Data Compression Conference

南京大學
NANJING UNIVERSITY

Lossless Point Cloud Attribute Compression Using Cross-scale, Cross-group, and Cross-color Prediction

Jianqiang Wang[1], Dandan Ding[2], Zhan Ma[1]
[1]Nanjing University, [2]Hangzhou Normal University

Introduction

- **Topic**
  - **Lossless Point Cloud Attribute Compression (PCAC)**

- **Method**
  - Multiscale Sparse Tensor based hierarchical structure
  - Neural network-based attribute probability prediction across layers
  - **Cross-scale, Cross-group, and Cross-color Prediction**

- **Contributions**
  - The first lightweight and generalized lossless PCAC approach that outperforms MPEG G-PCC.

Related Works: Point Cloud Attribute Compression (PCAC)

☐ Rules-based solutions

- Region-Adaptive Hierarchical Transform (RAHT)
- Predicting and Lifting Transforms
- Graph Fourier Transform

Rules-based solutions (**MPEG G-PCC**) has state-of-the-art PCAC efficiency.

☐ Learning-based approaches

- SparsePCAC[1], DeepPCAC[2], 3DAC[3],[4],[5] etc…

Most existing learning-based methods only support lossy coding, and are still inferior to MPEG G-PCC.

[1]. Wang, J., & Ma, Z. (2022). Sparse Tensor-based Point Cloud Attribute Compression. 2022 MIPR, 59-64.

[2]. Sheng, X., Li, L., Liu, D., Xiong, Z., Li, Z., & Wu, F. (2021). Deep-PCAC: An End-to-End Deep Lossy Compression Framework for Point Cloud Attributes. IEEE TMM, 24, 2617-2632.

[3]. Fang, G., Hu, Q., Wang, H., Xu, Y., & Guo, Y. (2022). 3DAC: Learning Attribute Compression for Point Clouds. 2022 IEEE/CVF CVPR, 14799-14808.

[4] M. Quach, G. Valenzise, and F. Dufaux, “Folding-based compression of point cloud attributes,” in IEEE ICIP, pp. 3309–3313, 2020.

[5] E. Alexiou, K. Tung, and T. Ebrahimi, “Towards neural network approaches for point cloud compression,” in Applications of digital image processing XLIII, vol. 11510, SPIE, 2020.

Related Works: Point Cloud Geometry Compression (PCGC)

□ **SparsePCGC [1]**

- Learning-based solution.
- Sparse tensor-based multiscale representation.
- State-of-the-art geometry compression performance.

This work extends the multiscale structure in SparsePCGC[1] to support PCAC by exhaustively exploiting cross-scale, cross-group, and cross-color correlations

1. Wang, J., Ding, D., Li, Z., Feng, X., Cao, C., & Ma, Z. (2021). Sparse Tensor-based Multiscale Representation for Point Cloud Geometry Compression. IEEE transactions on pattern analysis and machine intelligence.

Method: Outline

Core idea: Context-based entropy model.

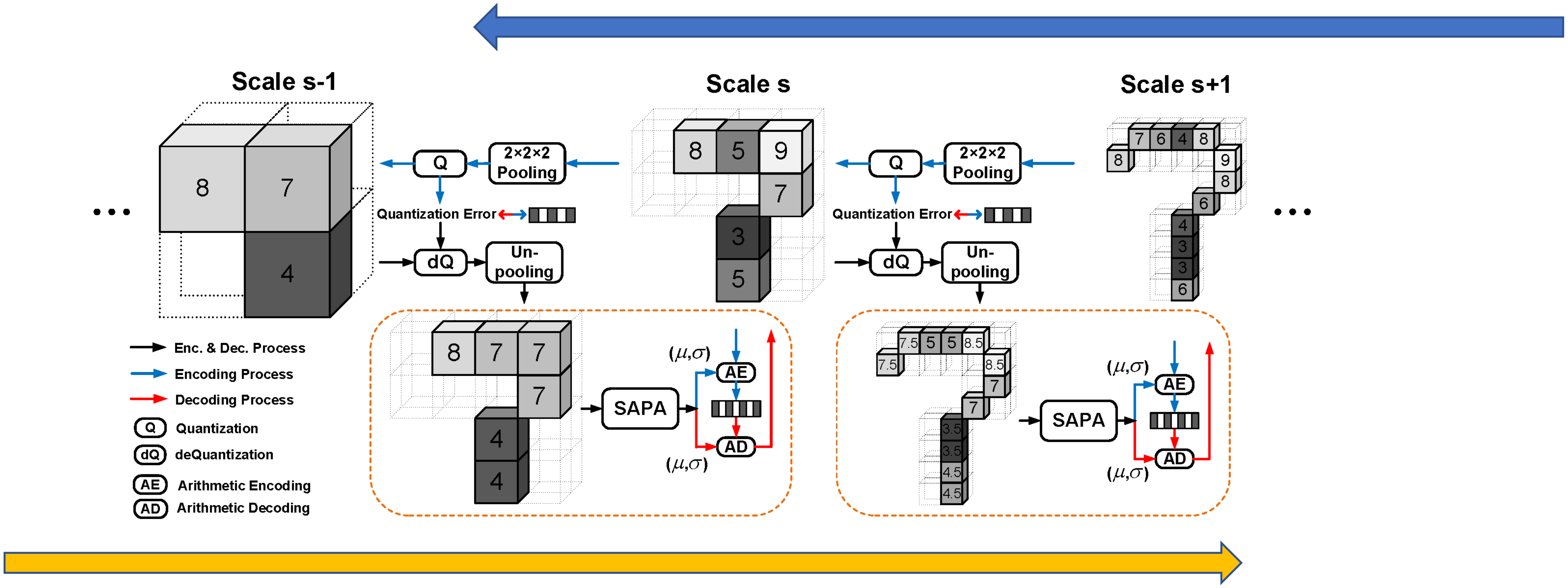
- ☐ **Cross-scale Prediction**
  - Multiscale Sparse Tensor based hierarchical structure
  - Neural Network based prediction and entropy model

- ☐ **Cross-group Prediction**

- ☐ **Cross-color Prediction**

Method: Construction of Multiscale Structure

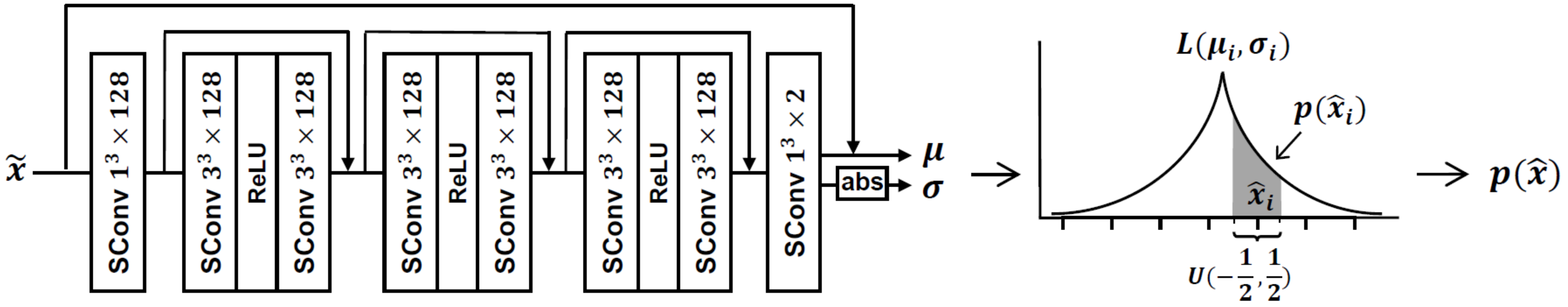
Progressive downscaling (average pooling) and quantization



Scale-wise prediction and entropy coding using neural network

Method: Neural Network based Prediction

- Sparse CNN-based Attribute Probability Approximation (SAPA) model

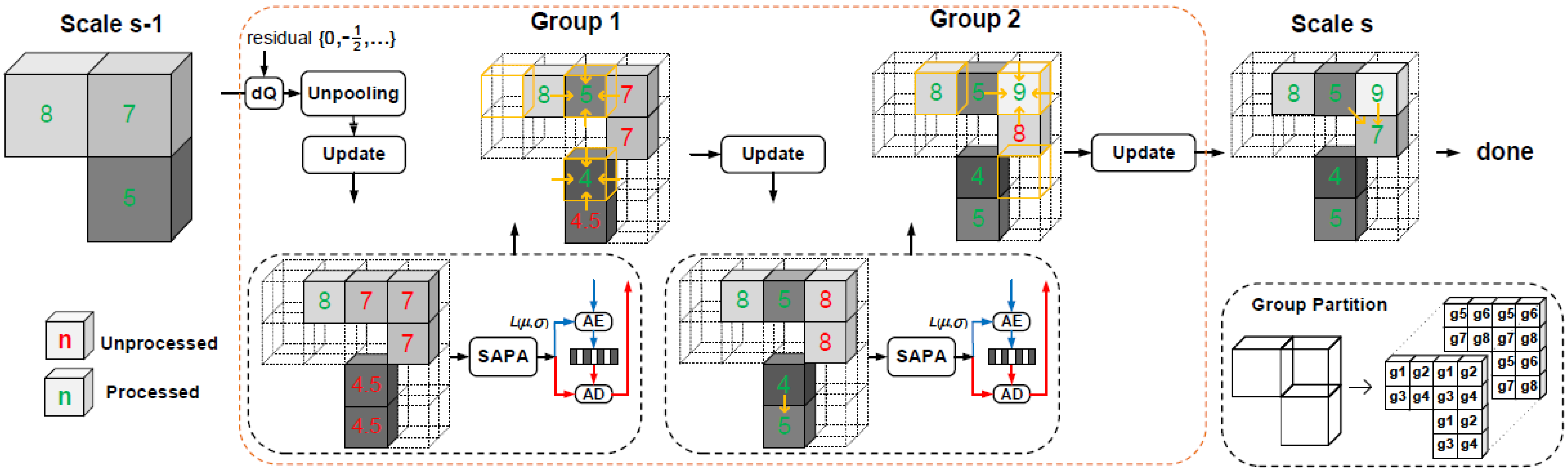


- Conditional Entropy Model

$$p(\{\hat{x}_i^{(s)}\}) = \prod_i \Big(\mathcal{L}(\mu_i, \sigma_i) * \mathcal{U}(-\tfrac{1}{2}, \tfrac{1}{2})\Big)(\hat{x}_i^{(s)}) \quad \text{with } \mu_i, \sigma_i = \mathrm{SAPA}(\{\tilde{x}_i^{(s)}\}),$$

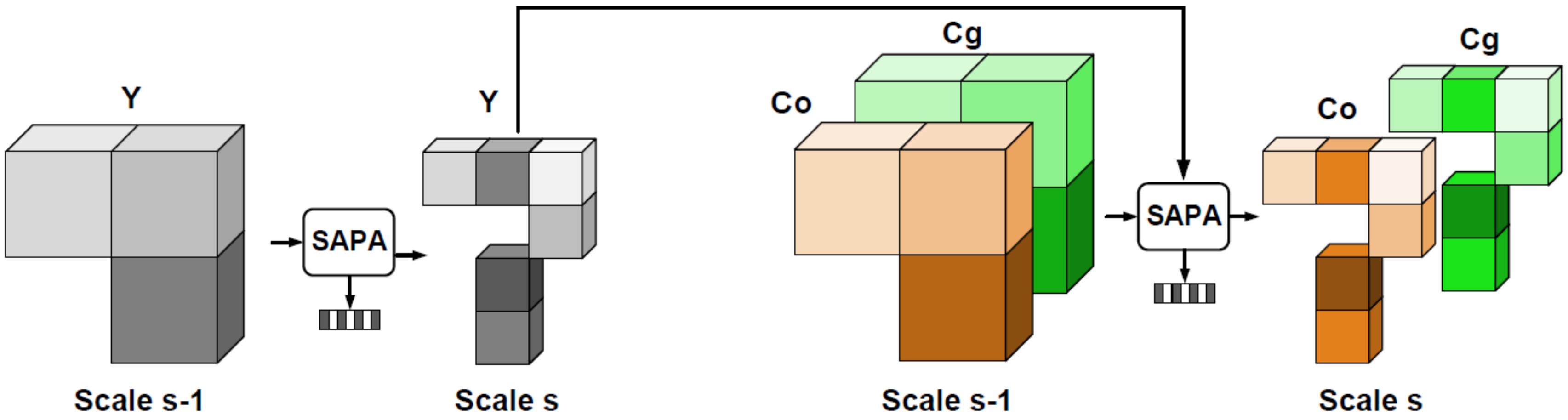
Method: Cross-scale Prediction

- Group Partition
- Predict
- Update



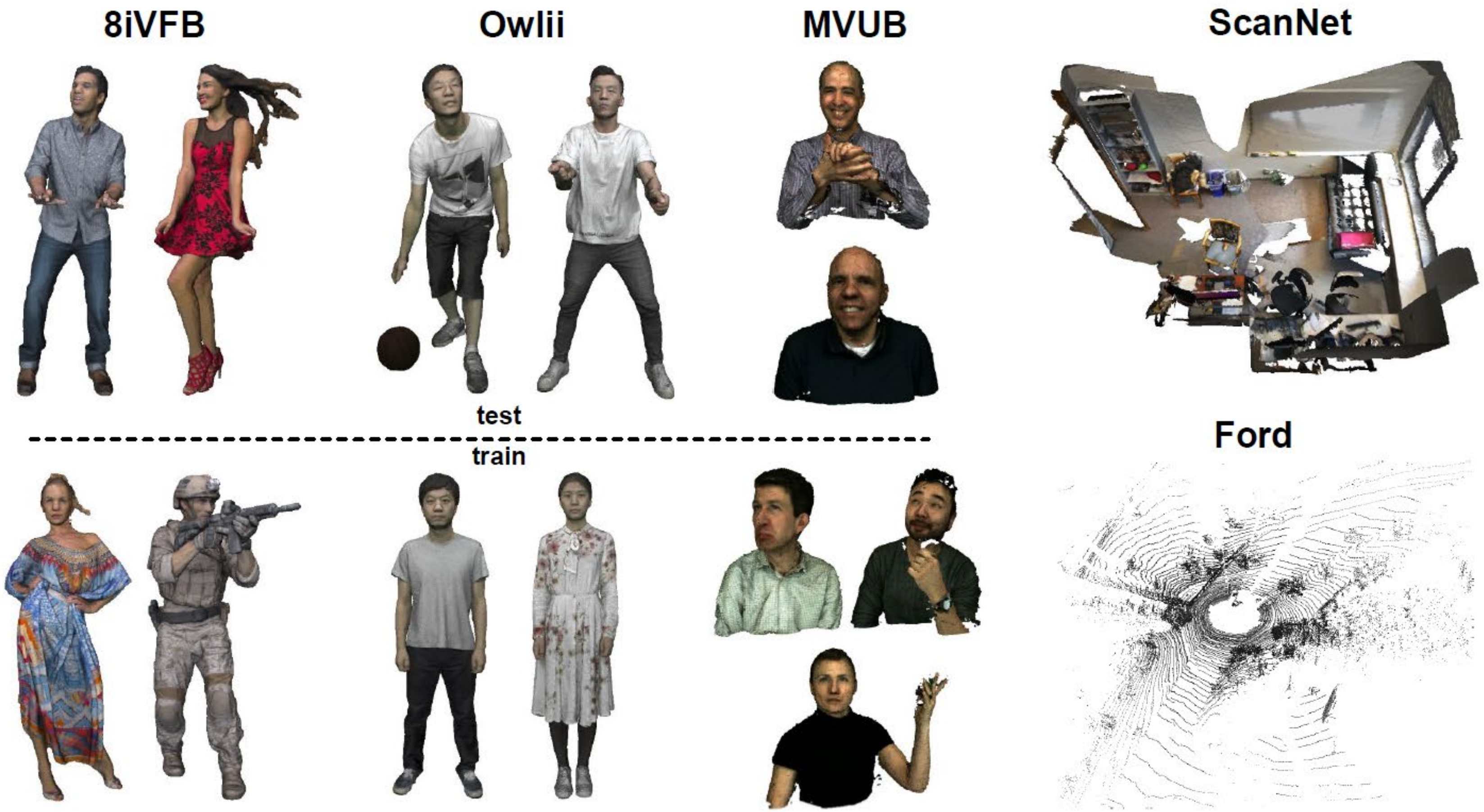
$$\mathrm{U}(\hat{x}_{init}^{(s,m)}) = \frac{k \times x^{(s-1)} - (\hat{x}^{(s,1)} + \hat{x}^{(s,2)} + \cdots + \hat{x}^{(s,g)})}{k-g} \quad \text{with } g+1 \leq m \leq 8$$

Method: Cross-color Prediction



Experimental Results: Datasets

- **Human Bodies**
  - 8iVFB[1]
  - Owlii[2]
  - MVUB[3]
- **ScanNet[4]**
- **Ford[5]**



[1] E. d'Eon, B. Harrison, T. Myers, and P. A. Chou, "8i voxelized full bodies - a voxelized point cloud dataset," ISO/IEC JTC1/SC29 Joint WG11/WG1 (MPEG/JPEG) WG11M40059/WG1M74006, 2017.
[2] Y. Xu, Y. Lu, and Z. Wen, "Owlii dynamic human mesh sequence dataset," ISO/IEC JTC1/SC29/WG11 (MPEG/JPEG) m41658, 2017.
[3] L. Charles, C. Qin, O. Sergio, and A. C. Philip, "Microsoft voxelized upper bodies – a voxelized point cloud dataset," ISO/IEC MPEG m38673, May 2016.
[4] A. Dai, A. X. Chang, M. Savva, et al., "ScanNet: Richly-annotated 3d reconstructions of indoor scenes," 2017 IEEE CVPR, pp. 2432–2443, 2017.
[5] WG 7, MPEG 3D Graphics Coding, "Common test conditions for G-PCC," ISO/IEC JTC1/SC29/WG11 N00106, 2021.

Experimental Results: Compression Gains over G-PCC

- **15.2%, 12.5%,** and **36.7%** bitrate reduction on 8iVFB, Owlii, and MVUB
- **13.2%** and **9.7%** gains on ScanNet with 5cm and 2cm precision.
- **6.0%** and **4.7%** gains on Ford with 2cm and 1mm precision

Table 1: Evaluation of compression efficiency and computational complexity.

PCs		G-PCC	Ours CS		Ours CS+CG		Ours CS+CG+CC	
		bpp	bpp	gain	bpp	gain	bpp	gain
8iVFB vox10	loot	6.19	6.26	1.1%	5.19	-16.1%	5.18	-16.4%
	red&black	9.39	10.20	8.6%	8.15	-13.2%	8.07	-14.1%
	average	7.79	8.23	4.8%	6.67	**-14.7%**	6.62	**-15.2%**
Owlii vox11	player	7.72	8.34	8.0%	7.13	-7.6%	6.78	-12.2%
	dancer	7.80	8.33	6.9%	7.11	-8.8%	6.80	-12.8%
	average	7.76	8.34	7.4%	7.12	**-8.2%**	6.79	**-12.5%**
MVUB vox10	Phil	10.27	10.13	-1.4%	7.33	-28.6%	6.78	-34.0%
	Ricardo	5.92	5.12	-13.6%	3.68	-37.9%	3.59	-39.4%
	average	8.10	7.62	-7.5%	5.50	**-33.3%**	5.19	**-36.7%**
ScanNet	**q5cm**	12.92	14.13	9.3%	11.47	**-11.2%**	11.21	**-13.2%**
	q2cm	13.13	15.04	14.6%	12.04	**-8.3%**	11.86	**-9.7%**
Ford	**q2cm**	5.32	7.05	32.5%	5.00	**-6.0%**	-	-
	q1mm	5.22	6.93	32.9%	4.97	**-4.7%**	-	-
Average Time (Eocoding Decoding) (s/frame)								
8iVFB_vox10		9.5	9.3	5.7	5.1	10.1	9.8	15.7	16.0
Owlii_vox11		32.8	32.0	17.5	15.3	37.0	35.9	56.0	58.1
MVUB_vox10		17.1	17.0	10.2	9.2	19.0	18.3	27.3	28.3
ScanNet_q2cm		2.0	2.0	1.2	1.2	3.6	3.5	6.3	6.3
Ford_q1mm		1.1	1.1	0.8	0.8	8.0	8.0	-	-

Experimental Results: Runtime Comparison

- Same level of encoding / decoding time.
- Tested on RTX 3090 GPU.

Because G-PCC and our method run on different platforms, these numbers are served as the reference for intuitive understanding.

Table 1: Evaluation of compression efficiency and computational complexity.

PCs		G-PCC		Ours CS		Ours CS+CG		Ours CS+CG+CC	
		bpp		bpp	gain	bpp	gain	bpp	gain
8iVFB vox10	loot	6.19		6.26	1.1%	5.19	-16.1%	5.18	-16.4%
	red&black	9.39		10.20	8.6%	8.15	-13.2%	8.07	-14.1%
	average	7.79		8.23	4.8%	6.67	**-14.7%**	6.62	**-15.2%**
Owlii vox11	player	7.72		8.34	8.0%	7.13	-7.6%	6.78	-12.2%
	dancer	7.80		8.33	6.9%	7.11	-8.8%	6.80	-12.8%
	average	7.76		8.34	7.4%	7.12	**-8.2%**	6.79	**-12.5%**
MVUB vox10	Phil	10.27		10.13	-1.4%	7.33	-28.6%	6.78	-34.0%
	Ricardo	5.92		5.12	-13.6%	3.68	-37.9%	3.59	-39.4%
	average	8.10		7.62	-7.5%	5.50	**-33.3%**	5.19	**-36.7%**
ScanNet	**q5cm**	12.92		14.13	9.3%	11.47	**-11.2%**	11.21	**-13.2%**
	q2cm	13.13		15.04	14.6%	12.04	**-8.3%**	11.86	**-9.7%**
Ford	**q2cm**	5.32		7.05	32.5%	5.00	**-6.0%**	-	-
	q1mm	5.22		6.93	32.9%	4.97	**-4.7%**	-	-
Average Time (Eocoding Decoding) (s/frame)									
8iVFB_vox10		9.5	9.3	5.7	5.1	10.1	9.8	15.7	16.0
Owlii_vox11		32.8	32.0	17.5	15.3	37.0	35.9	56.0	58.1
MVUB_vox10		17.1	17.0	10.2	9.2	19.0	18.3	27.3	28.3
ScanNet_q2cm		2.0	2.0	1.2	1.2	3.6	3.5	6.3	6.3
Ford_q1mm		1.1	1.1	0.8	0.8	8.0	8.0	-	-

Experimental Results: Ablation Studies

- Cross-Scale(CS):
  +4.8%
- Cross-Scale(CS)+Cross-Group(CG):
  -14.7%
- Cross-Scale(CS)+Cross-Group(CG)+Cross-Color(CS):
  -15.2%

(tested on 8iVFB)

Table 1: Evaluation of compression efficiency and computational complexity.

PCs		G-PCC	Ours CS		Ours CS+CG		Ours CS+CG+CC	
		bpp	bpp	gain	bpp	gain	bpp	gain
8iVFB vox10	loot	6.19	6.26	1.1%	5.19	-16.1%	5.18	-16.4%
	red&black	9.39	10.20	8.6%	8.15	-13.2%	8.07	-14.1%
	average	7.79	8.23	4.8%	6.67	**-14.7%**	6.62	**-15.2%**
Owlii vox11	player	7.72	8.34	8.0%	7.13	-7.6%	6.78	-12.2%
	dancer	7.80	8.33	6.9%	7.11	-8.8%	6.80	-12.8%
	average	7.76	8.34	7.4%	7.12	**-8.2%**	6.79	**-12.5%**
MVUB vox10	Phil	10.27	10.13	-1.4%	7.33	-28.6%	6.78	-34.0%
	Ricardo	5.92	5.12	-13.6%	3.68	-37.9%	3.59	-39.4%
	average	8.10	7.62	-7.5%	5.50	**-33.3%**	5.19	**-36.7%**
ScanNet	**q5cm**	12.92	14.13	9.3%	11.47	**-11.2%**	11.21	**-13.2%**
	q2cm	13.13	15.04	14.6%	12.04	**-8.3%**	11.86	**-9.7%**
Ford	**q2cm**	5.32	7.05	32.5%	5.00	**-6.0%**	-	-
	q1mm	5.22	6.93	32.9%	4.97	**-4.7%**	-	-
Average Time (Eocoding Decoding) (s/frame)								
8iVFB_vox10		9.5 / 9.3	5.7	5.1	10.1	9.8	15.7	16.0
Owlii_vox11		32.8 / 32.0	17.5	15.3	37.0	35.9	56.0	58.1
MVUB_vox10		17.1 / 17.0	10.2	9.2	19.0	18.3	27.3	28.3
ScanNet_q2cm		2.0 / 2.0	1.2	1.2	3.6	3.5	6.3	6.3
Ford_q1mm		1.1 / 1.1	0.8	0.8	8.0	8.0	-	-

Conclusion & Future Work

☐ Conclusion

- The **first lightweight and generalized lossless PCAC** approach that outperforms MPEG G-PCC.
- The outstanding compression performance comes with the neural network based **cross-scale, cross-group, and cross-color prediction**.
- The lightweight computation is due to the use of **sparse convolution** and **parallel processing** inherently supported by our design.

☐ Future Works

- Further improvement on sparse point clouds like LiDAR data.
- The support of lossy compression under the same framework.

Thank you for your attention!