CVPR SEATTLE WASHINGTON
JUNE 16-18 2020

PQ-NET:
A Generative Part Seq2Seq Network for 3D Shapes

Rundi Wu[1,4] Yixin Zhuang[1] Kai Xu[2] Hao Zhang[3,4] Baoquan Chen[1,4]

[1] Center on Frontiers of Computing Studies, Peking University

[2] National University of Defense Technology

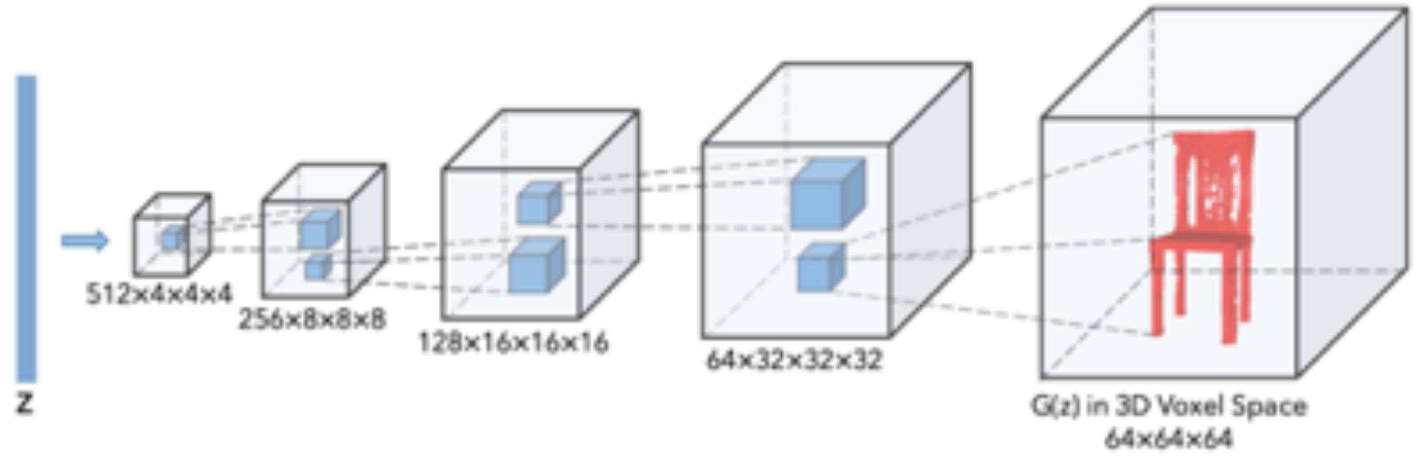
[3] Simon Fraser University

[4] AICFVE, Beijing Film Academy

3D shape generation

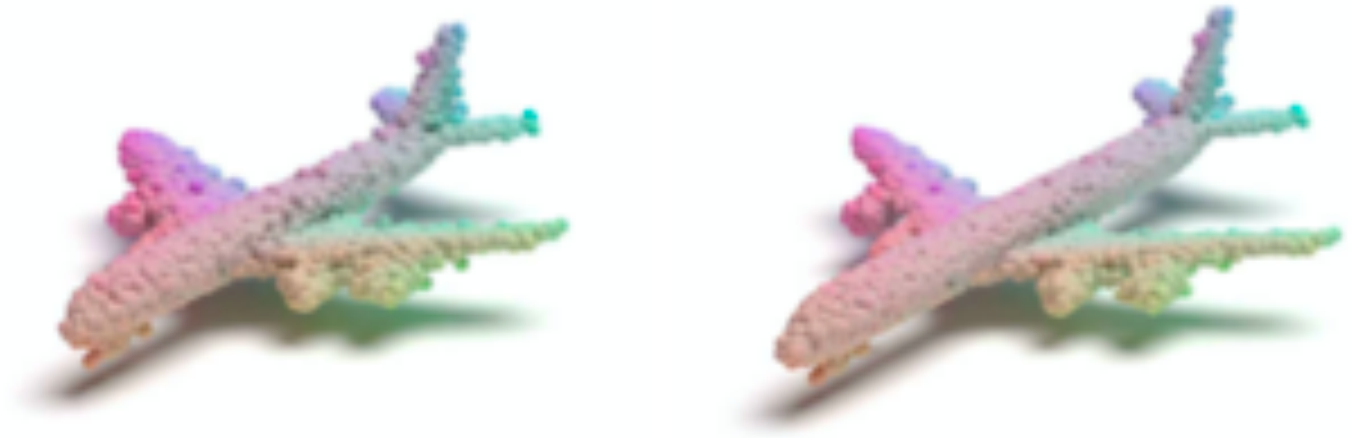
Voxel grid

[3DGAN, NIPS 2016]



Point cloud

[Pointflow, ICCV 2019]

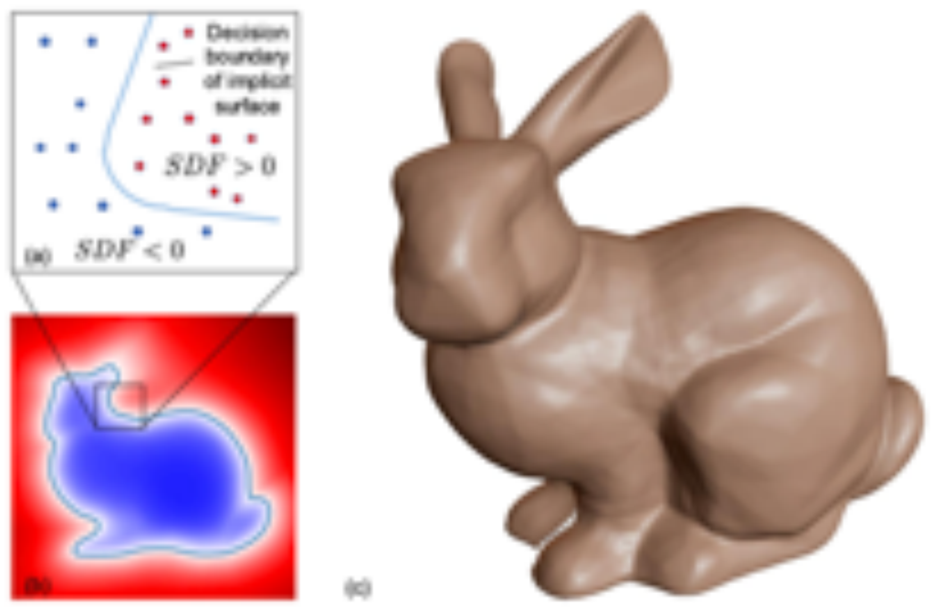


Mesh

[AtlasNet, CVPR 2018]

Implicit function

[DeepSDF, CVPR 2019]

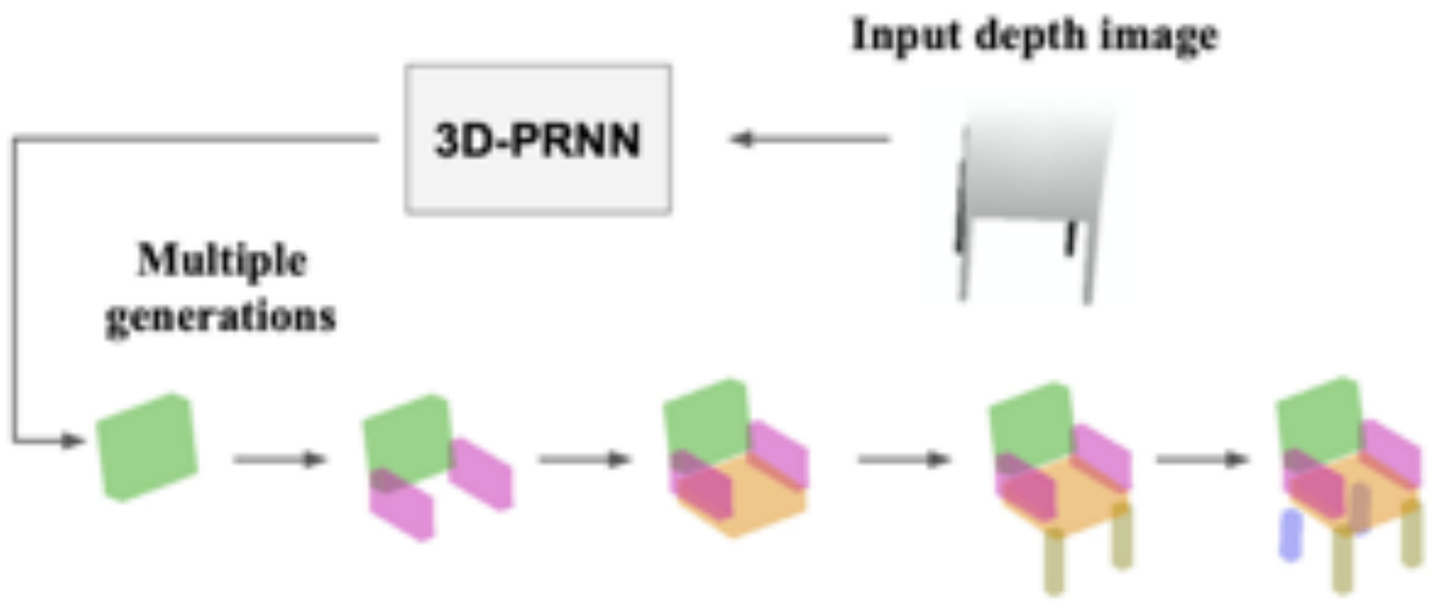


1. J. Wu, C. Zhang, T. Xue, B. Freeman, and J. Tenenbaum. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. In *Advances in Neural In- formation Processing Systems*, pages 82–90, 2016.
2. G. Yang, X. Huang, Z. Hao, M.-Y. Liu, S. Belongie, and B. Hariharan. Pointflow: 3d point cloud generation with con- tinuous normalizing flows. *2019 IEEE International Conference on Computer Vision (ICCV)*.
3. T. Groueix, M. Fisher, V. G. Kim, B. C. Russell, and M. Aubry. A papier-maˆche´ approach to learning 3d surface generation. In *Proc. CVPR*, pages 216–224, 2018.
4. J.J.Park,P.Florence,J.Straub,R.Newcombe,andS.Love- grove. DeepSDF: Learning continuous signed distance func- tions for shape representation. In *CVPR*, 2019.

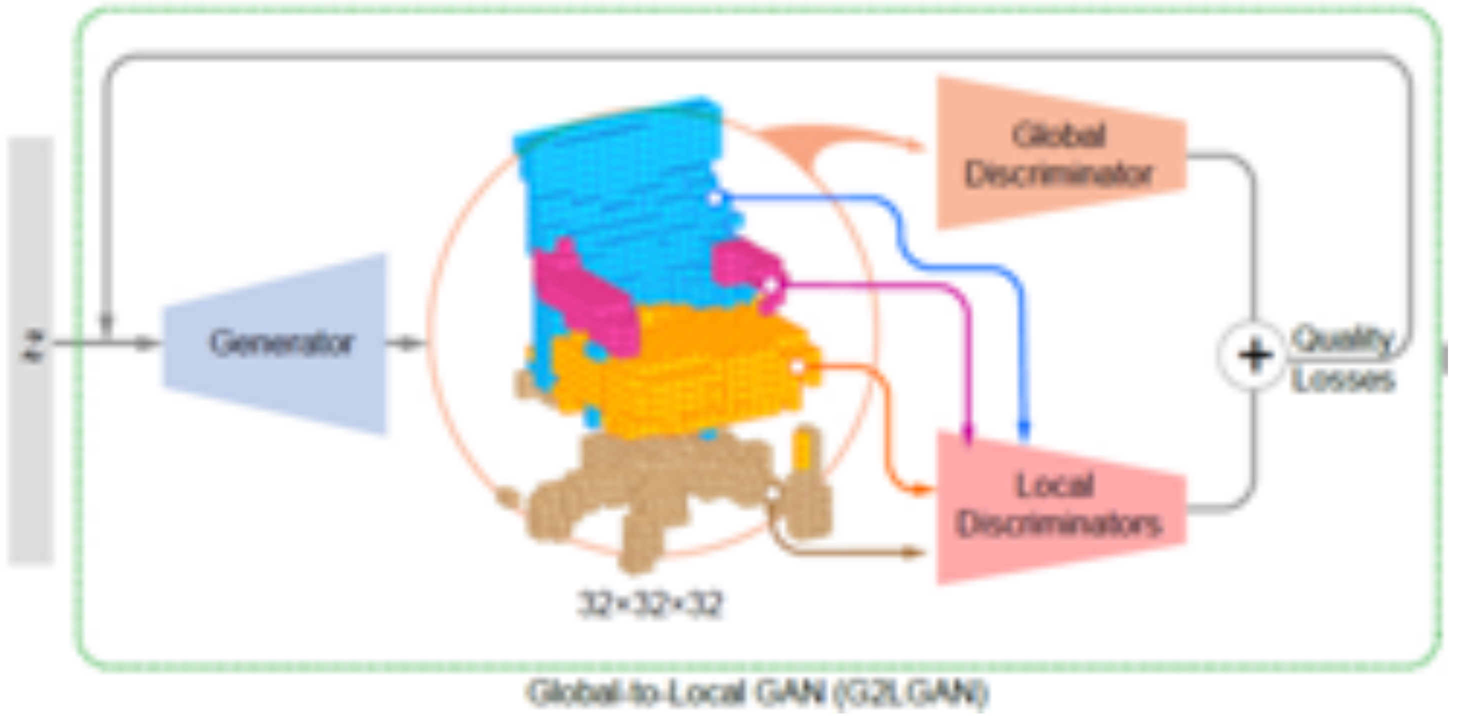
Structural 3D shape generation

[GRASS,[1] SIG 2017]

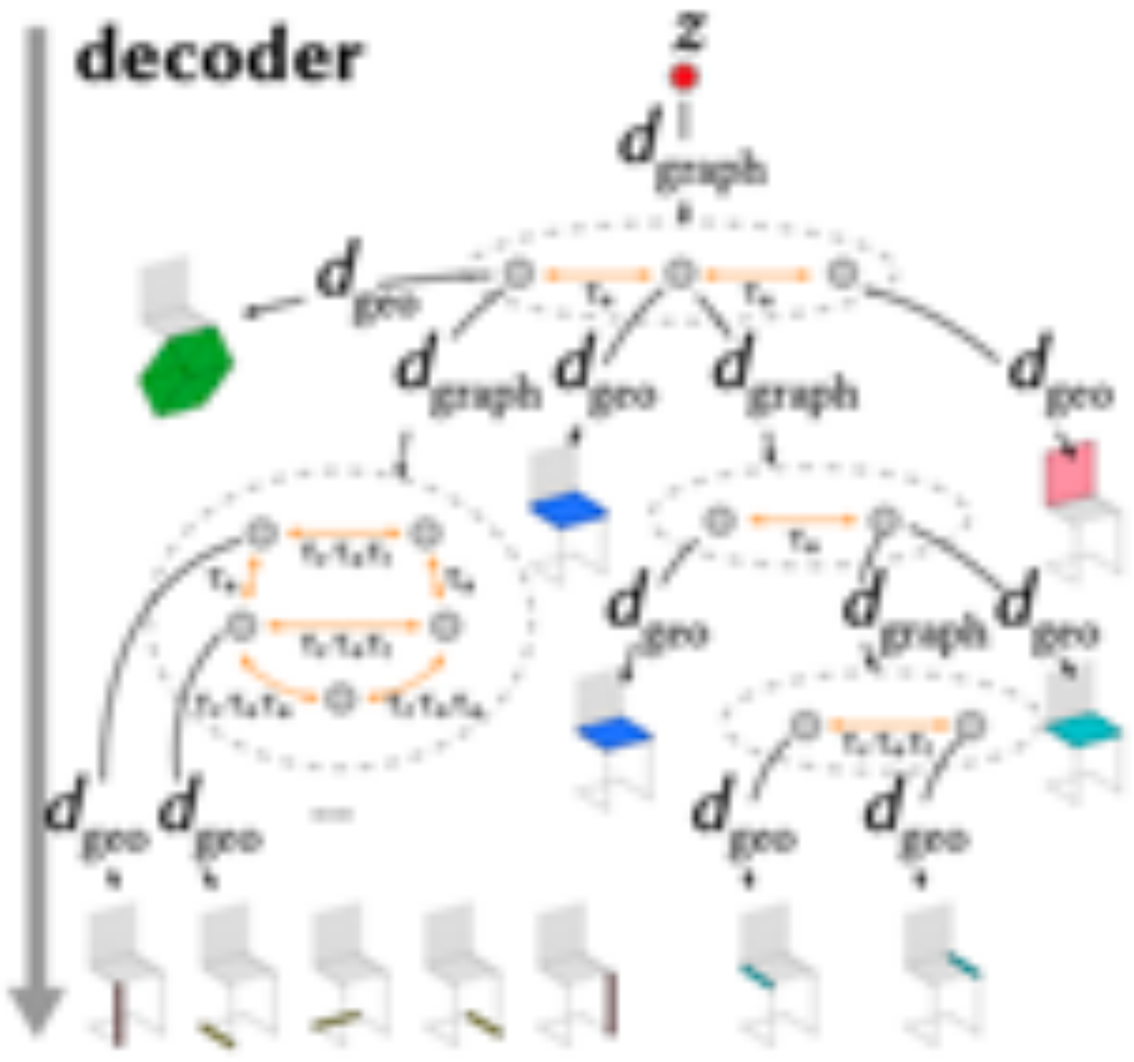
[3D-PRNN,[2] ICCV 2017]



[G2L,[3] SIGA 2018]



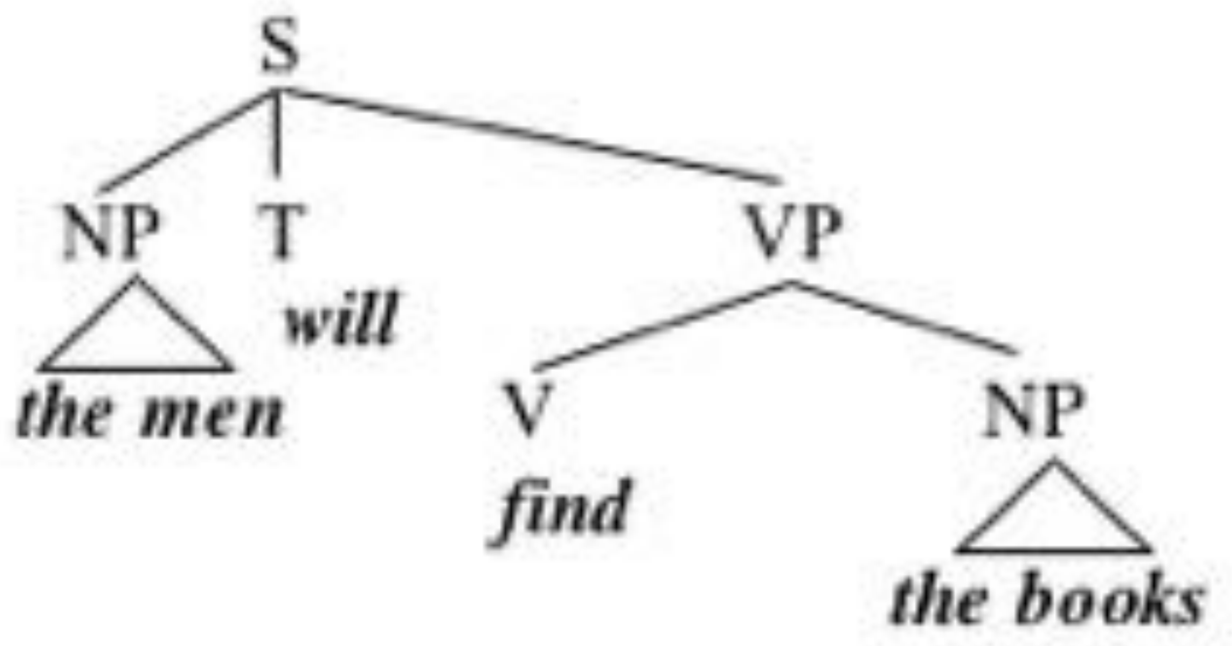
[StructureNet,[4] SIGA 2019]



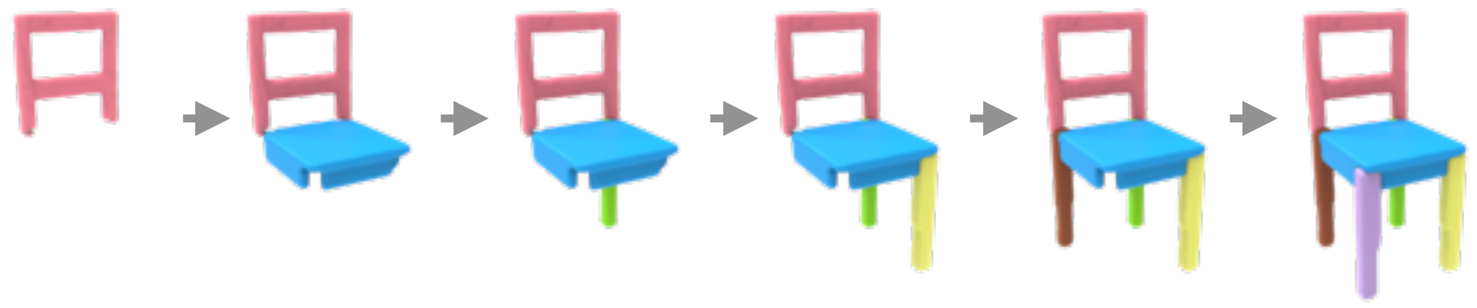
1. J. Li, K. Xu, S. Chaudhuri, E. Yumer, H. Zhang, and L. Guibas. Grass: Generative recursive autoencoders for shape structures. *ACM Trans. on Graph. (SIGGRAPH)*, 2017.
2. C. Zou, E. Yumer, J. Yang, D. Ceylan, and D. Hoiem. 3D- PRNN: Generating shape primitives with recurrent neural networks. *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.
3. H. Wang, N. Schor, R. Hu, H. Huang, D. Cohen-Or, and H. Huang. Global-to-local generative model for 3d shapes. *ACM Transactions on Graphics (Proc. SIGGRAPH ASIA)*, 37(6):214:1214:10, 2018.
4. K. Mo, P. Guerrero, L. Yi, H. Su, P. Wonka, N. Mitra, and L. J. Guibas. Structurenet: Hierarchical graph networks for 3d shape generation. *ACM Trans. on Graph. (SIGGRAPH Asia)*, 2019.

Shape structure presentations

① *hierarchical* part organization ≈ phrases nested in phrases



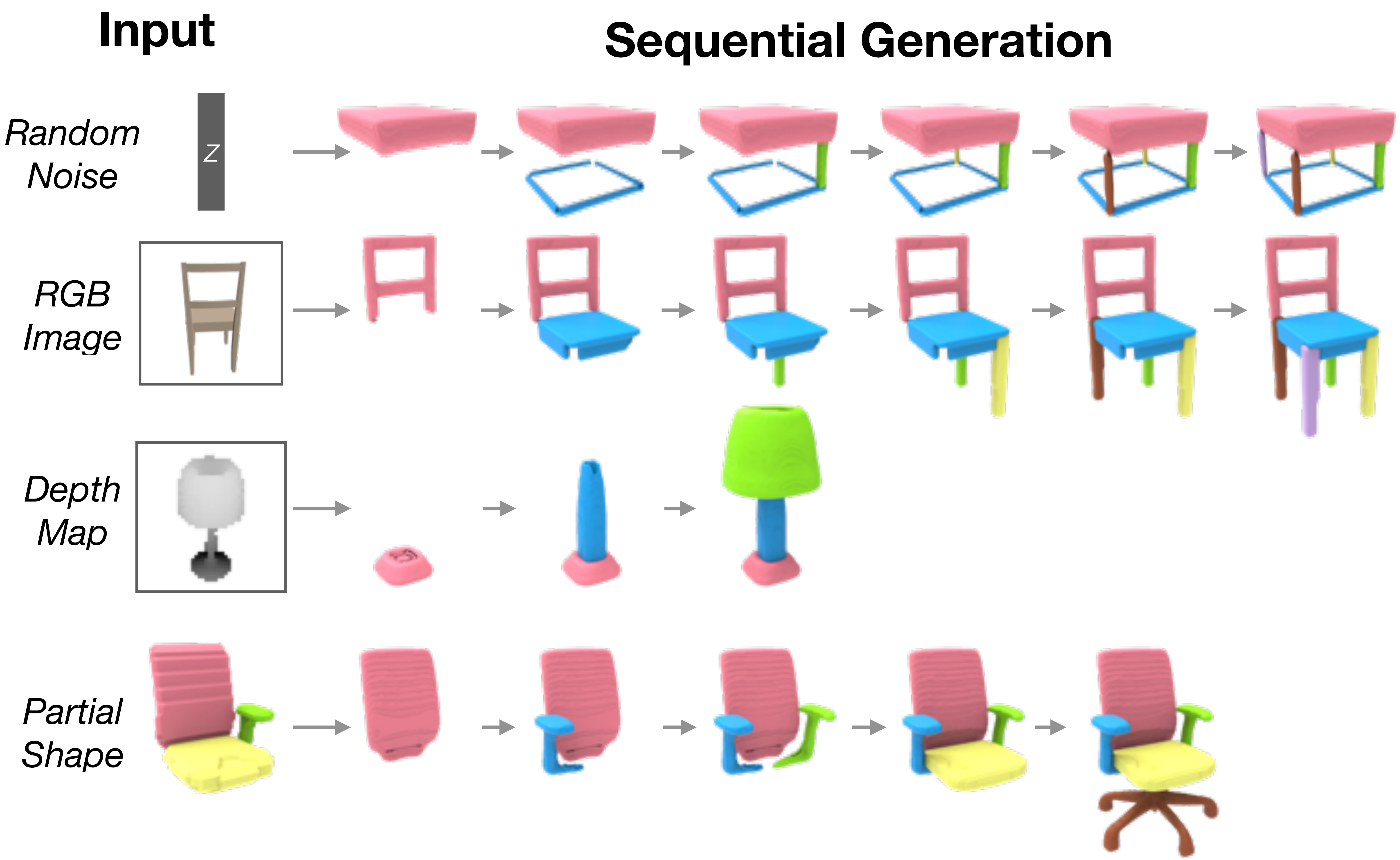
② ***linear* part order** ≈ linear string of words



“ *the men will find the books* ”

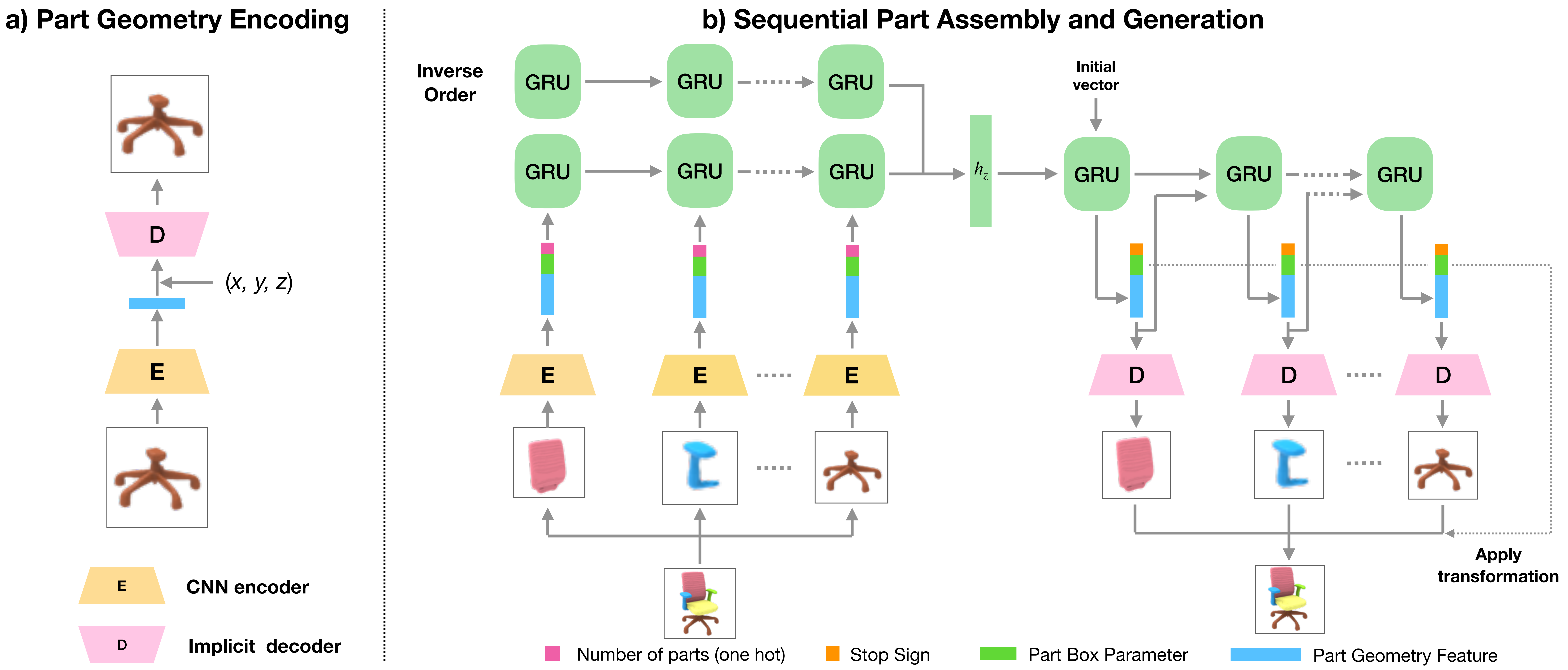
Generate as a sequence

- Our network, PQ-NET, learns 3D shape representation via *sequential part assembly*



Method

a. Apply IM-NET to encode each scaled part's geometry
b. Model sequential part assembly using a Sequence-to-Sequence Auto-encoder (Seq2Seq AE)



Method - Part geometry encoding

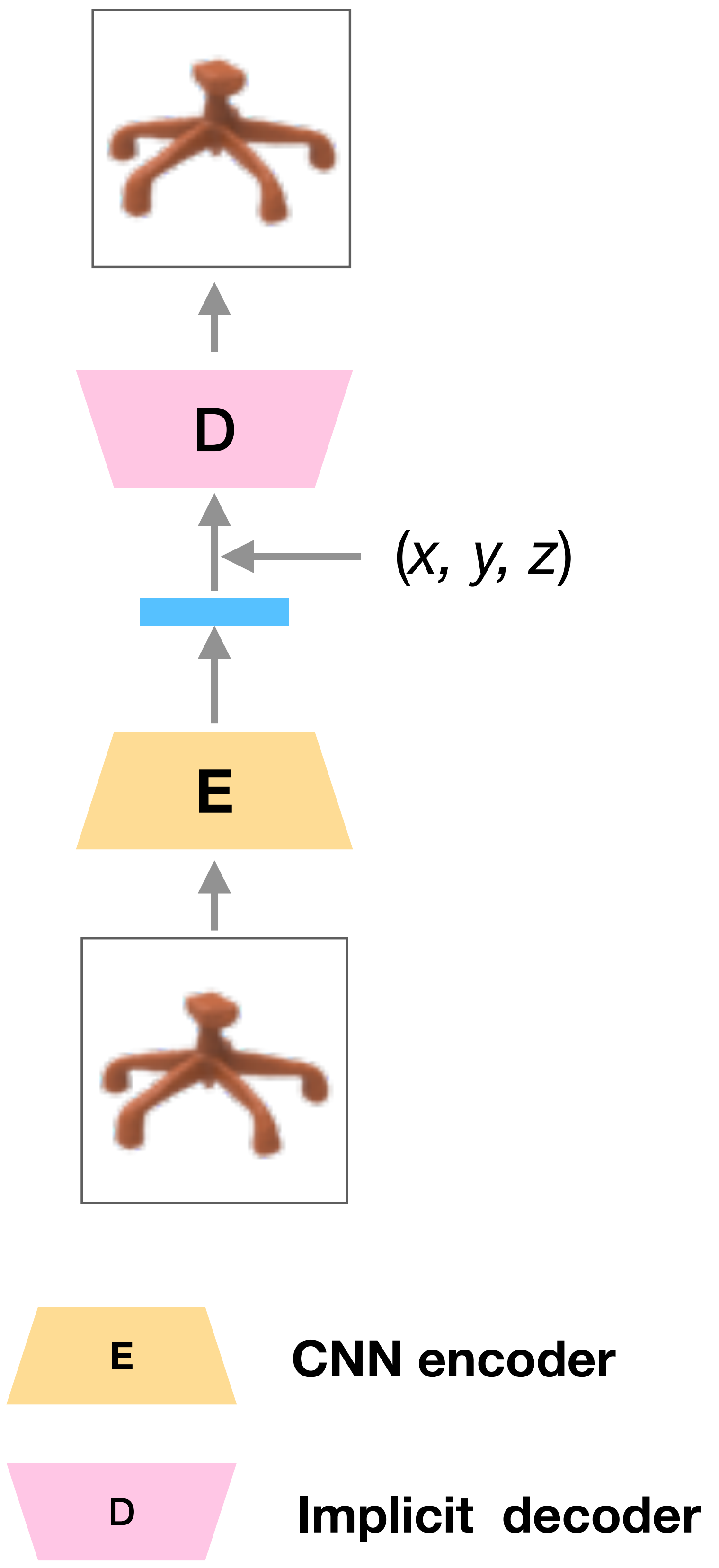
Similar architecture as IM-NET[1]:

- a CNN encoder e maps 64^3 voxelized part P to 128D vector
- a MLP decoder d that predicts the occupancy of a given point p

$$\mathcal{L}(P) = \mathbb{E}_{p \in T_P} |d(e(P), p) - \mathcal{F}(p)|^2$$

A set of sampled points from P (T_P)

ground truth signed function ($\mathcal{F}(p)$)



1. Z. Chen and H. Zhang. Learning implicit fields for generative shape modeling. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

Method - Seq2Seq AE

Encoder :

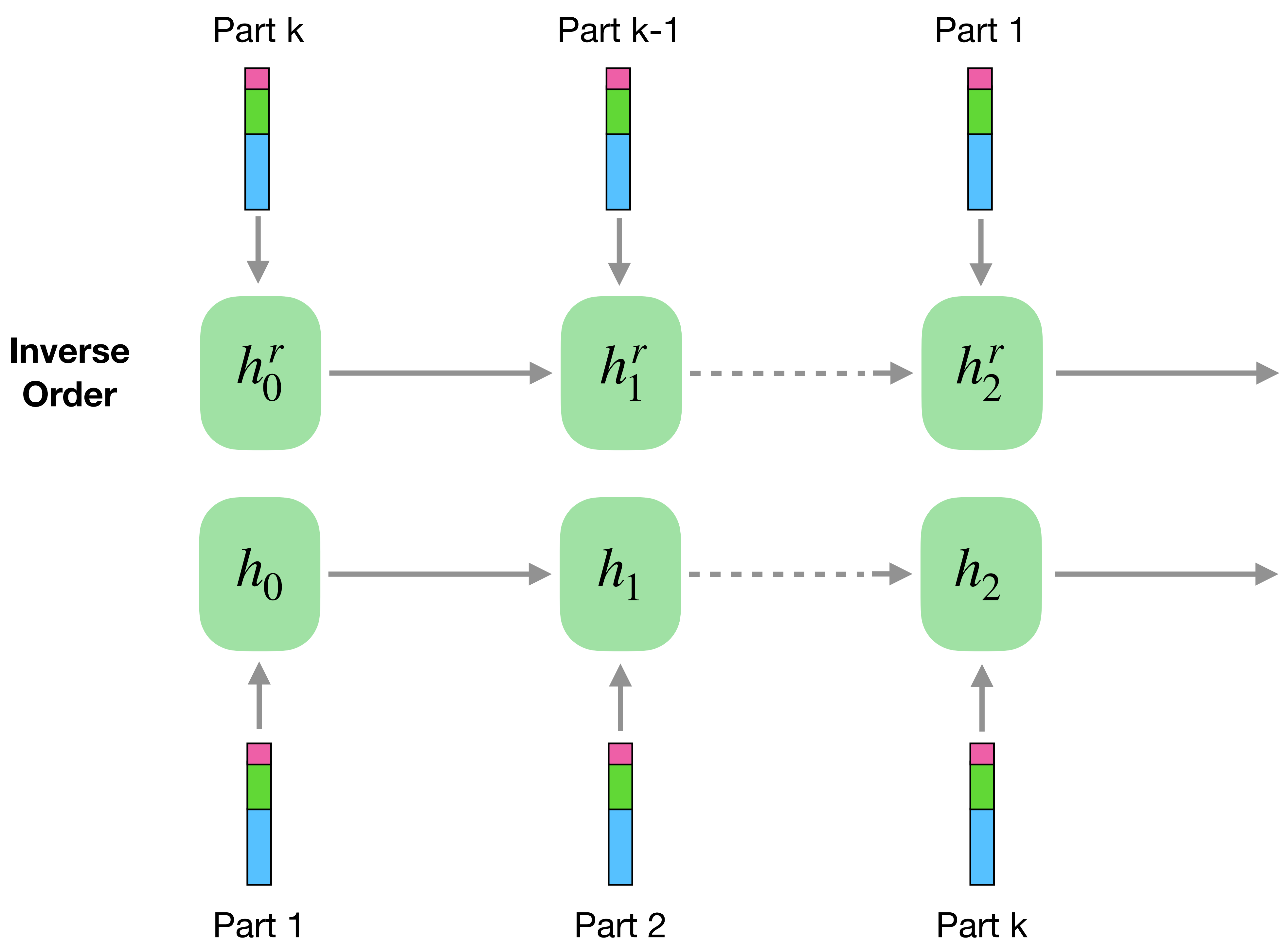
- a bidirectional stacked RNN to encode part sequence

Stacked GRU Cell

Number of parts in one-hot representation

Part Box Parameter : 6D, position + size

Part Geometry Feature : latent vector encoded by IM-NET



Method - Seq2Seq AE

Decoder :

- a stacked RNN to predict geometry and structure feature separately

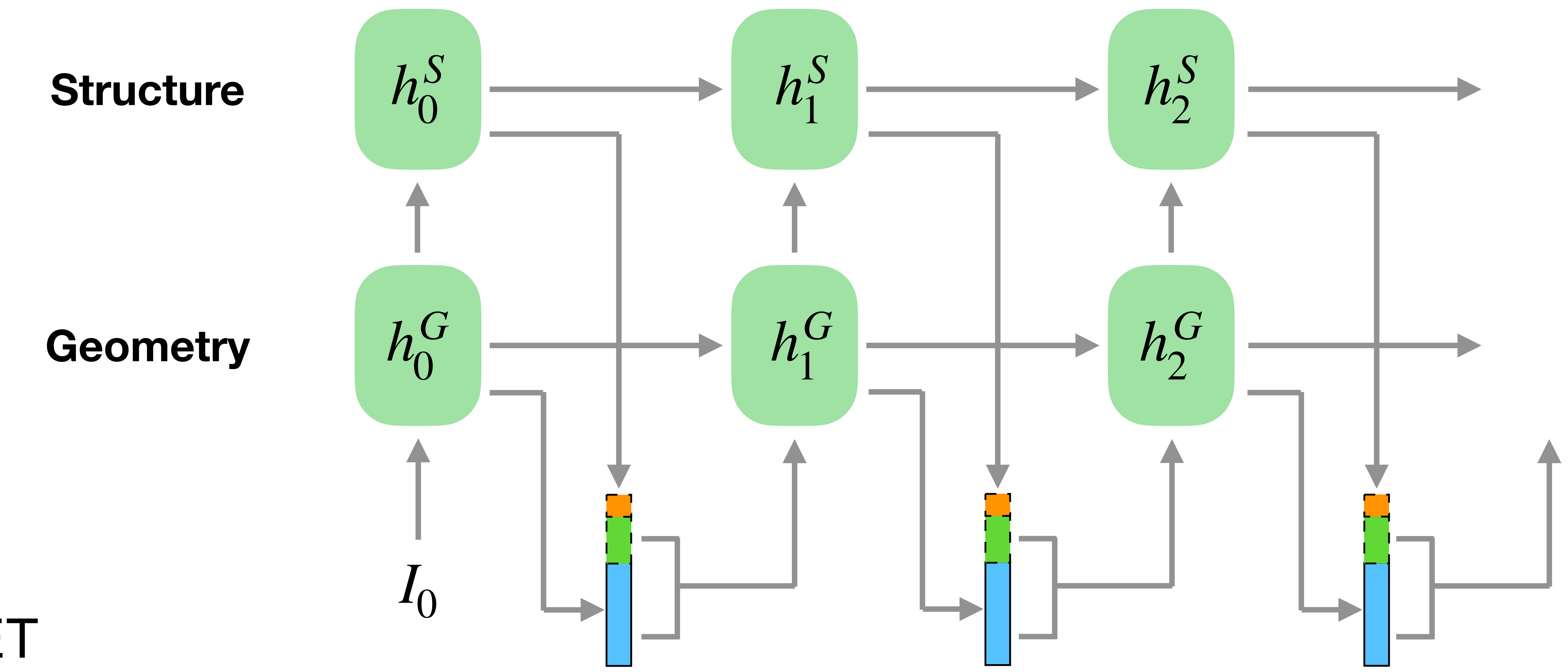
GRU Cell

I_0 Initial input: zero vector

Stop sign: a confidence value between 0~1

Part Box Parameter : 6D, position + size

Part Geometry Feature : latent vector to be decoded by IM-NET



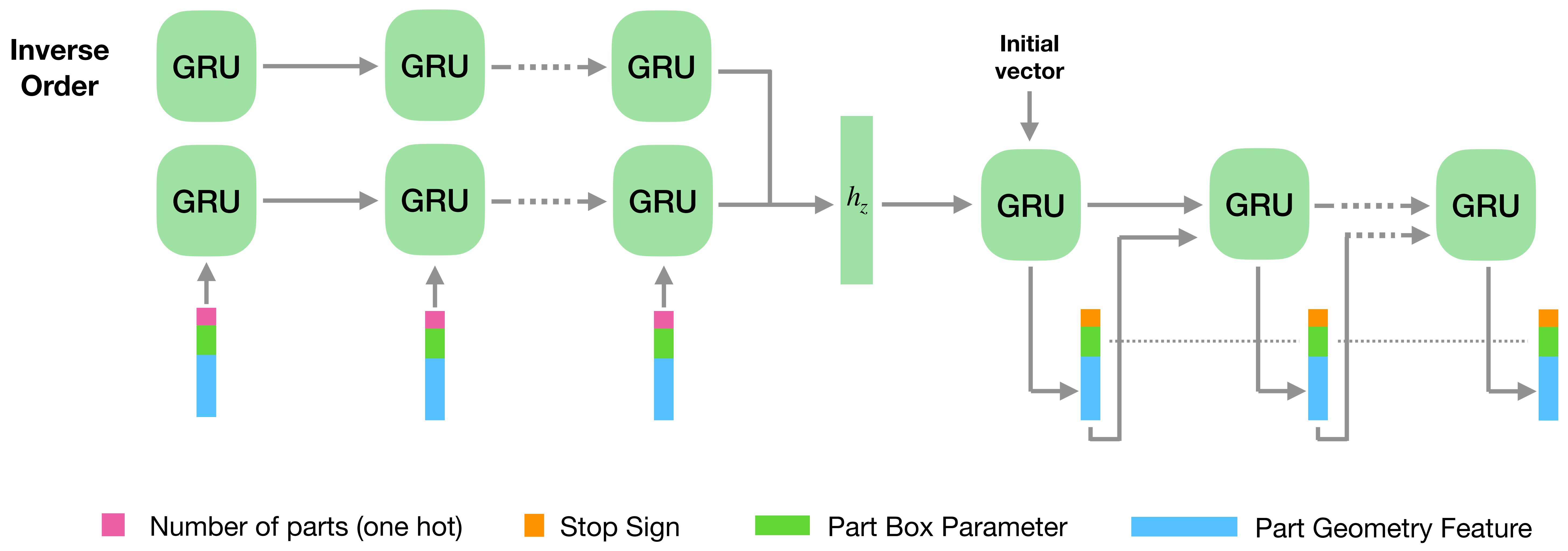
Method - Seq2Seq AE

Training losses

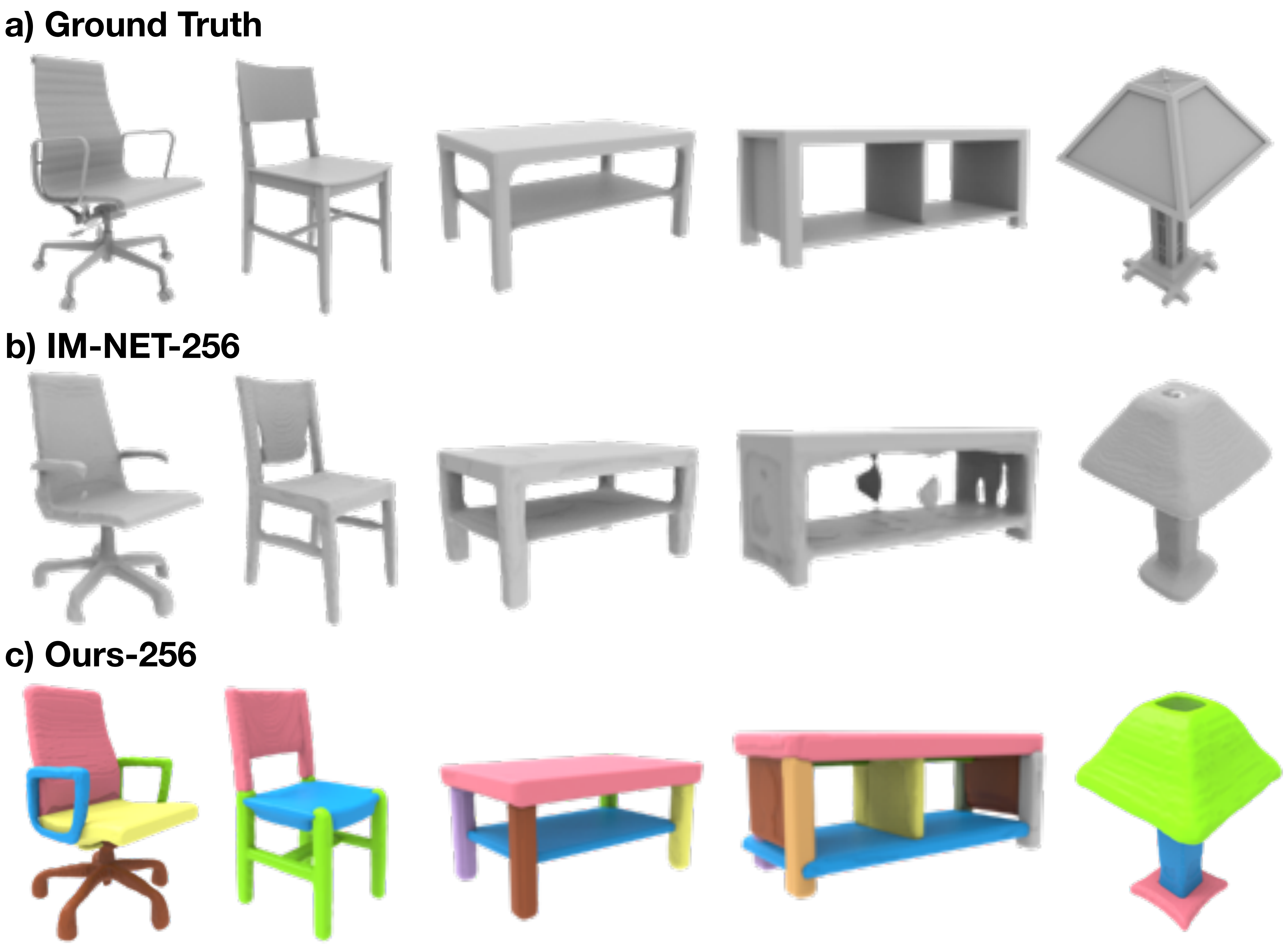
- *MSE* loss on the reconstruction of geometry feature g_i and structure feature b_i
- *Binary Cross Entropy* loss on the stop sign s_i predicted by decoder

$$\mathcal{L}_{\mathrm{r}}(S) = \frac{1}{k}\sum_{i=1}^{k}[\beta||g_i' - g_i||_2 + ||b_i' - b_i||_2]$$

$$\mathcal{L}_{\mathrm{stop}}(S) = \frac{1}{k}\sum_{i=1}^{k}[-s_i \log s_i' - (1 - s_i)\log(1 - s_i')]$$



Results : shape auto-encoding



Metrics	Method	Chair	Table	Lamp
IoU	Ours-64	**67.29**	47.39	39.56
	IM-NET-64	62.93	**56.14**	**41.29**
CD	Ours-64	3.38	5.49	11.49
	Ours-256	2.86	5.69	10.32
	Ours-Cross-256	**2.46**	**4.50**	**4.87**
	IM-NET-64	3.64	6.75	12.43
	IM-NET-256	3.59	6.31	12.19
LFD	Ours-64	2734	2824	6254
	Ours-256	**2441**	2609	5941
	Ours-Cross-256	2501	**2415**	**4875**
	IM-NET-64	2830	3446	6262
	IM-NET-256	2794	3397	6622

Results : shape generation

Category	Method	COV	MMD	JSD
Chair	Ours	**54.91**	8.34	**0.0083**
	IM-NET	52.35	**7.44**	0.0084
	StructureNet	29.51	9.67	0.0477
Table	Ours	56.51	7.56	0.0057
	IM-NET	**56.67**	**6.90**	**0.0047**
	StructureNet	16.04	14.98	0.0725
Lamp	Ours	**87.95**	**10.01**	**0.0215**
	IM-NET	81.25	10.45	0.0230
	StructureNet	35.27	17.29	0.1719

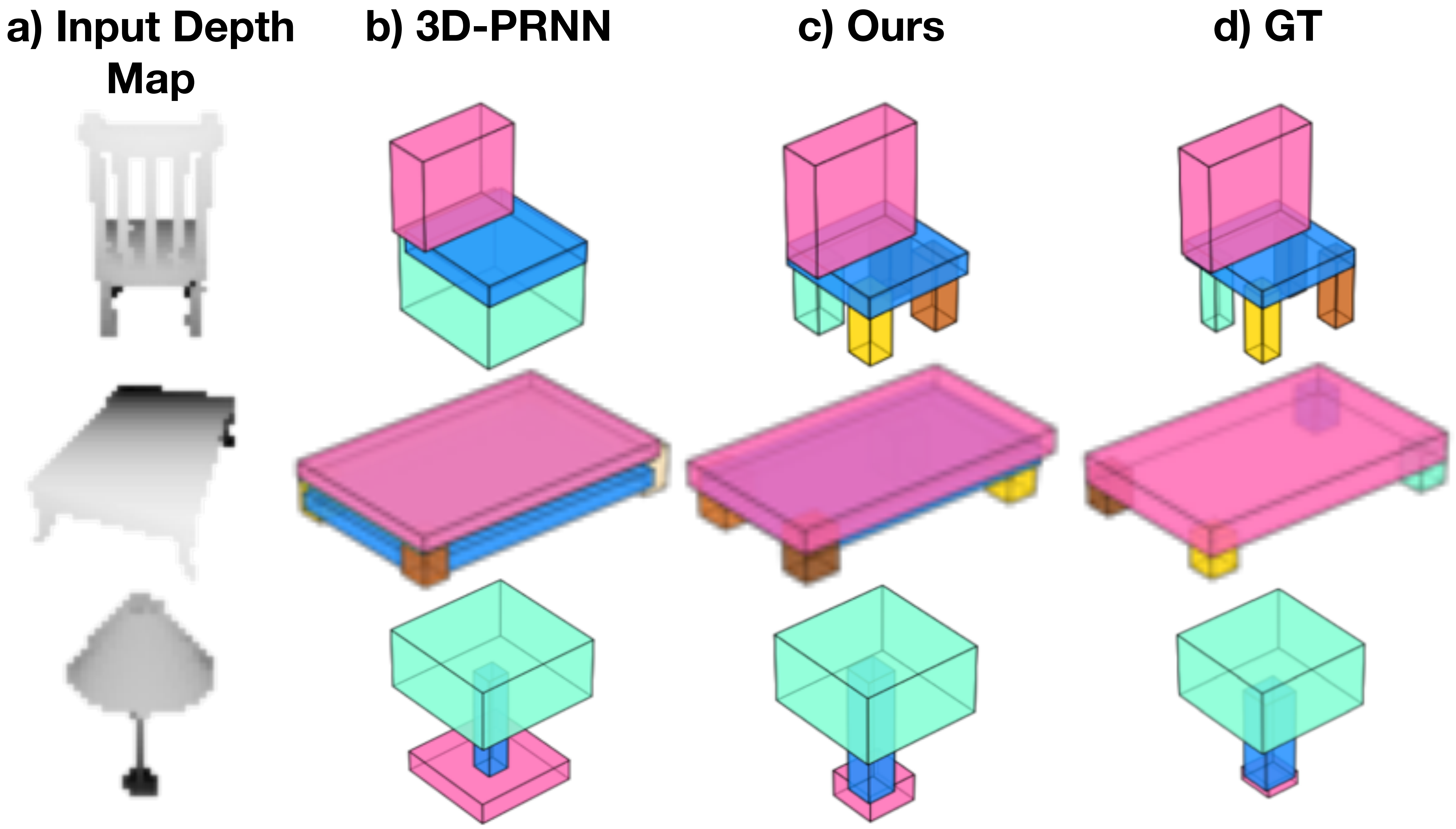
Results : shape generation

Results : latent space interpolation

Results : single view reconstruction

Results : comparison to 3D-PRNN

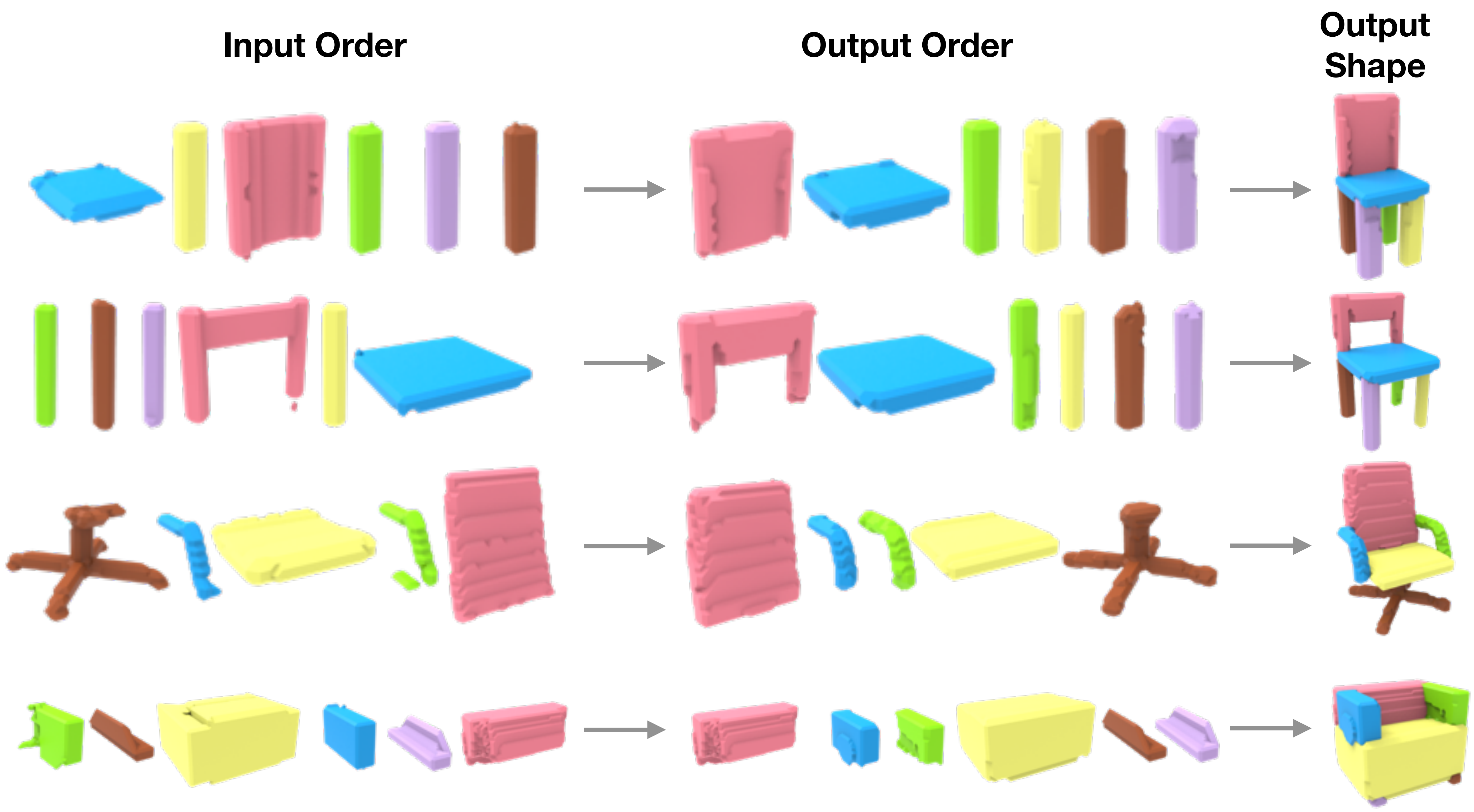
- Shape reconstruction from single depth image
- Compare on two orders: (A) PartNet default (B) enforced top-down



Method	Order	Chair	Table	Lamp	Average
Ours	A	**61.47**	**53.67**	**52.94**	**56.03**
	B	58.68	48.58	52.17	53.14
3D-PRNN	A	37.26	51.30	47.26	45.27
	B	36.46	51.93	43.83	44.07

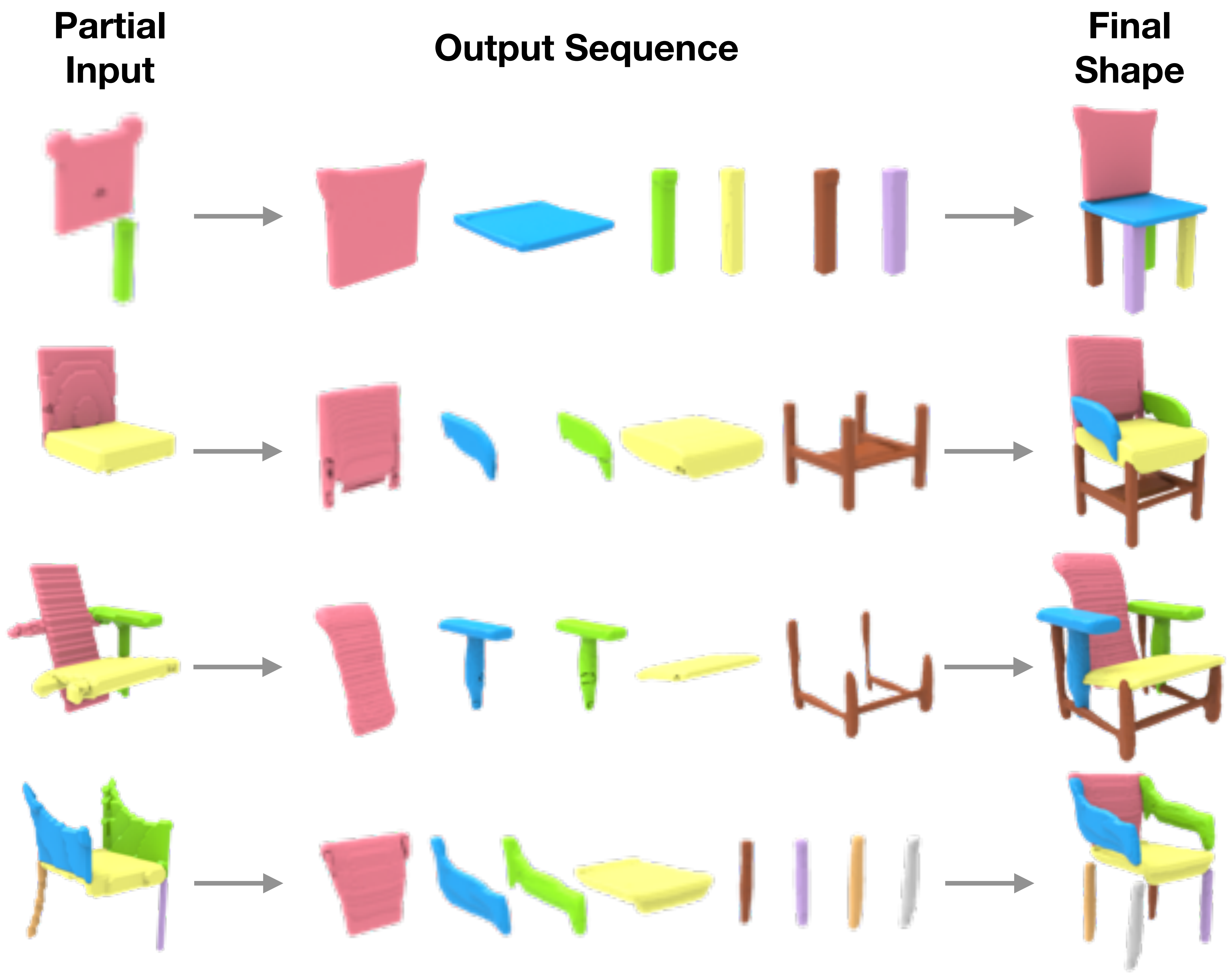
Results : applications

- Order denosing and part correspondence
  - Re-train the model the correct the input order



Results : applications

- Partial shape completion
  - Re-train the model to reconstruct from partial shape input



Limitation

- PQ-NET do not produce part relations
  - Comparing to prior works that seek to hierarchical representation
- The order of parts could affect the performance
  - A consistent part order over the dataset is required

CVPR SEATTLE WASHINGTON JUNE 16-18 2020

Thanks!

Code and data: https://github.com/ChrisWu1997/PQ-NET