



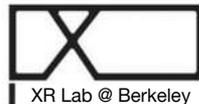
Generative Modeling in Spatial Computing

Mohammad Keshavarzi

Ph.D Candidate

FHL Vive Center for Enhanced Reality/ XR Lab

University of California, Berkeley







IMG: Angry Birds AR



actual footage shot through HoloLens











?



?



Contextual Spatial Computing



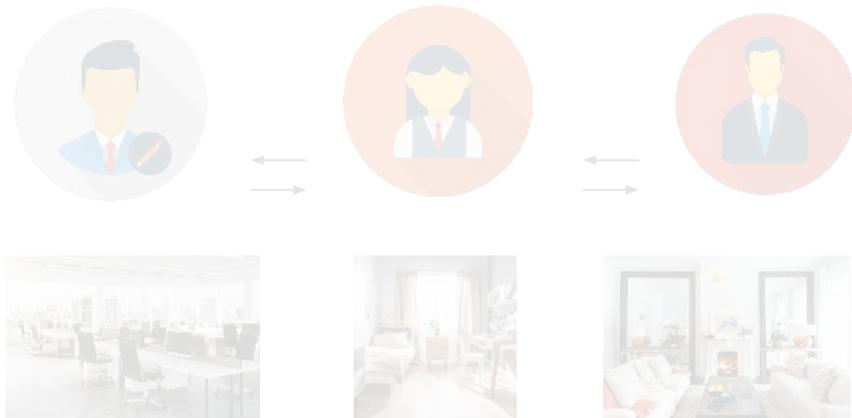
Content Augmentation



Mutual Space



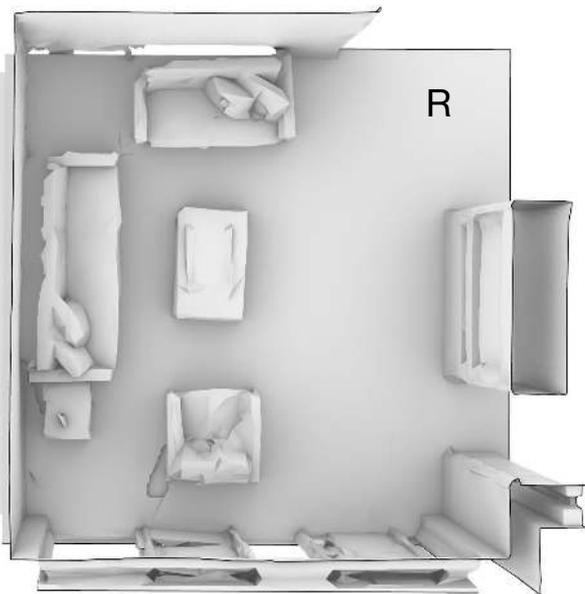
Content Augmentation



Mutual Space

Content Augmentation

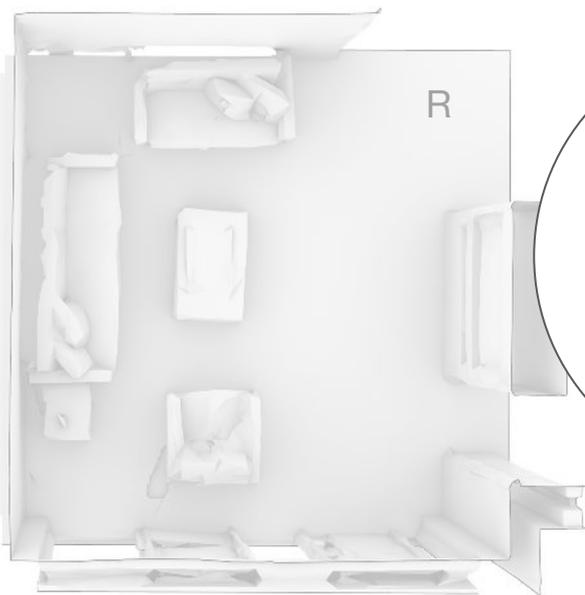
Constrained Placement



Target Scene Input



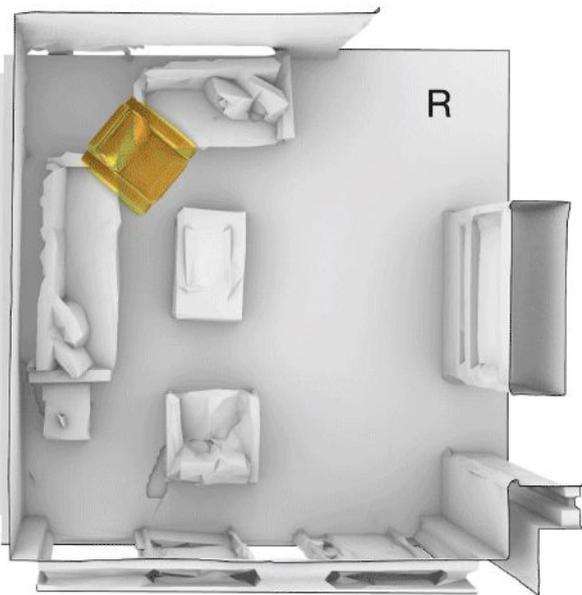
Constrained Placement



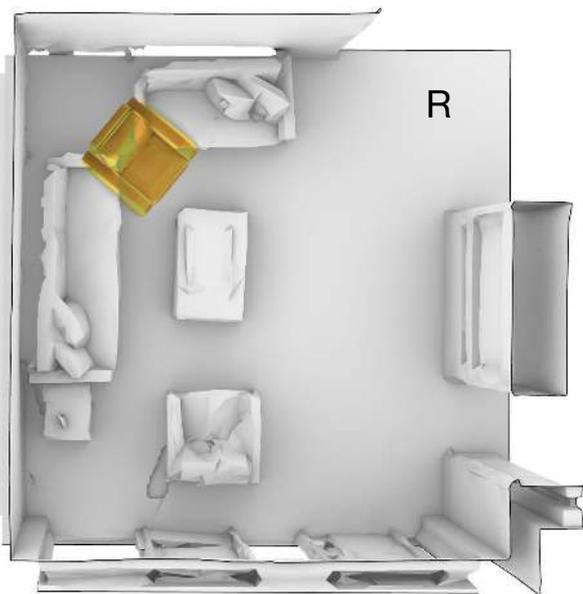
Target Scene Input



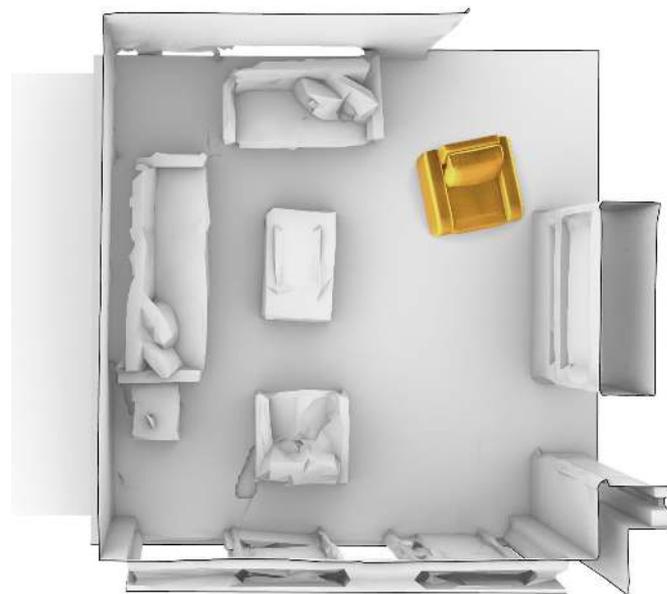
Constrained Placement



Constrained Placement

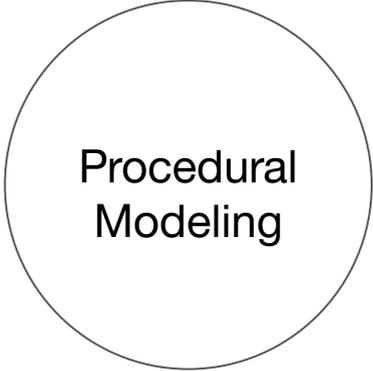


Conflicting Placement

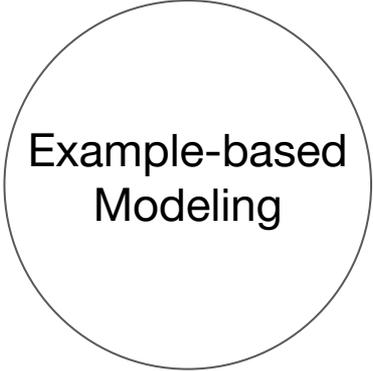


Inconsistent with context

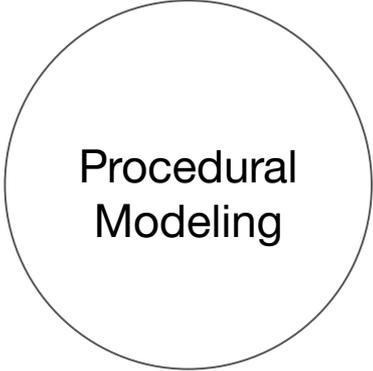
Background



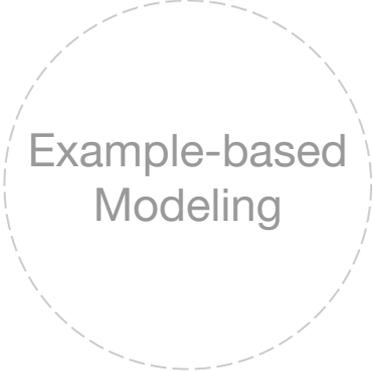
Procedural
Modeling



Example-based
Modeling



Procedural
Modeling



Example-based
Modeling

Procedural Modeling



RoboRaid, Microsoft HoloLens Demo

Procedural Modeling

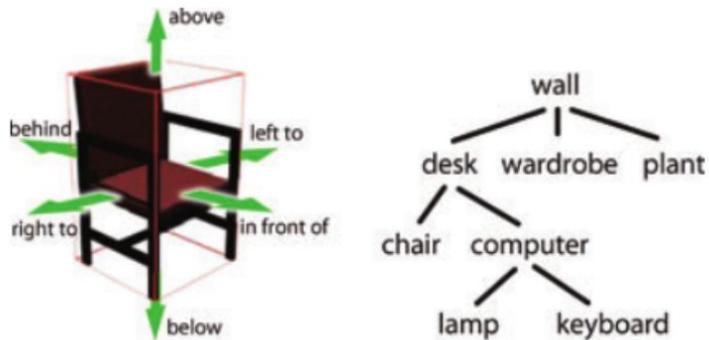
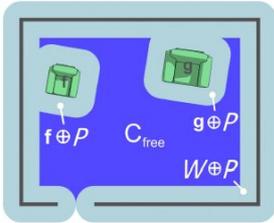
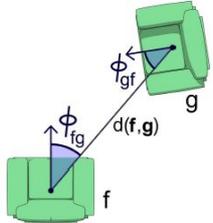


Figure 1: *Left: Objects are arranged based on their oriented bounding boxes (red) and their spatial relations (green). Right: Example of a hierarchy build by the agents.*

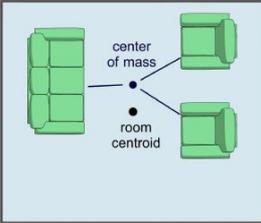
Procedural Modeling



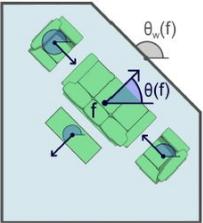
(a) circulation



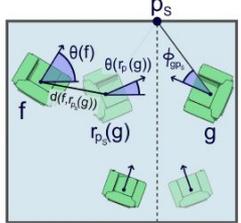
(b) conversation



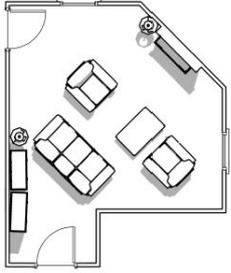
(c) visual balance



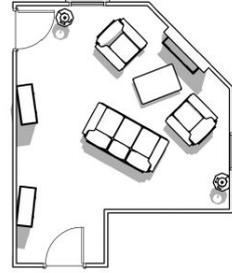
(d) alignment



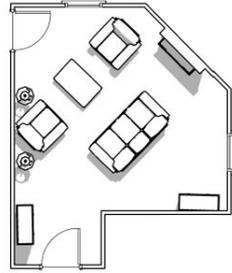
(e) emphasis



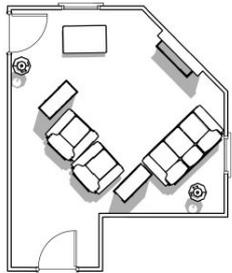
(a) Clearance and reachability term excluded



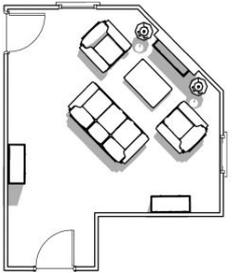
(b) Alignment term excluded



(c) Emphasis term excluded



(d) Conversation and pairwise terms excluded



(e) All terms included

Merrell, Paul, Eric Schkufza, Zeyang Li, Maneesh Agrawala, and Vladlen Koltun. "Interactive furniture layout using interior design guidelines." *ACM transactions on graphics (TOG)* 30, no. 4 (2011): 1-10.

Procedural
Modeling

Example-based
Modeling

Example-based Scene Synthesis

Input Scenes



Database

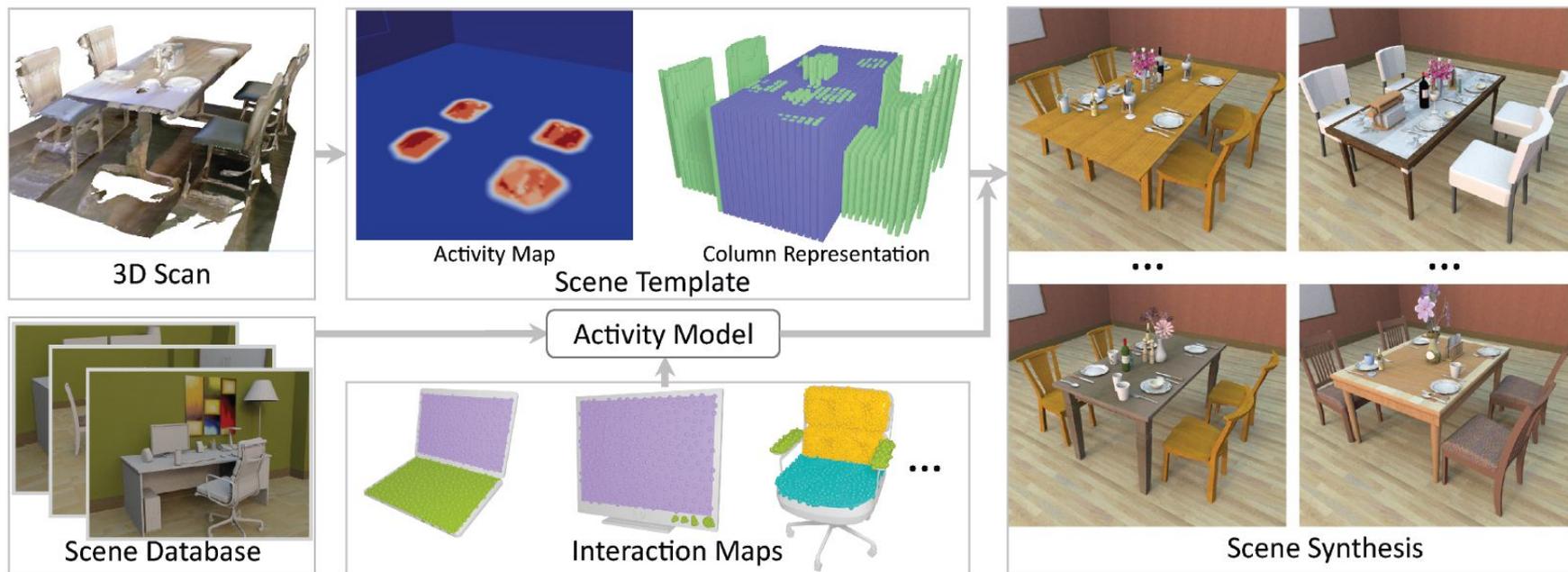


Synthesized Results



Fisher, Matthew, Daniel Ritchie, Manolis Savva, Thomas Funkhouser, and Pat Hanrahan. "Example-based synthesis of 3D object arrangements." *ACM Transactions on Graphics (TOG)* 31, no. 6 (2012): 1-11.

Example-based Scene Synthesis



Fisher, Matthew, Manolis Savva, Yangyan Li, Pat Hanrahan, and Matthias Nießner. "Activity-centric scene synthesis for functional 3D scene modeling." *ACM Transactions on Graphics (TOG)* 34, no. 6 (2015): 1-13.

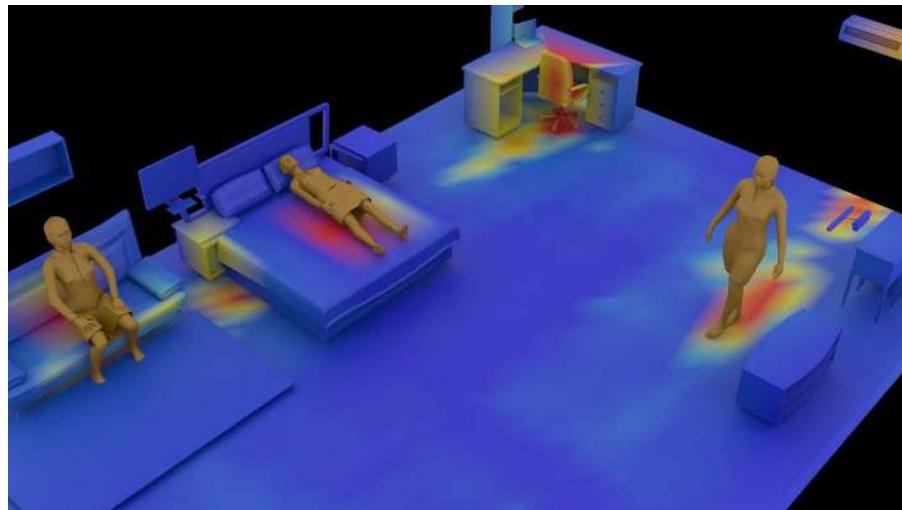
Example-based Scene Synthesis



(g) kitchen

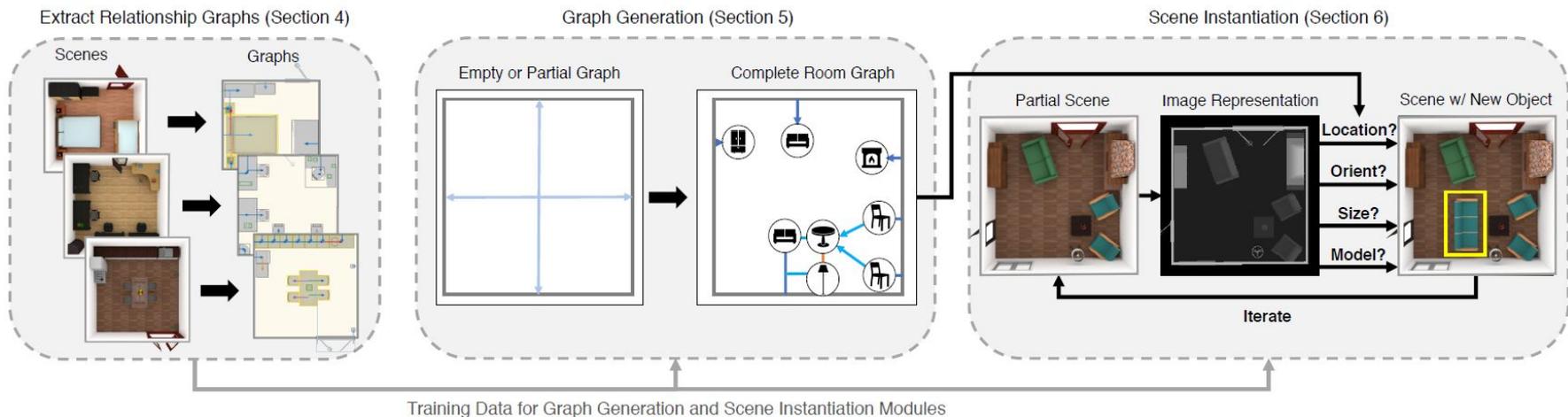
(h) living room

(i) office



Qi, Siyuan, Yixin Zhu, Siyuan Huang, Chenfanfu Jiang, and Song-Chun Zhu. "Human-centric indoor scene synthesis using stochastic grammar." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5899-5908. 2018.

Example-based Scene Synthesis



Planit: Planning and instantiating indoor scenes with relation graph and spatial prior networks; Wang, Kai, Yu-An Lin, Ben Weissmann, Manolis Savva, Angel X. Chang, and Daniel Ritchie. *ACM Transactions on Graphics (TOG)* 38, no. 4 (2019): 1-15.

Datasets

Synthetic Data (SUNCG)



Scanned Data (MatterPort3D)



3D Object Detection (PerspectiveNet)

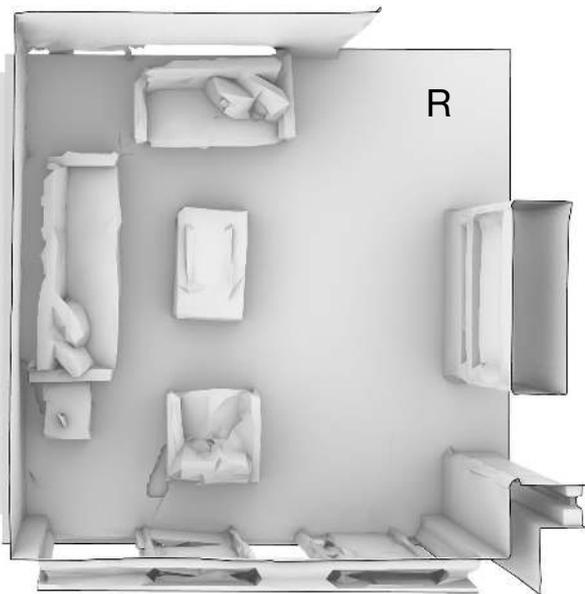


SceneGen:

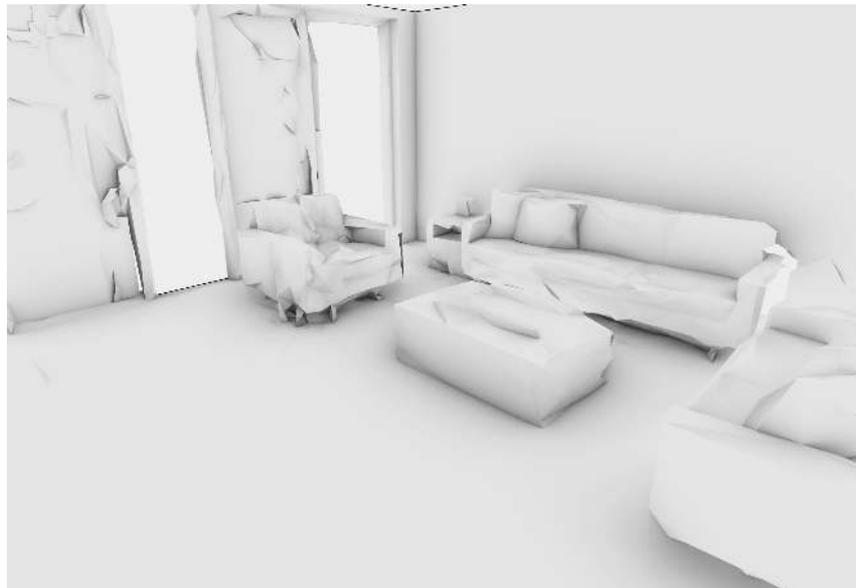
Generative Contextual Scene Augmentation using Scene Graph Priors

Mohammad Keshavarzi, Aakash Parikh, Xiyu Zhai, Melody Mao, Luisa Caldas, Allen Y. Yang

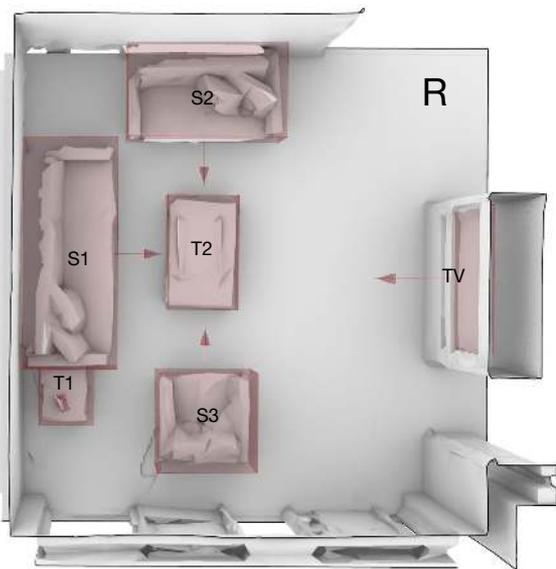
Spatial Scene Representation



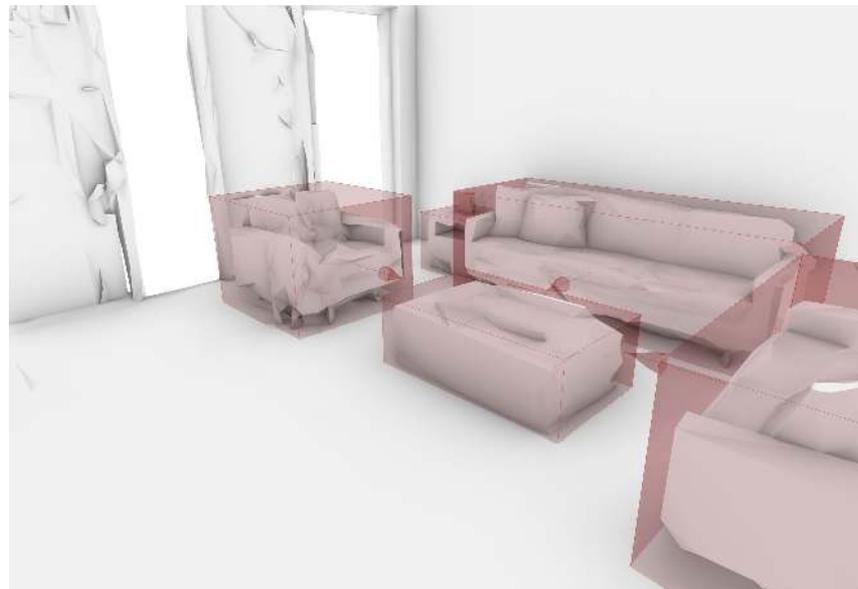
Target Scene Input



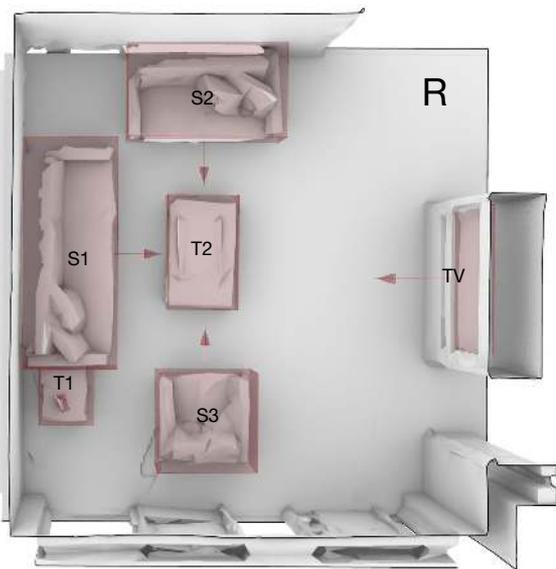
Spatial Scene Representation



Target Scene Input



Spatial Scene Representation



Target Scene Input

(S1) SOFA 1

(S2) SOFA 2

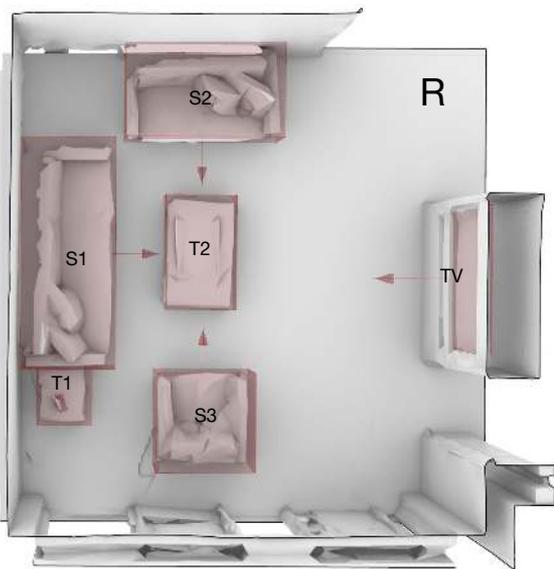
(S3) SOFA 3

(T1) TABLE 1

(T2) TABLE 2

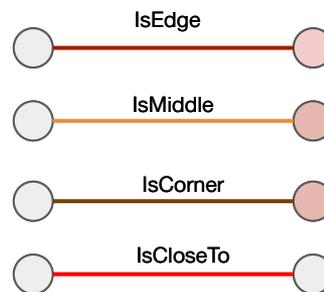
(TV) TV

Spatial Scene Representation

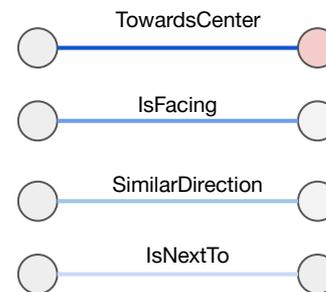


Target Scene Input

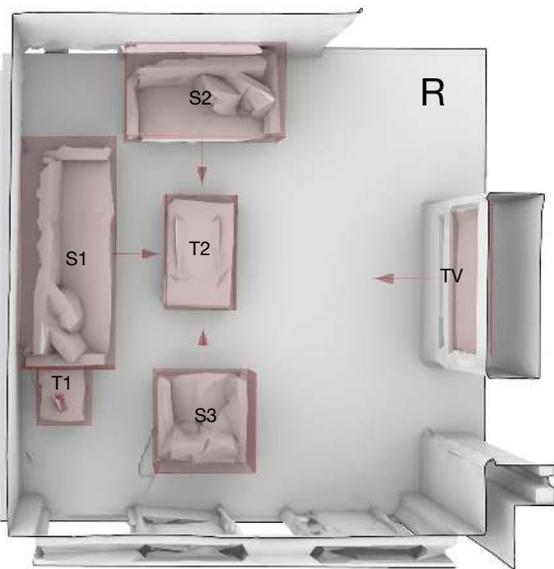
Positional



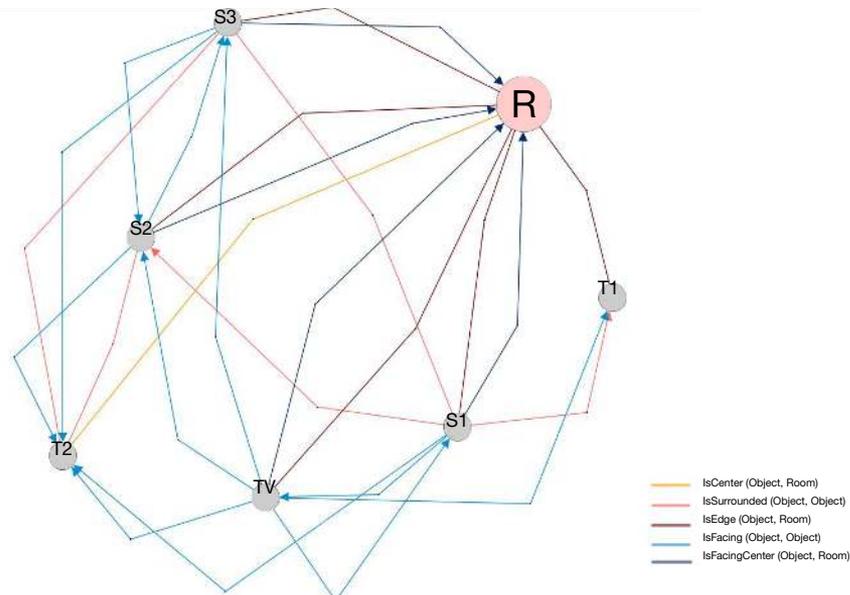
Orientalional



Spatial Scene Representation

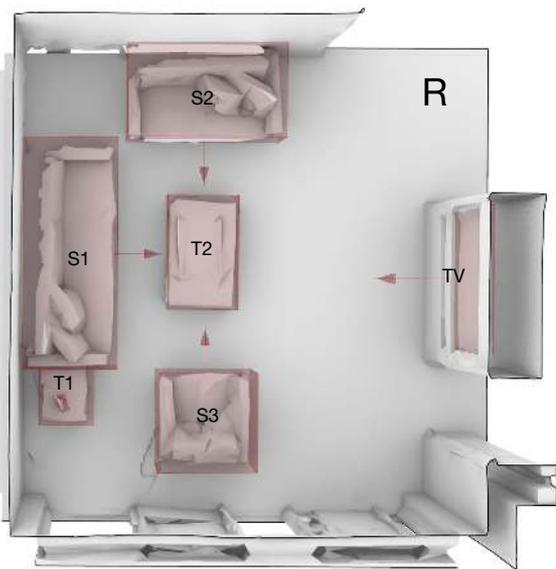


Target Scene Input

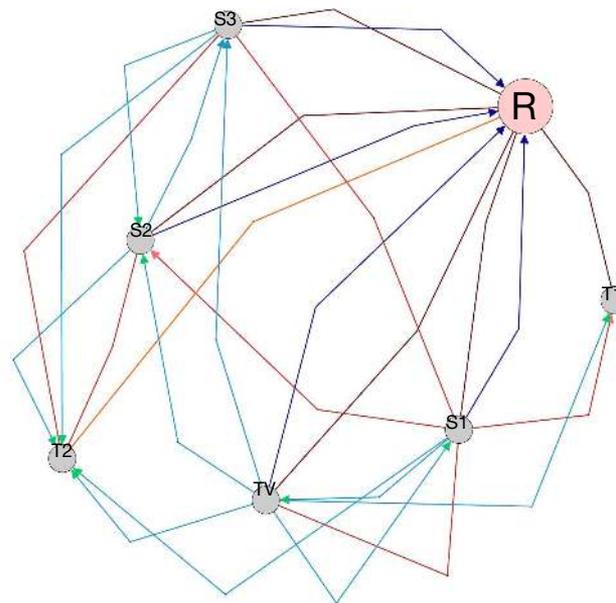


Spatial Representation Graph

Spatial Scene Representation

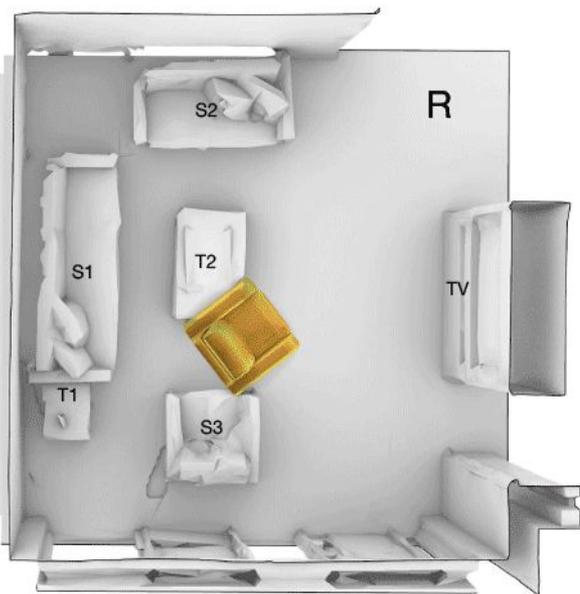


Target Scene Input

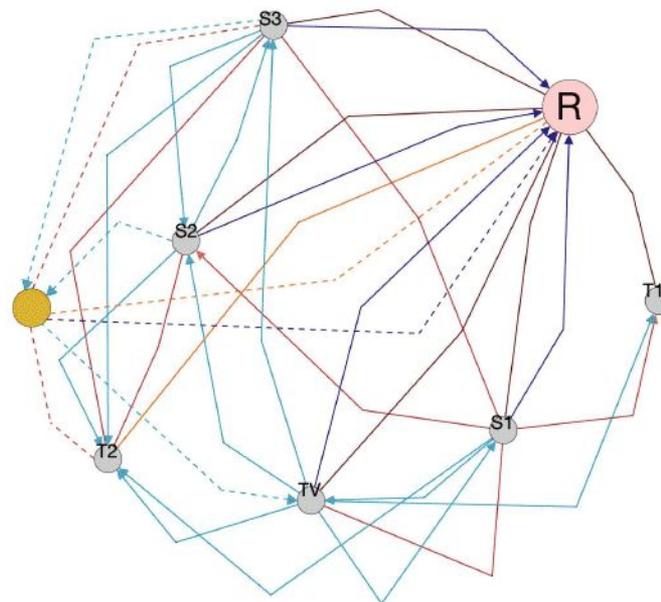


Spatial Representation Graph

Spatial Scene Representation

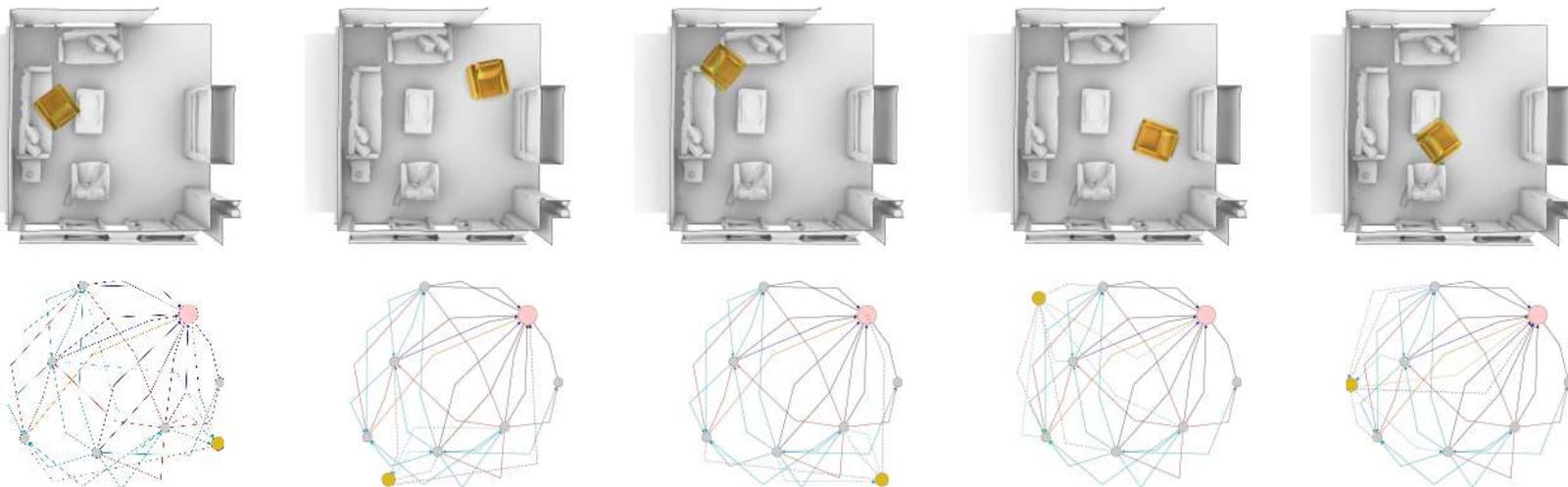


Target Scene Input

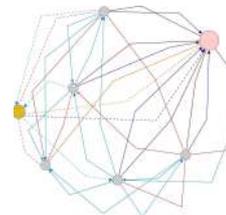
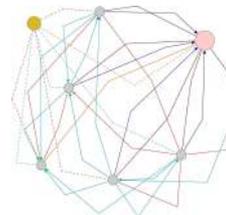
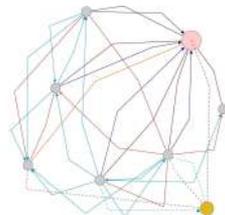
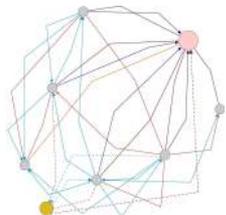
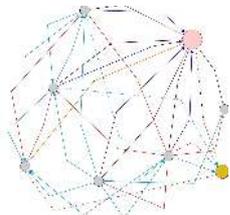
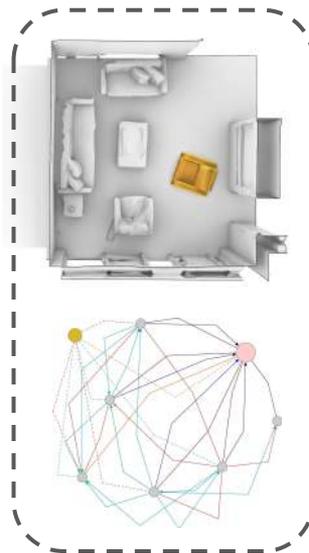
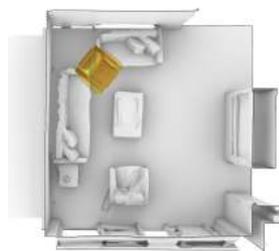


Spatial Representation Graph

Spatial Scene Representation



Spatial Scene Representation



Knowledge Model

Dataset

Datasets

Synthetic Data (SUNCG)



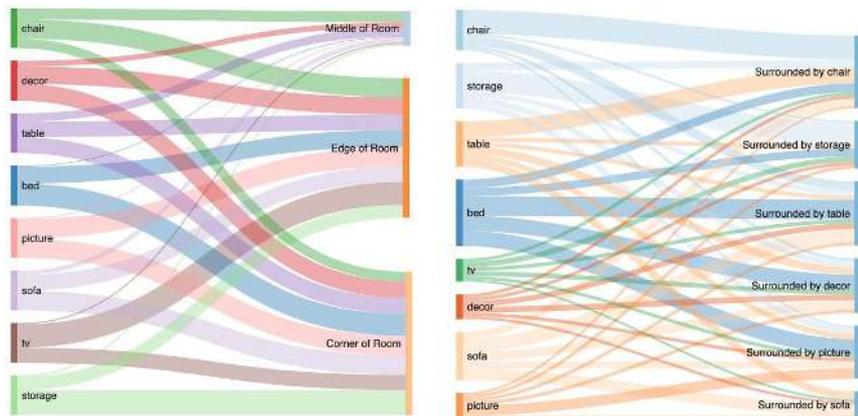
Scanned Data (MatterPort3D)



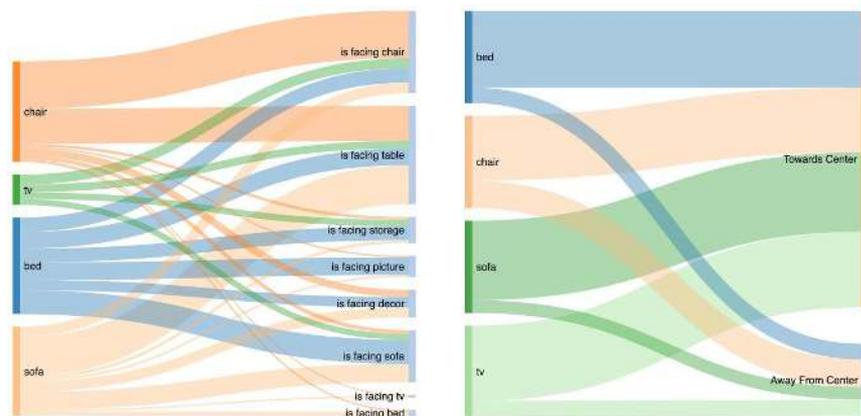
3D Object Detection (PerspectiveNet)



Knowledge Model



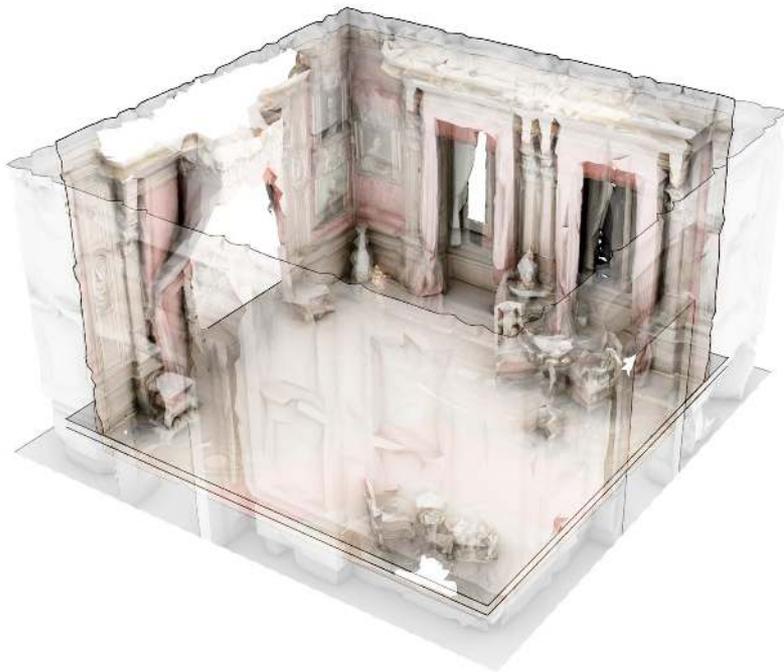
Positional Relationships



Orientalional Relationships

Experiments

Object Placement



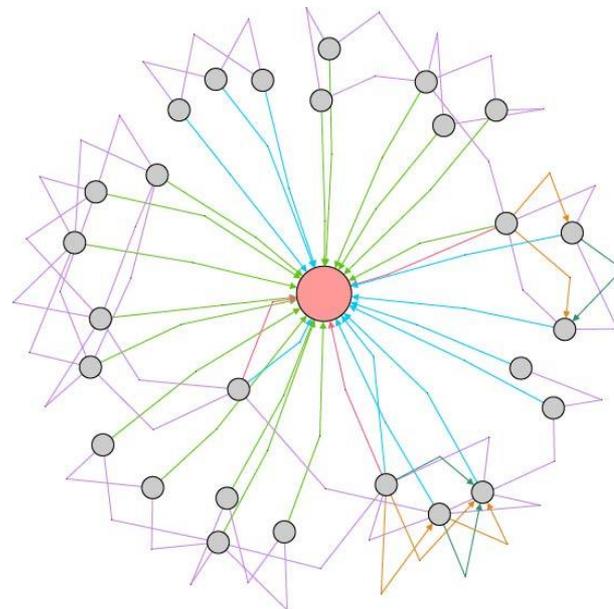
Object Placement



Object Placement



Target Scene



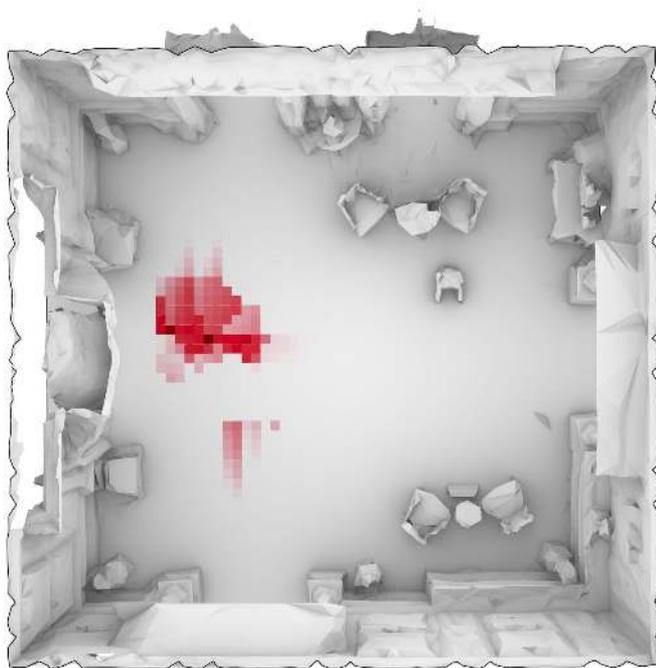
Spatial Representation Graph

IsCenter (Object, Room)
IsSurrounded (Object, Object)
IsEdge (Object, Room)
IsFacing (Object, Object)
IsFacingCenter (Object, Room)

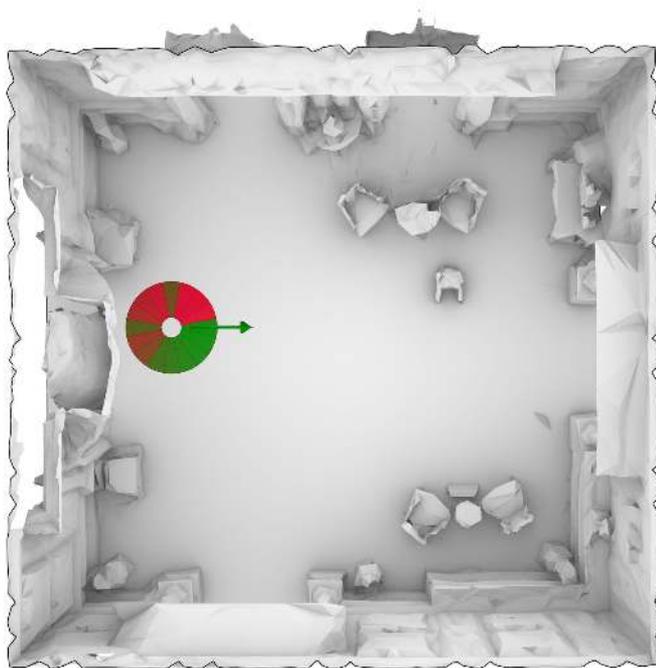
Object Placement



Object Placement



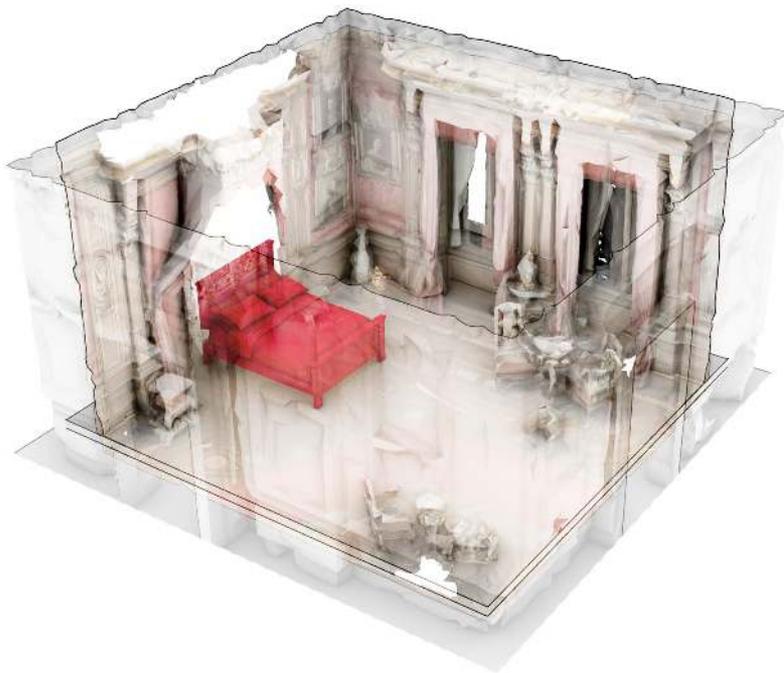
Object Placement



Object Placement



Object Placement



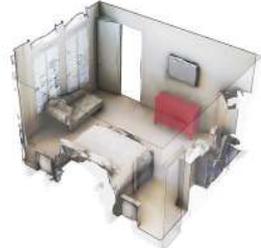
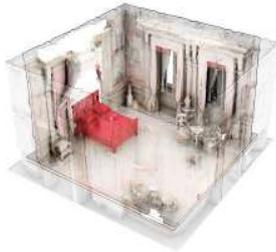
Target scene



SceneGen placement



SceneGen placement



(a)

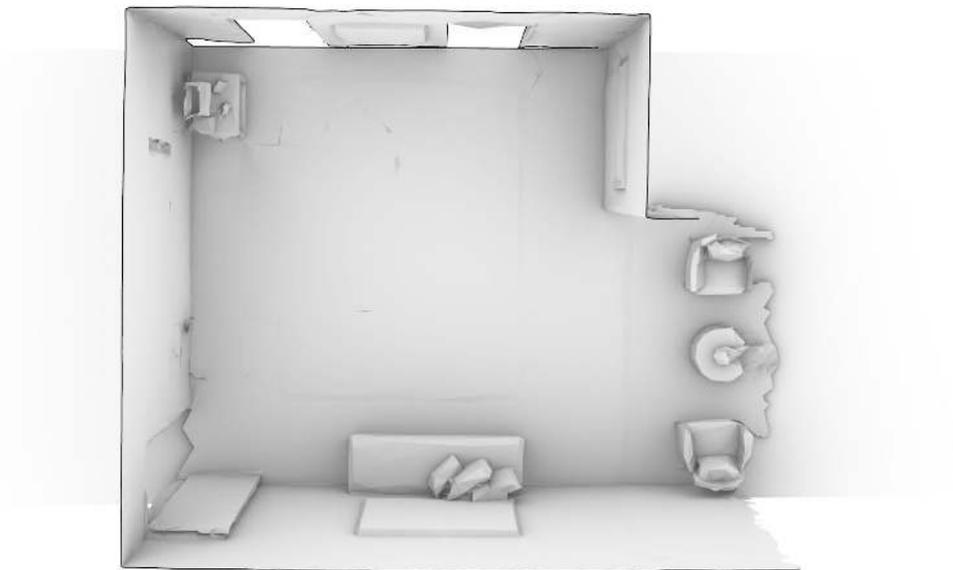
(b)

(c)

(d)

(e)

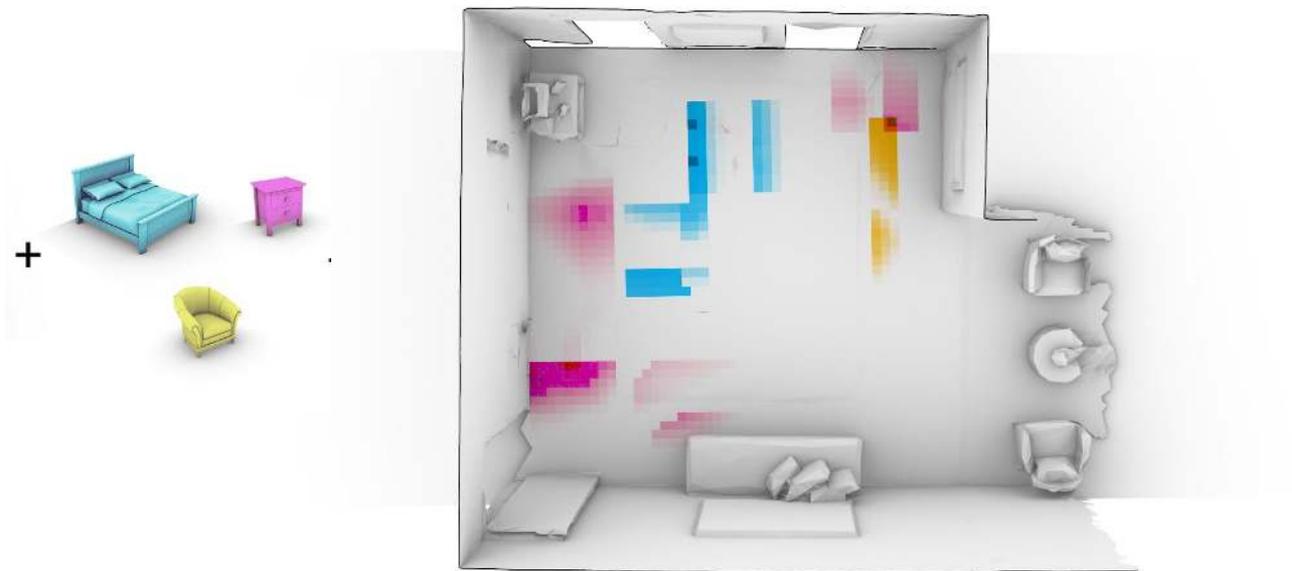
Multi-object Placement



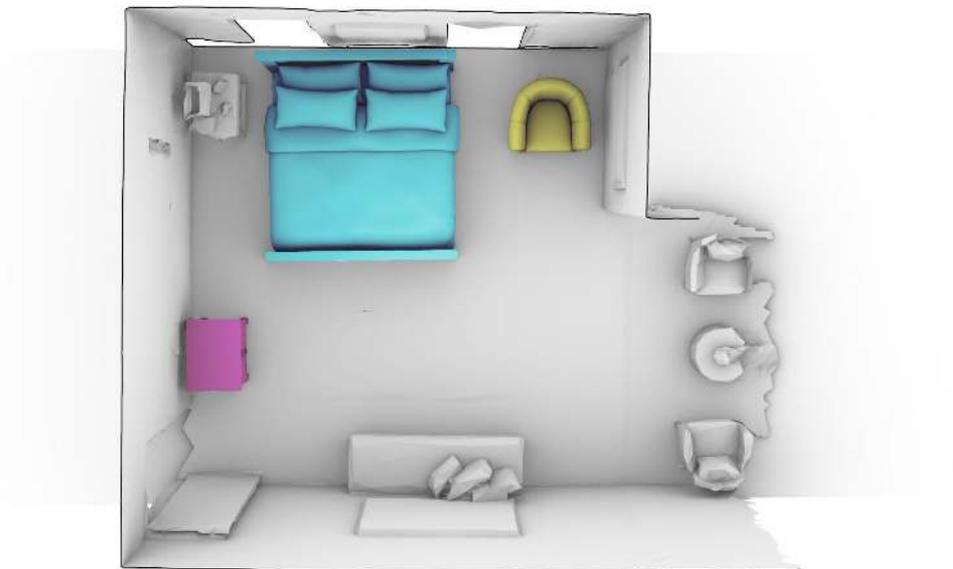
Multi-object Placement



Multi-object Placement



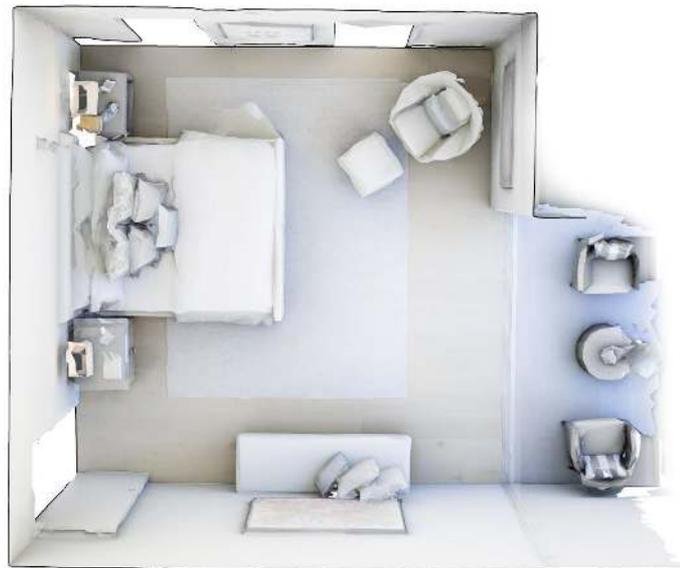
Multi-object Placement



Multi-object Placement

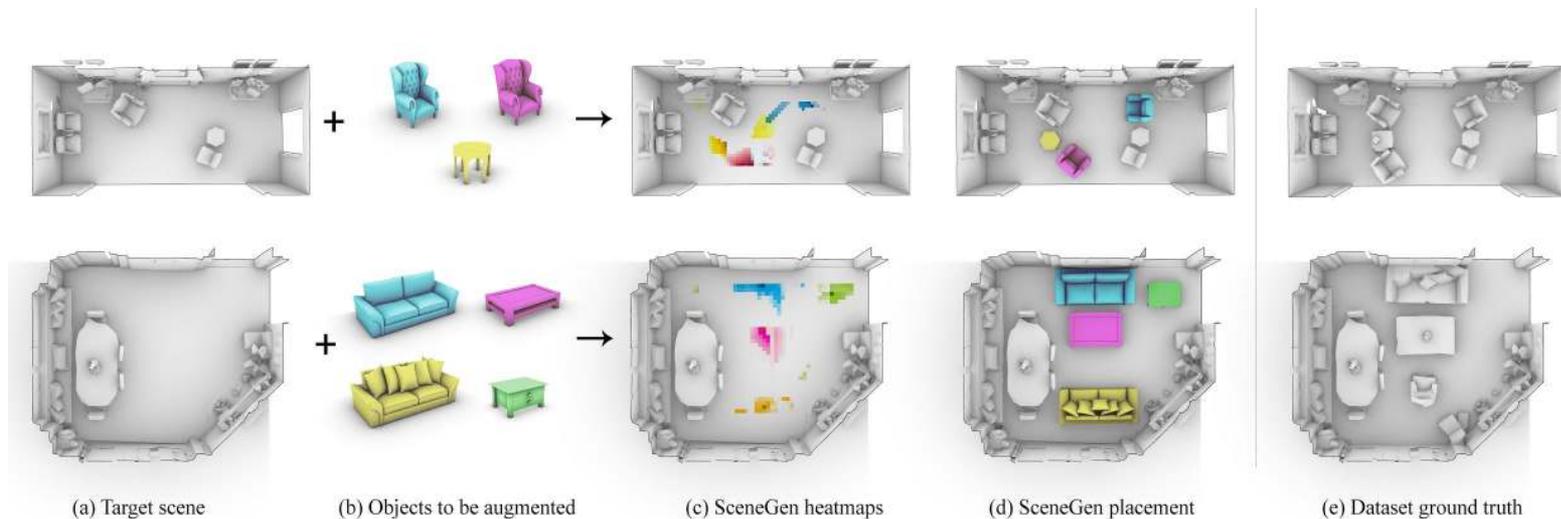


SceneGen



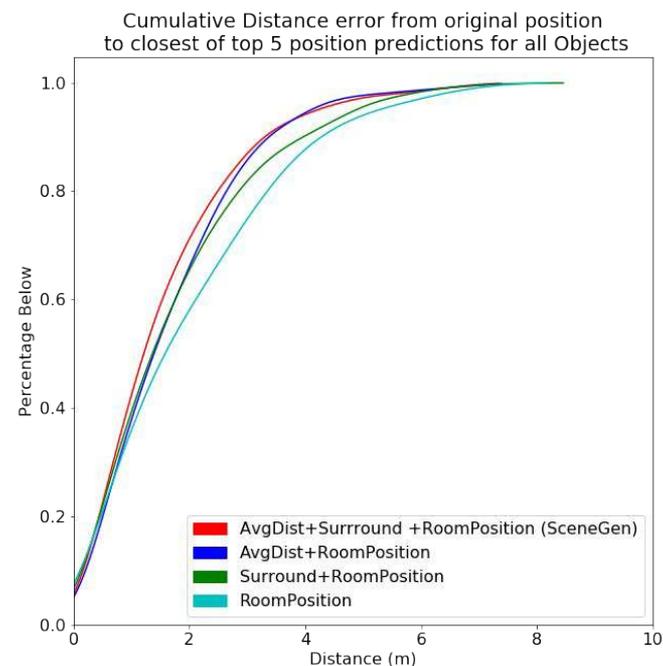
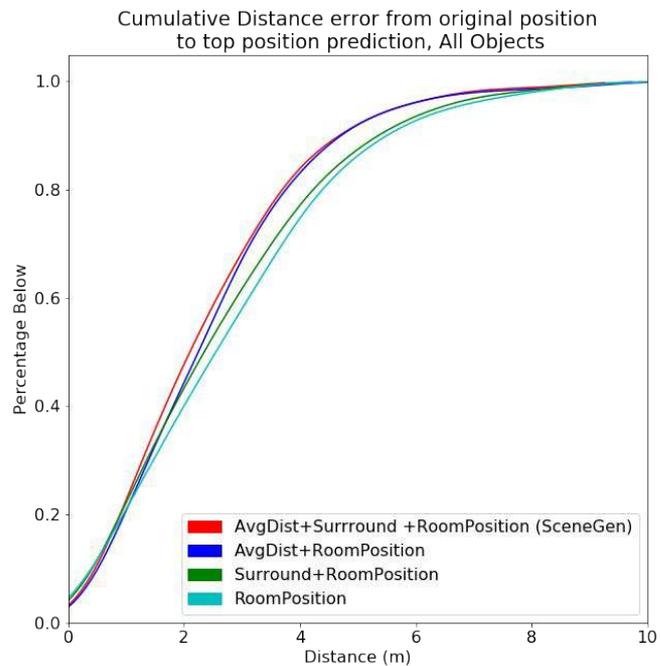
GroundTruth

Multi-object Placement



Numerical Experiments

Numerical Experiments



User Experiments

User Experiments



(a) Original MatterPort3D scene

USER STUDY
→



(b) Reconstructed simplified scene



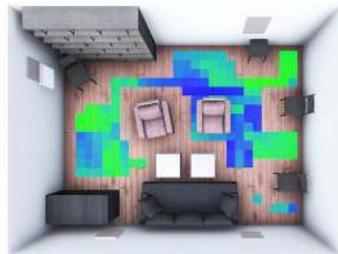
(c) Level 1 (Random)



(c) Level 2 (Open Random)



(e) Level 3 (SceneGen)

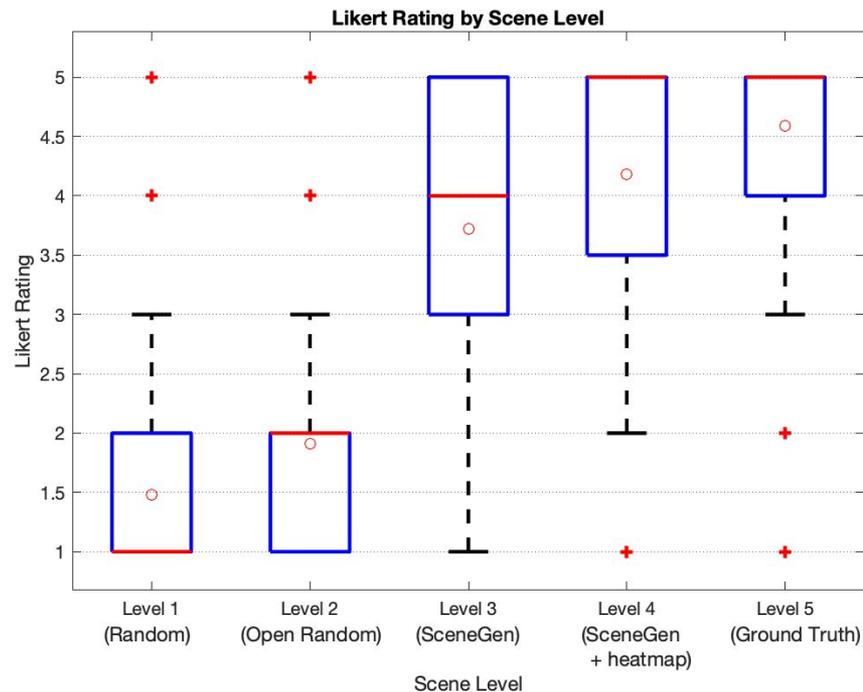
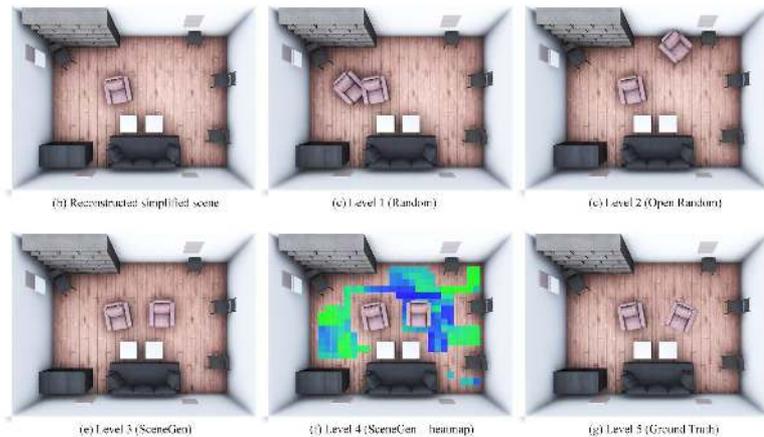


(f) Level 4 (SceneGen + heatmap)



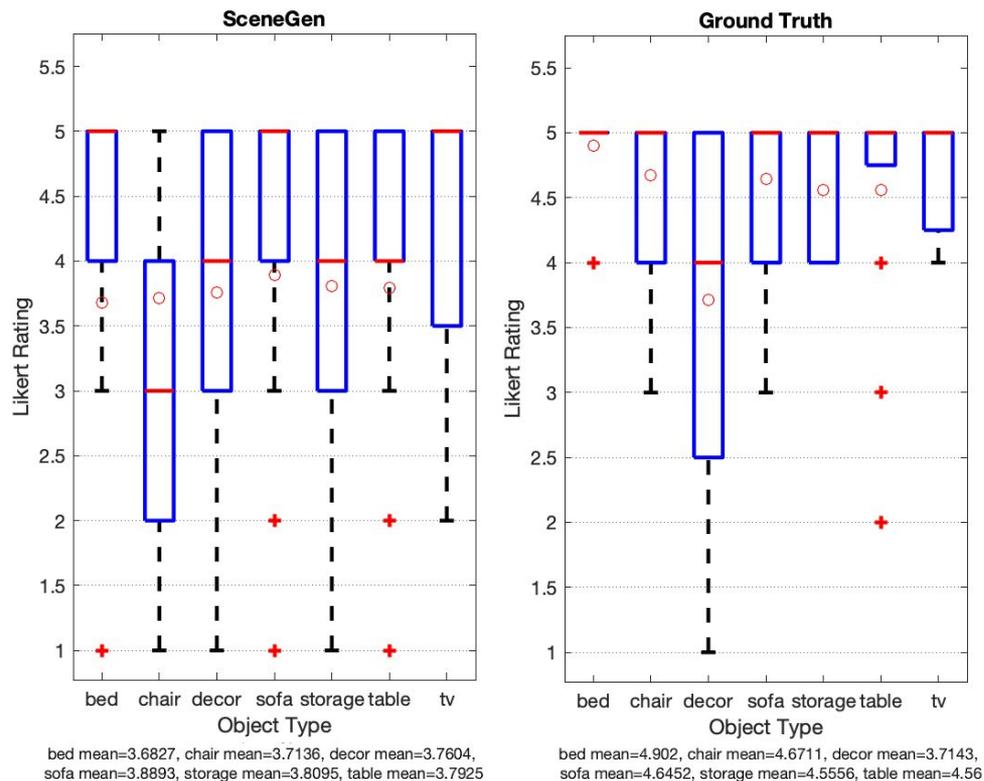
(g) Level 5 (Ground Truth)

User Experiments



Level 1 mean=1.48, Level 2 mean=1.91, Level 3 mean=3.72, Level 4 mean=4.185, Level 5 mean=4.5887

User Experiments



User Experiments



(a)



(b)



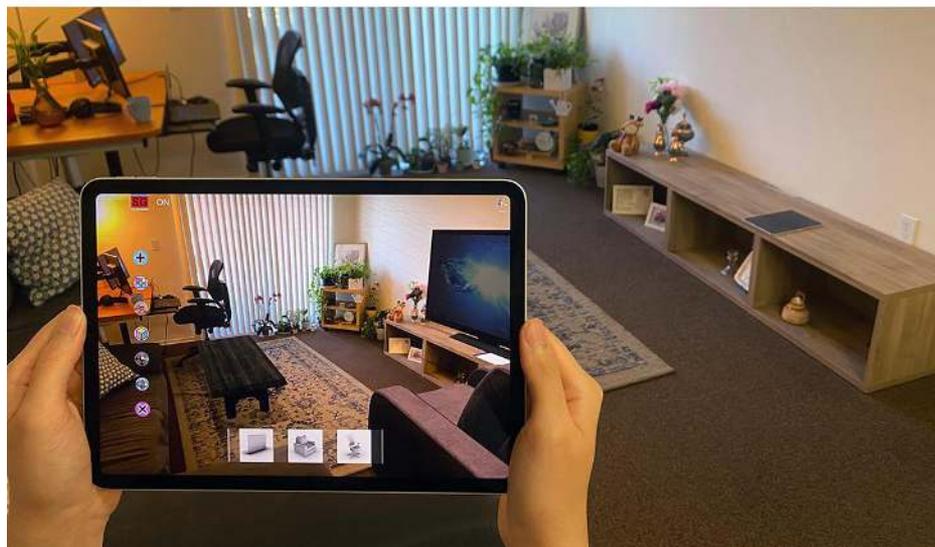
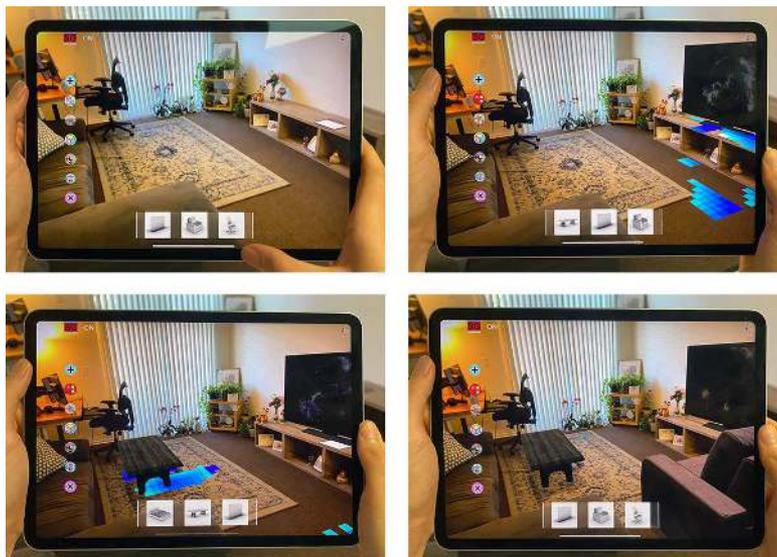
(c)



(d)

× Top prediction

Augmented Reality Application





Content Augmentation



Mutual Space



Content Augmentation



Mutual Space



actual footage shot
through HoloLens



actual footage shot
through HoloLens



actual footage shot through HoloLens







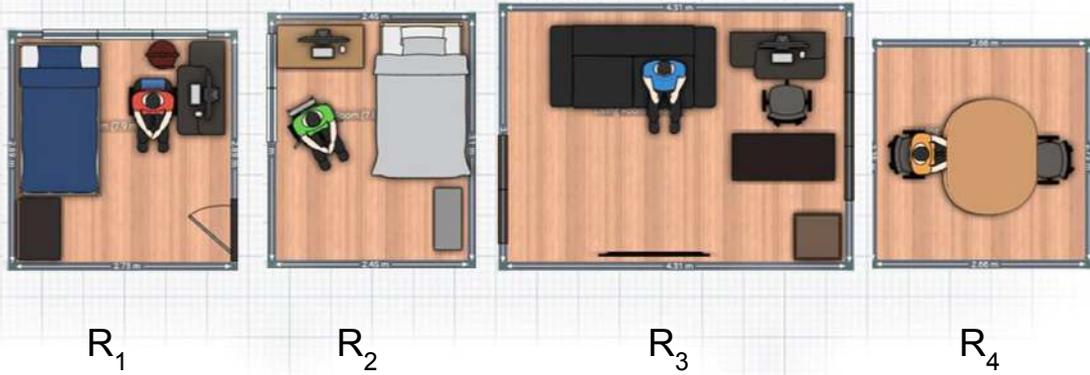
Spatial Computing has Spatial **Limitations**

How we address this
challenge?

Goal I

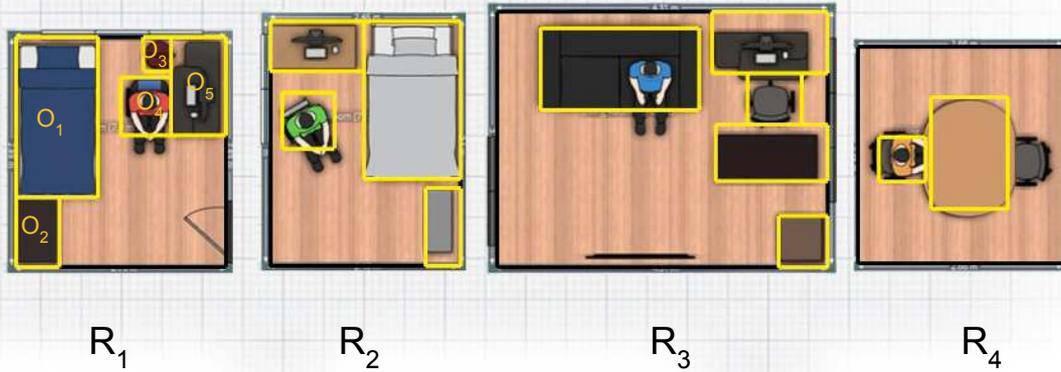
Locate the maximum mutual space

Input Data



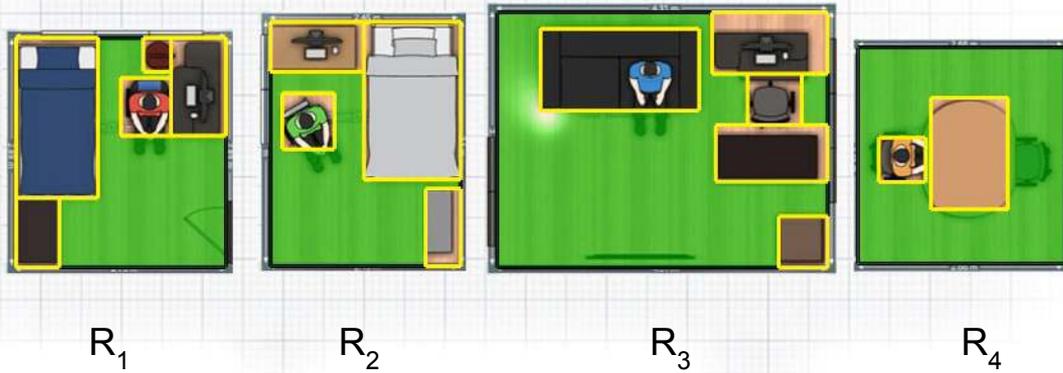
Input Data

Semantic Segmentation



Open Spaces

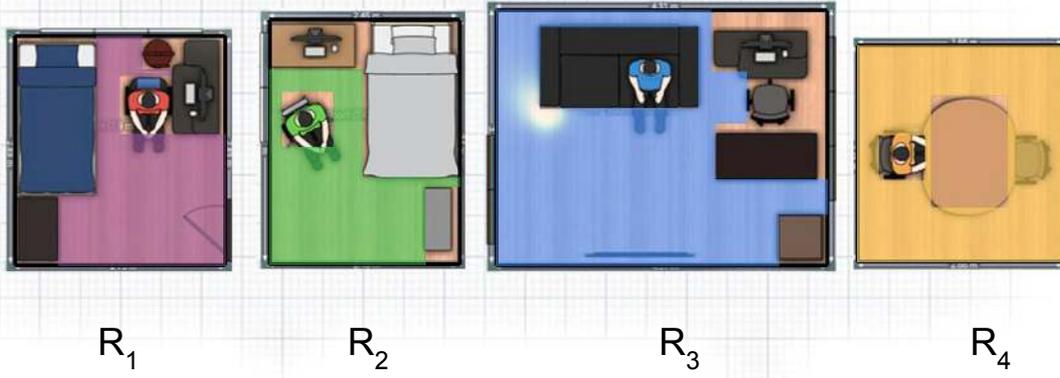
Standable Spaces



$$S_i = R_i - \bigcup_{k=1}^{n_i} O_{i,k}$$

Open Spaces

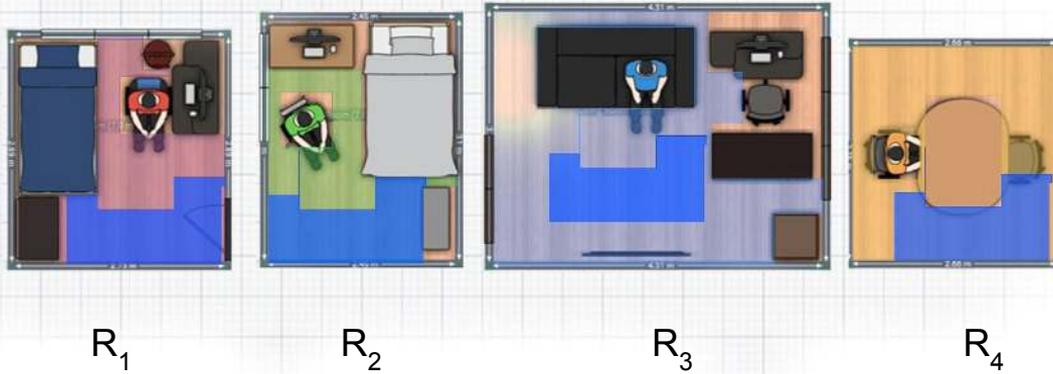
Standable Spaces



$$S_i = R_i - \bigcup_{k=1}^{n_i} O_{i,k}$$

Mutual Boundary Search

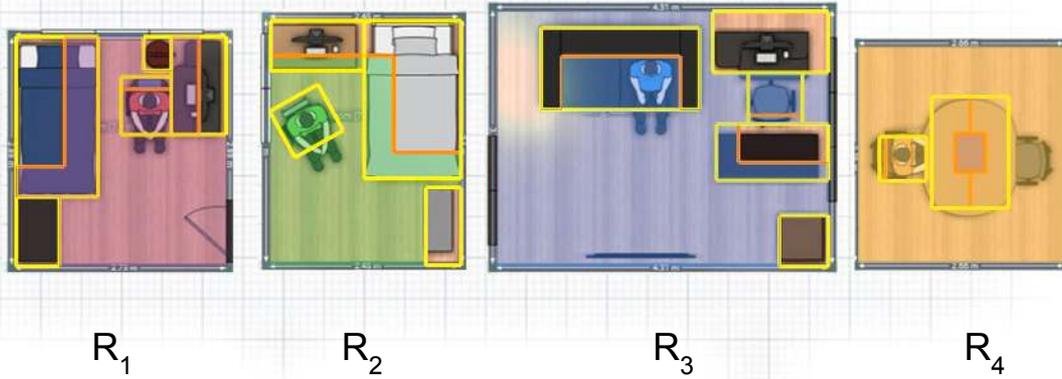
Standable Spaces



$$M_S(R_1, \dots, R_m)$$

Mutual Boundary Search

Sittable Space



$$A(R_i) = \bigcup_{k=1}^{n_i} A(O_{i,k}) + A(S_i).$$

$$A(R_1) = \bigcup_{k=1}^{n_1} A(O_{1,k}) + A(S_1).$$

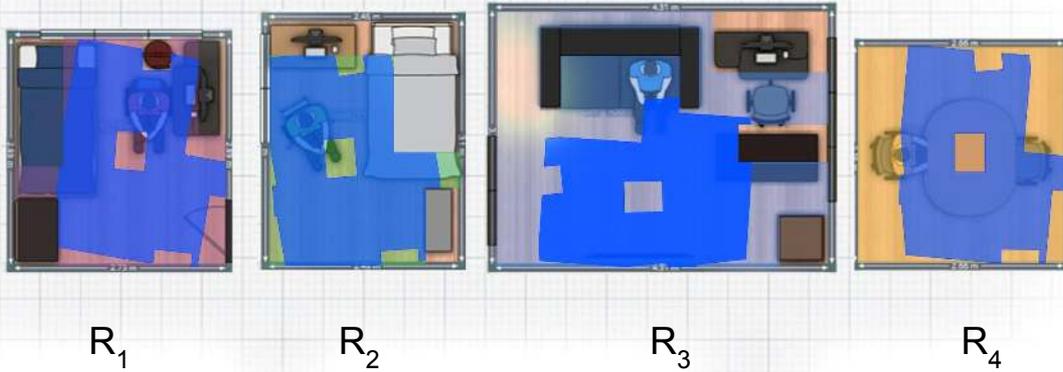
$$A(R_2) = \bigcup_{k=1}^{n_2} A(O_{2,k}) + A(S_2).$$

$$A(R_3) = \bigcup_{k=1}^{n_3} A(O_{3,k}) + A(S_3).$$

$$A(R_4) = \bigcup_{k=1}^{n_4} A(O_{4,k}) + A(S_4).$$

Mutual Boundary Search

Sittable Space



$$M_S(R_1, \dots, R_m)$$

Implementation on Real-world Datasets



familyroom_1pXnuDYAj8r



kitchen_1pXnuDYAj8r



9_familyroom_1pXnuDYAj8r



bedroom_1pXnuDYAj8r

-  Standable Spaces (S)
-  Non- standable Spaces



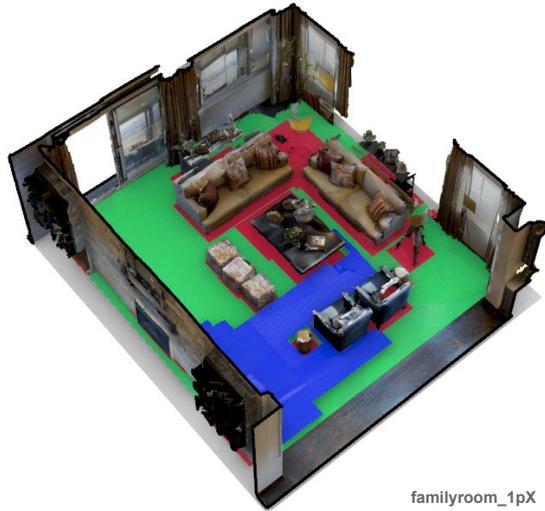
Standables Only

Standables + Sitables

-  Sittable Spaces (A)
-  Standable Spaces (S)
-  Non-standable Spaces

Mutual Boundary Search

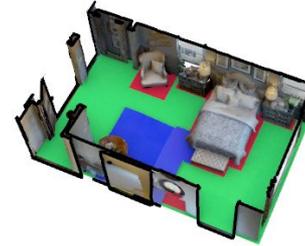
14Feb



familyroom_1pX
nuDYAj8r



bedroom_1pXnu
DYAj8r



bedroom_1pXnu
DYAj8r

-  Mutual Spaces (M_S)
-  Standable Spaces (S)
-  Non-standable Spaces

What if we need more mutual space?

Goal II

Rearrange Furniture so that:



Mutual Space

$$M_S(R_1, \dots, R_m)$$

$$E(R_i, \Theta_i) = \sum_{k=1}^{n_i} w_k \|G(O_{i,k}, \theta_{i,k})\|_t,$$

Effort

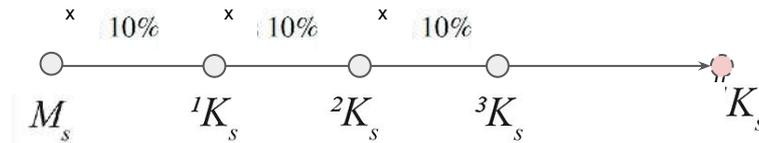


Increase Step

$$\min \sum_{i=1}^m E(R_i, \Theta_i^s) \quad \text{subj. to} \quad K^s \left(\bigcap_{i=1}^m G(S_i, \theta_i^s) \right) \text{ increases 10\%,}$$

Increase Step

$$\min \sum_{i=1}^m E(R_i, \Theta_i^s) \quad \text{subj. to} \quad K^s \left(\bigcap_{i=1}^m G(S_i, \theta_i^s) \right) \text{ increases 10\%,}$$



$K_{(\text{Area})}$

Furniture Layout Optimization



FamilyRoom

Area: $34.32 m^2$

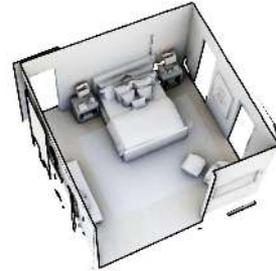
Objects: 10



Office

Area: $34.32 m^2$

Objects: 8

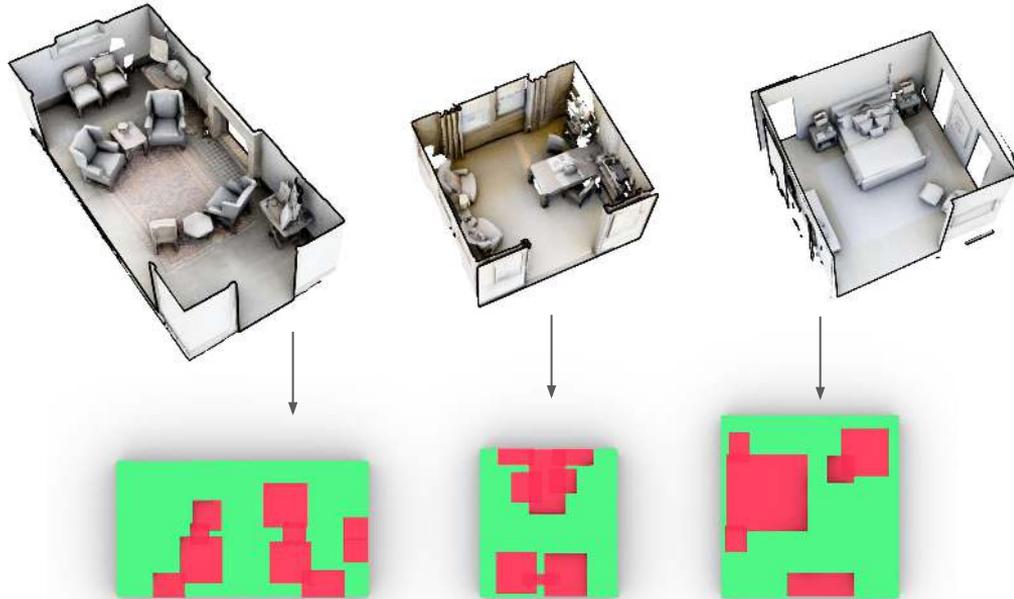


Bedroom

Area: $50.31 m^2$

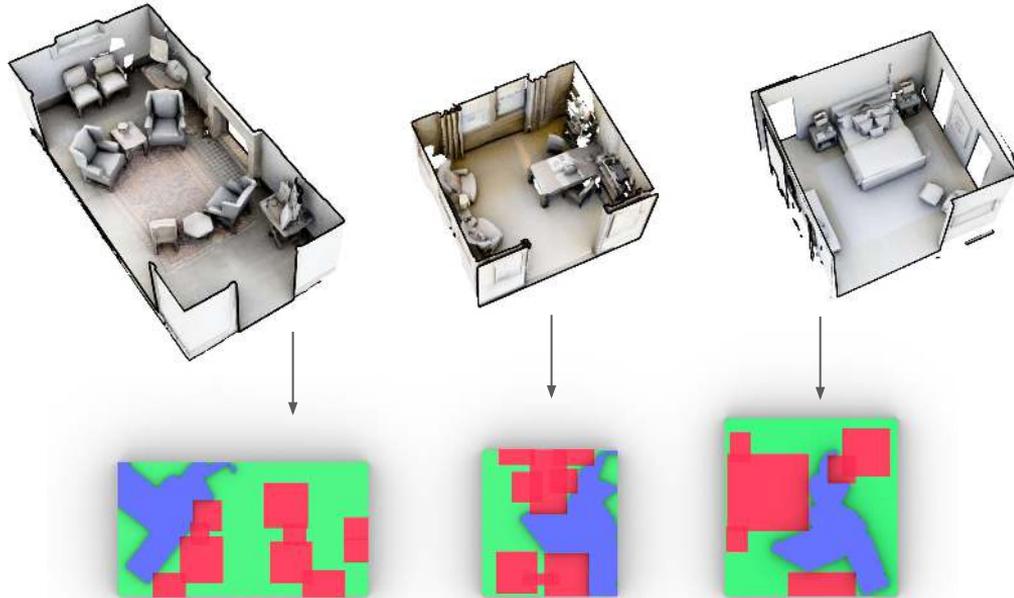
Objects: 6

Furniture Layout Optimization



■ Standable Spaces (S)
■ Non-standable Spaces

Furniture Layout Optimization

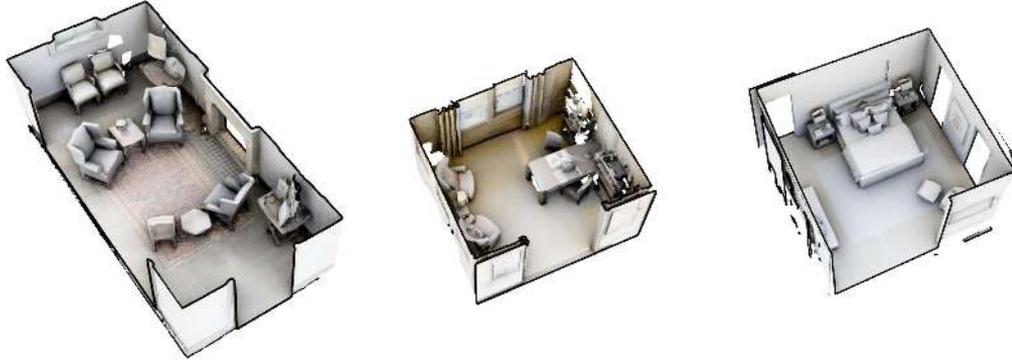


$$M_s = 10.88$$

- Mutual Spaces (M_s)
- Standable Spaces (S)
- Non-standable Spaces

Furniture Layout Optimization

Rearrange Furniture so that:



Mutual Space

$$M_S(R_1, \dots, R_m)$$

$$E(R_i, \Theta_i) = \sum_{k=1}^{n_i} w_k \|G(O_{i,k}, \theta_{i,k})\|_i,$$

Effort

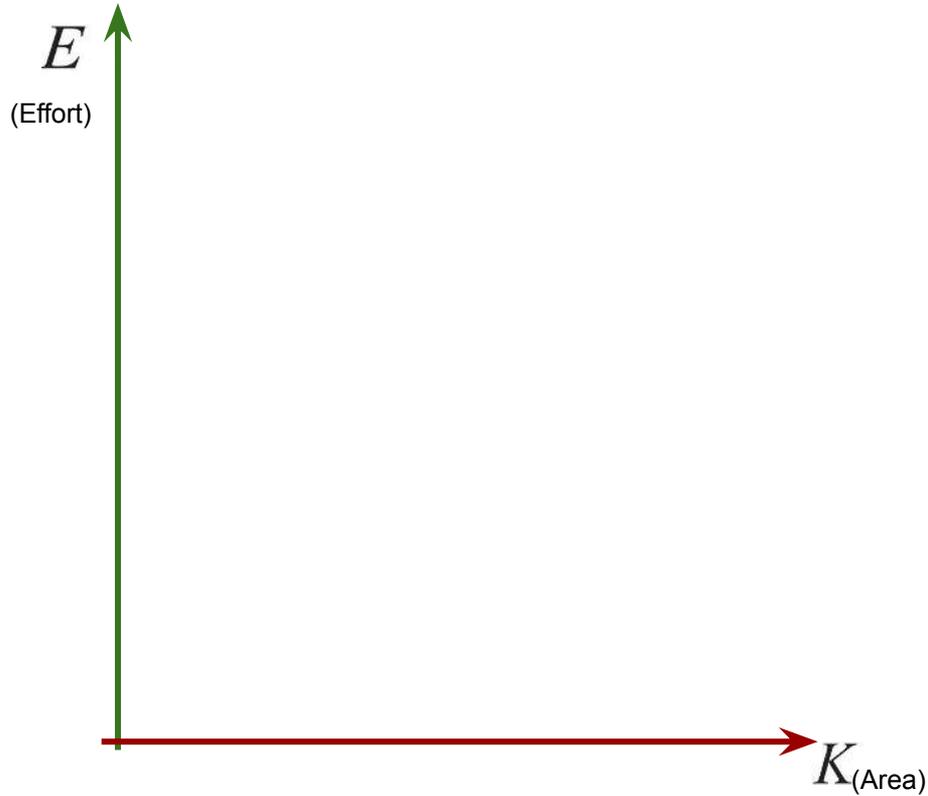


Furniture Layout Optimization

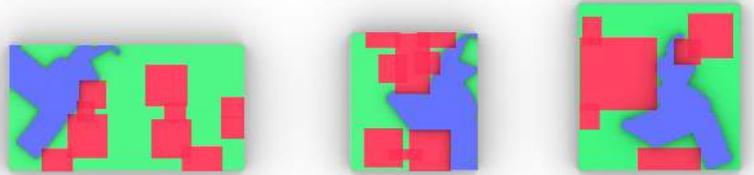
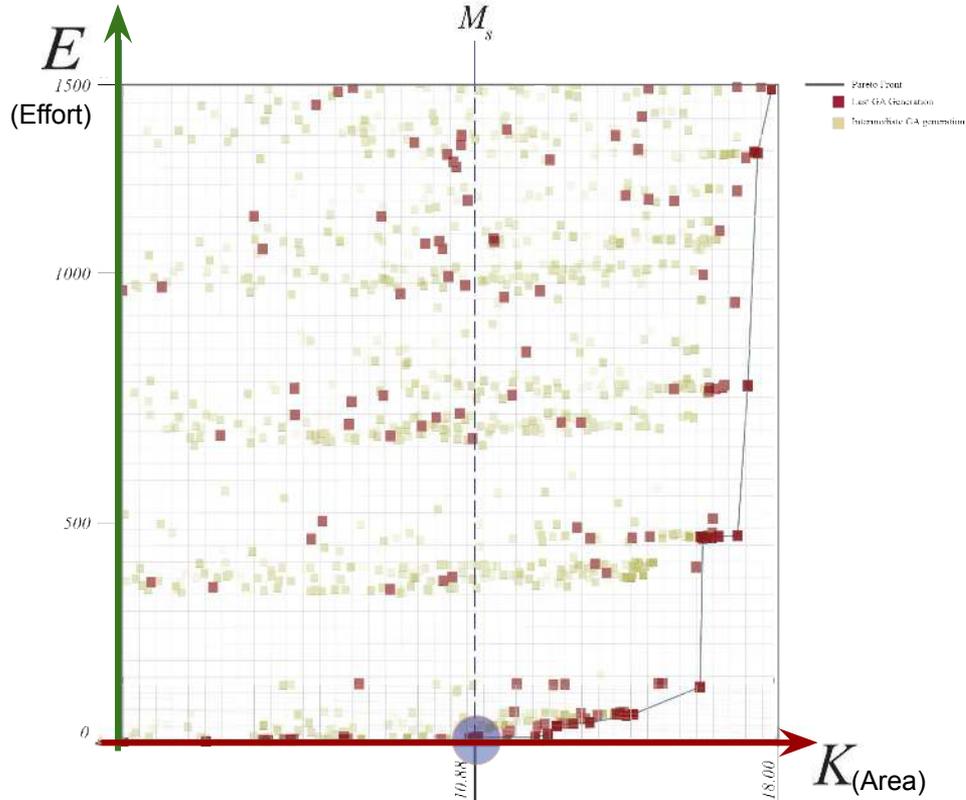
E
(Effort)



Furniture Layout Optimization



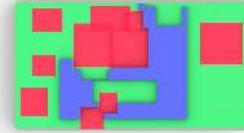
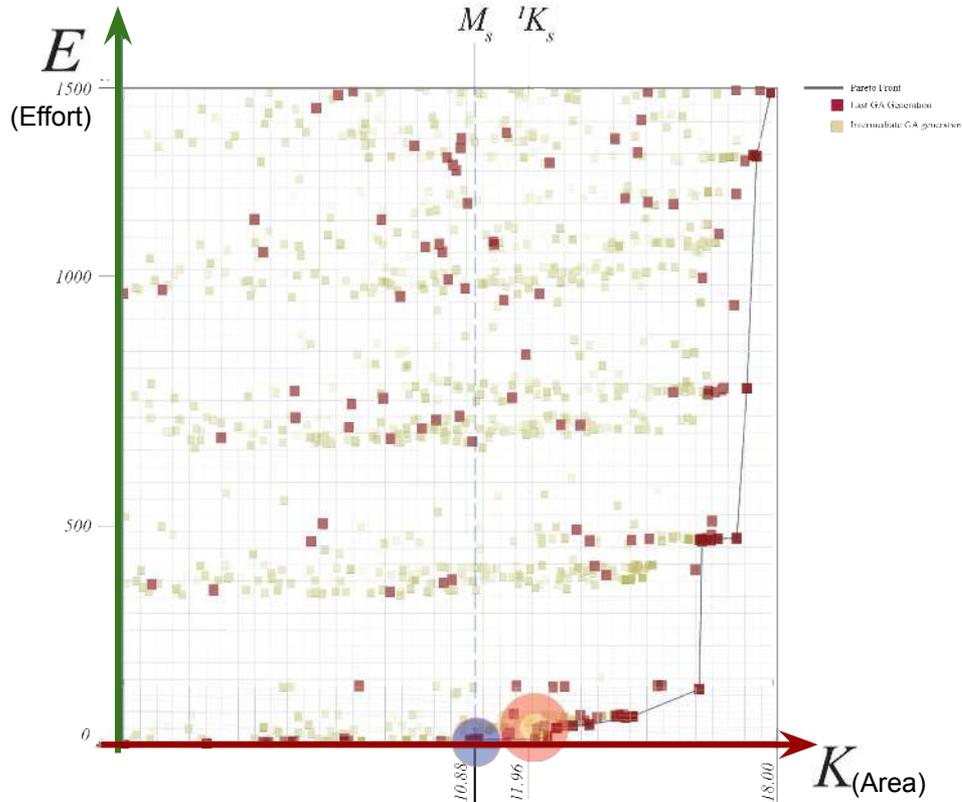
Furniture Layout Optimization



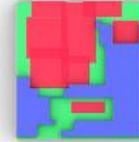
$$M_s = 10.88$$

- Mutual Spaces (M_s)
- Standable Spaces (S_s)
- Non-standable Spaces

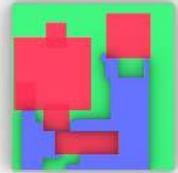
Furniture Layout Optimization



$E_1 = 52.51$



$E_2 = 9.50$



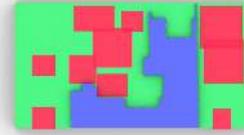
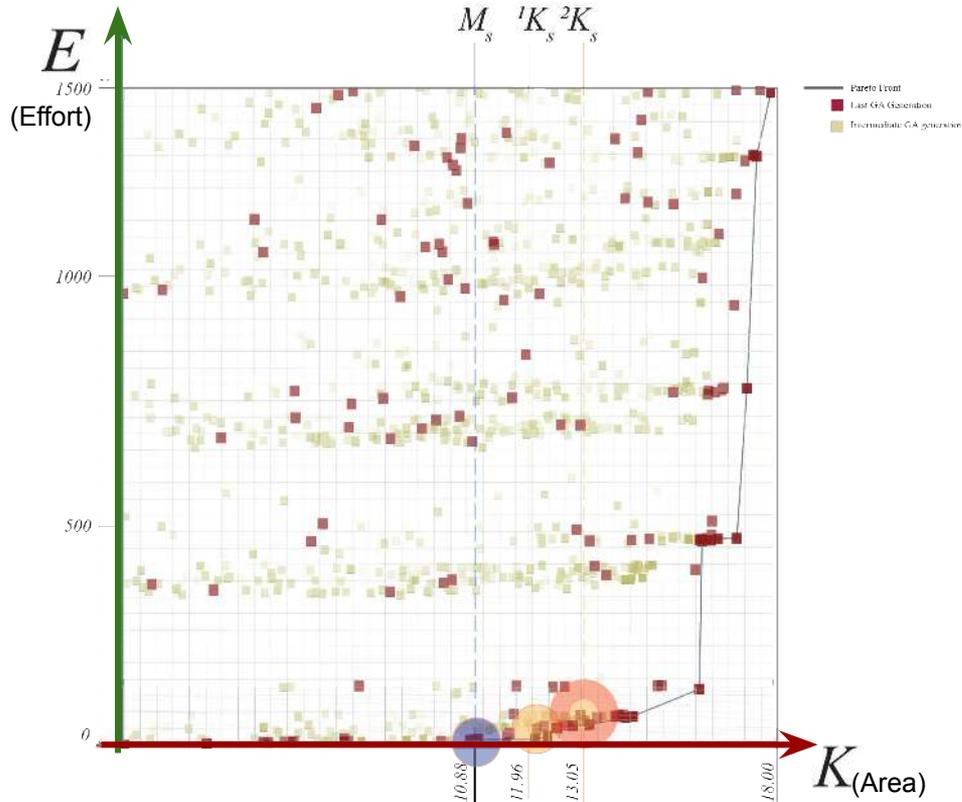
$E_3 = 15.66$

$${}^1K_s = 12.18$$

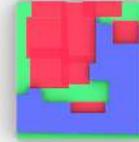
- Mutual Spaces (M_s)
- Standable Spaces (S_s)
- Non-standable Spaces

$$\min \sum_{i=1}^m E(R_i, \Theta_i^s) \quad \text{subj. to} \quad K^s \left(\bigcap_{i=1}^m G(S_i, \theta_i^s) \right) \text{ increases 10\%,}$$

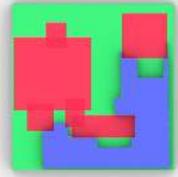
Furniture Layout Optimization



$E_1 = 60.34$



$E_2 = 9.50$



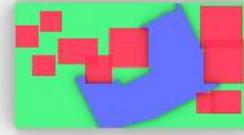
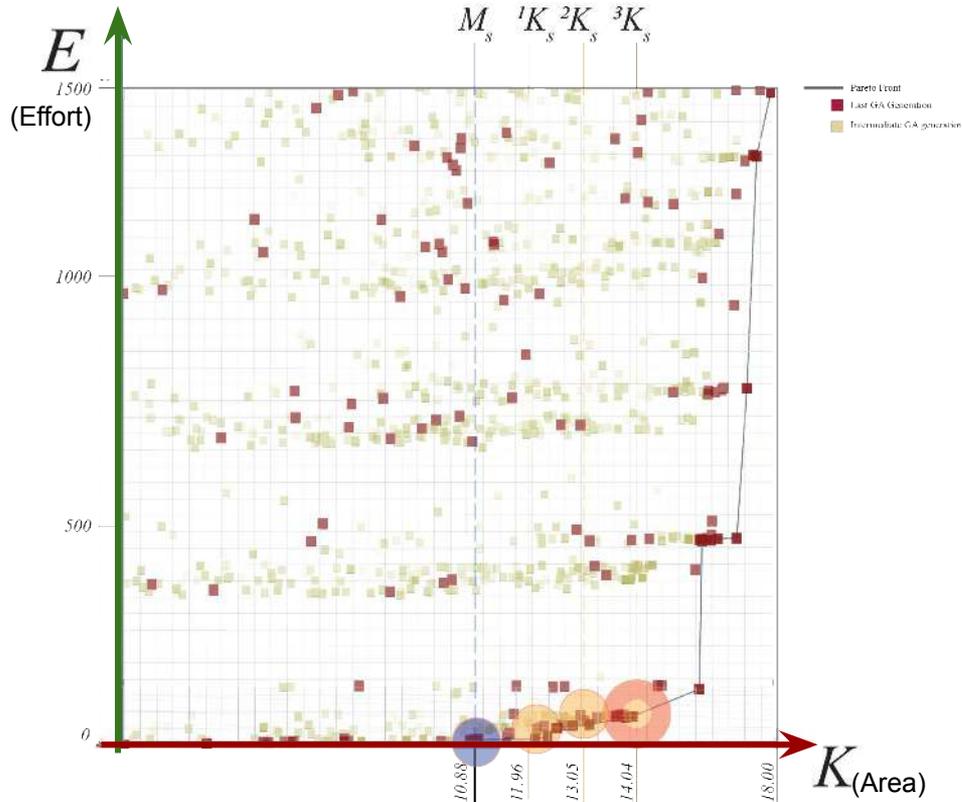
$E_3 = 17.80$

$${}^2K_s = 13.66$$

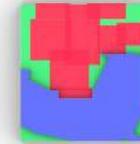
- Mutual Spaces (M_s)
- Standable Spaces (S_s)
- Non-standable Spaces

$$\min \sum_{i=1}^m E(R_i, \Theta_i^s) \quad \text{subj. to} \quad K^s \left(\bigcap_{i=1}^m G(S_i, \theta_i^s) \right) \text{ increases 10\%,}$$

Furniture Layout Optimization



$$E_1 = 74.86$$



$$E_2 = 10.43$$



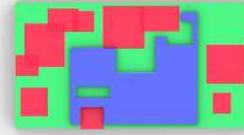
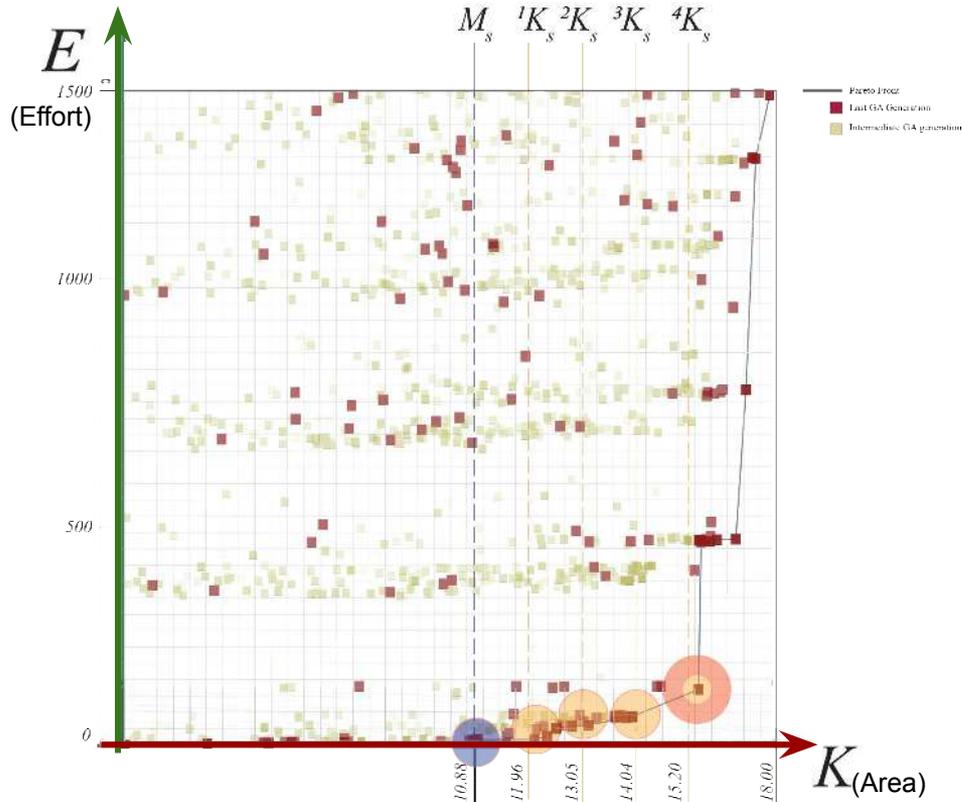
$$E_3 = 17.96$$

$$^3K_s = 14.33$$

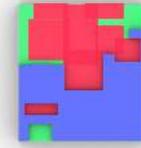
- Mutual Spaces (M_s)
- Standable Spaces (S_s)
- Non-standable Spaces

$$\min \sum_{i=1}^m E(R_i, \Theta_i^s) \quad \text{subj. to} \quad K^s \left(\bigcap_{i=1}^m G(S_i, \theta_i^s) \right) \text{ increases 10\%,}$$

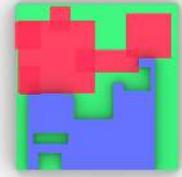
Furniture Layout Optimization



$E_1 = 57.09$



$E_2 = 11.02$



$E_3 = 70.43$

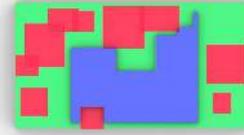
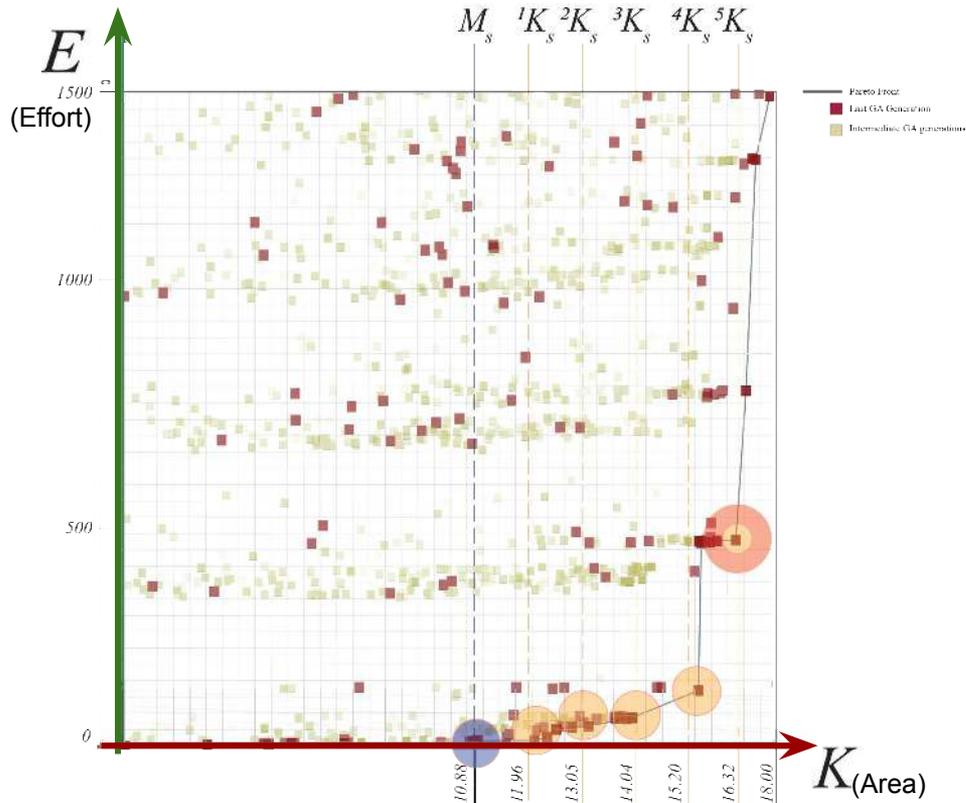
$${}^4K_s = 15.99$$

- Mutual Spaces (M_s)
- Standable Spaces (S_s)
- Non-standable Spaces

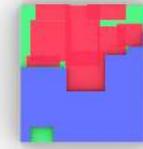
$$\min \sum_{i=1}^m E(R_i, \Theta_i^s) \quad \text{subj. to} \quad K^s \left(\bigcap_{i=1}^m G(S_i, \theta_i^s) \right) \text{ increases 10\%,}$$



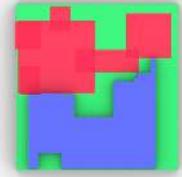
Furniture Layout Optimization



$$E_1 = 57.09$$



$$E_2 = 212.90$$



$$E_3 = 70.43$$

$$^5K_s = 17.05$$

- Mutual Spaces (M_s)
- Standable Spaces (S_s)
- Non-standable Spaces

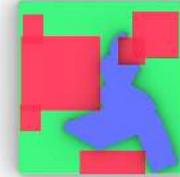
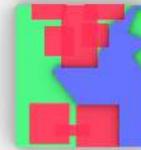
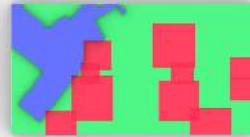
$$\min \sum_{i=1}^m E(R_i, \Theta_i^s) \quad \text{subj. to} \quad K^s \left(\bigcap_{i=1}^m G(S_i, \theta_i^s) \right) \text{ increases 10\%,}$$



Furniture Layout Optimization

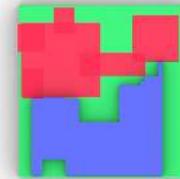
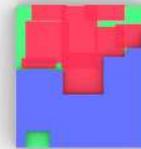
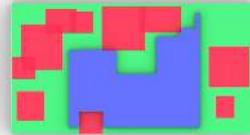
Initial Setting

$$M_s$$
$$=10.88$$



Furniture Rearrangement

$5K_s$
$$=17.05$$



-  Mutual Spaces (M_s)
-  Standable Spaces (S)
-  Non-standable Spaces

$$E_1 = 57.09$$

$$E_2 = 212.90$$

$$E_3 = 70.43$$

Hololens Visualization

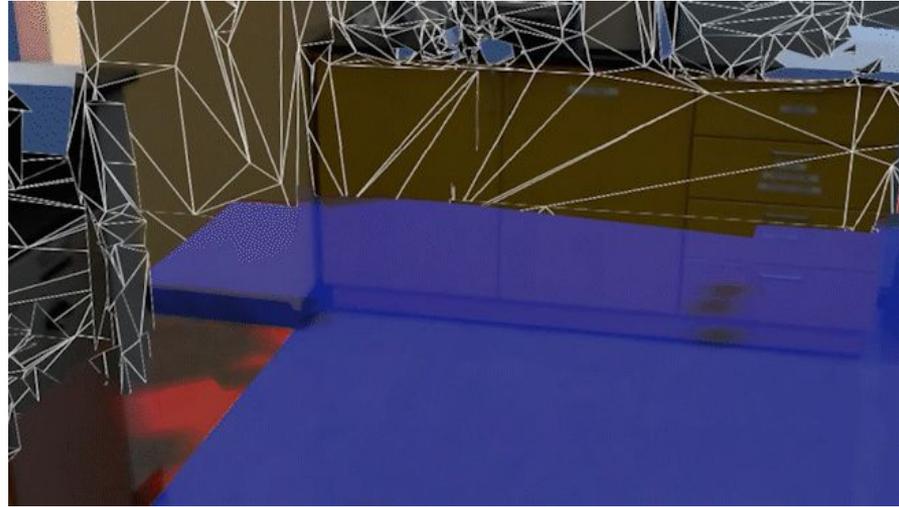
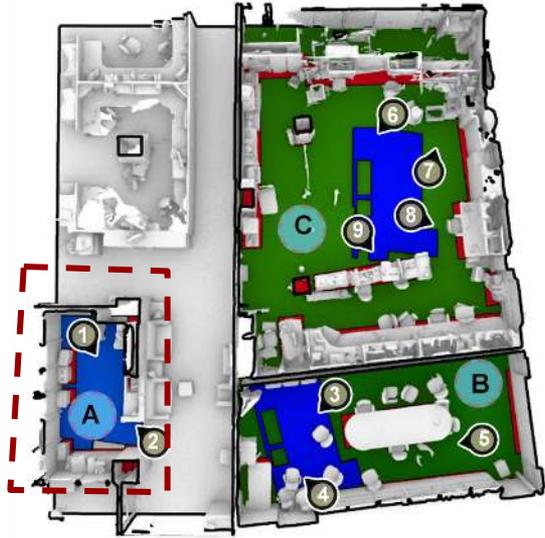
Hololens Visualization



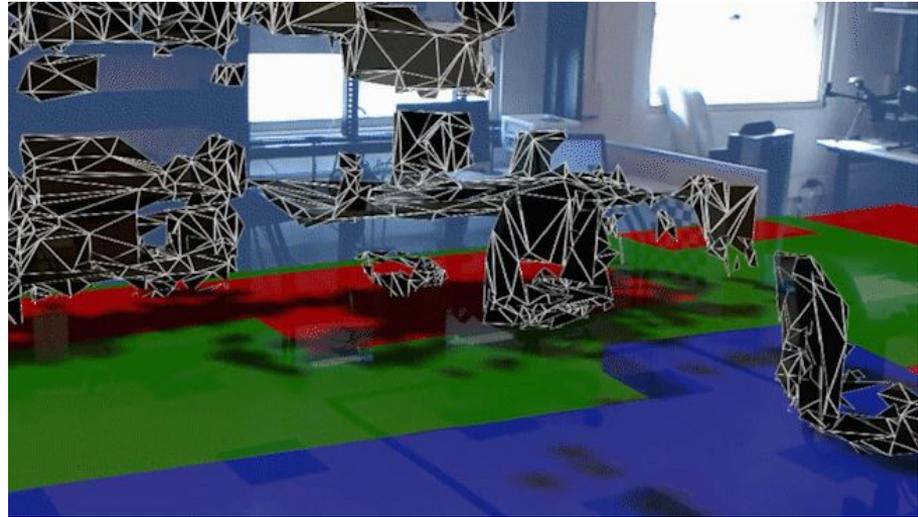
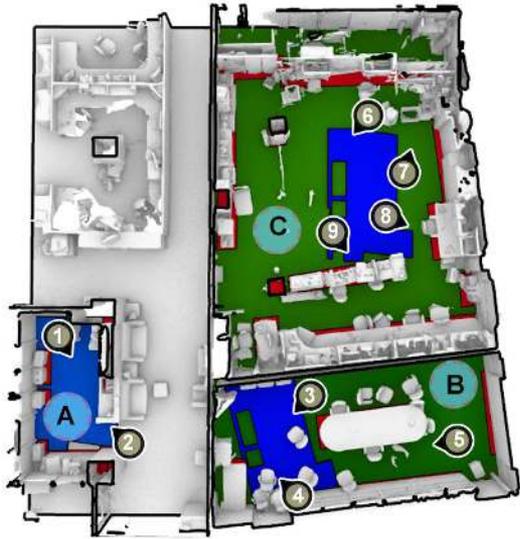
- Mutual Spaces (M_s)
- Standable Spaces (S)
- Non-standable Spaces



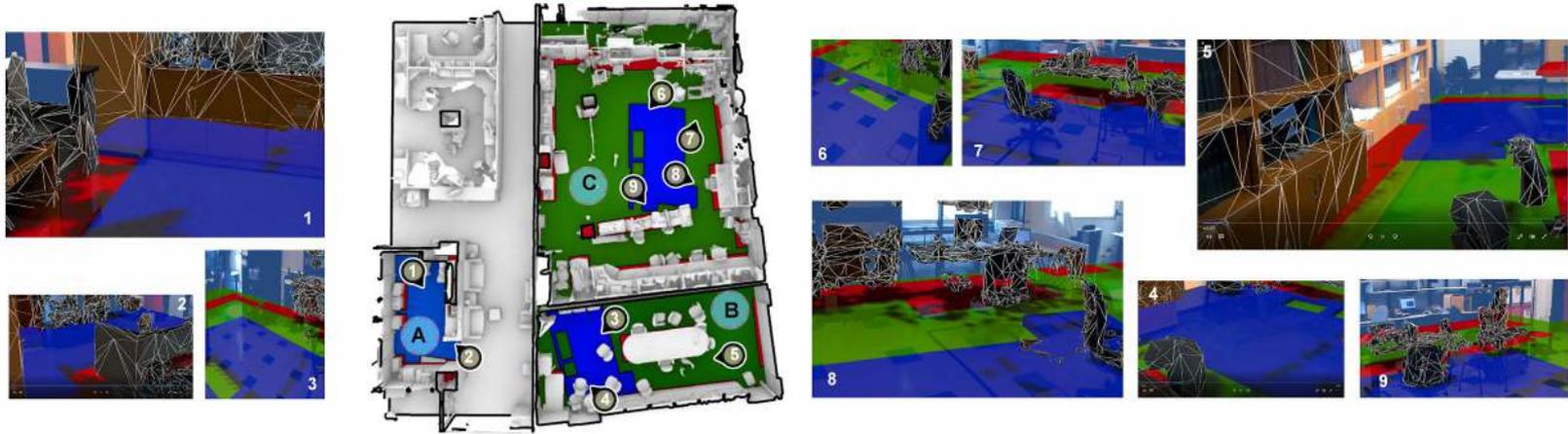
Hololens Visualization



Hololens Visualization



Hololens Visualization



Limitations

Limitations

- Furniture manipulation does not have continuous transformation in each increasing step
- Furniture with fixed positions are not automatically detected
- Weight is calculated based on standard assumptions.

Future Steps

Future Work

- Usability studies can be conducted on how to improve the visualization strategies so participants can experience the required tele-communication functionalities while preserving the mutual spatial ground.
- Cross-platform integration between various AR/VR platforms (mobiles, standalone, etc)
- Utilize robust floorplanning methods for furniture manipulation

Thank you!

References

- Keshavarzi, Mohammad, Allen Y. Yang, Woojin Ko, and Luisa Caldas. "Optimization and Manipulation of Contextual Mutual Spaces for Multi-User Virtual and Augmented Reality Interaction." In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 353-362. IEEE, 2020.
- Keshavarzi, Mohammad, Aakash Parikh, Xiyu Zhai, Melody Mao, Luisa Caldas, and Allen Y. Yang. "SceneGen: Generative Contextual Scene Augmentation using Scene Graph Priors." *arXiv preprint arXiv:2009.12395* (2020).
- Orts-Escolano, Sergio, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim et al. "Holoportation: Virtual 3d teleportation in real-time." In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 741-754. 2016.
- Fisher, Matthew, Manolis Savva, Yangyan Li, Pat Hanrahan, and Matthias Nießner. "Activity-centric scene synthesis for functional 3D scene modeling." *ACM Transactions on Graphics (TOG)* 34, no. 6 (2015): 1-13.
- Fisher, Matthew, Daniel Ritchie, Manolis Savva, Thomas Funkhouser, and Pat Hanrahan. "Example-based synthesis of 3D object arrangements." *ACM Transactions on Graphics (TOG)* 31, no. 6 (2012): 1-11.
- Merrell, Paul, Eric Schkufza, Zeyang Li, Maneesh Agrawala, and Vladlen Koltun. "Interactive furniture layout using interior design guidelines." *ACM transactions on graphics (TOG)* 30, no. 4 (2011): 1-10.
- Germer, Tobias, and Martin Schwarz. "Procedural Arrangement of Furniture for Real-Time Walkthroughs." In *Computer Graphics Forum*, vol. 28, no. 8, pp. 2068-2078. Oxford, UK: Blackwell Publishing Ltd, 2009.