

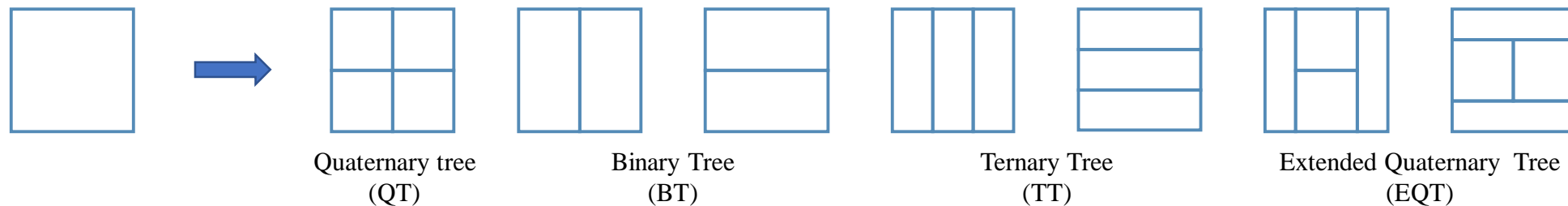
Intra Block Partition Structure Prediction via Convolutional Neural Network

Xu Han, Shanshe Wang, Yong Chen, Siwei Ma, Wen Gao

IEEE Data Compression Conference, 2021

Background

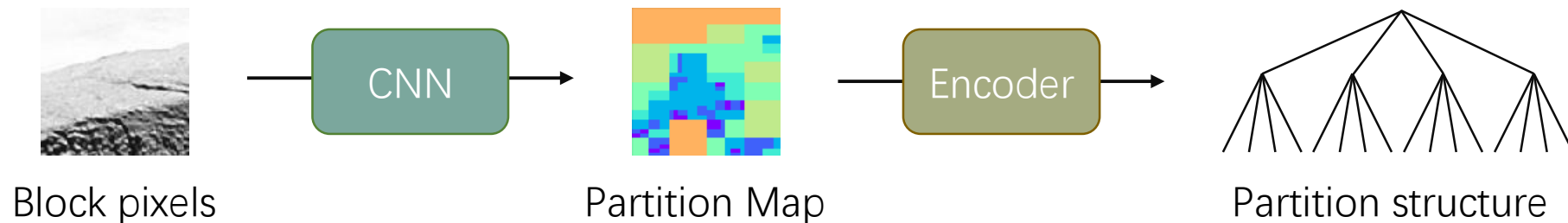
- Block partition decision
 - The most time-consuming process in video coding
 - Multiple partition mode are introduced (QT, BT, TT, EQT)
 - Reducing decision complexity is essential



- The drawback of previous methods
 - Limited acceleration with slightly performance loss
 - Affected by each other when multiple such methods are integrated

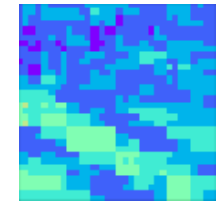
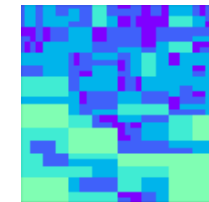
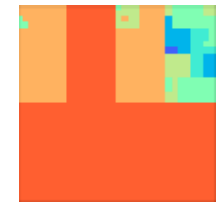
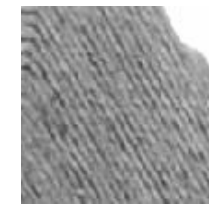
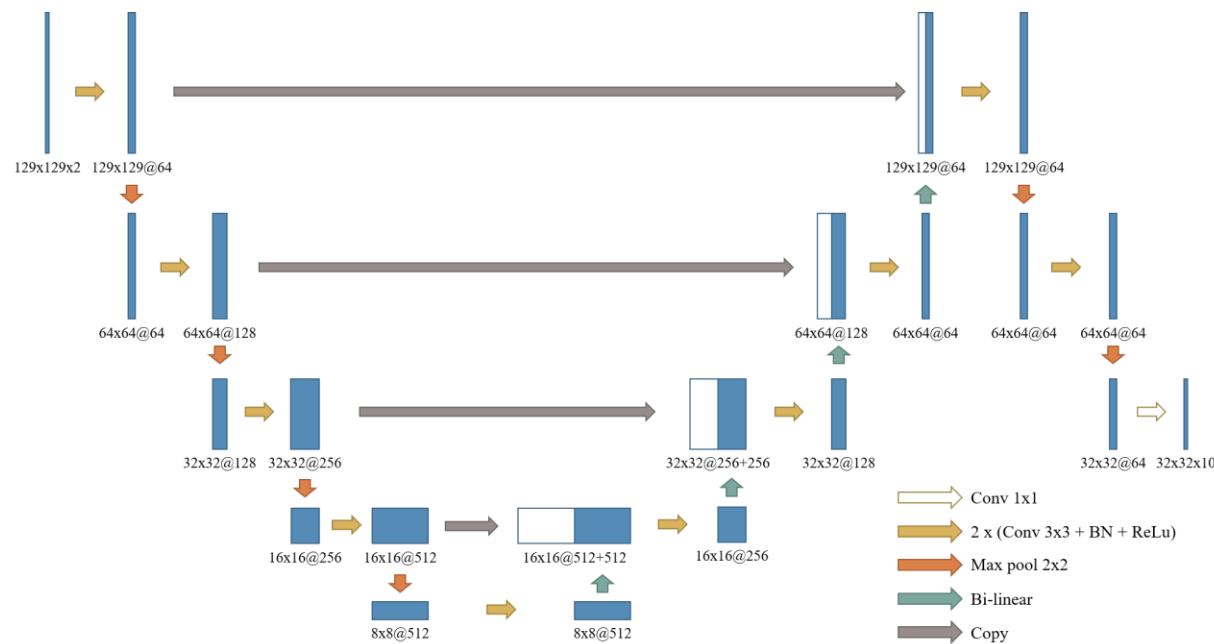
Method: Overview

- CNN-based partition structure prediction
 - It is difficult to predict the final coding unit tree directly
 - Predict partition map first
 - Derive the final partition structure in encoder



Method

- Stage 1: Partition map prediction
 - U-Net structure is adopted
 - Predict according to texture, adjacent edges and QP map
 - Predict optimal width and height respectively



Input

Ground Truth

Prediction

Method

- Stage 2: Partition structure searching
 - Similar to the RDO of block size
 - Use mean absolute error (MAE) between partition maps as distortion

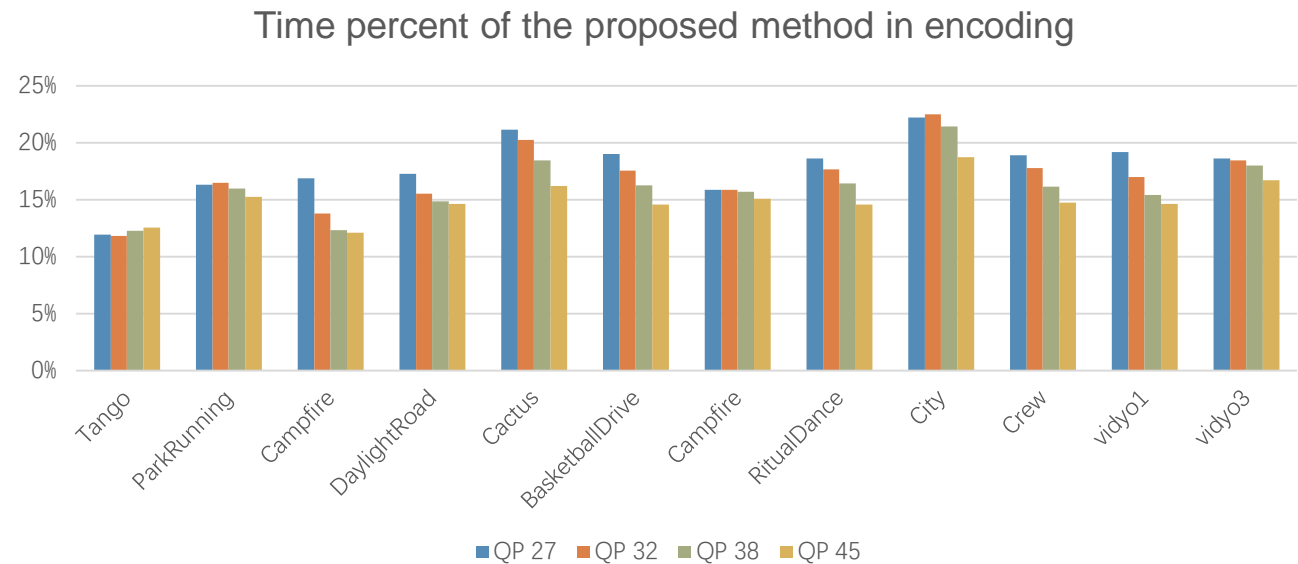
$$\mathcal{L} = \sum_{i,j} \text{abs}(w_{ij} - w_c) + \text{abs}(h_{ij} - h_c)$$

w_{ij}, h_{ij} : predicted width and height in partition maps
 w_c, h_c : width and height of current block

Experiments

- The proposed algorithm is implemented on HPM 4.0
 - About 97% time saving (56 times speedup) with 9% BD-rate loss
 - The overhead of the method is 16% on average
 - Fully parallel and no reconstructed pixels required

Class	Y	U	V	Speedup
4K	9.77%	19.75%	18.27%	43.9
1080P	8.65%	13.60%	15.43%	53.8
720P	9.28%	13.84%	13.90%	59.2
Average	8.97%	13.72%	14.67%	56.5



Conclusion

- A CNN-based partition modes prediction method is proposed to accelerate the intra coding of AVS3
- 97% time saving is achieved with 9% BD-rate loss
- In future, we will explore the method to decrease the performance loss and generalize the method to inter frame coding.

Thanks for watching