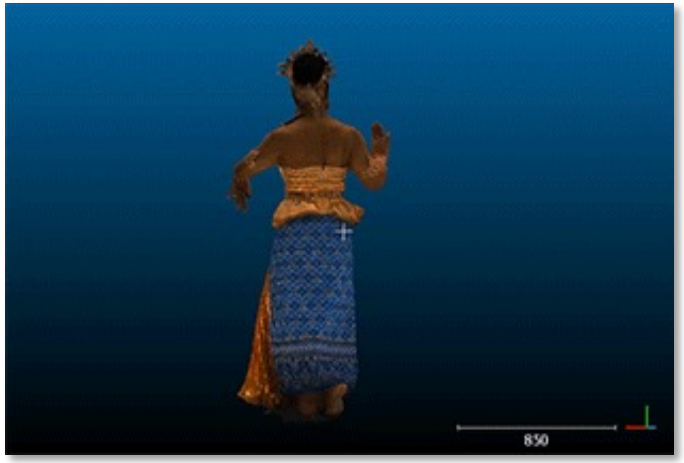
Occupancy Map Guided Attributes Deblocking for Video-based Point Cloud Compression

Peilin Chen[1], Shiqi Wang[1], and Zhu Li[2]

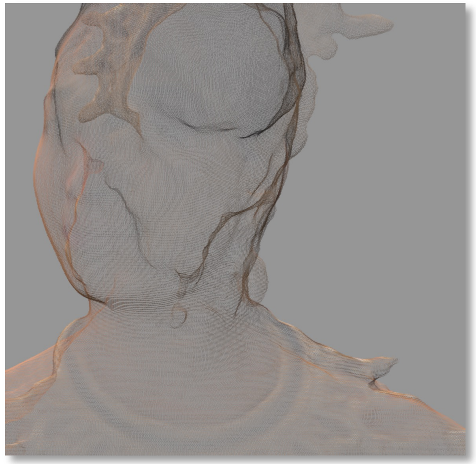
[1]Department of Computer Science, City University of Hong Kong, Hong Kong SAR, China.

[2]Department of Computer Science & Electrical Engineering, University of Missouri-KC, MO 64111, USA.

Introduction



- What is Point Cloud?
  - A collection of un-ordered points with
    - Geometry: expressed as [x, y, z]
    - Color Attributes
    - Additional info: normal, timestamp, etc.



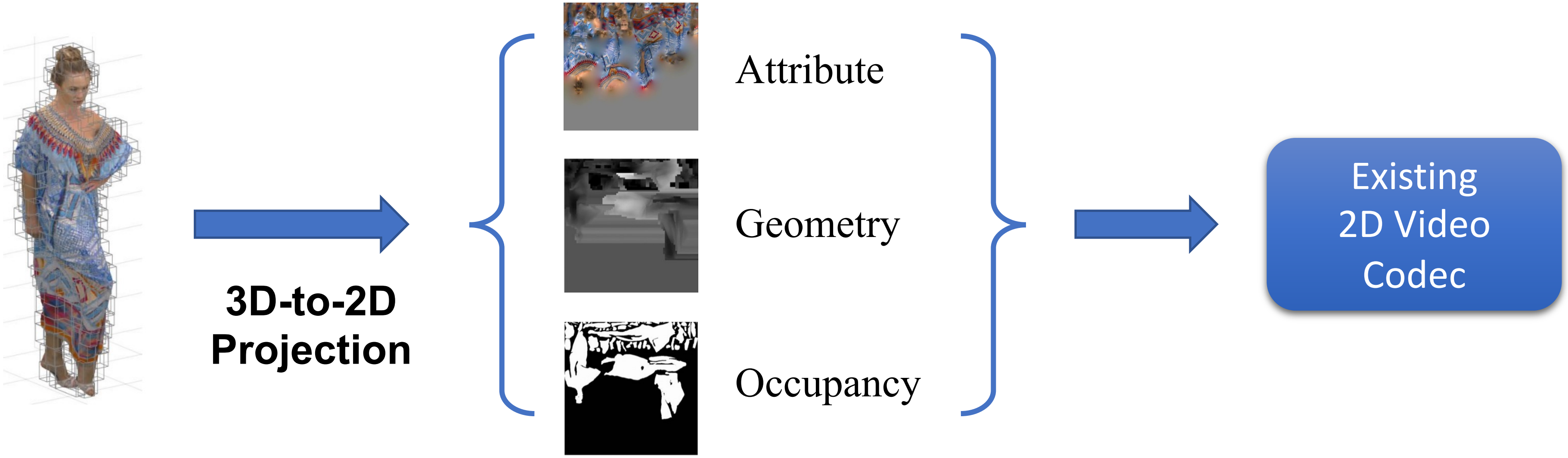
- Bringing **immersive** interactions:
  - Augmented reality
  - Telepresence conference
  - Cultural heritage documentation
  - …

- *But typically require massive storage and bandwidth*

  → ***Point cloud compression (PCC) is highly demanded***

Introduction

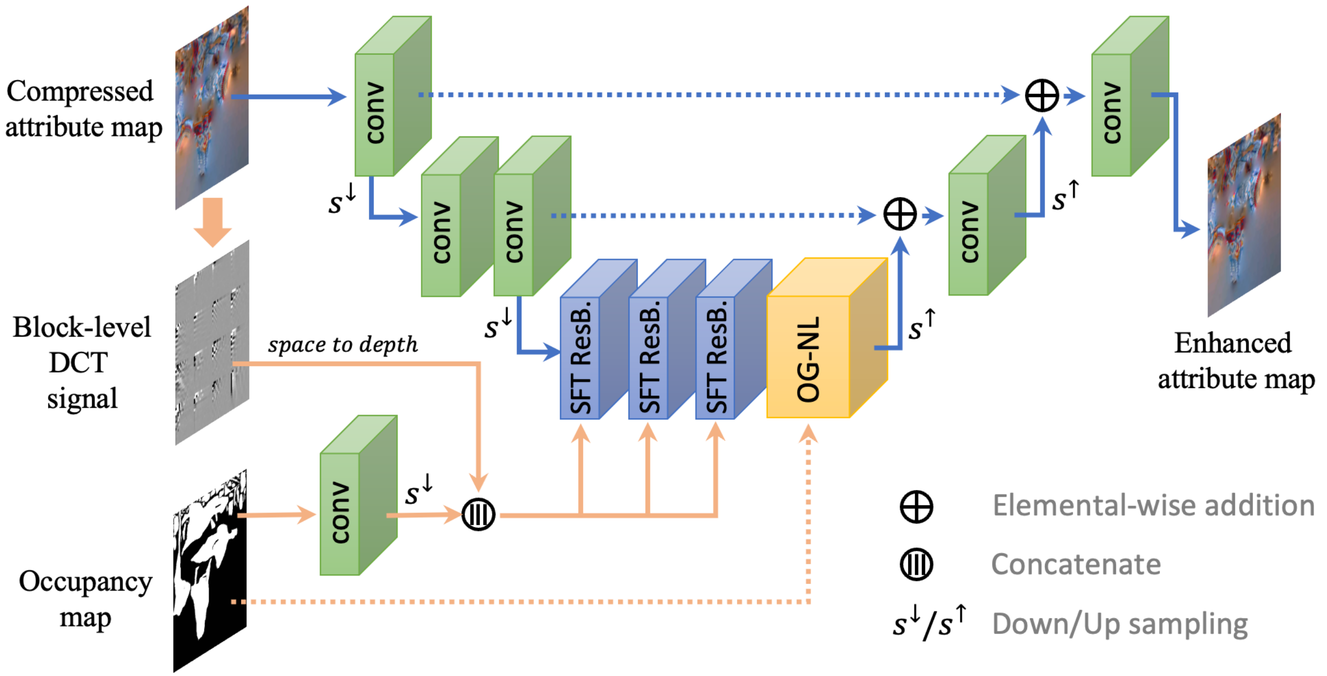
- Video-based Point Cloud Compression (V-PCC)

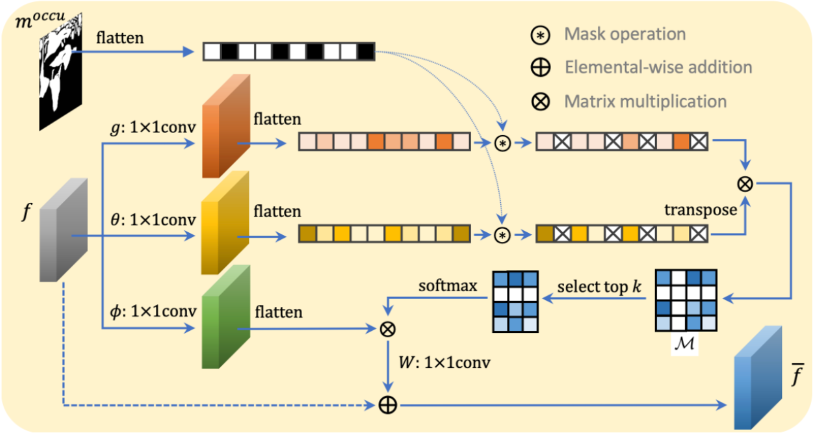


- **Lossy compression** introduce coding artifacts in attributes
  → ***degrade decoded PC quality***
  - ***challenge*:*** attributes are irregular mixtures without strong scene priors
  - ***opportunity*:*** occupancies are available, which can provide potential clues

We proposed occupancy map guided attributes deblocking for V-PCC

Proposed framework



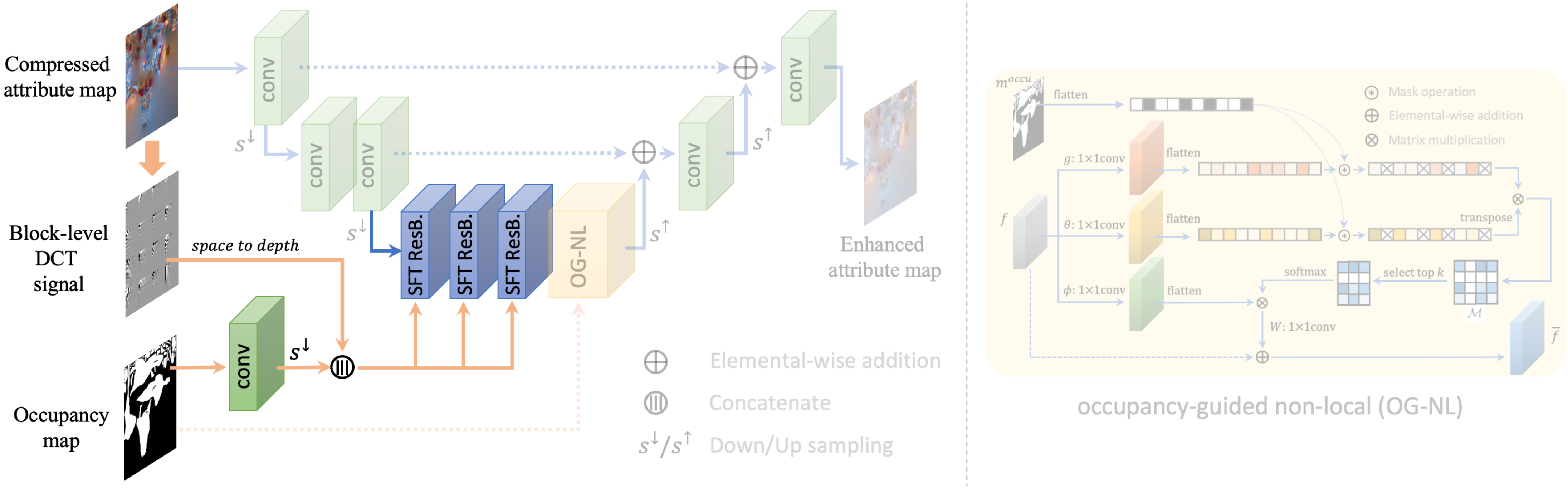


occupancy-guided non-local (OG-NL)

(a) The architecture of the proposed framework.

(b) overview of the OG-NL module.

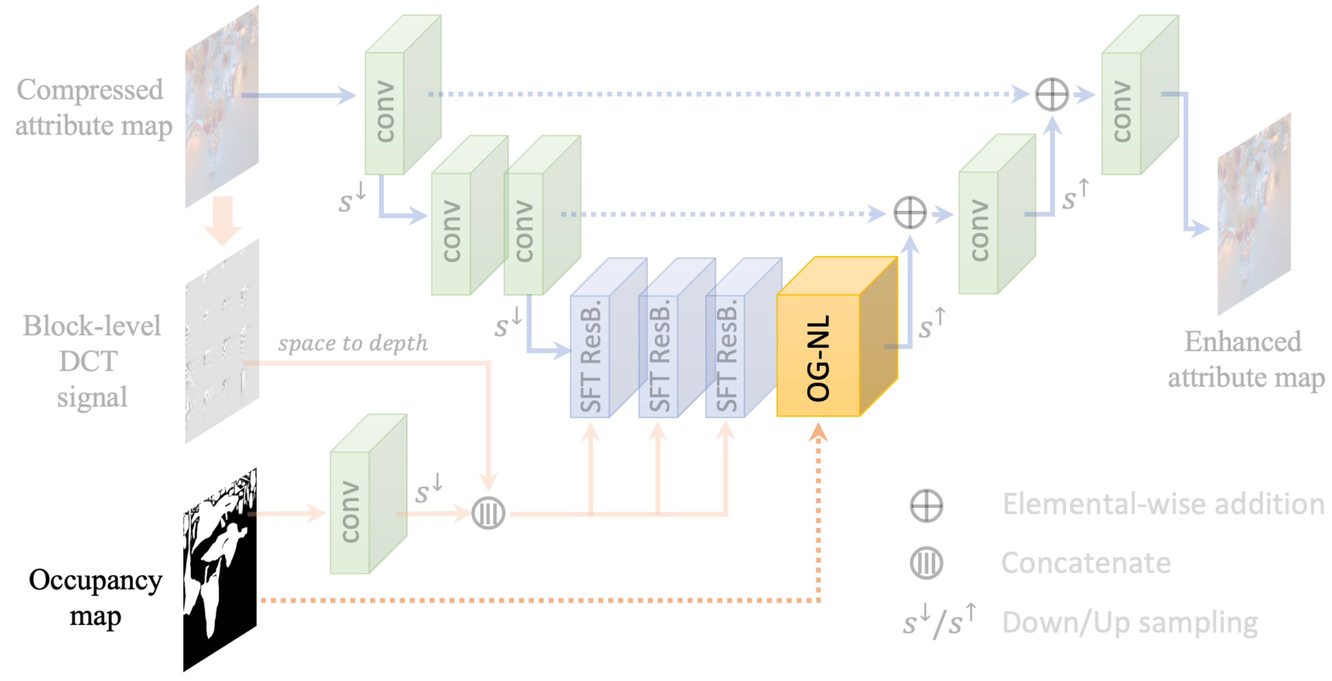
Proposed framework



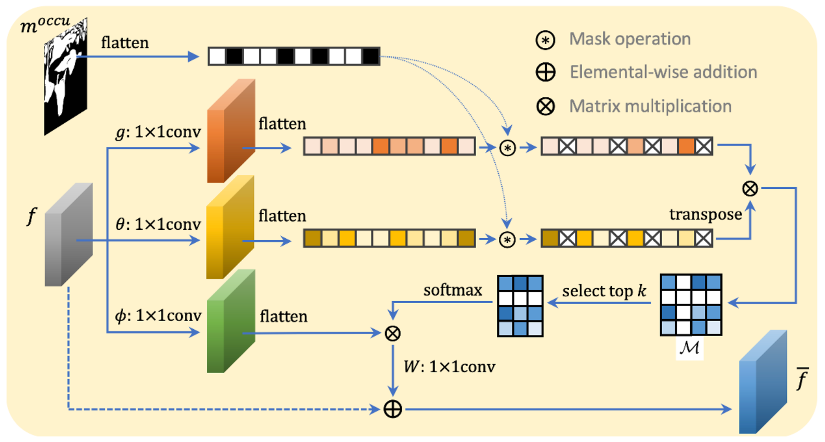
(a) The architecture of the proposed framework.

(b) overview of the OG-NL module.

Proposed framework



(a) The architecture of the proposed framework.



(b) overview of the OG-NL module.

Experiment settings

- **Codec:** version 18.0 of V-PCC reference software
- **Training:** Eight dynamic point cloud sequences[1] + Thuman2.0[2]
- **Patch size:** 256x256
- **Optimizer:** ADAM (initial lr: 1e-4)
- **Loss:** L1 with occupancy mask: $L(\Theta) = \frac{1}{N} \sum_{i=1}^{N} ||(m_i^{GT} - F(m_i^{attr}, m_i^{occu}|\Theta)) \odot m_i^{occu}||_1,$
- **Test:** Five sequences[3] from the common test condition (CTC)

[1] *soldier, queen, thaidancer, model, exercise, andrew, david and phil.*
[2] *Yu, Tao, et al. "Function4d: Real-time human volumetric capture from very sparse consumer rgbd sensors." CVPR. 2021.*
[3] *loot, redandblack, longdress, basketball player and dancer.*

Experiment results

Table 1: Overall BD-rate savings of the first 32 frames of each sequence with V-PCC reference software as the anchor under all-intra mode.

Class	Sequence	BD-AttrRate ↓			BD-TotalRate ↓		
		DCAD [10]	RNAN [16]	Proposed	DCAD [10]	RNAN [16]	Proposed
A	*loot*	1.7%	-0.9%	**-2.0%**	1.3%	-0.5%	**-1.5%**
	redandblack	-1.3%	-3.1%	**-3.9%**	-0.9%	-2.3%	**-3.1%**
B	*longdress*	-1.9%	-3.1%	**-3.3%**	-1.6%	-2.7%	**-2.8%**
C	*basketball player*	-2.6%	-5.5%	**-7.5%**	-1.6%	-3.6%	**-5.3%**
	dancer	-3.7%	-6.5%	**-8.5%**	-2.6%	-4.6%	**-6.2%**
	Average	-1.5%	-3.8%	**-5.0%**	-1.1%	-2.7%	**-3.8%**

Table 2: Illustration of model complexities. Results of FLOPs are measured with assumption that the input size is 128×128.

	DCAD [10]	RNAN [16]	Proposed
Parameters (M)	0.296	2.725	0.913
FLOPs (G)	4.851	37.784	1.357

Experiment results

(a) Original Point Cloud　(b) V-PCC Decoded Point Cloud　(c) Enhanced Result with Our Method

Thank you !