



Be Original.

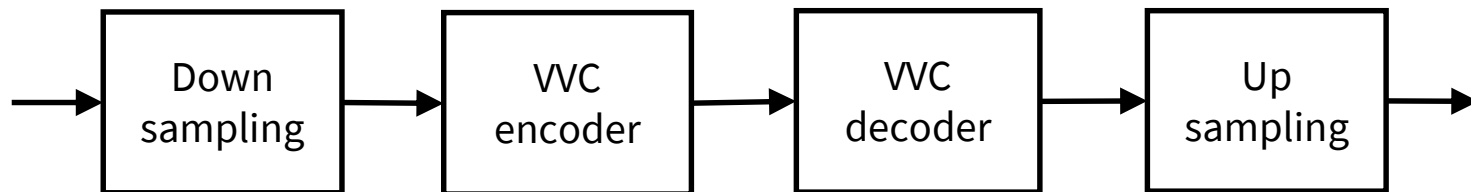
AHG9/AHG11 Level information  
for super-resolution neural network  
JVET-U0053

SHARP CORPORATION

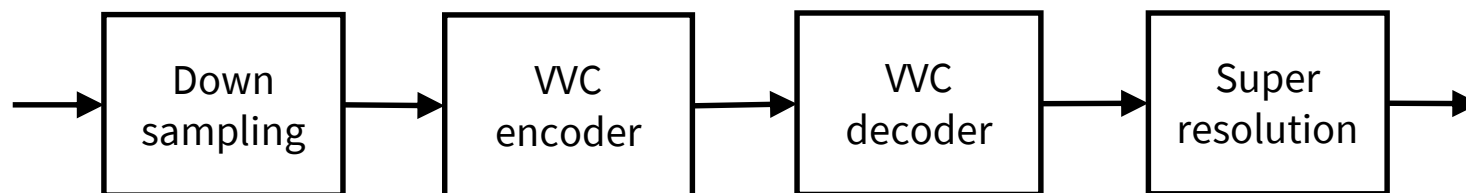
Takeshi Chujoh, Eiich Sasaki, Takuya Suzuki and Tomohiro Ikai

- Level information for a super-resolution neural network
  - At the previous meeting, JVET-T0092 proposed a framework for super-resolution post-filtering using a neural network.
  - Since the complexity of the processing for super-resolution might be too larger on the decoder side, it is necessary to indicate a certain criterion for the complexity.
- The size of compressed neural network parameters as the complexity for the neural networks
  - This is based on the idea of Kolmogorov complexity.

- In the case of high-resolution sequences and low bit-rate, there may be some coding gains because the size of CTB and transform has been limited.



		Random access Main10				
		Over VTM-11.0 (QP 32,37,42,47)				
		Y	U	V	EncT	DecT
Cass A1 4K	Tango2	-9.44%	-3.07%	2.49%	55%	29%
	FoodMarket4	-9.93%	2.52%	2.89%		
	Campfire	-13.09%	15.79%	7.71%		
Class A2 4K	CatRobot1	-0.34%	22.92%	29.23%	48%	29%
	DaylightRoad2	4.40%	3.05%	3.55%		
	ParkRunning3	-8.94%	121.76%	45.88%		



		Random access Main10				
		Over up-sampling filtering (QP 27,32,32,42)				
		Y	U	V	EncT	DecT
Cass A1 4K	Tango2	-0.37%	-4.72%	-2.67%	100%	581%
	FoodMarket4	-0.98%	-0.62%	-2.36%		
	Campfire	-2.65%	-1.78%	-4.27%		
Class A2 4K	CatRobot1	-4.96%	-4.09%	-4.63%	100%	528%
	DaylightRoad2	-5.99%	-3.13%	-2.48%		
	ParkRunning3	-2.44%	-4.07%	4.11%		

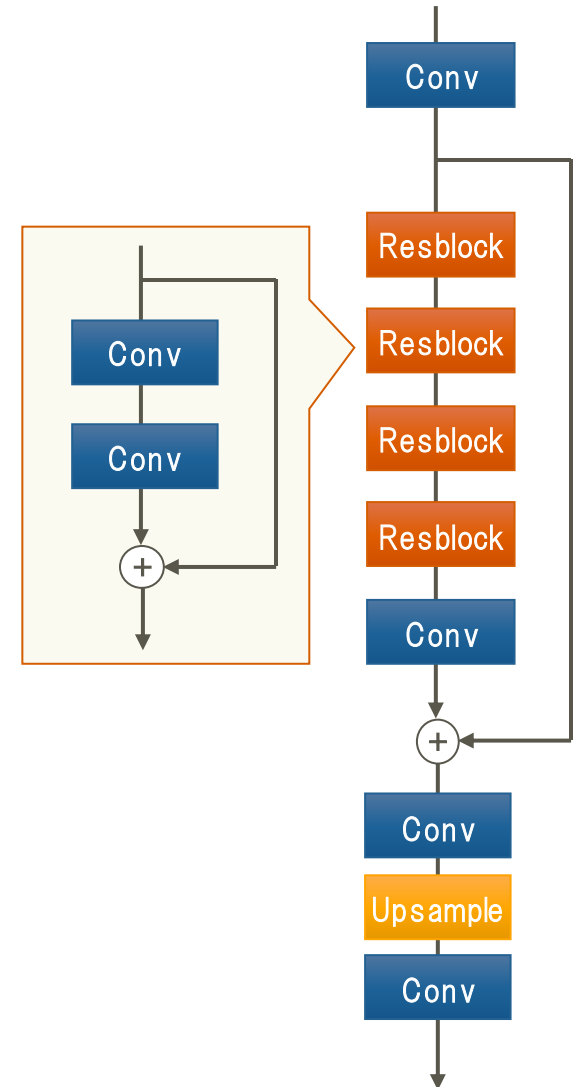
		Random access Main10				
		Over VTM-11.0 (QP 32,37,42,47)				
		Y	U	V	EncT	DecT
Cass A1 4K	Tango2	-9.78%	-7.64%	-0.21%	55%	170%
	FoodMarket4	-10.77%	2.04%	0.55%		
	Campfire	-15.34%	13.60%	3.29%		
Class A2 4K	CatRobot1	-4.63%	18.21%	23.62%	48%	154%
	DaylightRoad2	-0.57%	-0.03%	1.09%		
	ParkRunning3	-11.17%	114.24%	51.05%		

## • Training stage

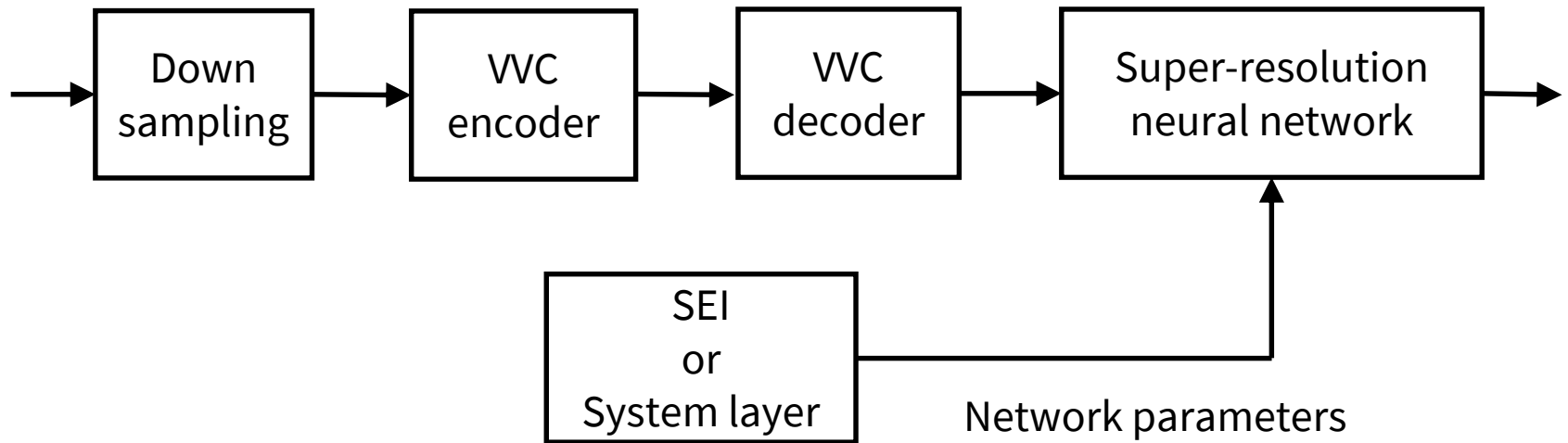
HW environment:	CPU: Intel(R) Core i7-9700K CPU @ 3.60GHz GPU: GeForce RTX 2080Ti
SW environment:	Ubuntu 18.04.4 LTS, CUDA 10.2, Nvidia Driver 440.100
Framework:	PyTorch v1.6.0
Epoch:	8 (100k iteration)
Batch size:	32
Training time:	5.5h
Training data information:	BVI-DVC 4K
Configurations for generating compressed training data	QP values: 22, 27, 32, 37, 42
Patch size	96x96
Learning rate:	2e-4
Optimizer:	ADAM
Loss function:	L1

## • Inference Stage

HW environment:	CPU: Intel(R) Core i7-9700K CPU @ 3.60GHz GPU: GeForce RTX 2080Ti
SW environment:	Ubuntu 18.04.4 LTS, CUDA 10.2, Nvidia Driver 440.100
Framework:	PyTorch v1.6.0
Total Conv. Layers	12
Total FC Layers	0
Total Parameter Number	About 500k

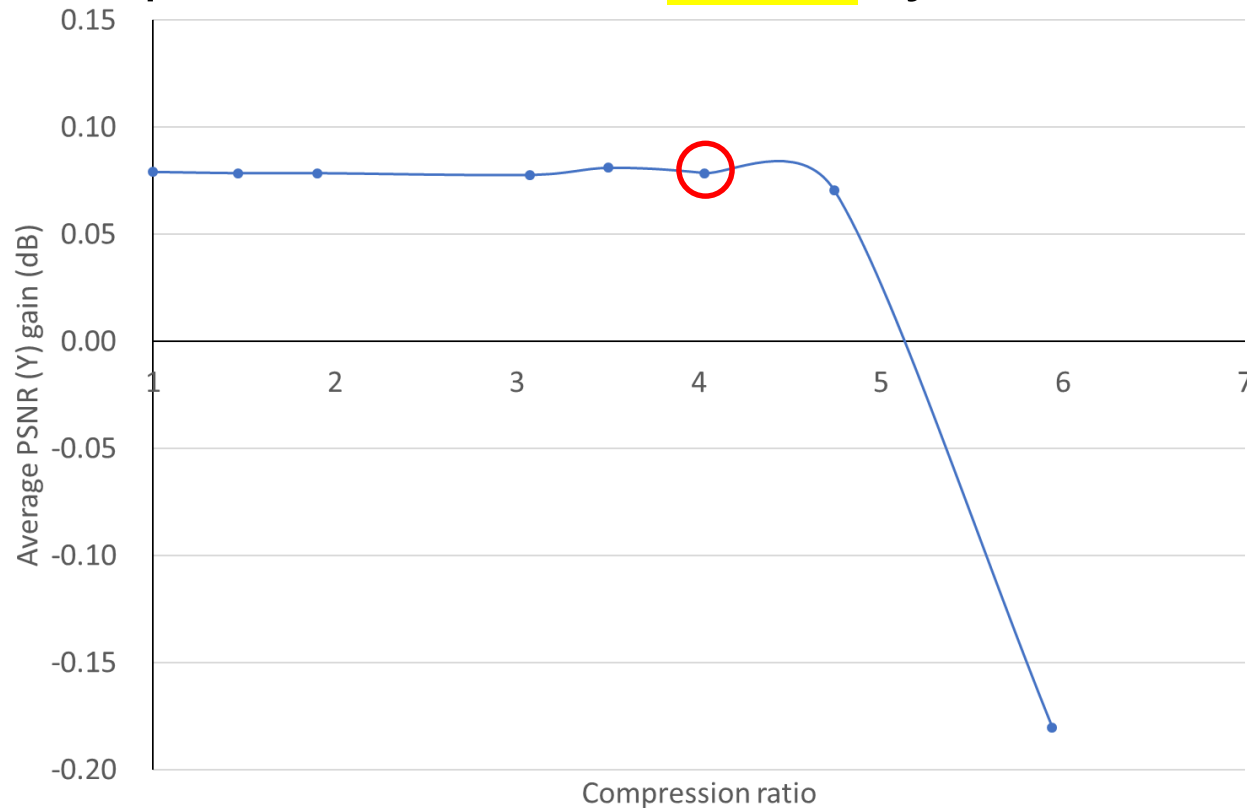


- The complexity of the processing for super-resolution might be too larger on the decoder side.
- A certain criterion for the complexity is needed.



- What is the complexity of neural network?

- The network parameters of the super-resolution have been compressed to about one quarter by MPEG NNR.
  - Original size : 1946523 bytes
  - Compressed size : **482902** bytes (NCTM 5.0, qp=-45)



- Level information based on the size of compressed neural network parameters.

nnr_level_sei_message( payloadSize ) {	Descriptor
<b>nnr_level_idc</b>	u(8)
<b>nnr_parameter_scale</b>	u(8)
}	

NNR_level	nnr_level_idc	Max_nnr_byte
1	16	32768
2	32	65536
2.1	35	74632
3	48	131072
3.1	51	149263
4	64	262144
4.1	67	298527
5	80	524288
5.1	83	597053
5.2	86	679917
6	96	1048576
6.1	99	1194106
6.2	102	1359835

**nnr\_level\_idc** specifies the level information for the neural network processing based on the maximum byte size of neural network representation and the variable NNR\_level which defines the complexity of neural network processing is set.

**nnr\_parameter\_scale** indicates the scaling factor for level information which is used to adjust the bit-depth of image pixel or the chroma format or the operational precision of neural network.



# SHARP

Be Original.

