Cross-Regional Attention Network for Point Cloud Completion

Hang Wu, Yubin Miao

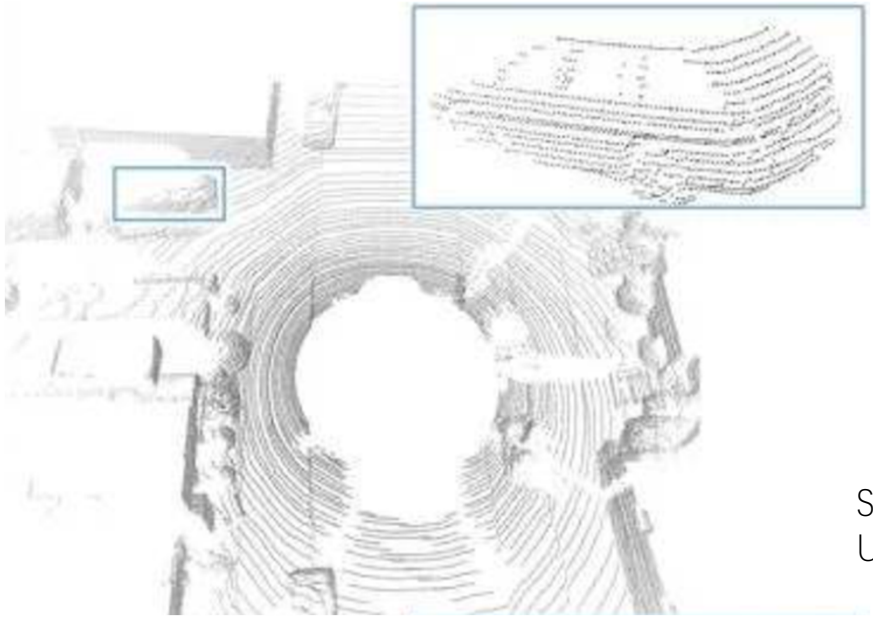
Shanghai Jiao Tong University, Shanghai, China

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

目录 Contents

Introduction

- Raw point clouds obtained from 3D sensors are always **not complete** due to occlusions, viewing angles and surface material properties. These partial point clouds are likely to cause **structural loss** of object shapes and bring about difficulties to further high-level 3D visions.
- In this paper, we propose a point cloud completion network using **cross-regional attention**.



Source: Ruihui Li, et al., PU-GAN: a Point Cloud Upsampling Adversarial Network

目录 Contents

3D shape generation

- L-GAN (P. Achlioptas , et al., ICML 2017):
  - Propose a pioneering point cloud auto-encoder.
  - Uses PointNet as encoder and fully connected layers as decoder.
  - Uses adversarial training.
- FoldingNet (Y. Yang , et al., CVPR 2018) & AtlasNet (T. Groueix , et al., CVPR 2018) :
  - Improve the decoders by borrowing the idea of manifold.
  - FoldingNet learns a mapping function of a canonical 2D grid and 3D surface.
  - AtlasNet learns the mapping of a set of 2D squares to 3D shape .

Point cloud completion

- PCN (W. Yuan, et al., 3DV 2018):
  - First network specialized for point cloud completion.
  - Uses a modified PointNet-based encoder.
  - Two-stage decoder (coarse->fine).
- PF-Net (Z. Huang , et al., CVPR 2020):
  - A feature-points-based multi-scale generating (coarse->fine).
  - Predict only missing part of partial point cloud.
  - Defines the missing parts by hierarchically removing certain points from complete point cloud.

Purpose of this paper

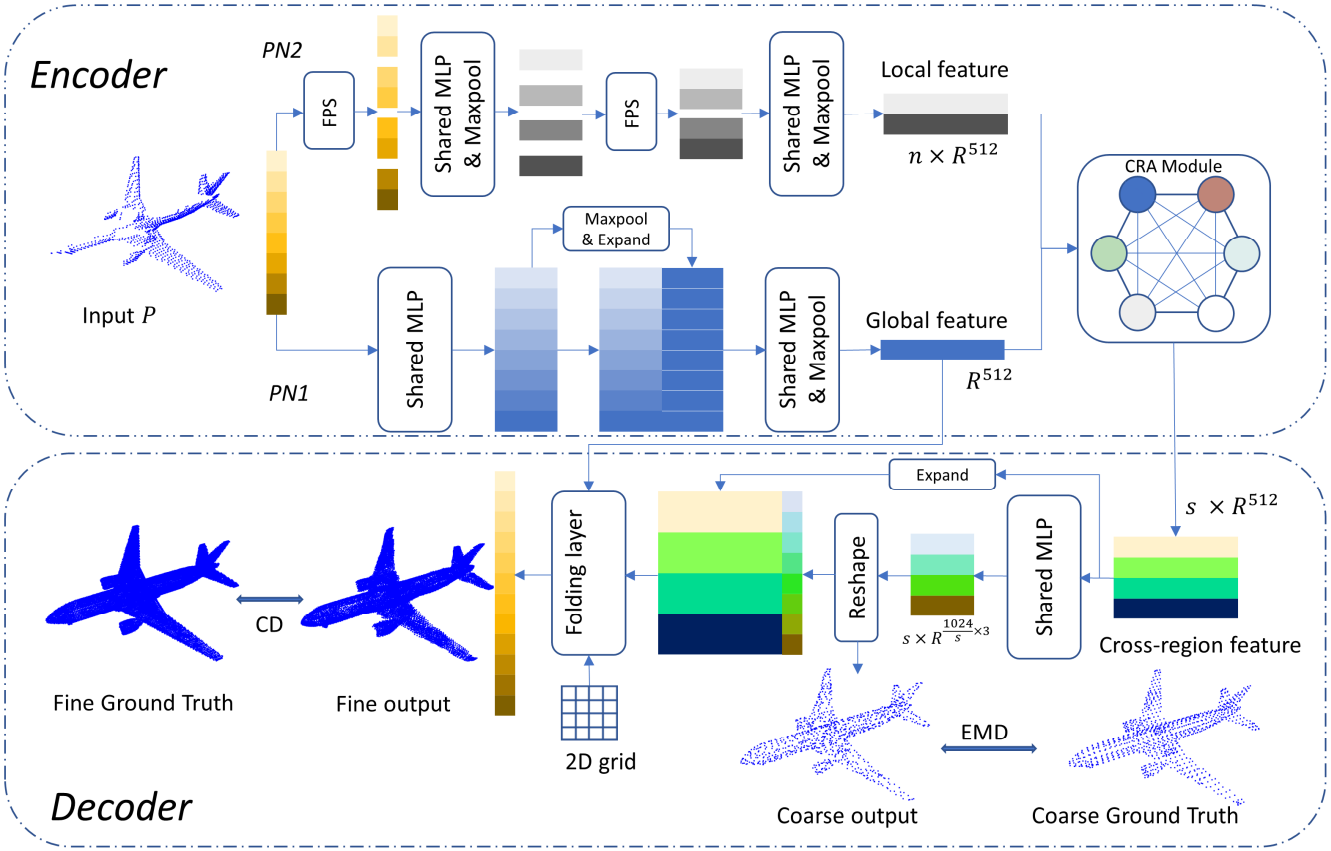
- All the above-mentioned methods only extract global features in the encoder, which might make the network lose the ability of on local detail awareness.

- <u>Main purpose of this paper:</u>

- Apply local feature awareness module in point cloud completion to generate point clouds with **more details and smoother shapes**, and avoid potential instability caused by local region sampling.

目录 Contents

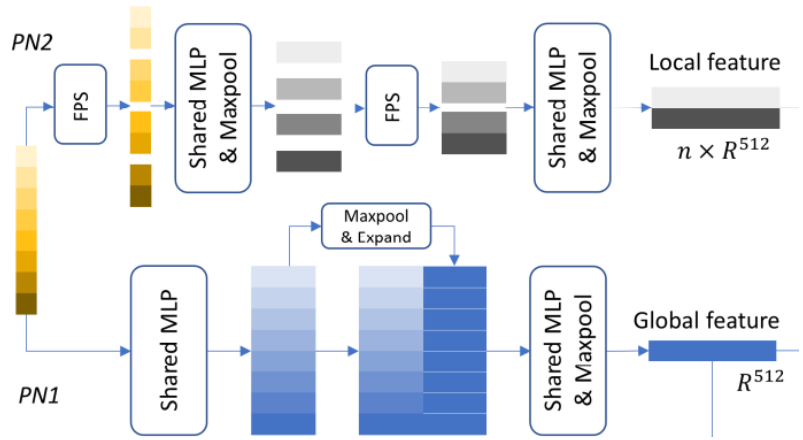
Network Structure



Encoder

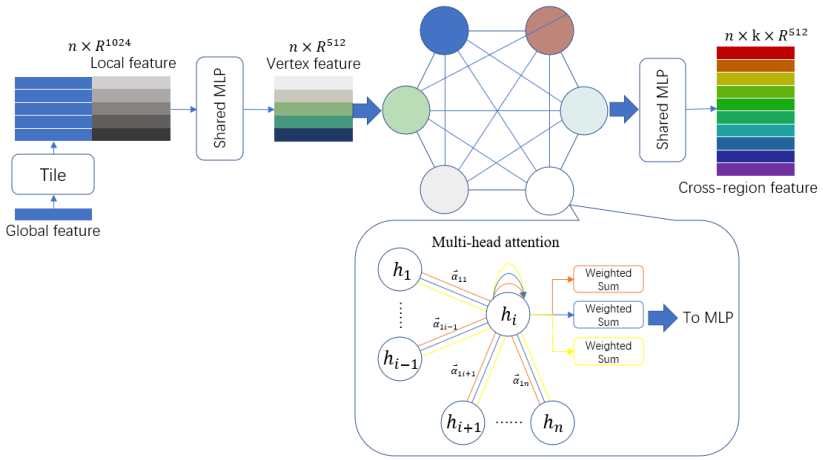
Multiscale Feature Extraction module (MFE):

Independently extract both global and local features from the inputs.



Cross-Regional Attention module (CRA) :

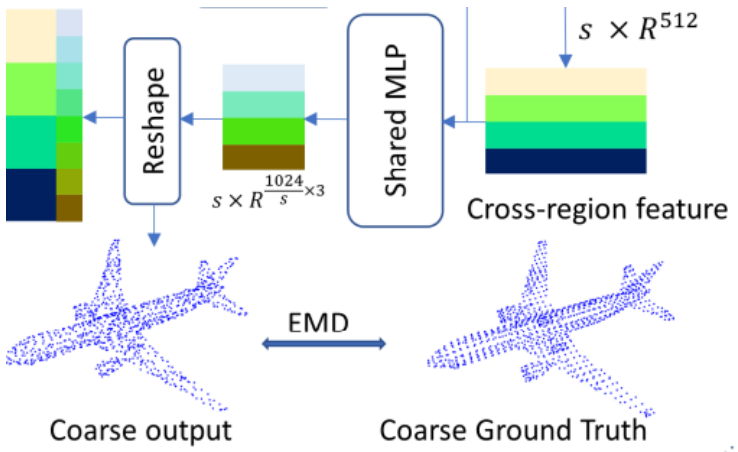
Interpret local features under the condition of global features, learn the relationships among these conditional features.
It is realized by convolution with attention on a fully-connected feature graph.



Two-step decoder

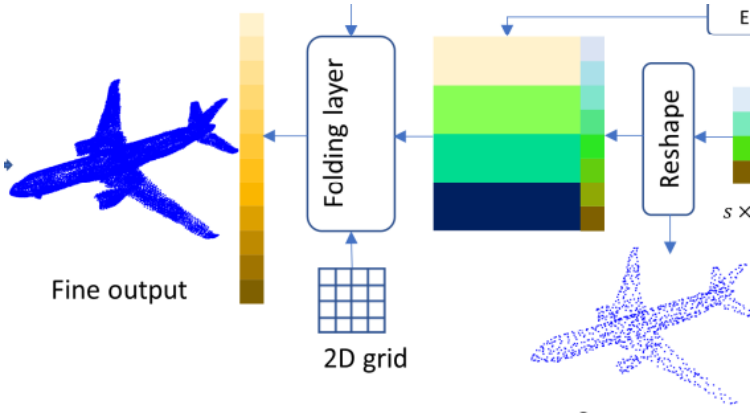
Coarse decoder:

Decode each cross-region feature vector to a point set with fixed number of points. All these point sets will form the skeleton of complete point cloud.



Fine decoder:

For each point in coarse point cloud, reuse its corresponding cross-region feature and global feature to fold a 2D grid.



Loss function

- Earth Moving Distance (EMD): Used for guide coarse point cloud generation.

$$EMD(S_1, S_2) = \min_{\phi: S_1 \to S_2} \sum_{x \in S_1} \min_{y \in S_2} \|x - \phi(x)\|_2$$

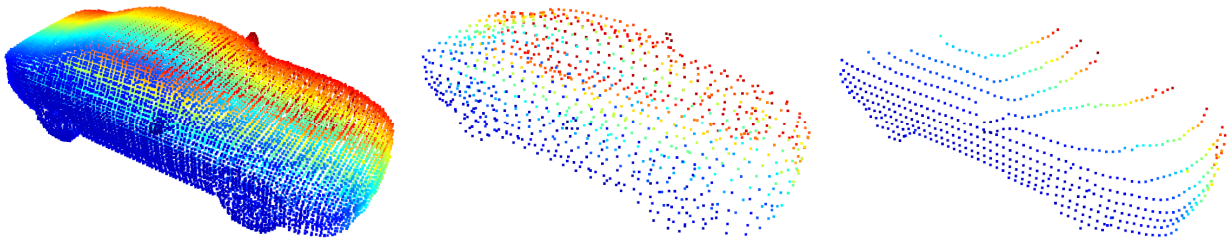
- Chamfer Distance (CD): Used for guide fine point cloud generation.

$$CD(S_1, S_2) = \frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2 + \frac{1}{|S_2|} \sum_{y \in S_2} \min_{x \in S_1} \|y - x\|_2$$

目录 Contents

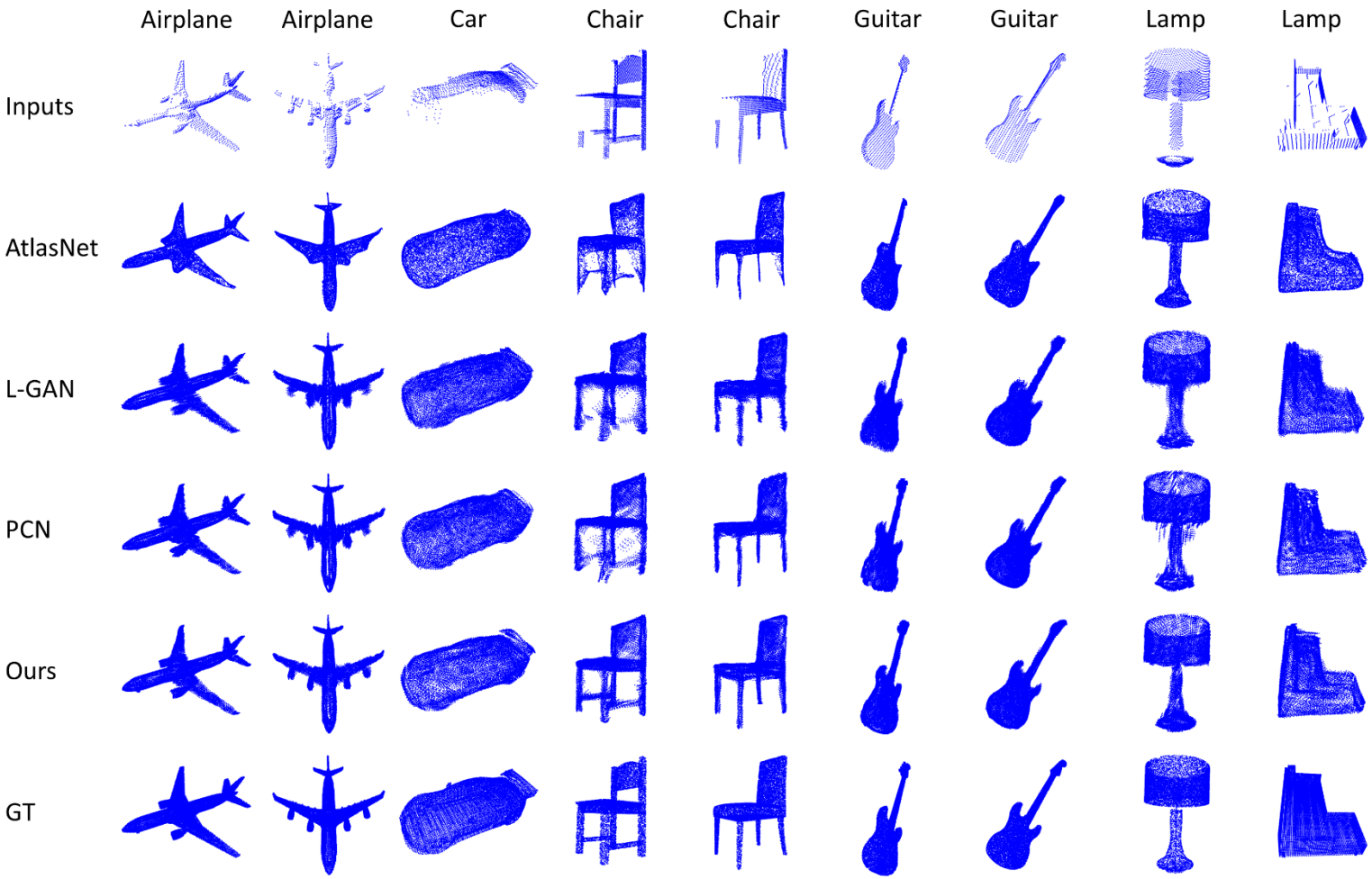
Experiments and Results

- We train and test our network, as well as other existing point cloud generate methods, on **six** categories of objects in *ModelNet* dataset.



An Example of complete (ground truth) and partial (input) point clouds in our dataset. (Left) A complete point cloud with 16384 points. (Middle) A complete point cloud with 1024 points. (Right) A partial point cloud generated by a virtual 64-channel LiDAR from a near view point.

Experiments and Results



Experiments and Results

TABLE I. QUANTITATIVE COMPARISON OF POINT CLOUD COMPLETION IN TERMS OF EARTH MOVING DISTANCE ($EMD \times 100$)

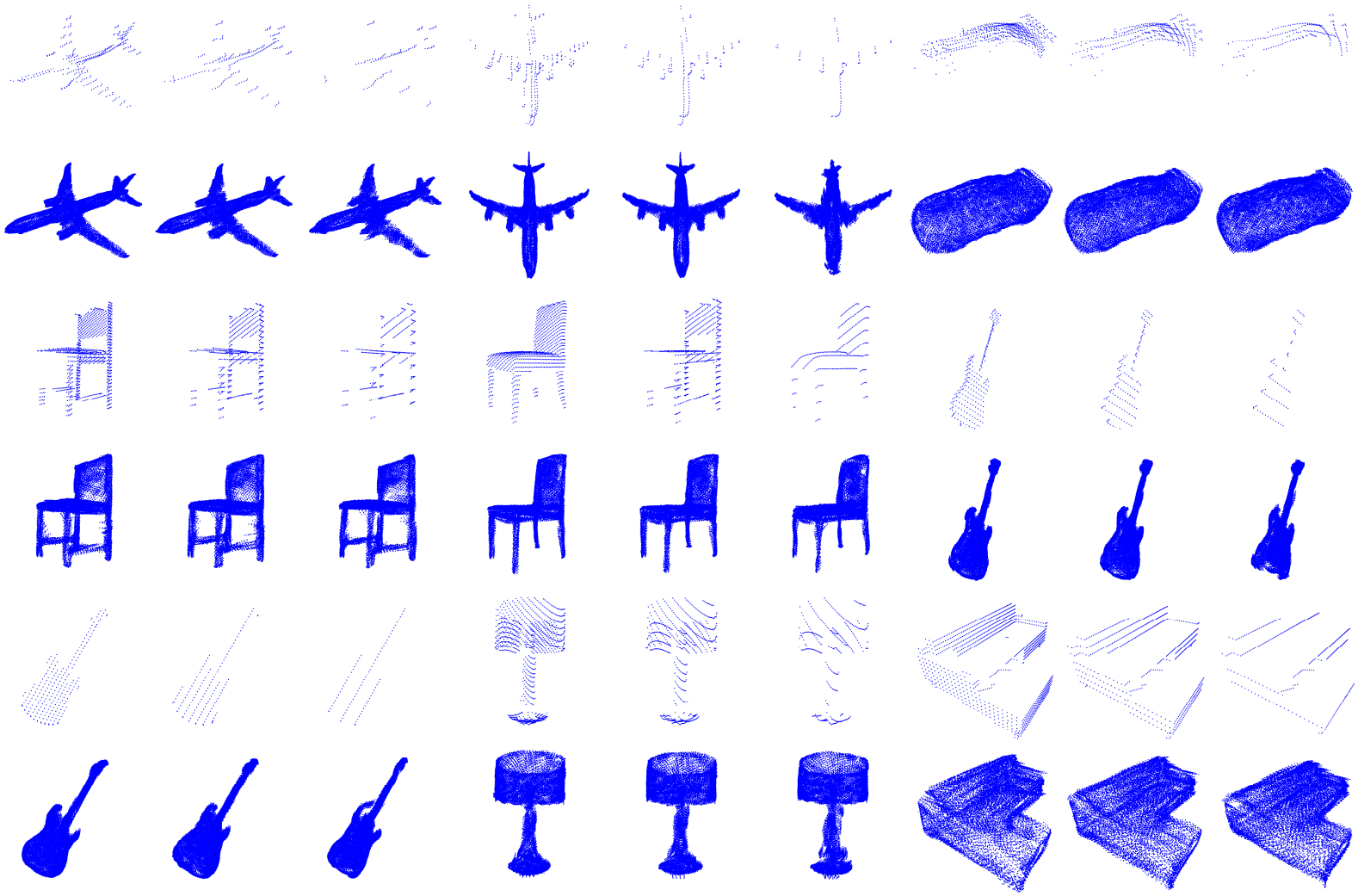
Methods	***Average***	***Airplane***	***Car***	***Chair***	***Guitar***	***Lamp***	***Sofa***
L-GAN	13.49	8.52	8.09	16.79	11.01	25.88	10.65
AtlasNet	15.20	10.16	10.80	19.00	11.43	26.38	13.45
PCN	12.84	8.20	**7.82**	16.22	10.37	24.40	10.03
Ours	**11.70**	**7.58**	7.94	**15.12**	**9.15**	**20.53**	**9.89**

TABLE II. QUANTITATIVE COMPARISON OF POINT CLOUD COMPLETION IN TERMS OF CHAMFER DISTANCE ($CD \times 100$)

Methods	***Average***	***Airplane***	***Car***	***Chair***	***Guitar***	***Lamp***	***Sofa***
L-GAN-Folding	3.06	1.68	2.01	3.93	1.80	6.04	2.92
AtlasNet	5.24	3.30	4.19	6.74	2.88	8.90	5.43
PCN	2.89	1.58	1.91	3.73	1.68	5.72	2.89
Ours	**2.70**	**1.56**	**1.85**	**3.62**	**1.54**	**5.06**	**2.61**

*Note that the values in this table are about 2-4 times of those reported in PCN [28], this is because out ground truth points are normalized to [-1,1], while in PCN it is [-0.3,0.3]

More Results from Virtual LiDAR Scanning



目录 Contents

Conclusion

- We propose *a new point cloud completion network* that is able to reconstruct high-density complete point clouds from partial point clouds.
- Through *parallel multiscale feature extraction*, *cross-regional feature fusion* and *two-stage feature decoding*, we are able to introduce local feature extraction in generative tasks.
- The test results on the ModelNet dataset show that our network can not only output smooth global shapes, but can also provide details as many as possible.

- We hope that our method could contribute to further 3D environment perceptions.

Thank You!