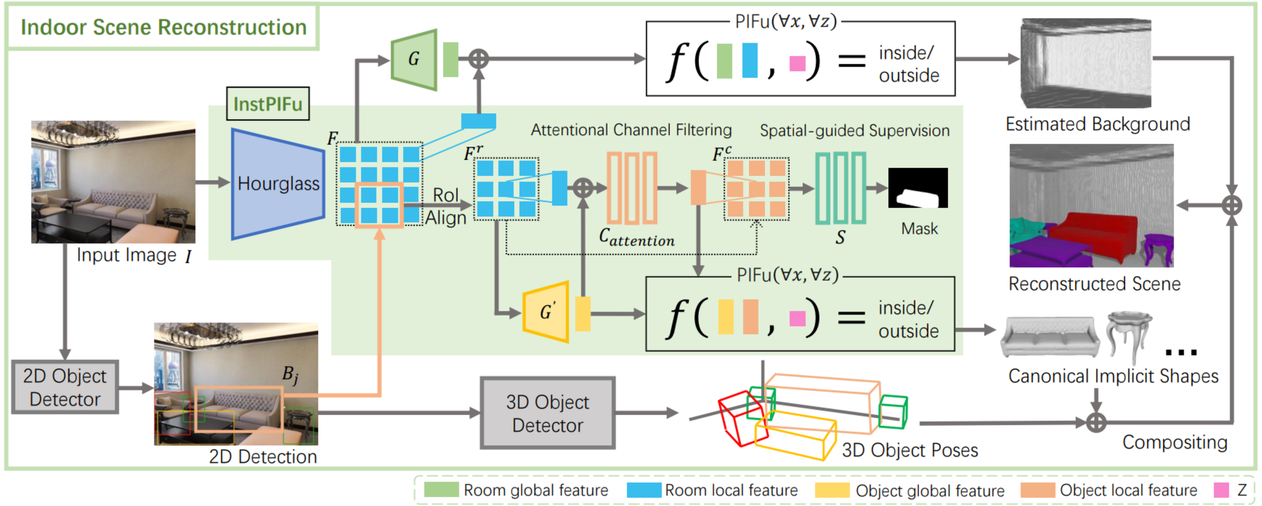
Towards High-Fidelity Single-view Holistic Reconstruction of Indoor Scenes

Haolin Liu*, Yujian Zheng*, Guanying Chen, Shuguang Cui, and Xiaoguang Han✉
*Joint first authors
✉Contact: hanxiaoguang@cuhk.edu.cn

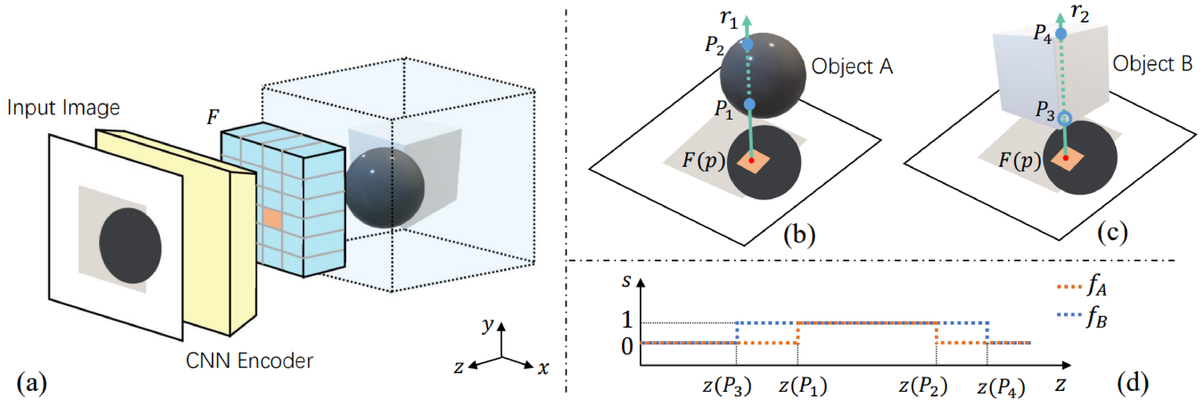
香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

FNii 未来智联网络研究院



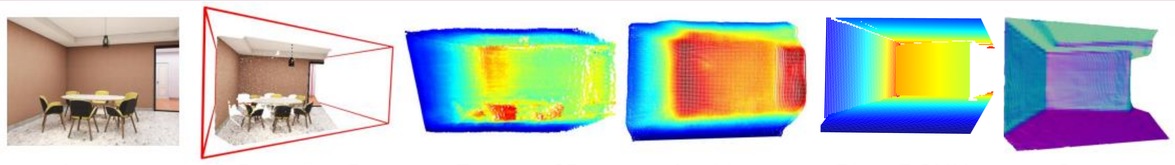
- The first system that uses **pixel-aligned feature** to recover the 3D indoor scene from a single view
- First to attempt to reconstruct the room **background** via **implicit** representation
- Better to **solve the occlusion** between indoor objects by introducing **InstPIFu**

Issue of occlusion



Occlusion causes **local feature ambiguity** among different objects. (a) A scene contains two objects, and F is the extracted local feature from the image. (b)-(c) Object reconstruction in canonical coordinate system, where points along the rays are projected at p to sample local feature $F(p)$. (d) Variations of occupancy s with depth z along the ray r_1 and r_2 for f_A and f_B.

Background recon.



Method	Factored3D [49]	Adabins [1]	Im3D [57]	PlaneRCNN [27]	Ours
CD on 3D-FRONT ↓	0.697	0.573	1.974	0.717	**0.481**

Object recon.

From left to right of every quintuplet: Input images, results from MGN, LDIF, Ours, and Ground truth.

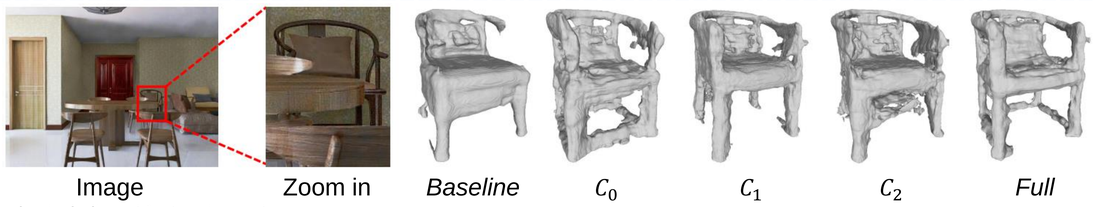
3D-FUTURE

Method	mean ↓ / ↑
MGN [34]	**14.07** / 55.64
LIEN [56]	28.52 / 45.63
Ours	14.46 / **61.32**

Pix3D

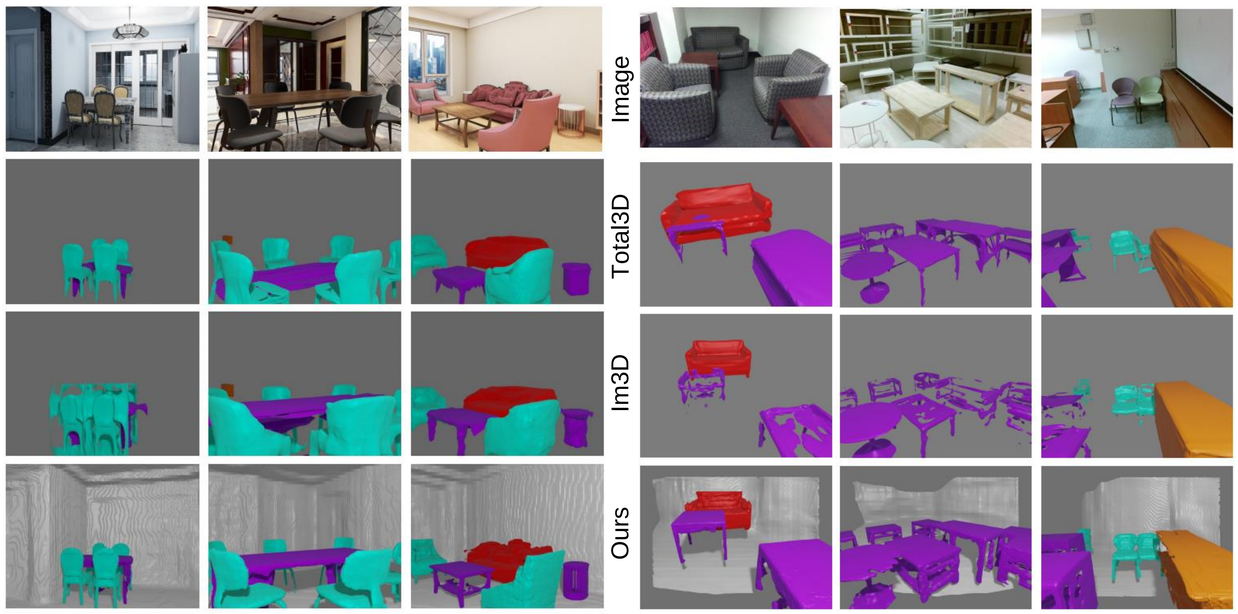
Split in [34]	mean ↓ / ↑
MGN [34]	6.84 / **73.18**
LIEN [56]	**6.72** / 63.96
Ours	13.60 / 56.07
Non-overlapped Split	mean ↓ / ↑
MGN [34]	44.32 / 36.20
LIEN [56]	51.31 / 31.45
Ours	**24.65** / **45.62**

Ablation



Baseline: only pixel-aligned feature is used.
C_0 : *Baseline* + global instance feature.
C_1 : C_0 + attentional channel filtering.
C_2 : C_0 + spatial-guided supervision.
Full: C_1 + C_2 .

Method	***Baseline***	C_0	C_1	C_2	***Full***
CD ↓	17.95	16.42(-1.53)	15.54(-2.41)	15.28(-2.67)	**14.46**(-3.49)
F-Score ↑	56.98	58.62(+1.64)	60.23(+3.25)	60.56(+3.58)	**61.32**(+4.34)



Scene reconstruction results on 3D-FRONT

Scene reconstruction results on SUN RGB-D

ECCV EUROPEAN CONFERENCE ON COMPUTER VISION TEL AVIV 2022