Combining Intra Block Copy and Neighboring Samples Using Convolutional Neural Network for Image Coding

Zhaobin Zhang*, Yue Li♭, Li Li*, Zhu Li* and Shan Liu♮

*University of Missouri Kansas City, ♭University of Science and Technology of China, ♮Tencent America

Abstract

We propose an intra prediction method by combining intra block copy and neighboring samples using convolutional neural networks. A novel CNN is developed to further exploit the spatial correlation. Instead of only considering local information, the proposed method can infer the current block via fusing the non-local recurrent features, which is captured by intra block copy, with the local samples located at the left and above boundaries of current block. We also investigate how the performance is affected by the way of fusing IBC and reference boundary pixels. In additional, training data pre-processing is studied to enable the CNN with a better learning capability. Simulation results yield promising coding gain and indicate great potential ability that CNN can be used for future work.

Introduction

- Existing HEVC intra prediction scheme performs well in predicting local, continuous and directional image features, but still not powerful enough to characterize the varied textures in visual content.
- Intra block copy generates the intra prediction by preforming block matching over the reconstructed area and is especially good at predicting non-local, recurrent image patterns.
- Recently, deep leaning has shown superior capability in various tasks and it is expected to achieve better performance in image coding task.
- We proposed a CNN Intra Prediction model (CIP), by combining the power of intra block copy and conventional intra prediction scheme to better characterize the correlation between the reconstructed pixels and current block.

CNN Intra Prediction

The proposed CIP architecture is illustrated in Fig. 1. The framework consists of two parts, fully-connected (FC) layers for reference boundary pixels feature extraction and convolutional neural networks for fusion pixels reconstruction.

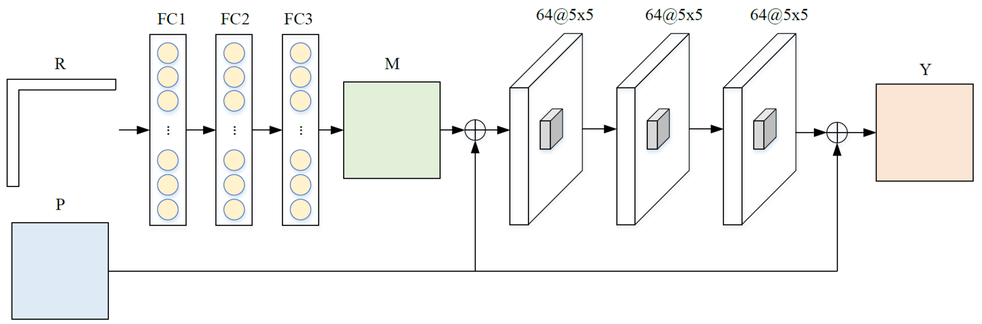

Figure 1. Framework of CNN Intra Prediction.

For current ground truth block Y_0 with size $N \times N$, the proposed model aims at learning a projection from $\{R, P\}$ to Y_0, where R is the $2N + 1$ boundary reference pixels and P is the intra block copy prediction. The input of the FC layers is R and outputs a vector contains N^2 pixel values which will be reshaped to a $N \times N$ block. Element-wise summation is applied on M and intra block copy P followed by convolutional neural networks. Residual learning is adopted for better learning performance.

Integration to HEVC

As shown in Fig. 2, the proposed method is plugged after the existing intra prediction process. Rate-distortion optimization is used to choose the optimal model. A binary flag will be transmitted to indicate whether CIP is adopted.

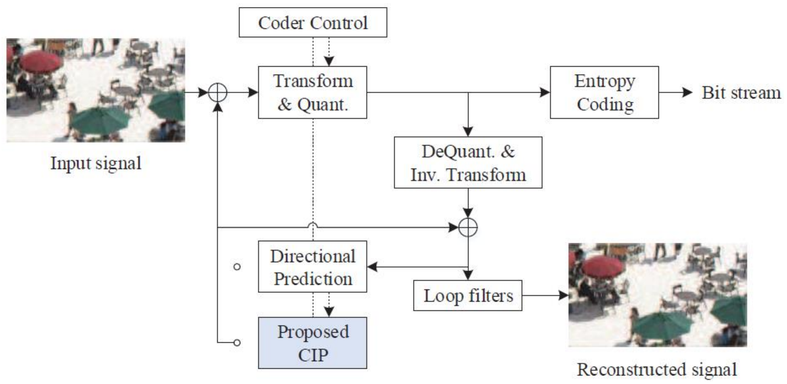

Figure 2. Integration into HEVC.

Results

The experiments are conducted under HEVC common test conditions using All Intra (AI) configuration. Only 16×16 block size is implemented in this paper. The BD-Rate reduction of all the test sequences are listed in Table I.

TABLE I
THE BD-RATE RESULTS OF THE PROPOSED METHOD.

Sequence		BD-Rate		
		Y	U	V
Class A	Traffic	-1.5%	-0.8%	-0.7%
	PeopleOnStreet	-2.1%	-2.8%	-2.6%
	Nebuta	-0.4%	-0.7%	-0.4%
	SteamLocomotive	-0.1%	0.9%	-0.1%
Class B	Kimono	0.3%	-0.2%	-0.3%
	ParkScene	-0.6%	-1.1%	-1.0%
	Cactus	-2.0%	-2.7%	-1.8%
	BQTerrace	-3.6%	-2.5%	-4.5%
	BasketballDrive	-2.8%	-5.7%	-4.2%
Class C	BasketballDrill	-2.8%	-4.5%	-3.6%
	BQMall	-0.8%	0.1%	-0.7%
	PartyScene	-0.7%	0.2%	0.1%
	RaceHorsesC	-0.2%	-0.4%	0.2%
Class D	BasketballPass	-0.7%	-0.7%	-3.4%
	BQSquare	-0.9%	-0.3%	0.1%
	BlowingBubbles	0.2%	-1.0%	-1.0%
	RaceHorses	0.3%	1.3%	0.9%
Class E	FourPeople	-1.3%	0.5%	-0.3%
	Johnny	-2.7%	-2.3%	-3.9%
	KristenAndSara	-2.1%	-4.9%	-2.5%
Class A		-1.1%	-0.8%	-1.0%
Class B		-1.8%	-2.3%	-2.4%
Class C		-1.2%	-1.2%	-1.0%
Class D		-0.2%	-0.2%	-0.8%
Class E		-2.0%	-2.2%	-2.6%
Average		-1.3%	-1.4%	-1.6%
Enc Time		794%		
Dec Time		150%		

Conclusion

This paper proposes a novel CNN-based intra prediction scheme by combining neighboring pixels and intra block copy. To make the proposed CNN more tractable and efficient, training data is refined by empirically setting a threshold to remove distractors. The experimental results demonstrate the effectiveness of the proposed scheme.