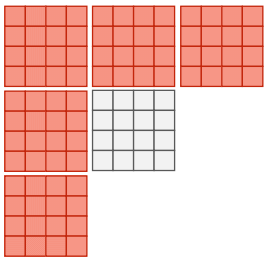
Enhanced Intra Prediction with Recurrent Neural Network in Video Coding

Yueyu Hu[1], Wenhan Yang[1], Sifeng Xia[1], Wen-Huang Cheng[2], and Jiaying Liu[1*]

[1]Institute of Computer Science and Technology, Peking University, Beijing, China
[2]Research Center for Information Technology Innovation (CITI), Academia Sinica, Taiwan

In this paper, an intra prediction scheme based on spatial recurrent neural network (RNN) with blocking referencing is proposed to improve the performance of intra coding in HEVC. The prediction signals are progressively generated to address the asymmetric prediction problem. Local connections in the network make it easier to train compared with fully-connected networks. The embedded spatial RNN performs the progressive signal generation, reducing the interference of the large area on the right bottom which contains little information. To further enhance the modeling capacity, a guiding CNN is designed for robust feature capturing. A multi-scale scheme using is utilized for global feature extraction. We implement the network into HEVC test model (HM) 16.15 and it achieves much gain in BD-Rate, as is shown in Figure 1(b).



(a)

Sequence	BD-Rate		
	Y	U	V
Class B (1080p)	-0.2%	0.4%	0.7%
Class C (WVGA)	-0.2%	0.0%	-0.1%
Class D (WQVGA)	-0.1%	-0.2%	0.0%
Class E (720p)	-0.8%	0.0%	-0.6%
Average	-0.3%	0.1%	0.0%

(b)

Figure 1: (a) Illustration of the blocking referencing scheme (b) Experimental result on CTC, with QP set to $[22, 27, 32, 37]$.

High-resolution images are encoded to generate training context. The common test conditions (CTC) is used for testing. The anchor and proposed method only allow CU size of 16×16 and forced to do a $N \times N$ split. Settings of different QP share the same models. Two different models are utilized for complete context (5 adjacent CUs) and incomplete context(3 adjacent CUs) respectively. Our proposed model brings better BD-Rate performance than HEVC, especially for Class E, which is composed of 720p HD video sequences.

* Corresponding Author. This work was supported by National Natural Science Foundation of China under contract No. 61772043 and the Ministry of Science and Technology of Taiwan under Grants MOST-105-2628-E-001-003-MY3 and MOST-106-3114-E-002-009. We gratefully acknowledge the support of NVIDIA Corporation with the GPU for this research.

2375-0359/18/$31.00 ©2018 IEEE
DOI 10.1109/DCC.2018.00066