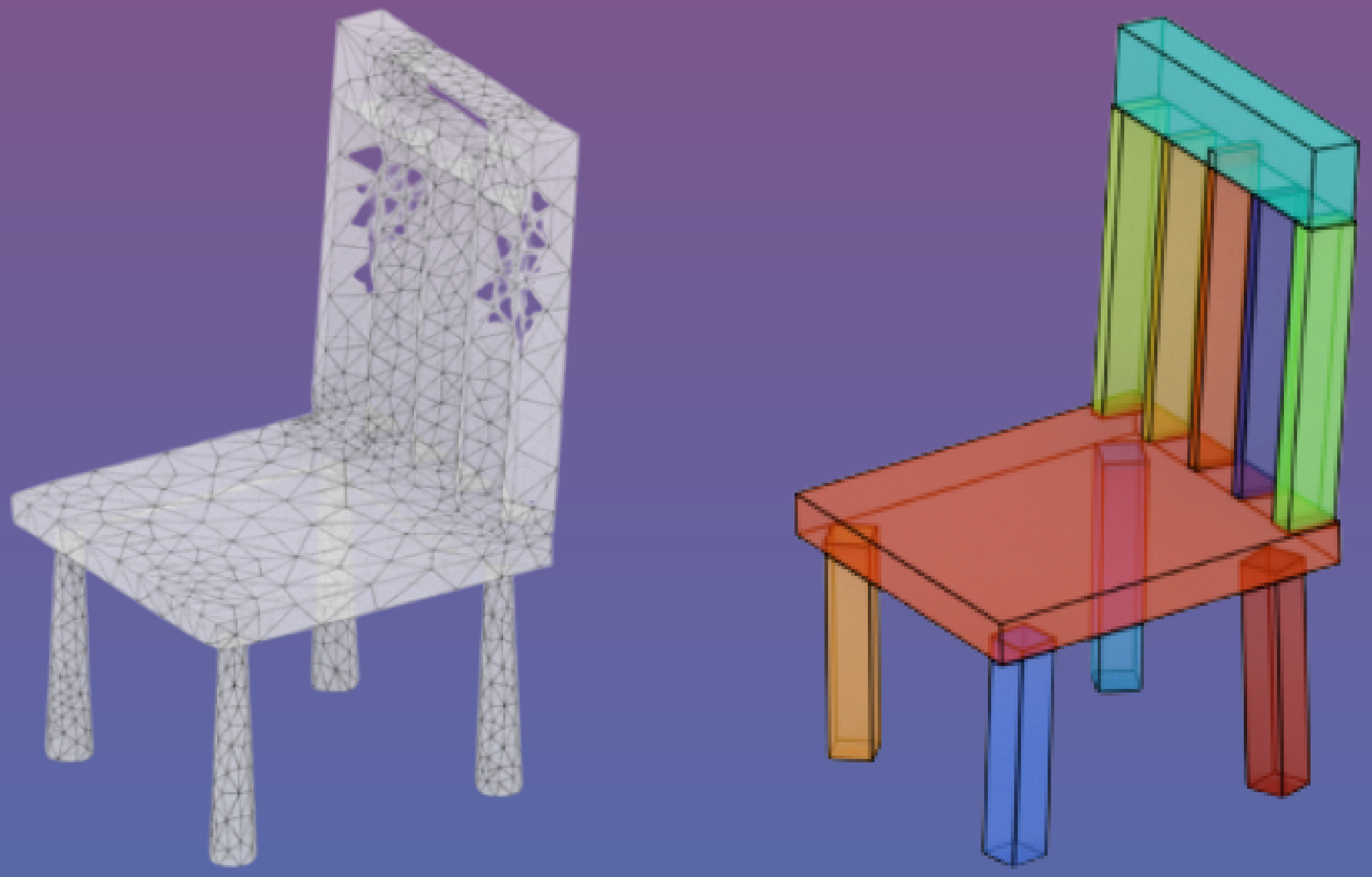
CS492 Deep RL Final Presentation

CubeFit: Minimizing Cuboids and Maximizing Coverage in 3D Shape via MCTS



Team 9

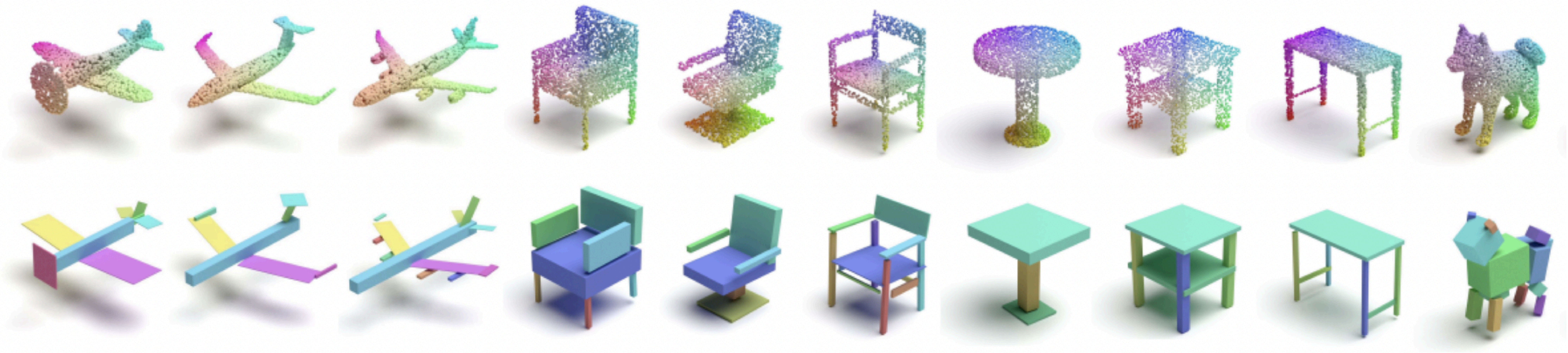
Inhwa Song

Chanhyeok Park

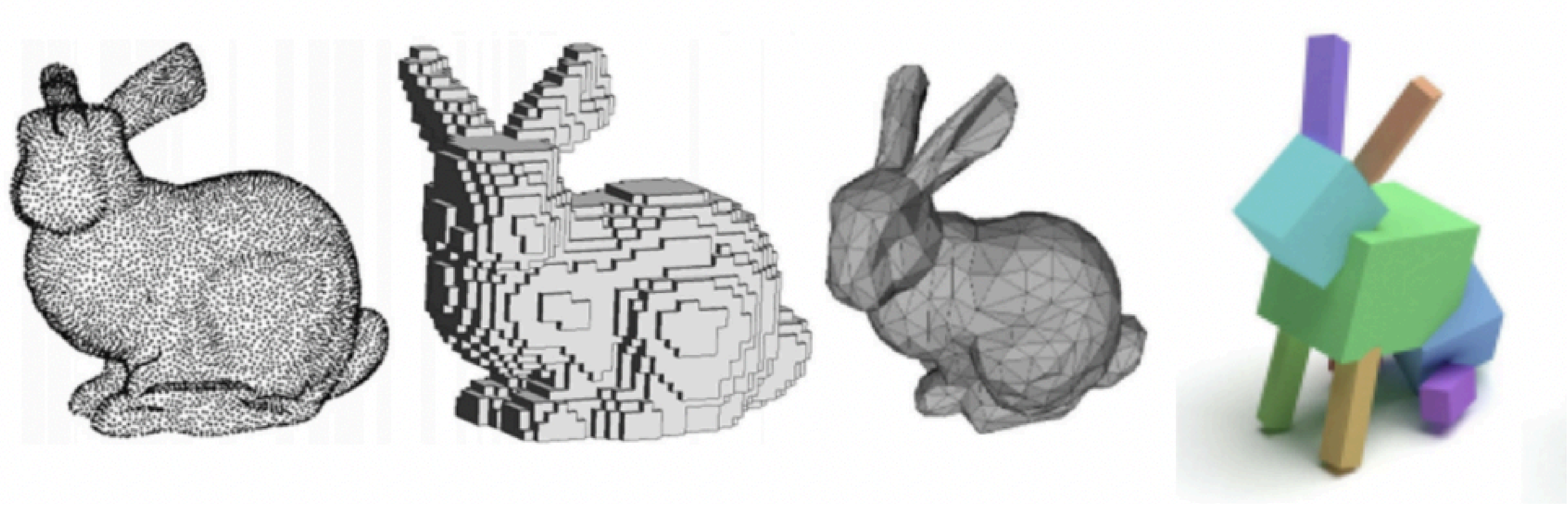
Sangmin Lee

Representing 3D Shapes as Boxes

We want to approximate / segment the 3D shapes using set of boxes.



Why is it Beneficial?



Compact Representation

3D Segmentation

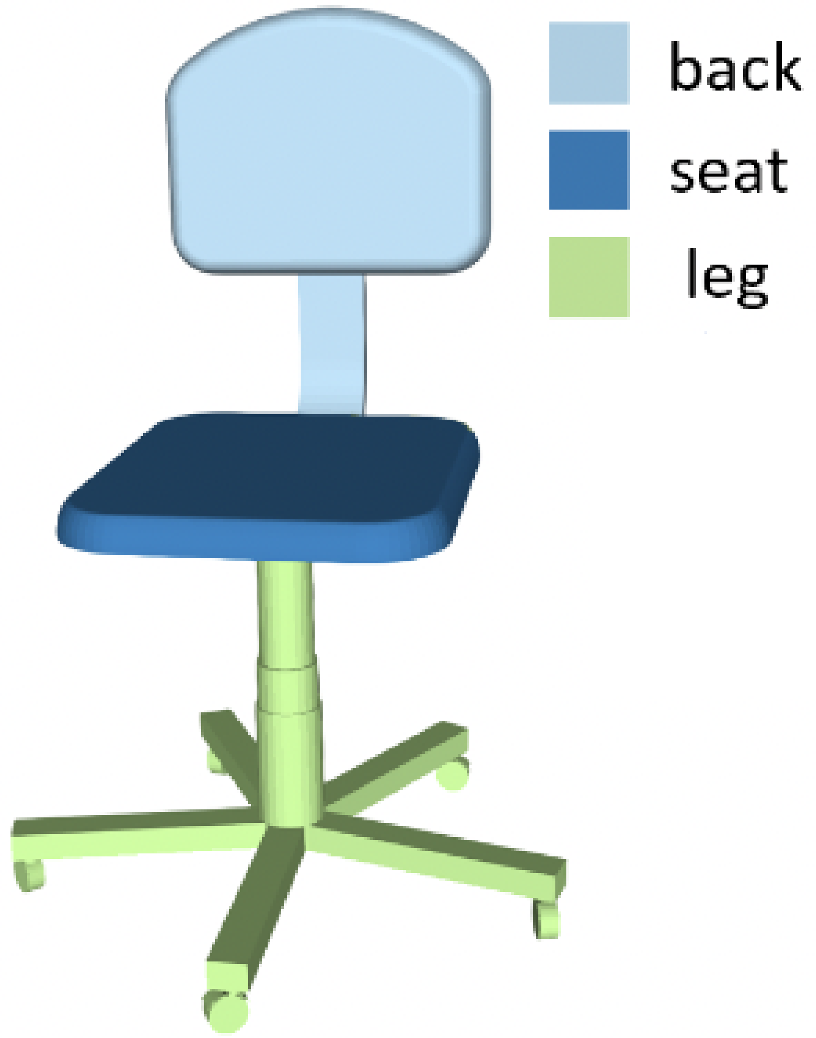
Ray Tracing

Fast 3D Geometric Proximity Queries

Intersection Tests

Why is it Beneficial?

Compact Representation



3D Segmentation

Ray Tracing

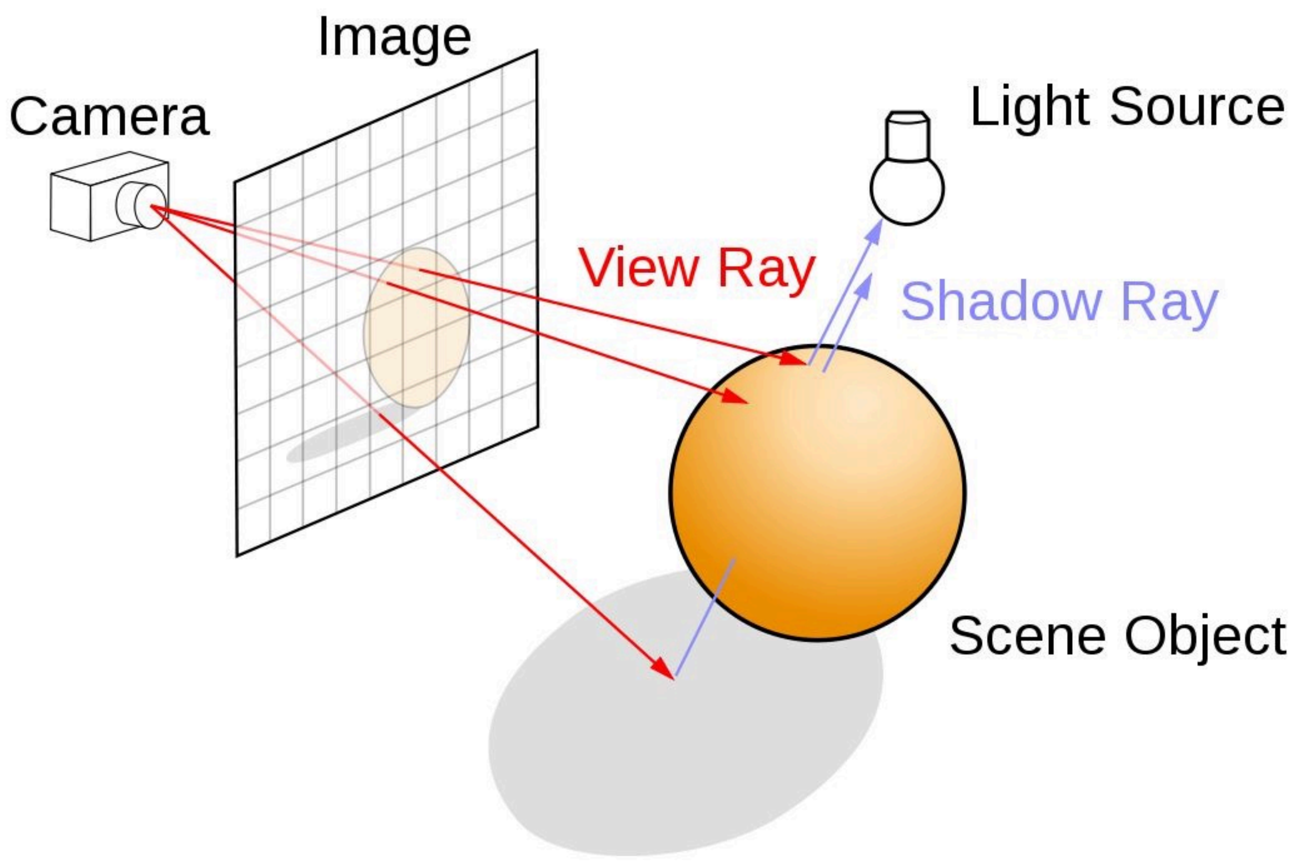
Fast 3D Geometric Proximity Queries

Intersection Tests

Why is it Beneficial?

Compact Representation

3D Segmentation



Ray Tracing

Fast 3D Geometric Proximity Queries

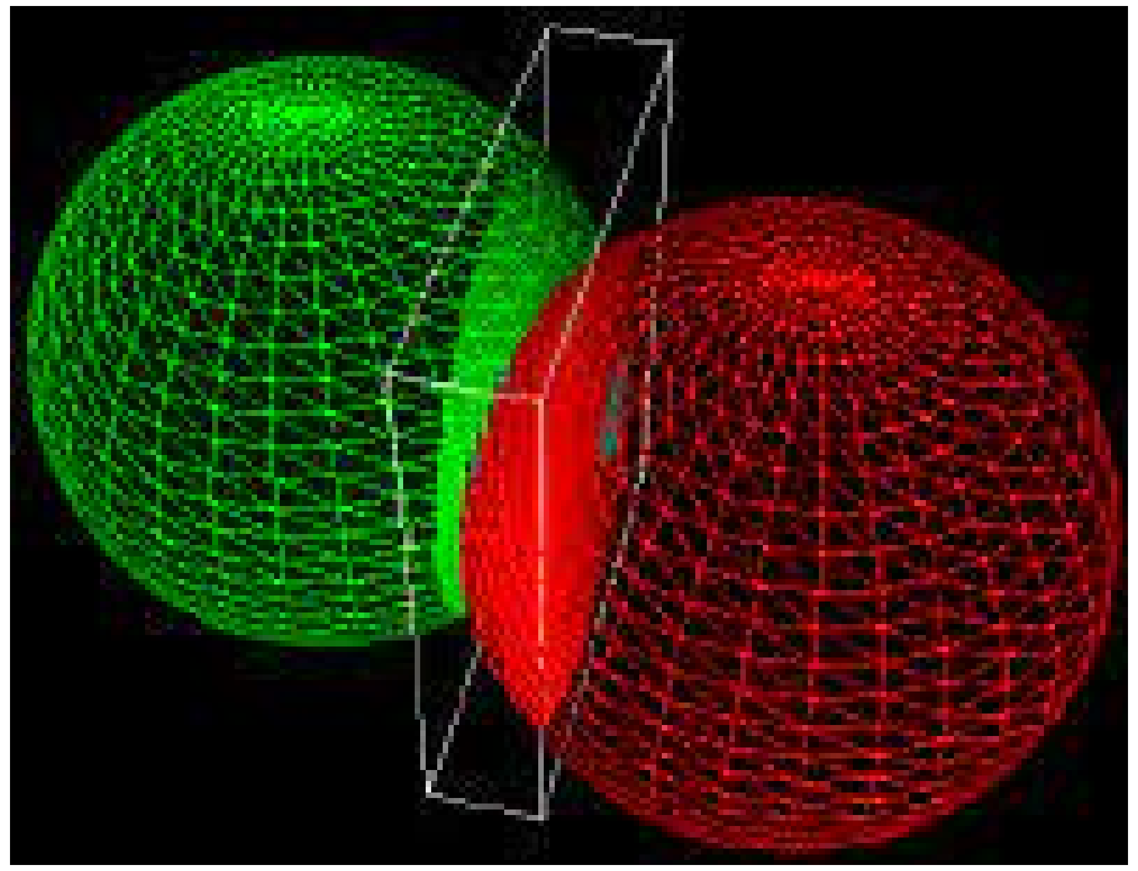
Intersection Tests

Why is it Beneficial?

Compact Representation

3D Segmentation

Ray Tracing



Fast 3D Geometric Proximity Queries

Intersection Tests

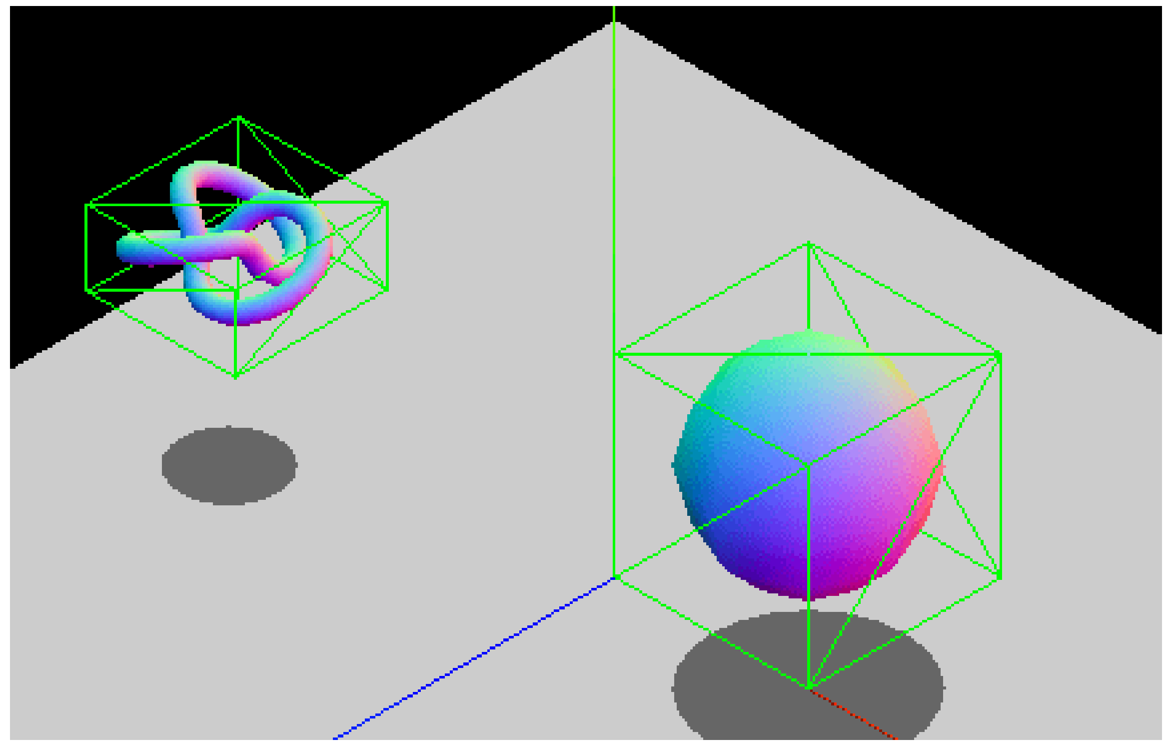
Why is it Beneficial?

Compact Representation

3D Segmentation

Ray Tracing

Fast 3D Geometric Proximity Queries

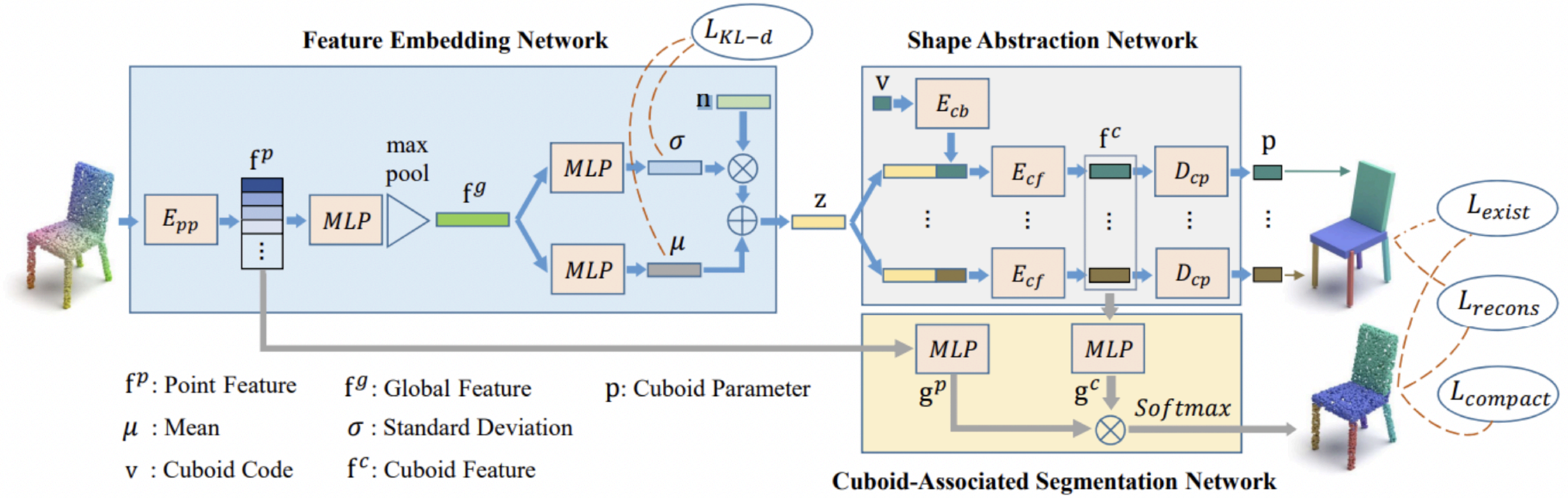


Intersection Tests

Related Work: Method

NN based learning methods

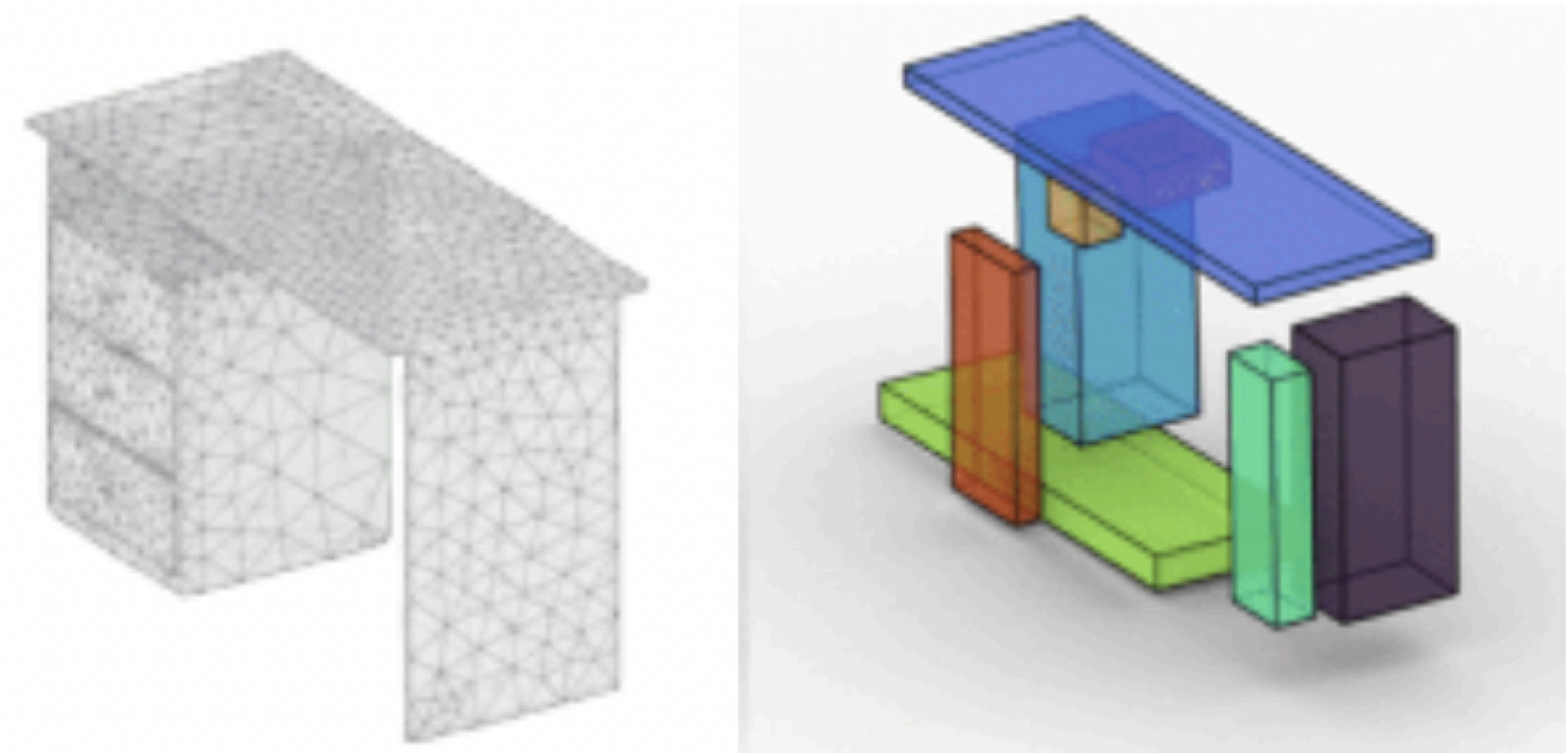
Unsupervised Learning for Cuboid Shape Abstraction via Joint Segmentation from Point Clouds



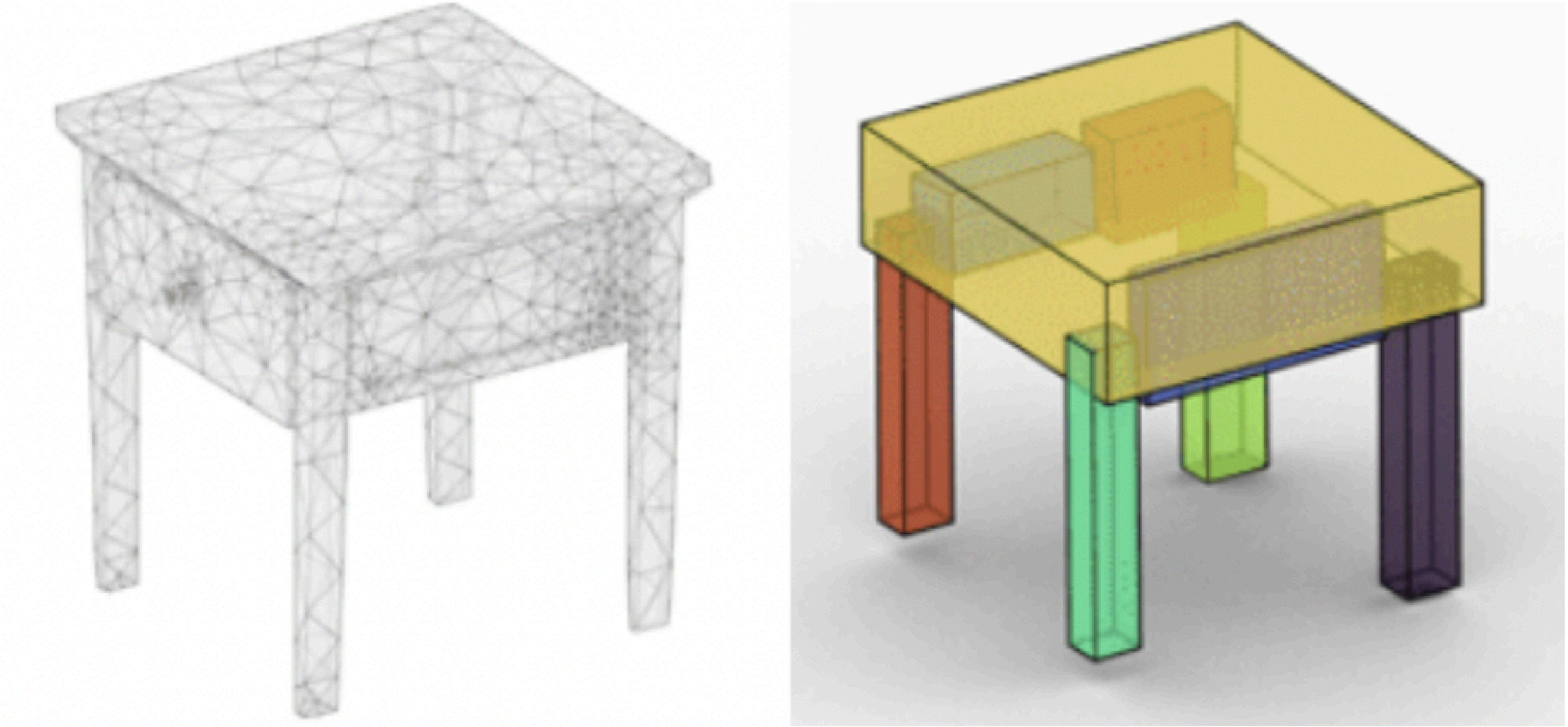
Yang, K., & Chen, X. (2021). Unsupervised learning for cuboid shape abstraction via joint segmentation from point clouds. ACM Transactions on Graphics (TOG), 40(4), 1-11.

Related Work: Problems

1. Boxes those not guarantee to cover the whole shape
2. Overlaps exist
3. Boxes are not tightly fitted to the shape



Not fully covered, Not Tightly Fitted

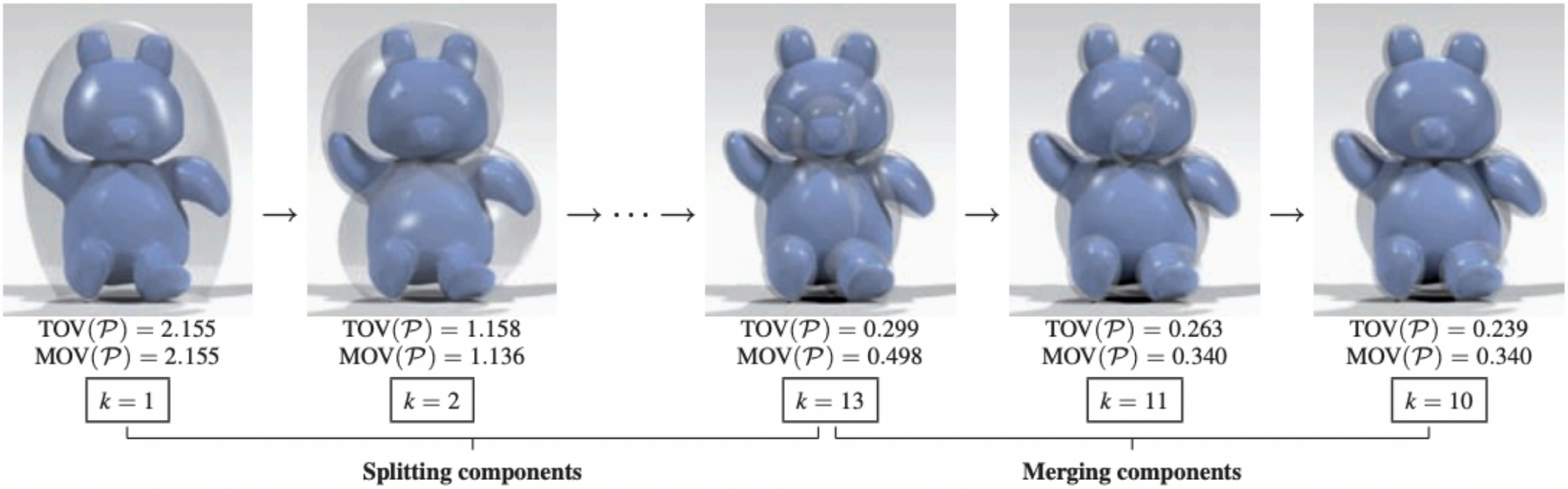


Overlaps, Meaningless Boxes

Related Work: Method

Energy Minimization Method

Variational 3D Shape Segmentation for Bounding Volume Computation



Lu, L., Choi, Y. K., Wang, W., & Kim, M. S. (2007, September). Variational 3D shape segmentation for bounding volume computation. In Computer Graphics Forum (Vol. 26, No. 3, pp. 329-338). Oxford, UK: Blackwell Publishing Ltd.

Related Work: Problems

- Heavily dependent on user input
- Long time: Point level perturbation
- Bad initialization
- Uncertain criteria: Not tightly Fitted

$\mathrm{TOV}(\mathcal{P}) = 0.292$, $\mathrm{MOV}(\mathcal{P}) = 0.497$
(d) 38 ellipsoids

$\mathrm{TOV}(\mathcal{P}) = 0.486$, $\mathrm{MOV}(\mathcal{P}) = 0.897$
(e) 18 OBBs

$\mathrm{TOV}(\mathcal{P}) = 0.380$, $\mathrm{MOV}(\mathcal{P}) = 1.302$
(f) 16 OBBs

Problem Definition

🎯 Goal

Binding 3D shapes tight with boxes by optimizing 2 geometric criterias

Problem Definition - Defining Concepts

Coverage

Tightness

Problem Definition - Defining Concepts

$B : Set\ of\ Boxes\ (B = \{B_i\}_{i=1}^{k})$

$S : Target\ 3D\ Shape$

Coverage

$$Cov(B) = 1 - \frac{vol(S - \cup_i B_i)}{vol(S)}$$

Tightness

$$BVS(B) = \frac{\sum_{i=1}^{k} vol(B_i)}{vol(S)}$$

Problem Definition - Geometric Criterias

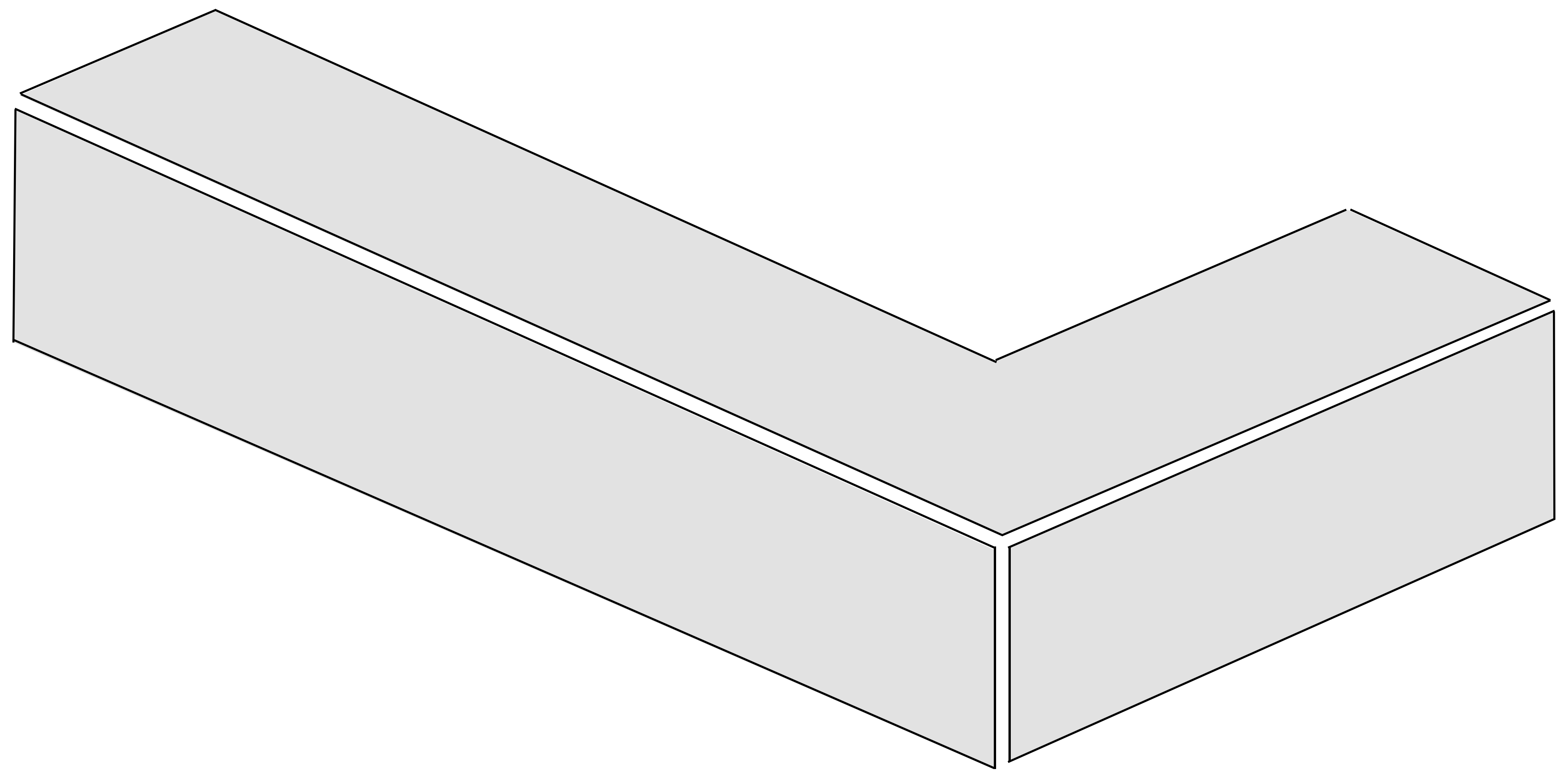
🎯 **Goal : Binding 3D shapes tight with boxes by optimizing 2 geometric criterias**

📌 **Full Coverage** : Boxes should not leave empty space of the target 3D shape

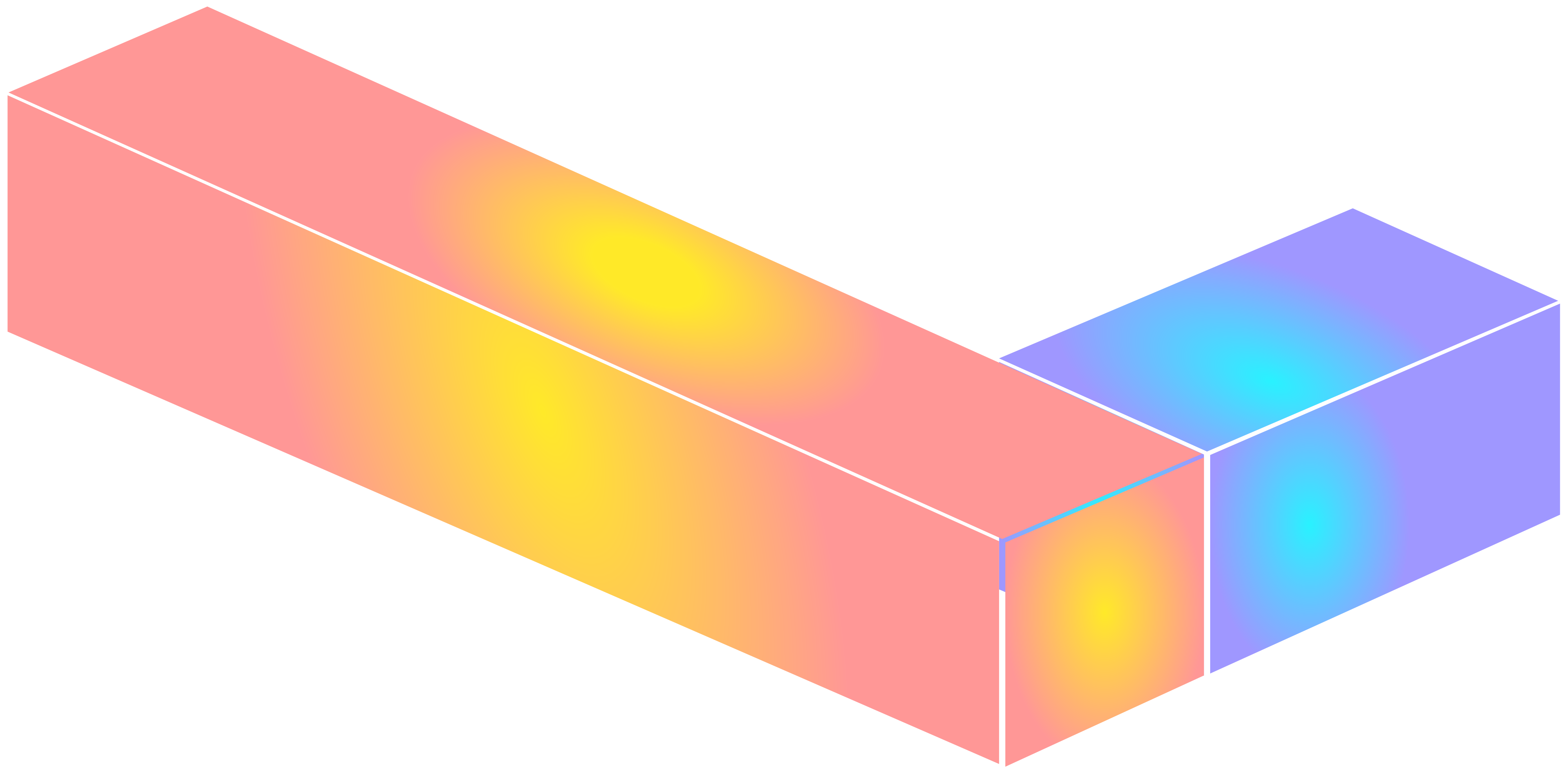
📌 **Tightness** : Boxes should minimize spare volume exceeding the target 3D shape

$$\underset{\{B_i\}_{i=1}^{k}}{\operatorname{argmin}} \; BVS(B) \quad s.t. \quad Cov(B) = 1$$

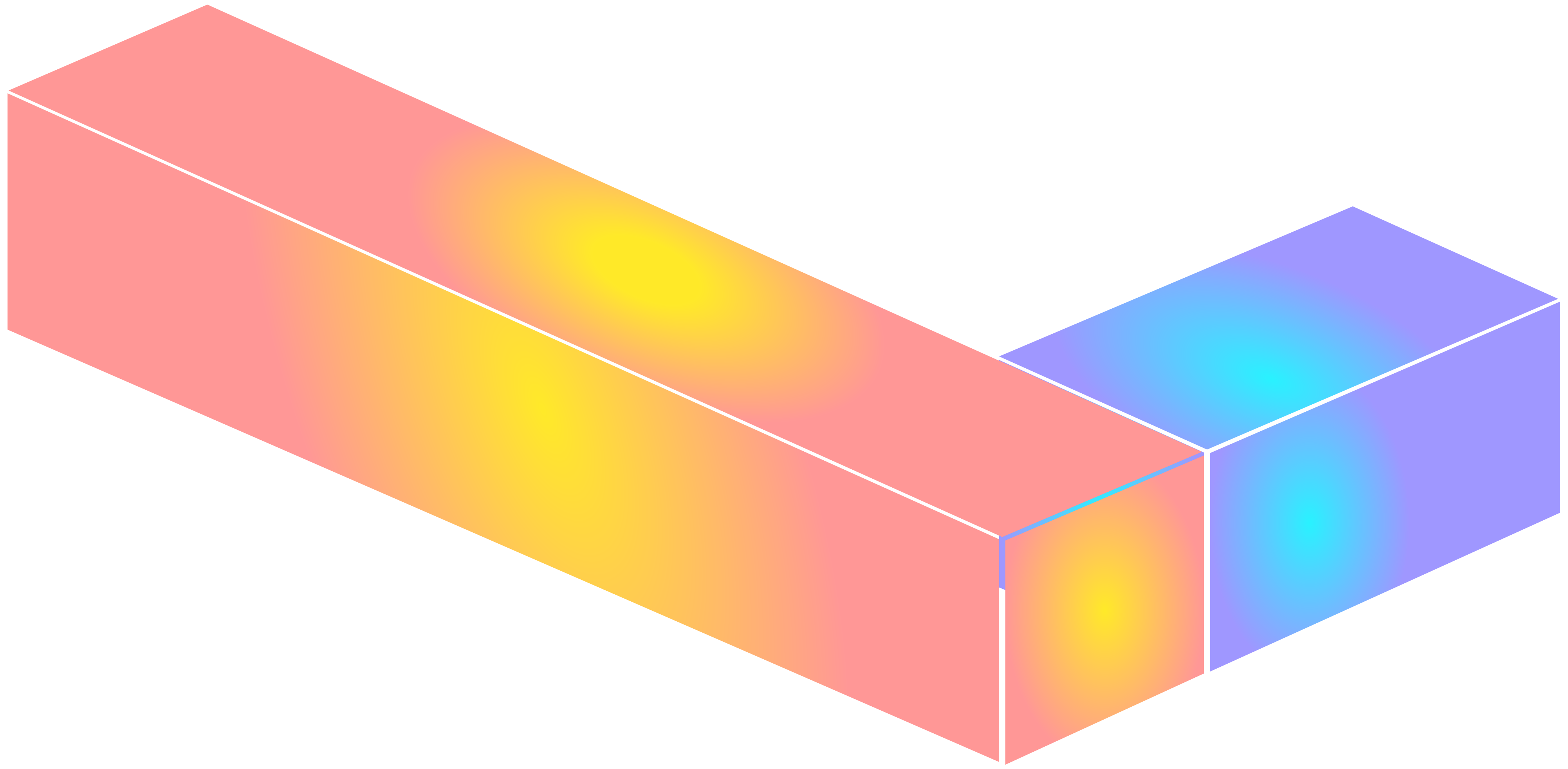
Consideration - Adequate Number of Boxes

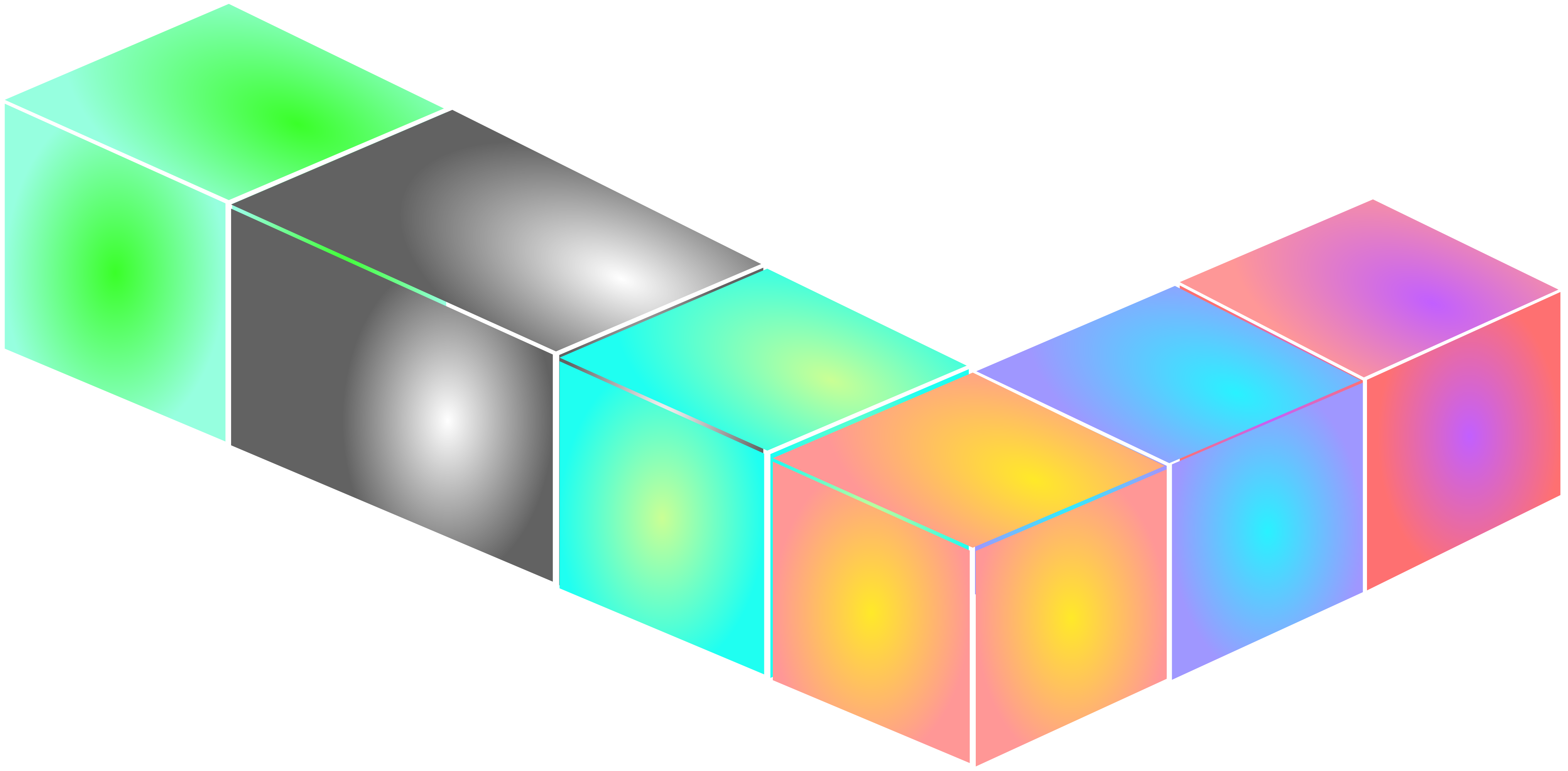


Consideration - Adequate Number of Boxes

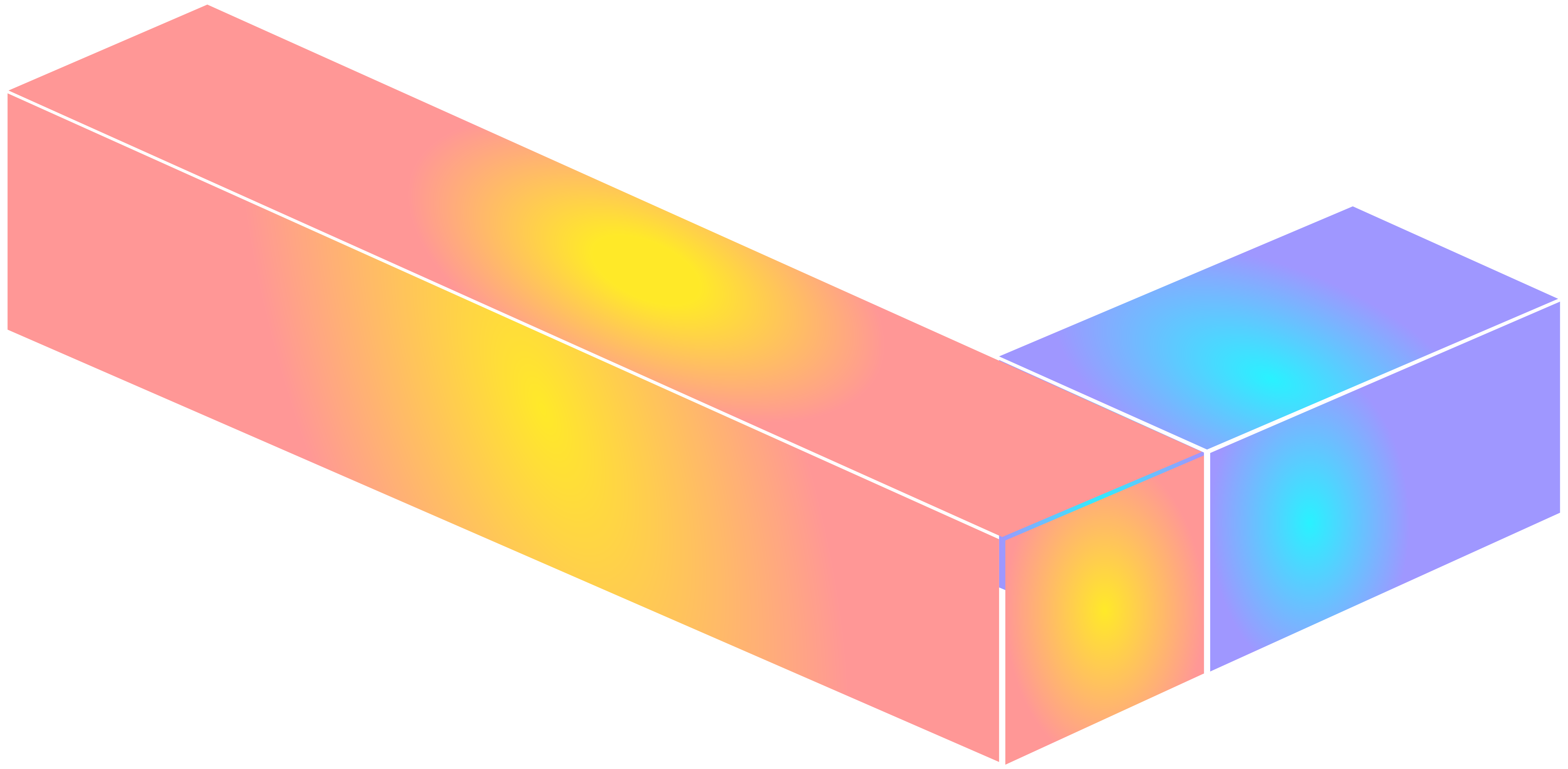


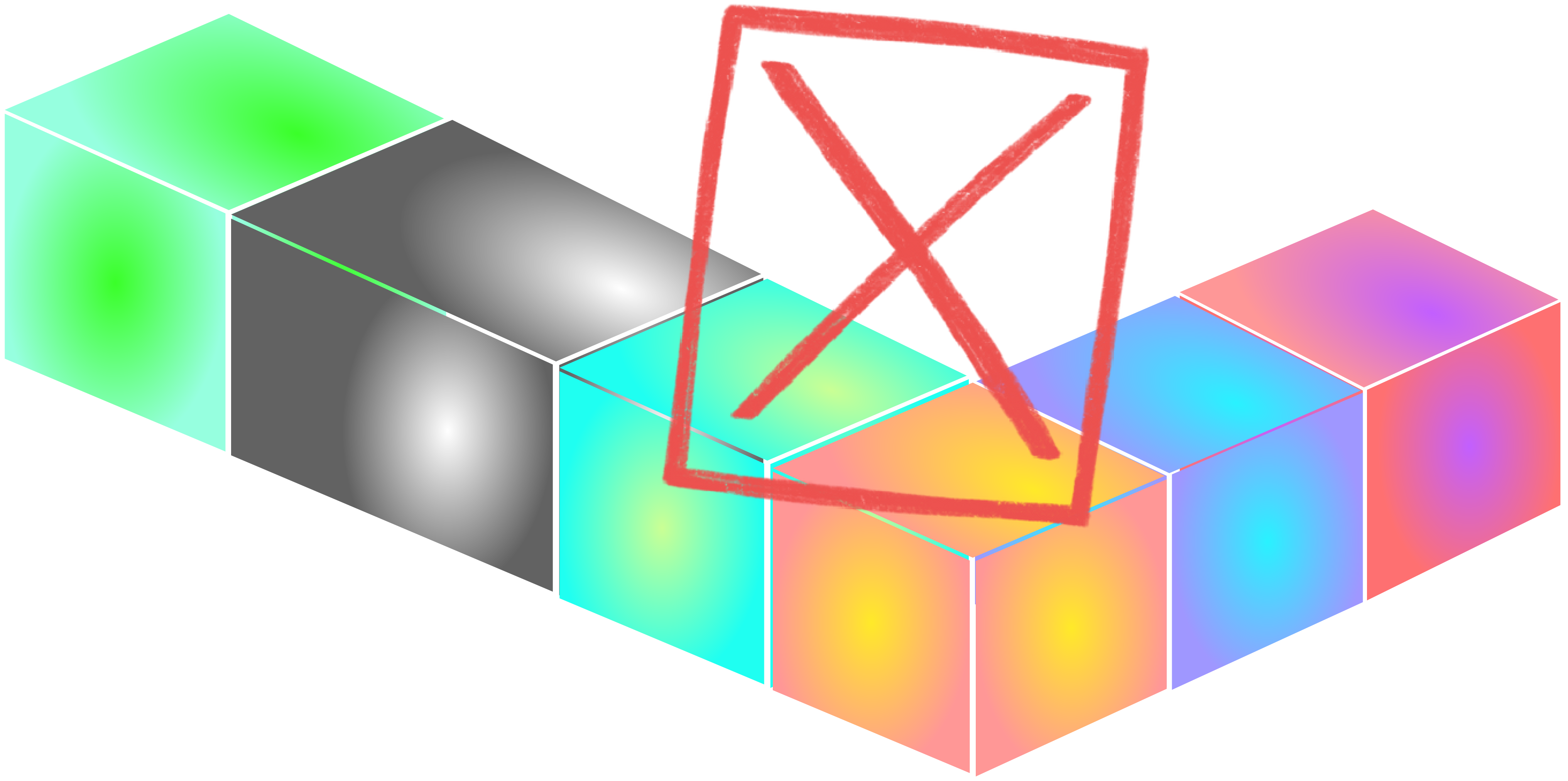
Consideration - Adequate Number of Boxes





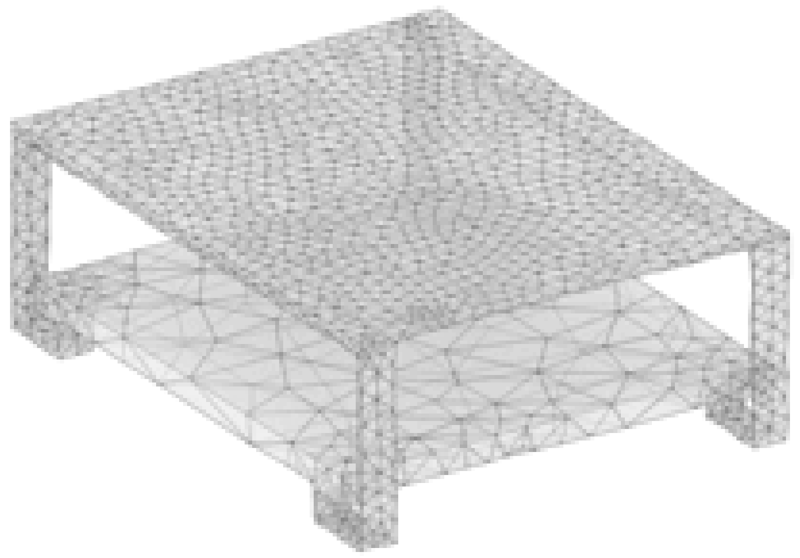
Consideration - Adequate Number of Boxes





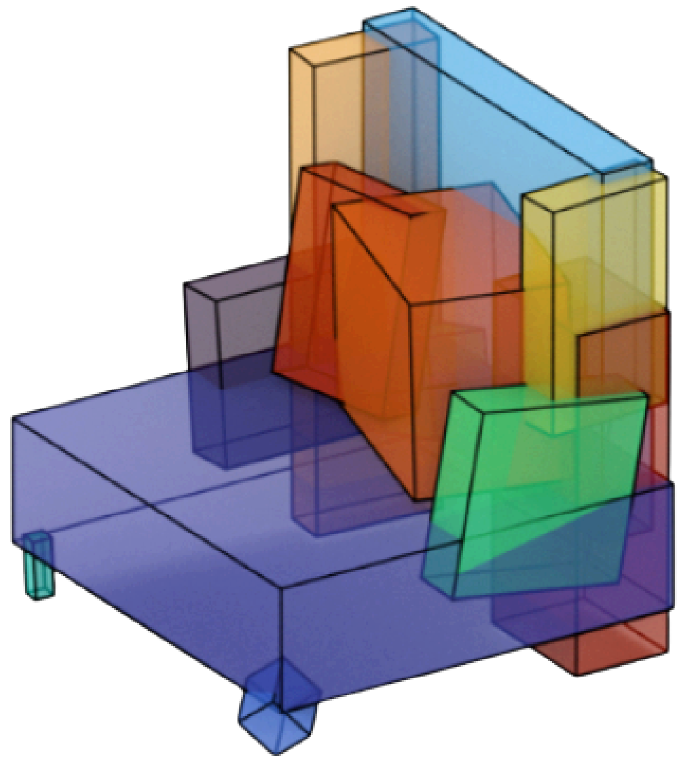
Starting Point - Initial Bounding Boxes

Initial Bounding Boxes : Preprocessing CoACD(Wei et al.) with sufficient number of boxes



→

→



Input

CoACD

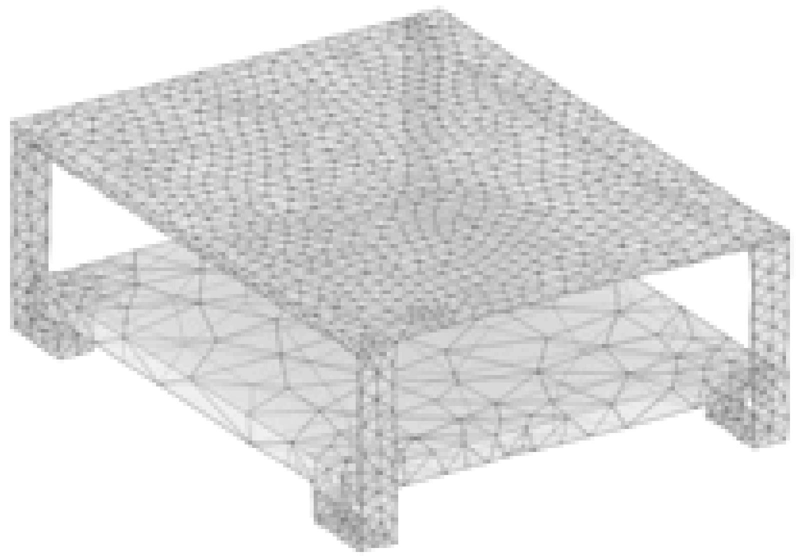
Preprocessed
(Starting Point)

Wei, X., Liu, M., Ling, Z., & Su, H. (2022). Approximate convex decomposition for 3d meshes with collision-aware concavity and tree search. ACM Transactions on Graphics (TOG), 41(4), 1-18.

Starting Point - Initial Bounding Boxes

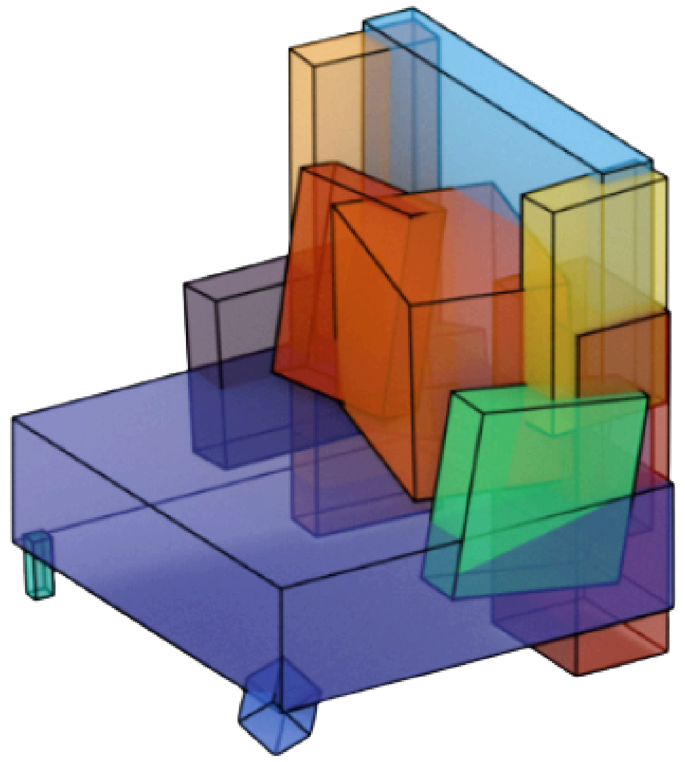
Initial Bounding Boxes : Preprocessing CoACD(Wei et al.) with sufficient number of boxes

↳ Doesn't satisfy our geometric criterias !! (Full Coverage & Tightness)



→

→



Input

CoACD

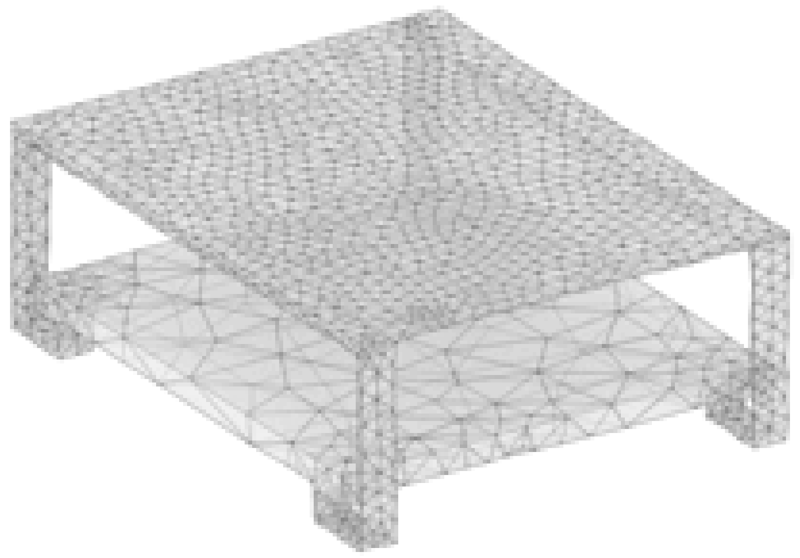
Preprocessed
(Initial Boxes)

Wei, X., Liu, M., Ling, Z., & Su, H. (2022). Approximate convex decomposition for 3d meshes with collision-aware concavity and tree search. ACM Transactions on Graphics (TOG), 41(4), 1-18.

Starting Point - Initial Bounding Boxes

Initial Bounding Boxes : Preprocessing CoACD(Wei et al.) with sufficient number of boxes

↳ Doesn't satisfy our geometric criterias !! (Full Coverage & Tightness)

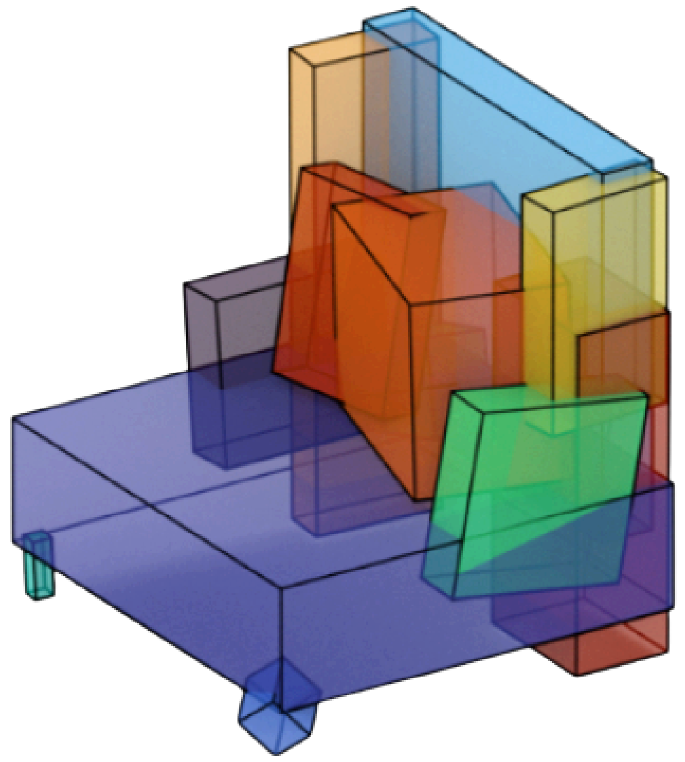


Input

→

CoACD

→



Preprocessed (Initial Boxes)

→

$$\underset{\{B_i\}_{i=1}^{k}}{\operatorname{argmin}} \; BVS(B) \quad s.t. \quad Cov(B) = 1$$

Wei, X., Liu, M., Ling, Z., & Su, H. (2022). Approximate convex decomposition for 3d meshes with collision-aware concavity and tree search. ACM Transactions on Graphics (TOG), 41(4), 1-18.

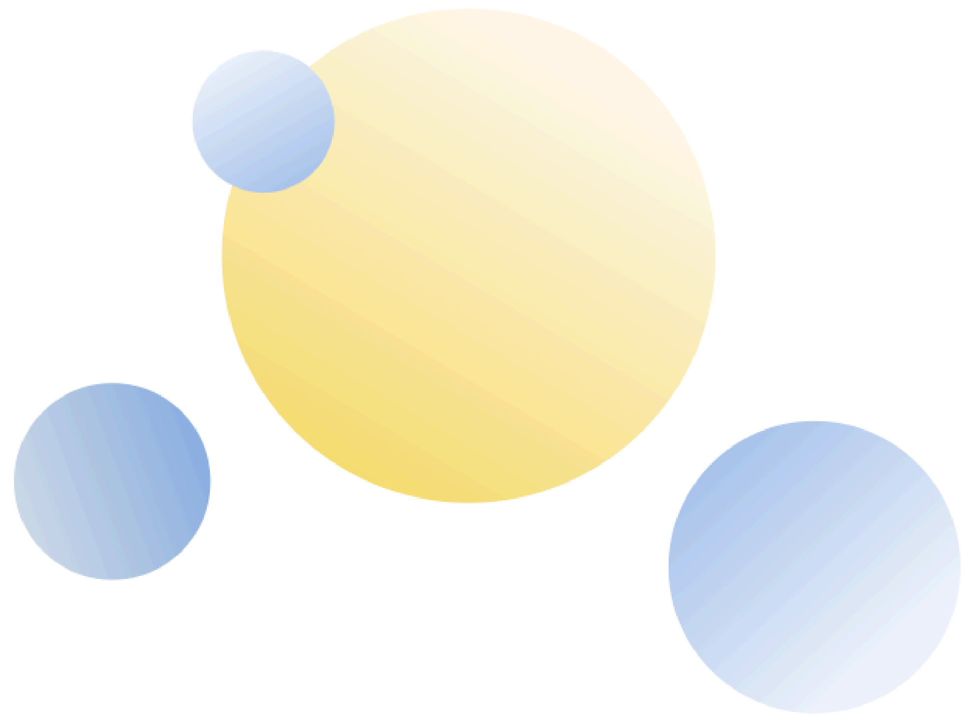
Briefing of Approach

Get tighter bounding boxes by making adjustment operations from the initial bounding boxes

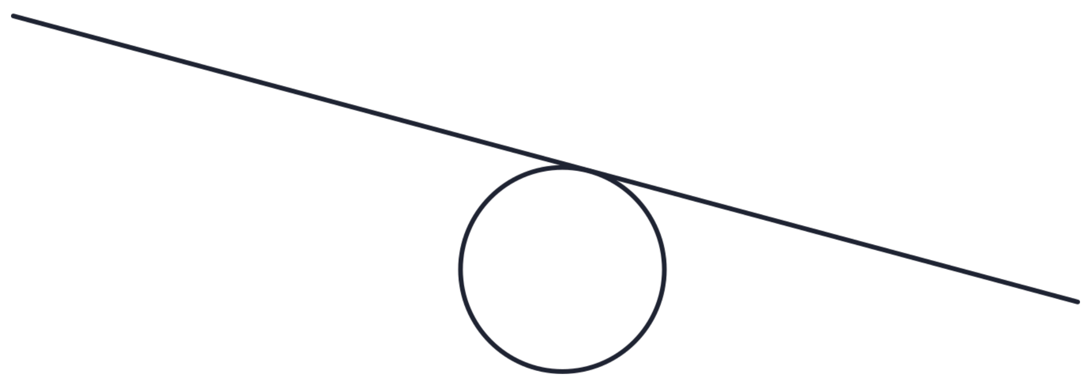
Method of modifying the bounding boxes iteratively using discrete unit actions

Complex Optimization Problem → **Simple Search Problem**

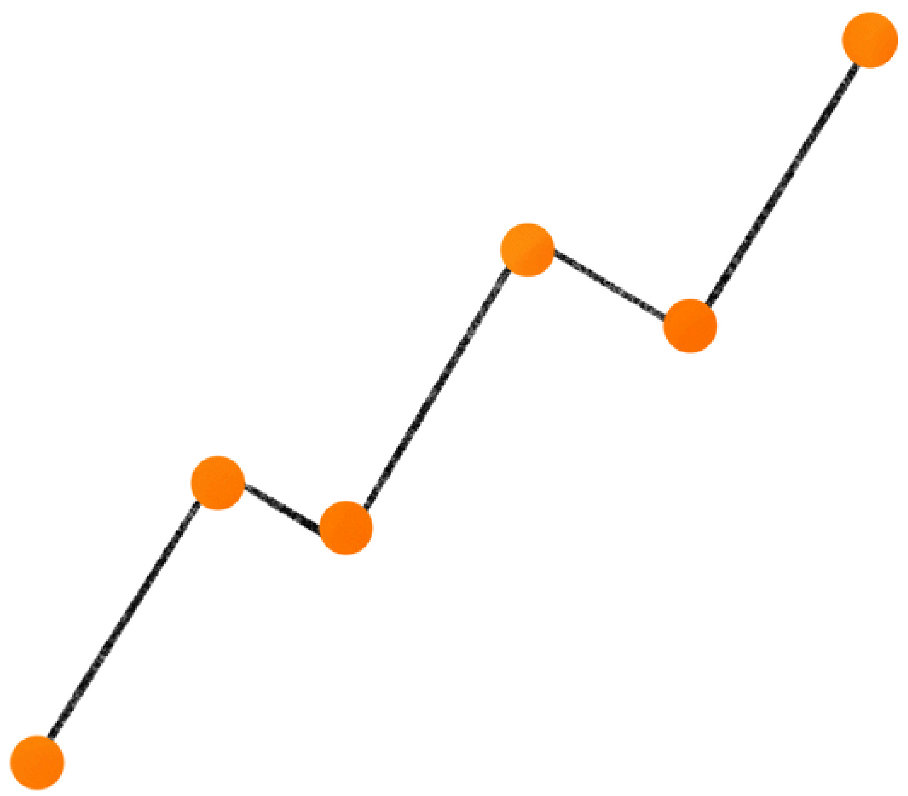
Challenges to Breakthrough



Generalization
&
Robustness



Trade-off between
Coverage & Tightness



Search Space and
Optimization

Our Approaches

- Redefined Objective Function
- Monte Carlo Tree Search (MCTS)

Comparision with Other Approaches

- Neural network-based learning approaches
- Iterative approach using Lloyd sampling
- Reinforcement learning approaches

Rationale for Outperforming Existing Approaches

Improved Coverage

Our approach aims to achieve full coverage of the 3D shape,
allowing for more accurate reconstruction and preservation of details

Enhanced Tightness

By refining the bounding boxes using MCTS, we expect to obtain tighter results compared to the initial bounding boxes

Human-like Perception

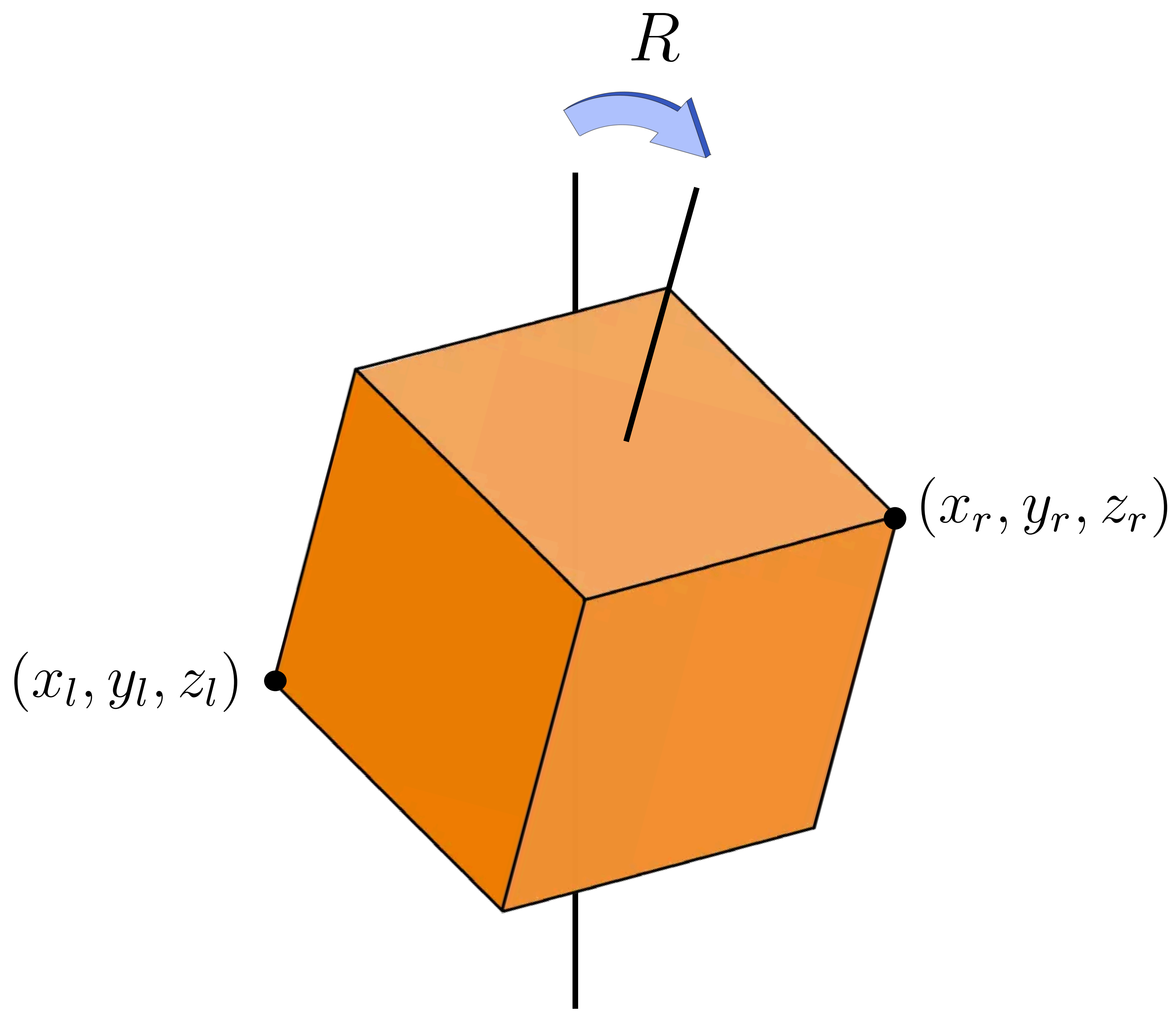
Guided by geometric criteria and MCTS, is more aligned with human perception

Generalization of Unseen Examples

Existing NN-based approaches often struggle when presented with unseen examples

Definition of State and Action - State

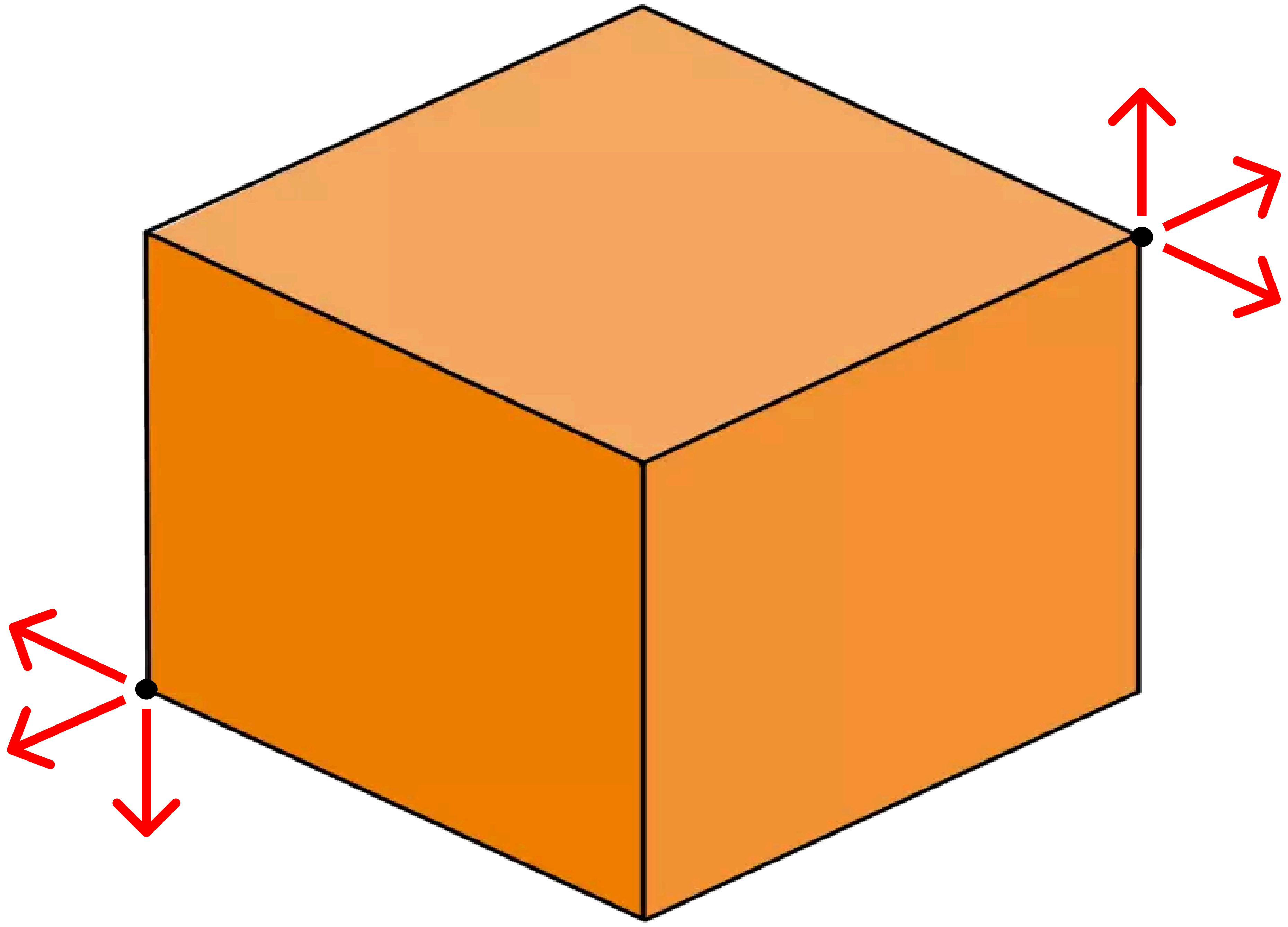
$$B_t = \{(B_{t_i})\}_{i=1}^{N} = \{(x_l, y_l, z_l, x_r, y_r, z_r, R)\}_{i=1}^{N}$$



Definition of State and Action - Action

- Adding or subtracting a predefined unit scale : 6 × 2
- Changing the rotation matrix to orient the bounding box correctly : 1

Total 13 (6x2+1) Actions



Definition of Objective Function

Redefine Objective Function

Hard Constraint

$$\underset{\{B_i\}_{i=1}^{k}}{\operatorname{argmin}} \; BVS(B) \quad s.t. \quad Cov(B) = 1$$

Soft Constraint

$\alpha \sim 100$

$$\underset{\{B_{t_i}\}_{i=1}^{N}}{\operatorname{argmin}} \; \{BVS(B) - \alpha Cov(B)\}$$

Definition of Objective Function

Redefine Objective Function

Hard Constraint

$$\underset{\{B_i\}_{i=1}^{k}}{\operatorname{argmin}} \; BVS(B) \quad s.t. \quad Cov(B) = 1$$

↓

Reward

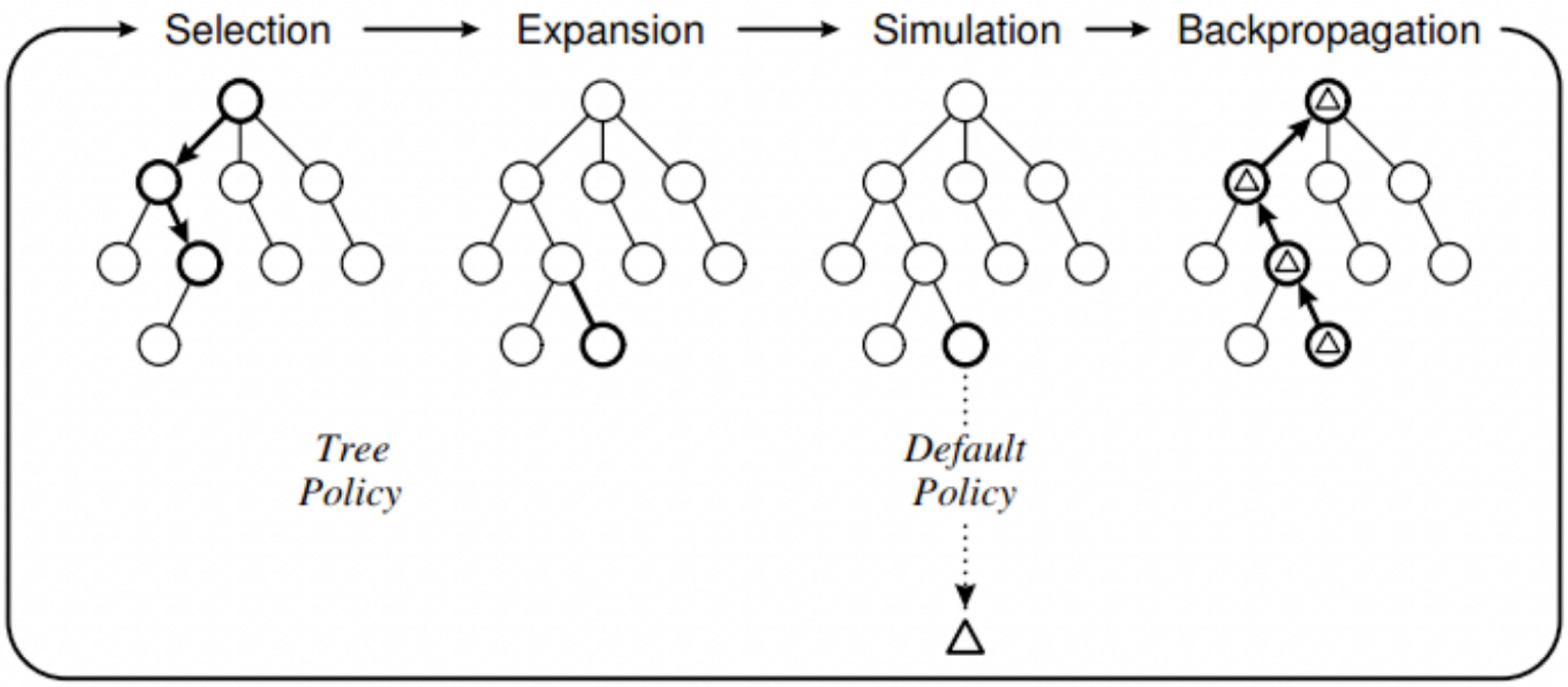
Soft Constraint

$\alpha \sim 100$

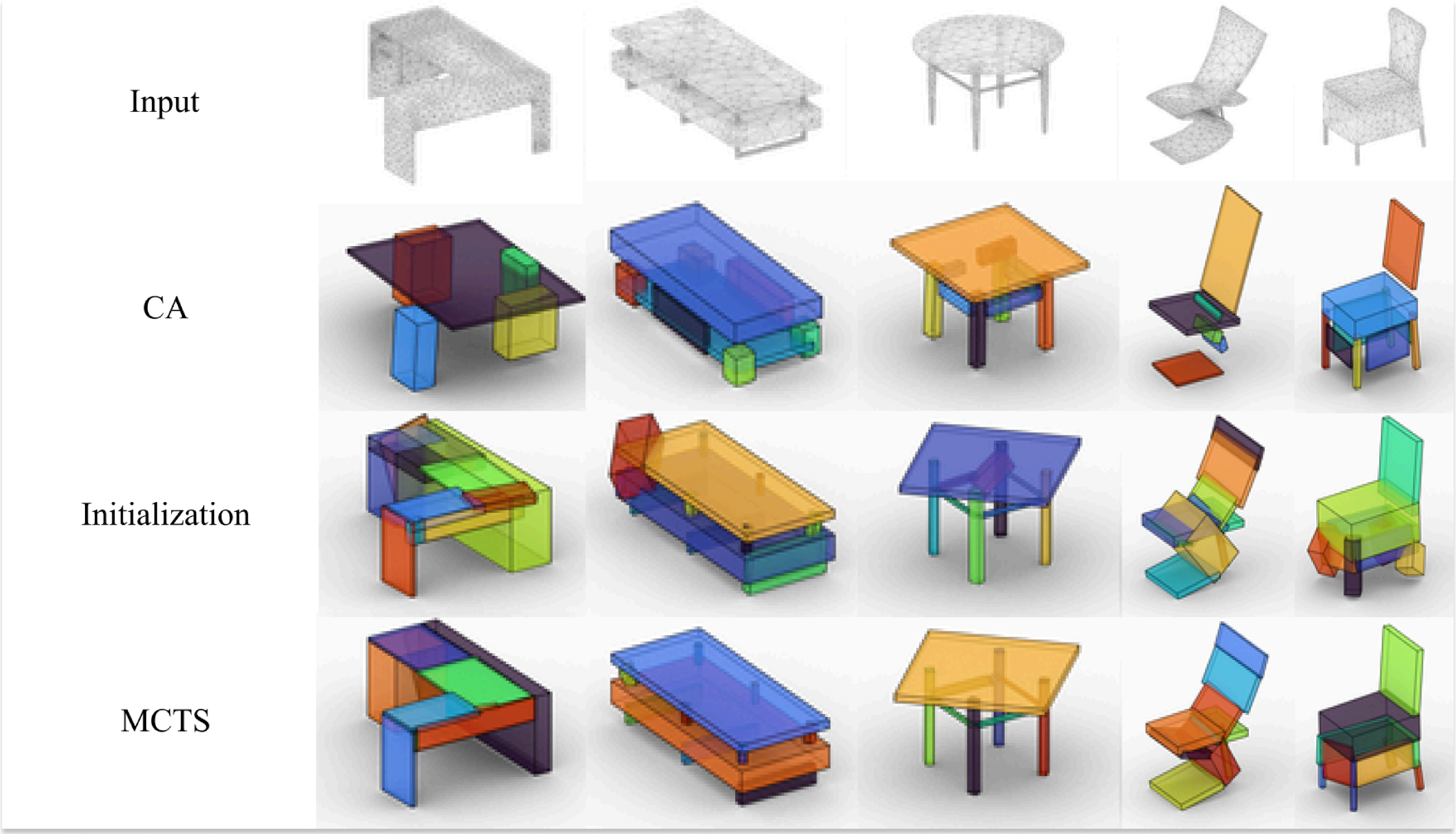
$$\underset{\{B_{t_i}\}_{i=1}^{N}}{\operatorname{argmin}} \; \{BVS(B) - \alpha Cov(B)\}$$

Monte Carlo Tree Search (MCTS)

- Node : State
- Edge : Action
- Reward : $BVS(B) - \alpha Cov(B)$



Qualitative Comparison Results



Initial Mesh

Result for NN based learning approach

CoACD + Preprocessing

Our Result

Evaluation Metrics

1. Maximum Local Outside Volume $$\mathrm{MOV}(B, \{S_i\}) = \max_i \frac{\mathrm{vol}(B_i \setminus S)}{\mathrm{vol}(S_i)}$$
2. Total Outside Volume $$\mathrm{TOV}(B) = \frac{\mathrm{vol}(\bigcup_i B_i \setminus S)}{\mathrm{vol}(S)}$$
3. Chamfer Distance $$d_{CD}(S_1, S_2) = \sum_{x \in S_1} min_{y \in S_2} ||x - y||_2^2 + \sum_{y \in S_2} \min_{x \in S_1} ||x - y||_2^2$$
4. Volumetric Intersection over Union $$\mathrm{VIoU}(B) = \frac{\mathrm{vol}(S \cap \bigcup_i B_i)}{\mathrm{vol}(S \cup \bigcup_i B_i)}$$

Quantitative Comparison Results

	Table				Chair			
Metrics	BVS↓	Cov↑	MOV↓	TOV↓	BVS↓	Cov↑	MOV↓	TOV↓
CA [5]	1.08	0.61	-	-	1.10	0.72	-	-
Initialization	2.21	**1.00**	3.95	1.04	2.09	**1.00**	4.29	0.83
MCTS	**1.72**	**1.00**	**2.82**	**0.69**	**1.64**	**1.00**	**2.96**	**0.58**

Table 1: Comparison of tightness of bounding boxes between our initialization, MCTS, and baselines.

	Table		Chair	
Metrics	CD↓	VIoU↑	CD↓	VIoU↑
CA [5]	1.19	0.45	**0.812**	0.56
Initialization	0.757	0.60	0.962	0.63
MCTS	**0.641**	**0.67**	0.844	**0.69**

Table 2: Comparison of reconstruction performance between initialization, MCTS, and baseline. Note that all CD is scaled by 1000.

What We've Learned

Chanhyeok Park

Learning is not always the answer. Efficient exploration can give us good solution if combined with well defined criterias (Rewards).

Inhwa Song

Many design decisions required throughout the pipeline can be some selection among the options, but also generation. For successful decision making sequences, it's important to define the problem concretely and precisely.

Sangmin Lee

Defining a proper objective function, and sometimes softening the constraints can make better search

Thank You :)