE1-4.1 Test 1 (SADL, fp32)

	Random Access				
	BD-rate Over VTM-11_nnvc-5.0 (NN tool off)				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	-4.79%	-7.73%	-6.55%	134%	2003%
Class A2	-5.00%	-9.73%	-6.47%	143%	2485%
Class B	-4.24%	-10.67%	-9.48%	127%	2300%
Class C	-4.53%	-11.59%	-9.78%	127%	2860%
Class E				138%	2025%
Overall	-4.58%	-10.14%	-8.37%	132%	2340%
Class D	-5.67%	-11.27%	-10.99%	120%	2927%

Test 2 Results

Test 2 :

- Loop filter enabled (SADL fp32) and NN IntraPred enabled
- Anchor is NNVC-5.0 (NN IntraPred and NNLF LOP enabled - SADL int16)
- Encoder run times are not reliable, decoder run times are reliable

Table 3 AI performance of EE1-4.1 Test 2 (SADL, fp32)

	All Intra				
	BD-rate Over NNVC-5.0 (SADL int16)				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	0.61%	-2.75%	-2.06%	101%	54%
Class A2	0.06%	-2.79%	-0.95%	103%	54%
Class B	0.28%	-2.36%	-2.39%	103%	53%
Class C	0.39%	-2.15%	-1.70%	103%	51%
Class E	0.48%	-3.21%	-3.13%	100%	56%
Overall	0.35%	-2.59%	-2.07%	102%	53%
Class D	0.35%	-2.72%	-2.51%	106%	52%

Table 4 RA performance of EE1-4.1 Test 2 (SADL, fp32)

	Random Access				
	BD-rate Over NNVC-5.0 (SADL int16)				
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT
Class A1	0.38%	-4.04%	-2.06%	68%	49%
Class A2	0.35%	-4.38%	-2.15%	68%	50%
Class B	0.47%	-3.18%	-2.81%	74%	49%
Class C	0.66%	-2.05%	-1.35%	81%	48%
Class E					
Overall	0.48%	-3.29%	-2.14%	73%	49%
Class D	0.71%	-2.79%	-2.40%	85%	49%

Conclusion

- Salient features of proposed low complexity NN loop filter
 - Split Luma-Chroma Architecture
 - 16x SIMD friendly convolution depths
- Complexity and performance trade off
 - ~40% reduction in kmac and model size over NNLF LOP filter
 - ~0.5% luma drop with more than 2% coding gain in chroma over NNLF LOP filter
- It is proposed to further study reduced complexity architecture in EE for
 - Training stabilization
 - SADL int16 quantization performance

THANKS TO ERICSSON FOR CROSS-CHECK