ECM-OPCC: EFFICIENT CONTEXT MODEL FOR OCTREE-BASED POINT CLOUD COMPRESSION

Yiqi Jin[1,2], Ziyu Zhu[3], Tongda Xu[1], Yuhuan Lin[1,2], Yan Wang[1]

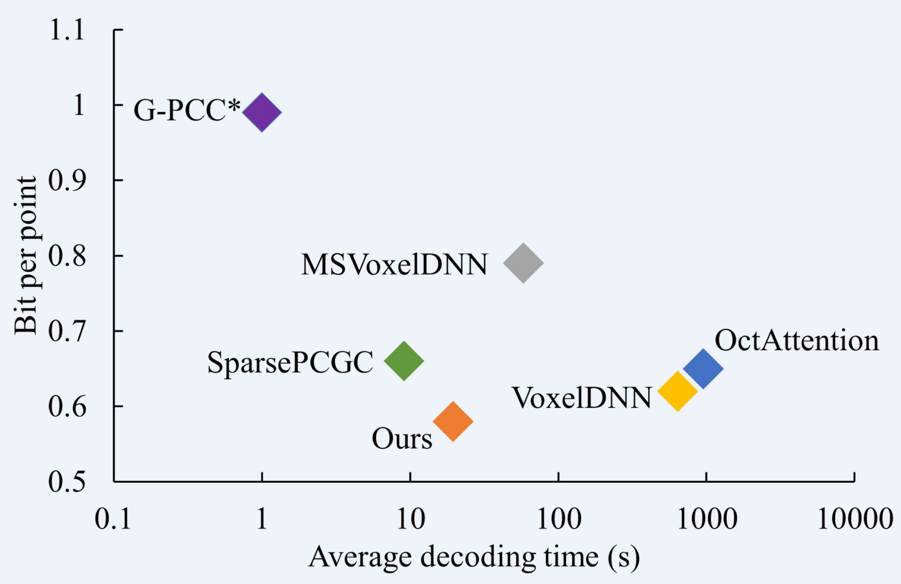
[1]Institute for AI Industry Research (AIR), [2]Dept. of Automation, [3]Dept. of Computer Science, Tsinghua University

1 Introduction

A major issue of octree-based methods is the decoding complexity due to the autoregressive context model, which hinders the practical deployment of such methods.

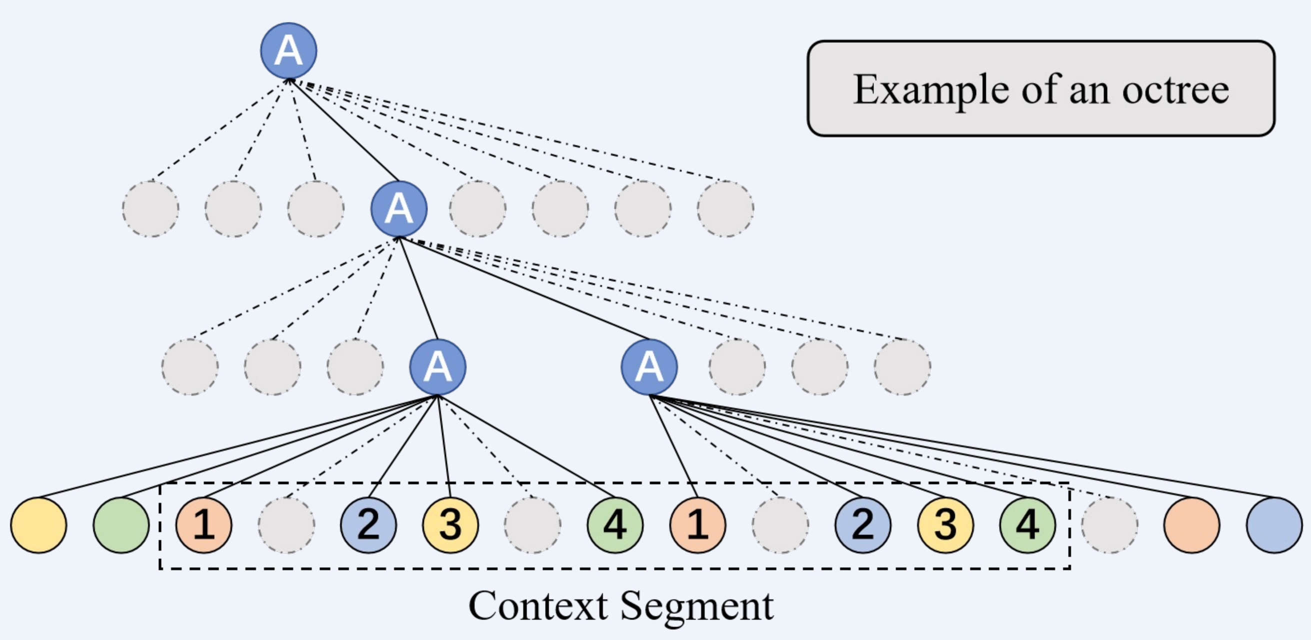
Inspired by parallelizable autoregressive model in density estimation and image compression, we propose to solve DPCC problem by multi-group coding strategy. Specifically:

- We propose a new **segment-constrained multi-group coding strategy** that enables parallel decoding of nodes inside each group, accelerating decoding process while maintaining compression performance;
- We propose a novel **dual transformer architecture with level-parallel and group-parallel branch** to support our coding strategy and better extract context information from ancestors and siblings;
- Results show that our model achieves competitive compression performance and significantly reduces decoding time, which makes the practical deployment of octree-based DPCC possible.



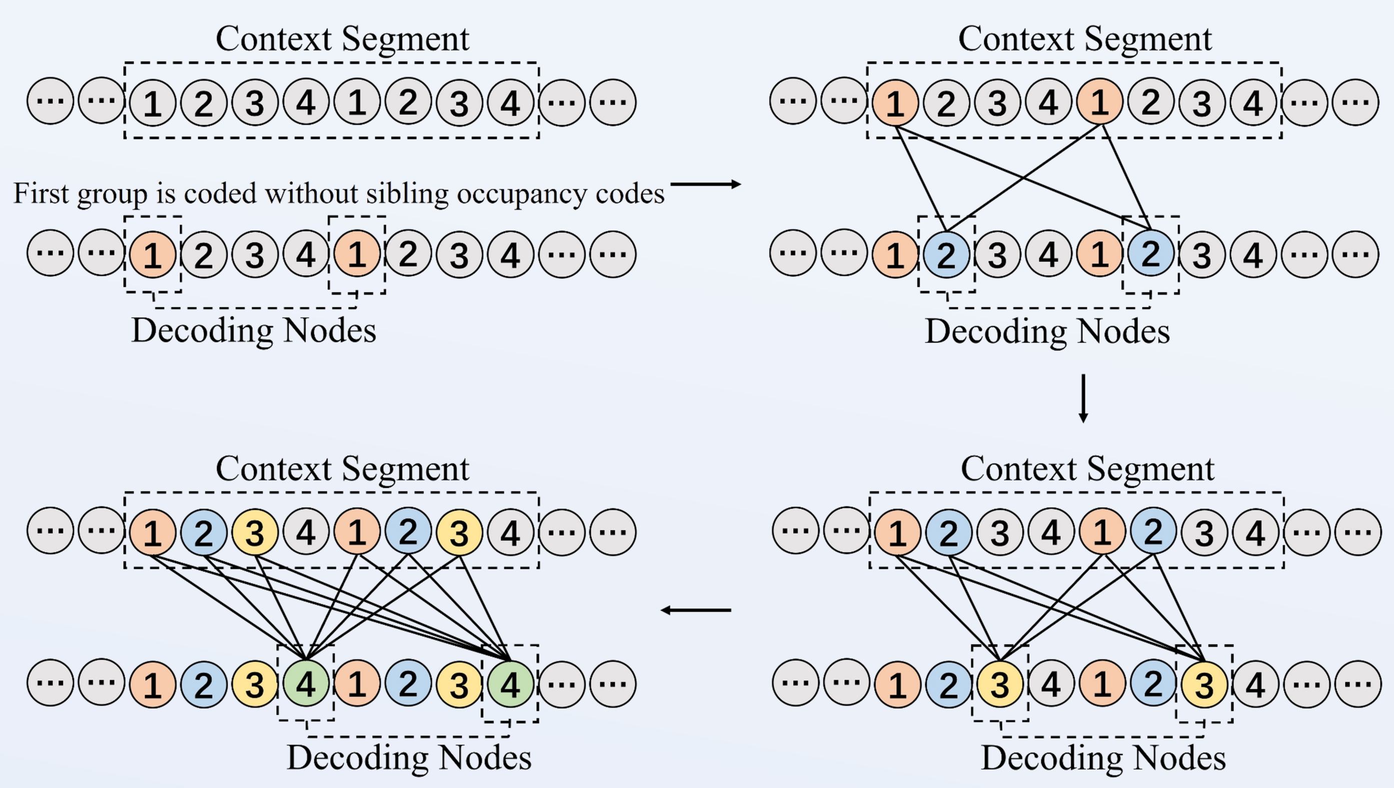
- The bpp-decoding time of different methods.

2 Methodology



Multi-Group Coding Strategy:

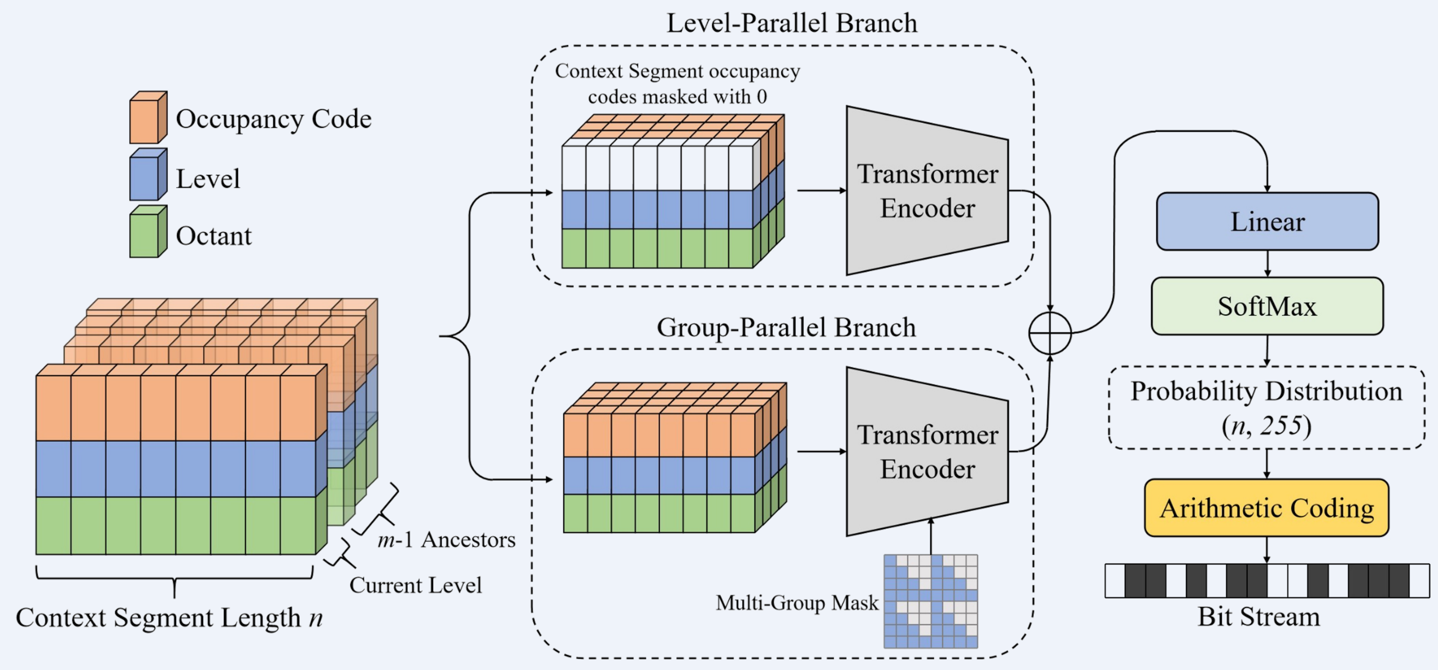
- We divide the nodes in each layer into **parallel coding context segments**. Then we group the nodes in each context segment and autoregressively code these groups. The **nodes in each group are coded in parallel**, and the occupancy codes of previous groups serve as context for latter groups.



Dual Transformer Structure:

- We design a dual branch transformer structure to support our coding strategy. The **level-parallel branch** is designed to fully use all available information from previous layers. The occupancy codes of current encoding/decoding level are masked with 0 to prevent information leakage. The **group-parallel branch** leverages context information carried by occupancy code of sibling nodes. We use a multi-group mask matrix to modulate the attention layer.

2



- Dual transformer structure of our model

3 Experiments

We train and test our method on **LiDAR** and **object** dataset for comprehensive evaluation.

Training Datasets:

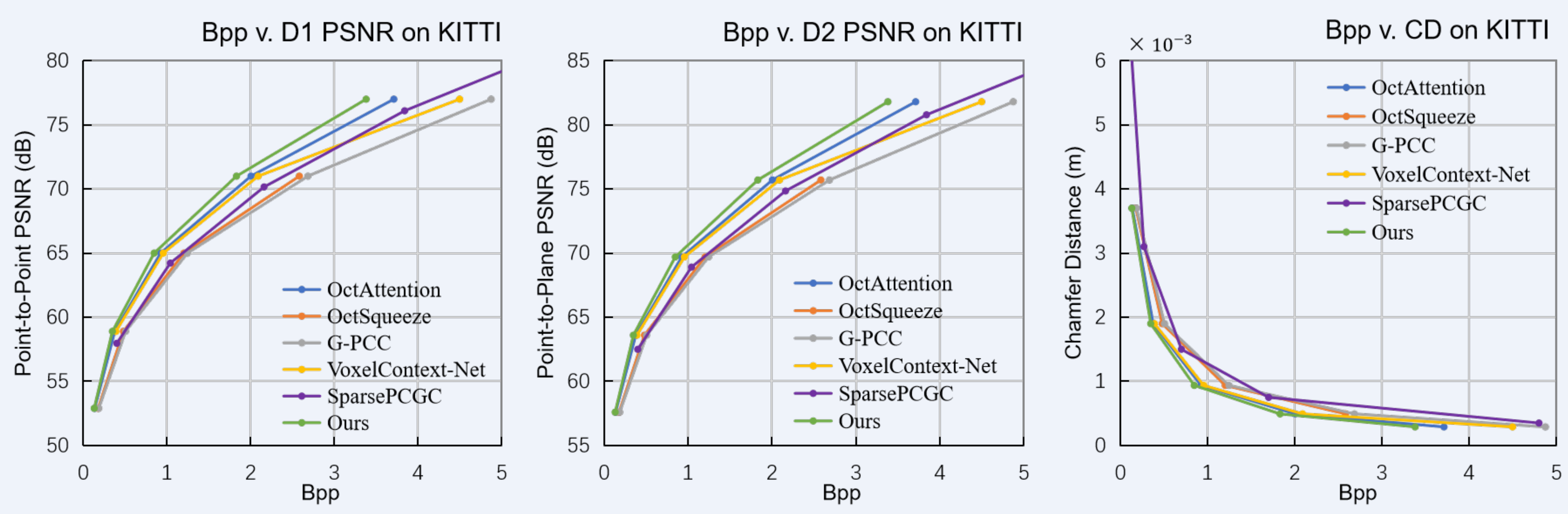
- **Lidar**: SemanticKITTI (sequence 00 - 10)
- **Object**: MPEG 8i (Soldier10, Longdress10), MVUB (Andrew10, David10, Sarah10)

Evaluation Datasets:

- **Lidar**: SemanticKITTI (sequence 11 - 21)
- **Object**: MPEG 8i (Loot10, RedandBlack10, Boxer10, Thaidancer10), MVUB (Phil, Ricardo)

Results:

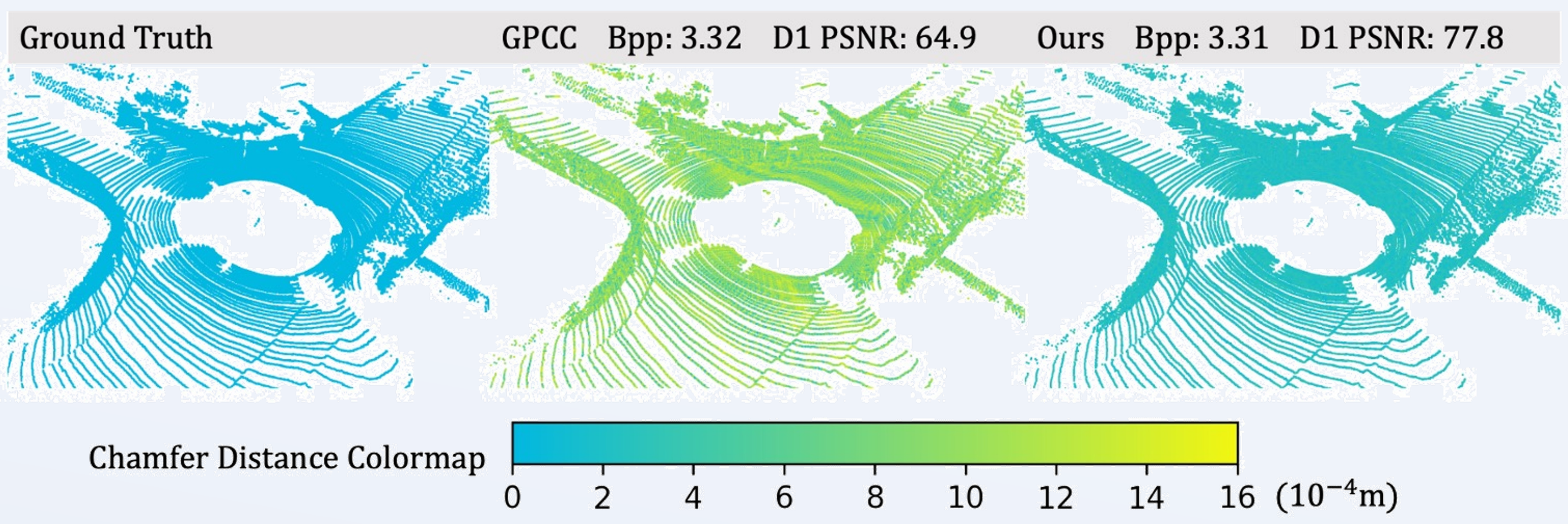
- Results of different methods on SemanticKITTI:



- Bpp and coding time results on MPEG 8i dataset compared with G-PCC, voxel-based and octree-based methods:

Point Cloud	Traditional	Voxel-based			Octree-based	
	G-PCC	SparsePCGC	VoxelDNN	MSVoxelDNN	OctAttention	Ours
loot_vox10 (bpp)	0.95	0.63	0.58	0.73	0.62	0.55
redandblack_vox10 (bpp)	1.09	0.72	0.66	0.87	0.73	0.66
boxer_viewdep_vox10 (bpp)	0.94	0.60	0.55	0.70	0.59	0.51
Thaidancer_viewdep_vox10 (bpp)	0.99	0.67	0.68	0.85	0.65	0.58
Average bpp	0.99	0.66	0.62	0.79	0.65	0.58
Average Gain over G-PCC	-	33.8%	37.6%	20.5%	34.6%	41.9%
Average Encoding Time (s)	4.0	9.5	885	54	0.80	1.92
Average Decoding Time (s)	1.0	9.1	640	58	948	19.5

- Visualized chamfer distance between normalized ground truth and reconstructed point cloud:



4 Key References

1. Lila Huang, Shenlong Wang, Kelvin Wong, Jerry Liu, and Raquel Urtasun, "Octsqueeze: Octree-structured entropy model for lidar compression," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.
2. Zizheng Que, Guo Lu, and Dong Xu, "Voxelcontextnet: An octree based framework for point cloud compression," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 6042–6051.
3. Chunyang Fu, Ge Li, Rui Song, Wei Gao, and Shan Liu, "Octattention: Octree-based large-scale contexts model for point cloud compression," arXiv preprint arXiv:2202.06028, 2022.
4. Jianqiang Wang, Dandan Ding, Zhu Li, Xiaoqing Feng, Chuntong Cao, and Zhan Ma, "Sparse tensor-based multiscale representation for point cloud geometry compression," arXiv preprint arXiv:2111.10633, 2021.