

Combinative Matching for Geometric Shape Assembly

Nahyuk Lee^{1*}

Juhong Min^{1,2*}

Junhong Lee¹

Chunghyun Park¹

Minsu Cho^{1,3}

¹ Pohang University of Science and Technology (POSTECH)

² Samsung Research America

³ RLWRLD

* Equal Contribution



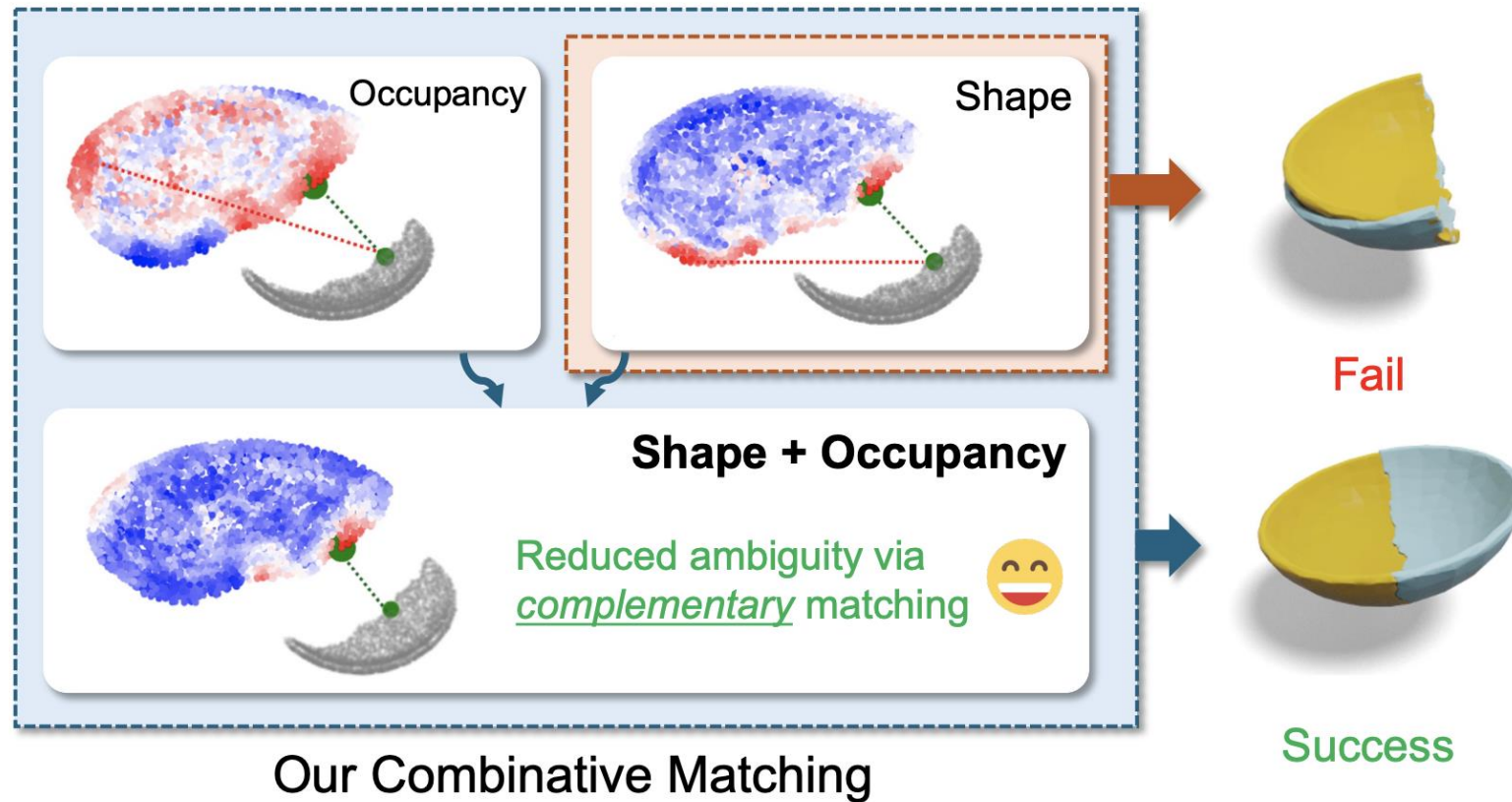
Highlight Presentation ✨

POSTECH
POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

SAMSUNG

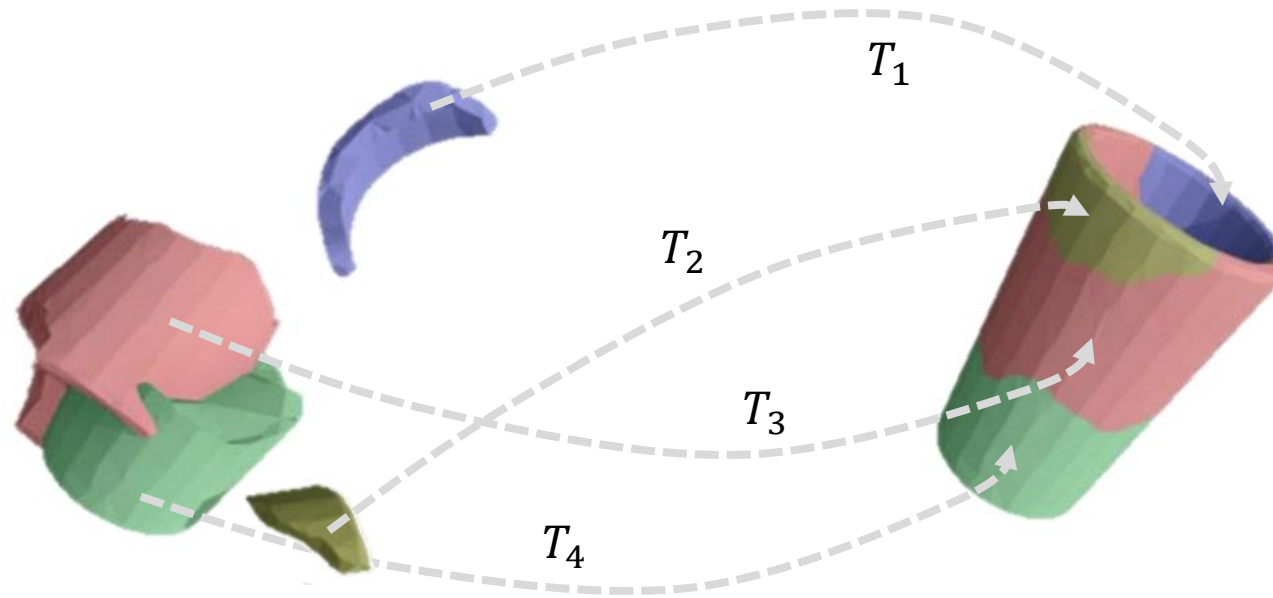
RLWRLD

Combinative Matching for Geometric Shape Assembly



3D Geometric Shape Assembly

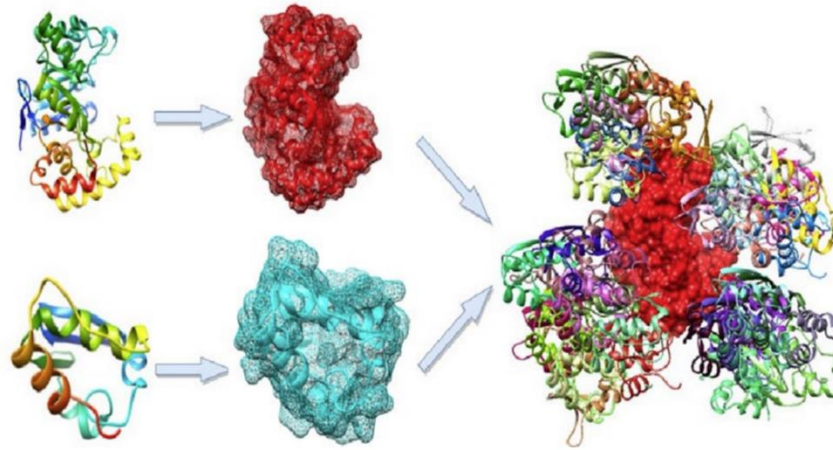
Given a set of fractured pieces $\mathcal{P} = \{P_1, P_2, \dots, P_N\}$, our goal is to recover the 6-DoF pose $\{T_1, T_2, \dots, T_N\}$, in $SE(3)$ for each piece and restore the underlying object.



3D Geometric Shape Assembly



Archaeology

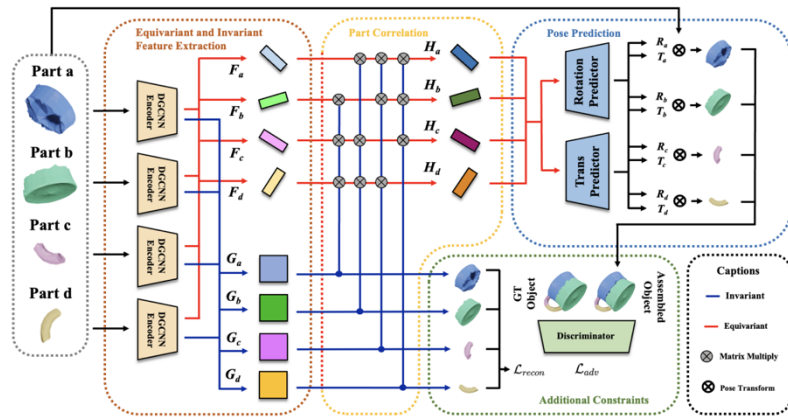


Scientific domain
(e.g., Protein docking)



Robotics

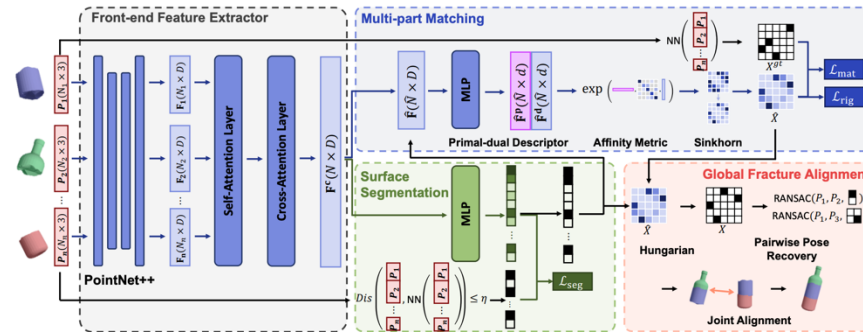
State-of-the-art methods



Wu et al.

(Wu et al., ICCV 2023)

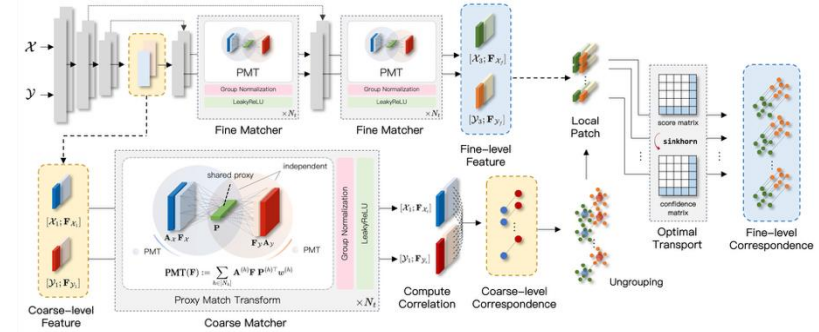
- 😊 Leveraging SE(3) equivariance for multi-part correlations
- 😞 Pose regression with limited spatial information



Jigsaw

(Lu et al., NeurIPS 2023)

- 😊 Effective matching between mating surfaces of parts
- 😞 Heavily rely on segmentation of mating surfaces



PMTR

(Lee et al., ICML 2024)

- 😊 Approximate high-order conv. for dense matching of parts
- 😞 Often fails to establish good correspondences between mating surfaces

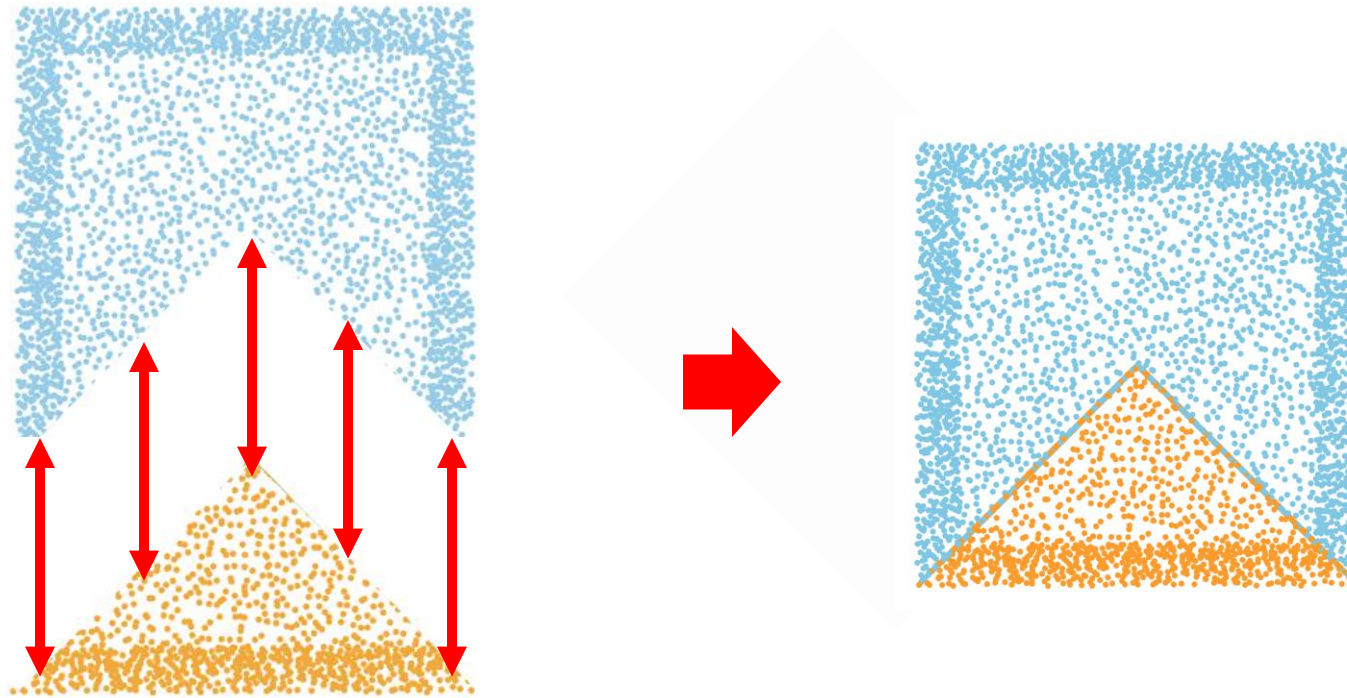
Wu et al., "Leveraging SE(3) Equivariance for Learning 3D Geometric Shape Assembly", ICCV 2023

Lu et al., "Jigsaw: Learning to Assemble Multiple Fractured Objects", NeurIPS 2023

Lee et al., "3D Geometric Shape Assembly via Efficient Point Cloud Matching", ICML 2024

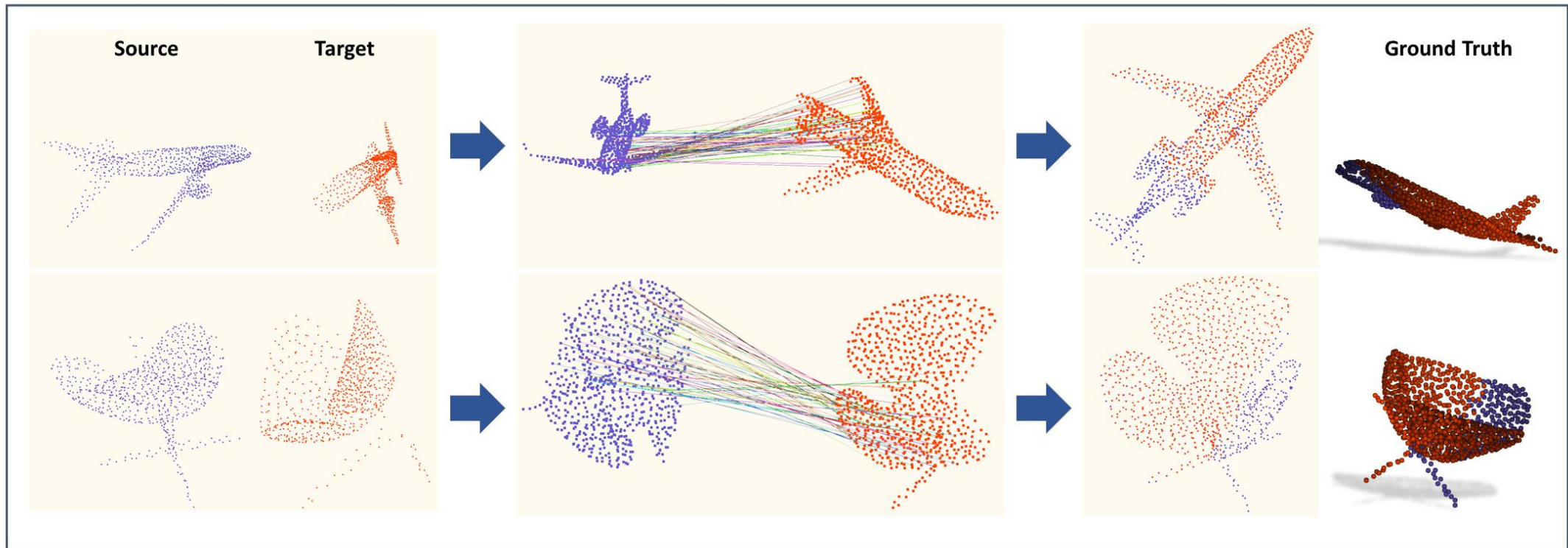
Previous Approach

Equative matching: establishing correspondences based on visual similarity alone
under the assumption of the same appearance between mating parts.



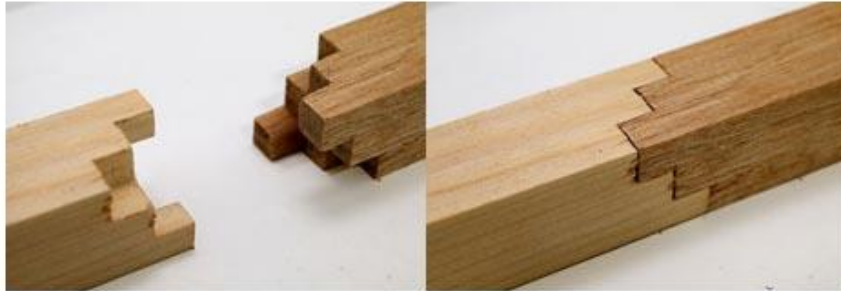
Previous Approach

Equative matching: establishing correspondences based on visual similarity alone
under the assumption of the same appearance between mating parts.



Point Cloud Registration

Beyond Equative Matching



a) I-axial



b) I-perpendicular*



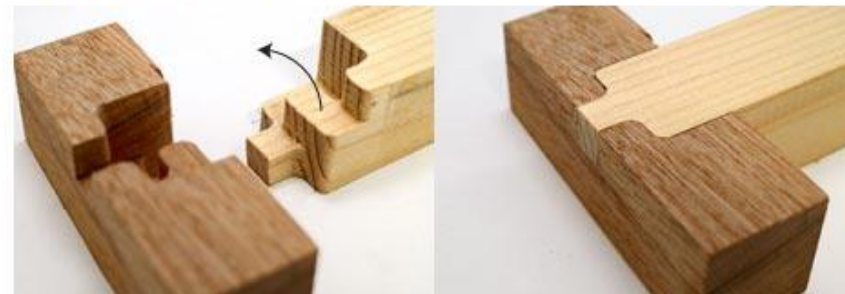
c) L-axial



d) L-perpendicular*

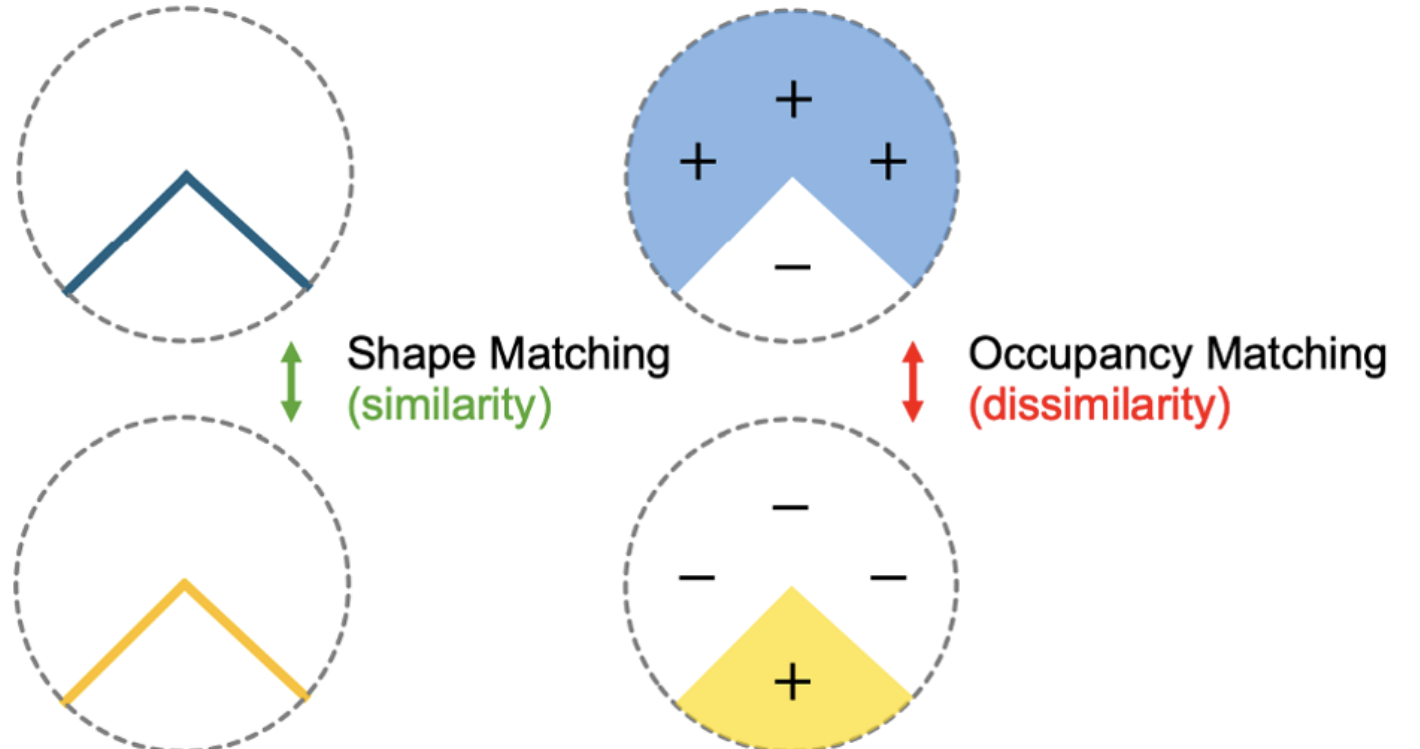
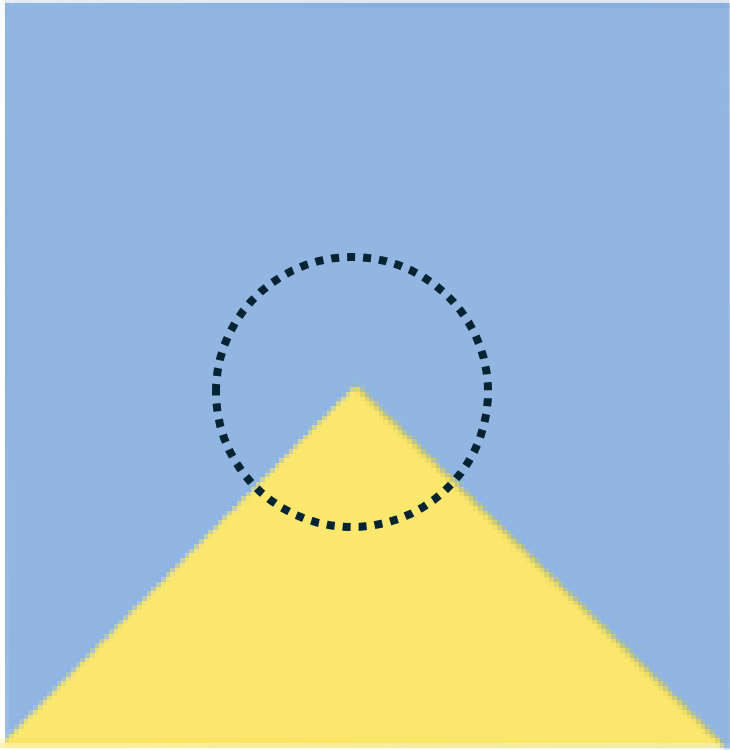


e) T-axial

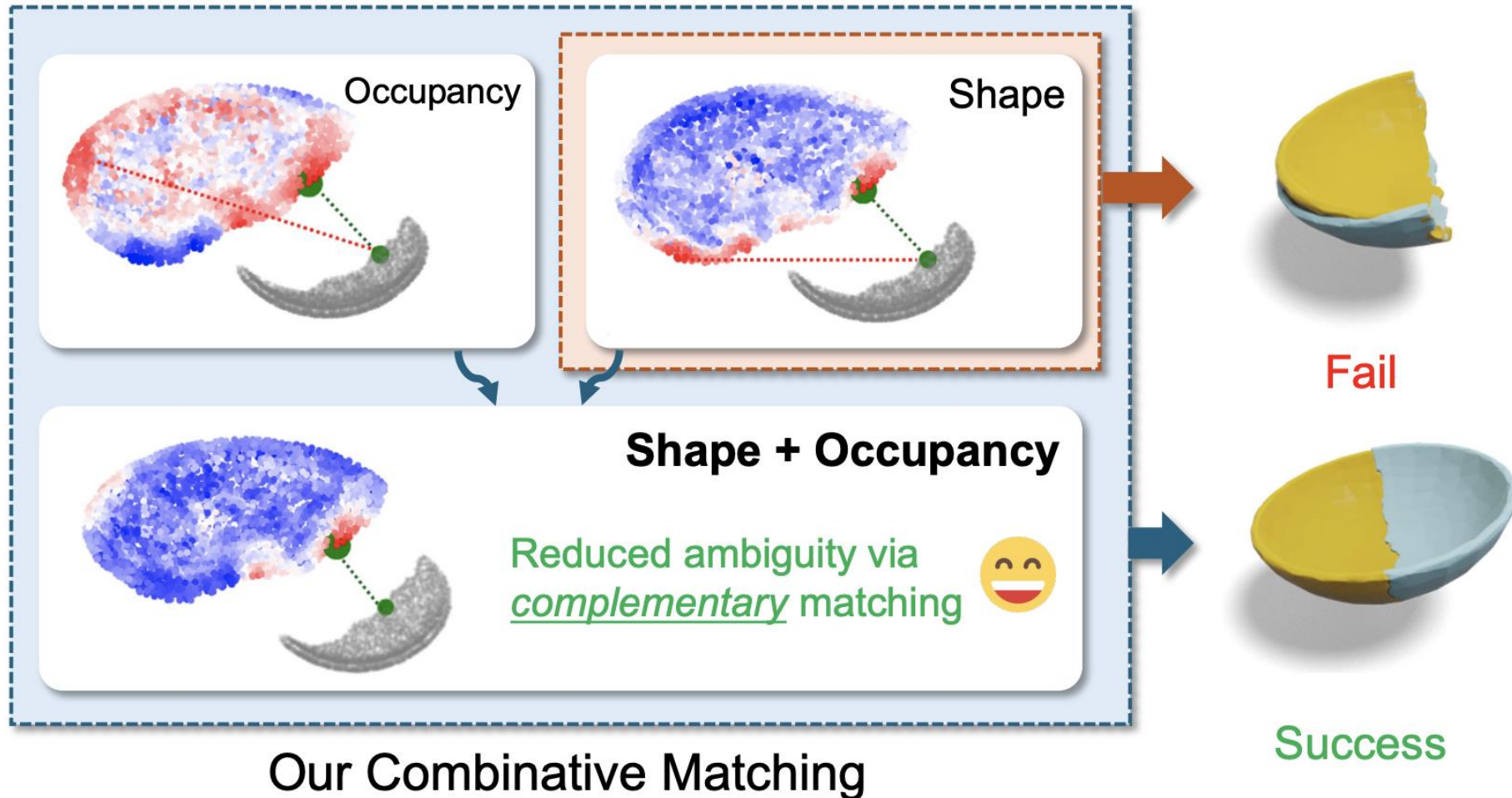


f) T-perpendicular*

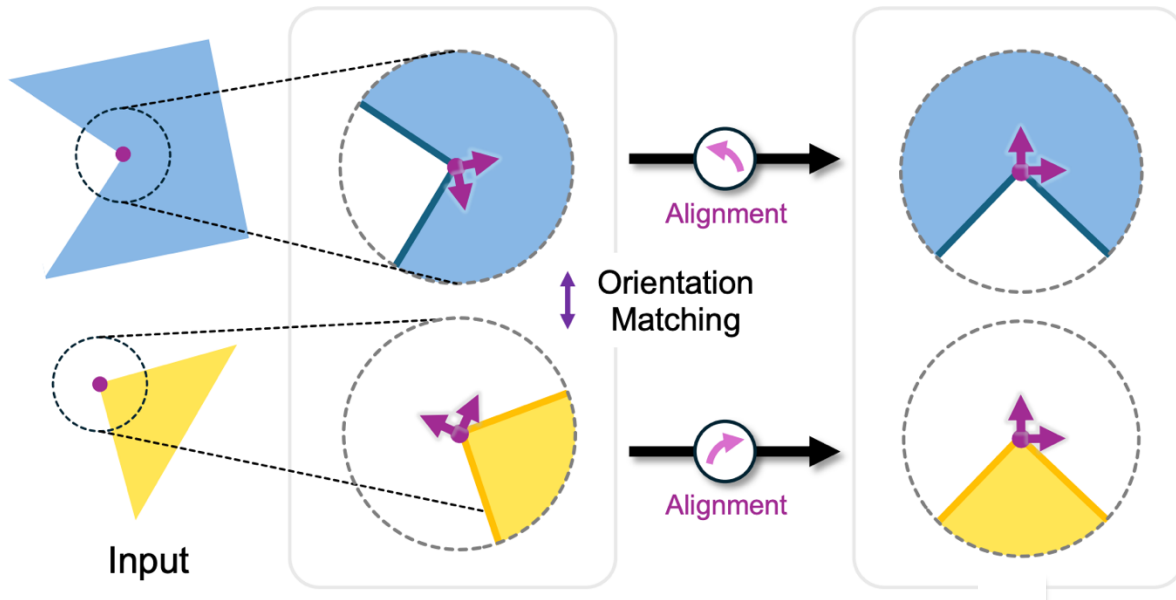
Beyond Equative Matching



Combinative Matching: ‘combination’ or ‘joining’ of elements form the basis for matching

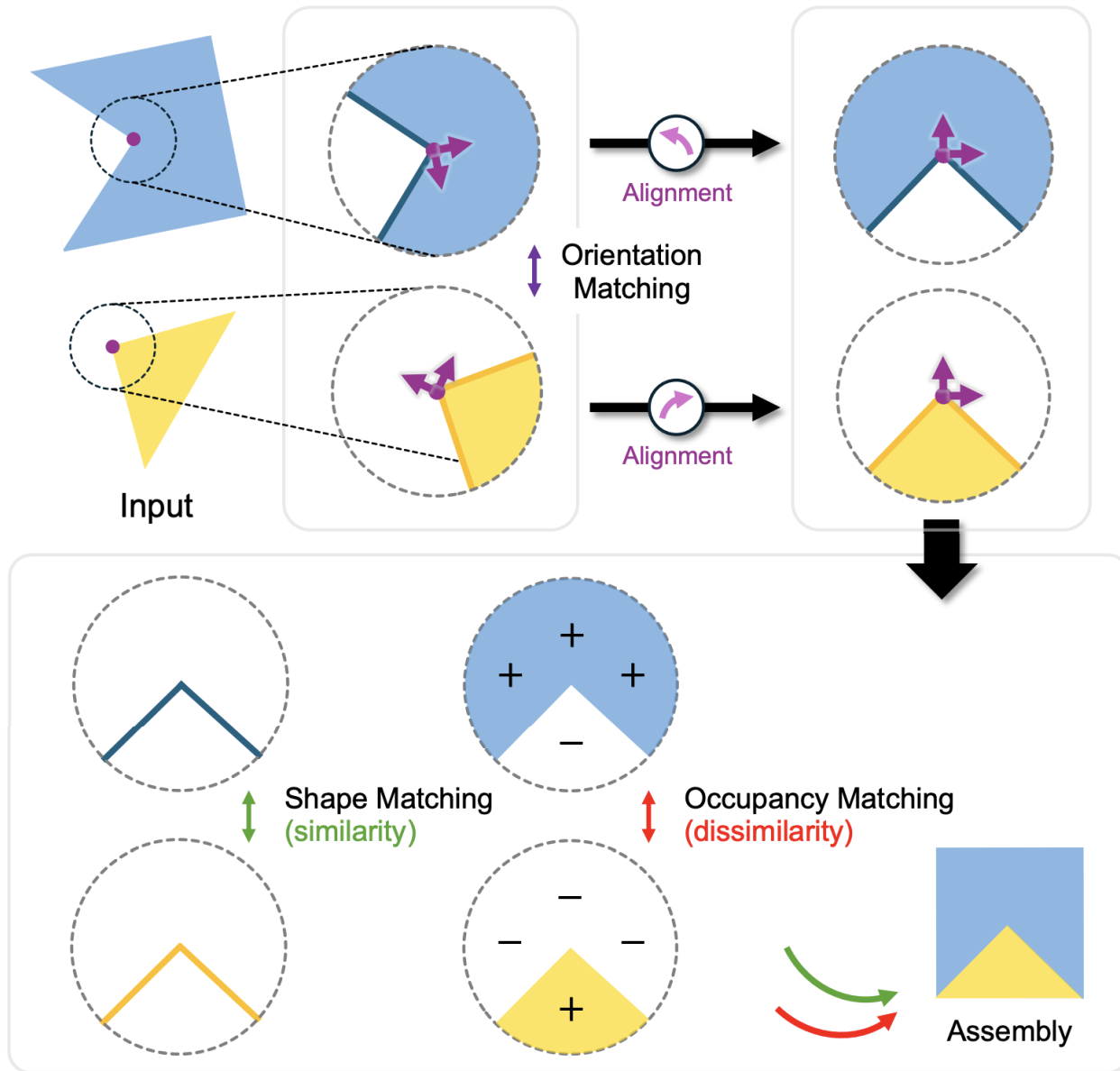


Objectives for Combinative Matching



\mathcal{L}_d : maximize 'orientation' consistency
(in $SO(3)$)

Objectives for Combinative Matching

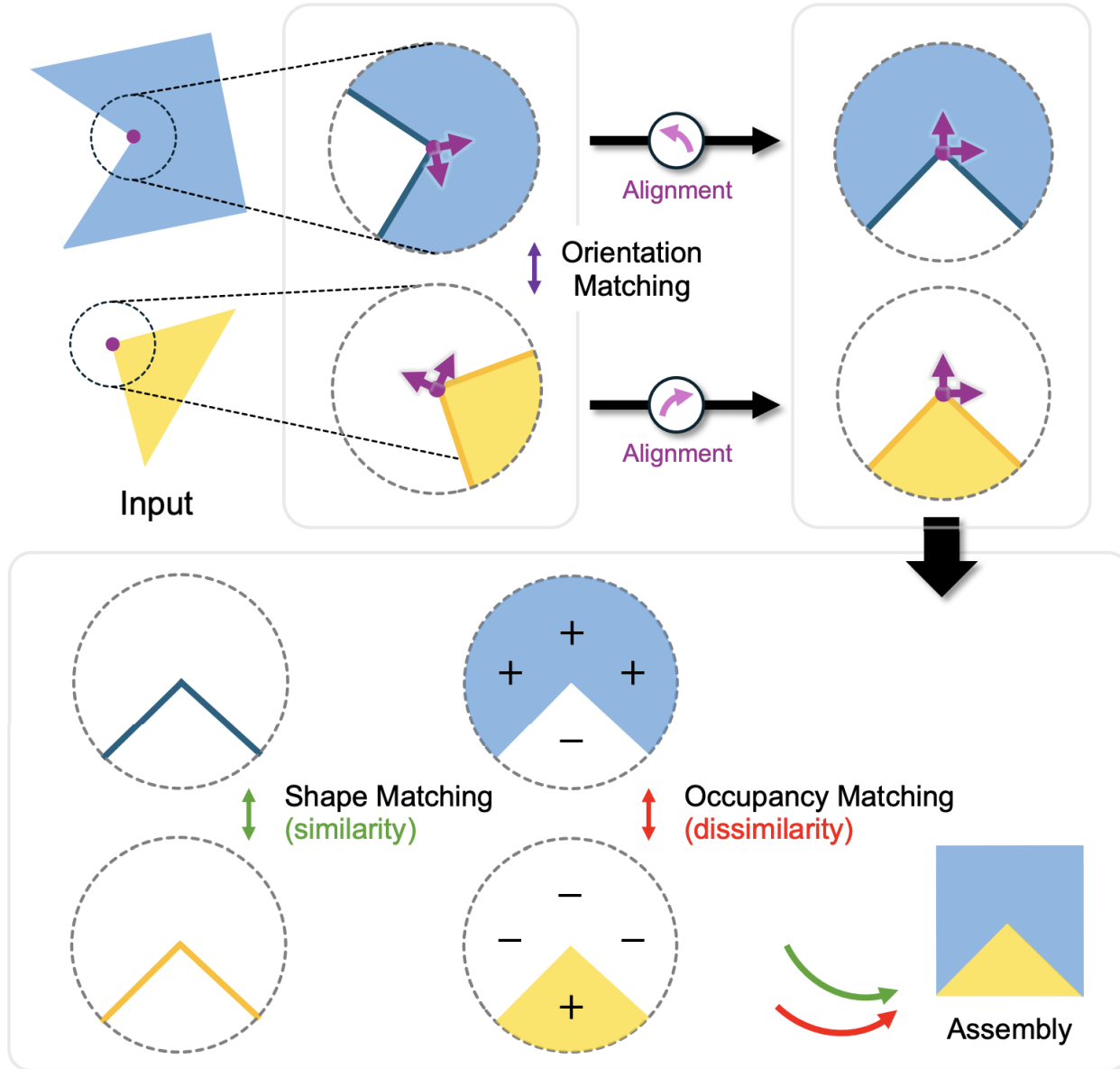


\mathcal{L}_d : maximize 'orientation' consistency
(in $SO(3)$)

\mathcal{L}_s : maximize 'visual' feature similarity

\mathcal{L}_o : maximize 'occupancy' feature dissimilarity

Objectives for Combinative Matching



\mathcal{L}_d : maximize 'orientation' consistency



$$\mathcal{L}_d = \frac{1}{|\mathcal{C}|} \sum_{(i,j) \in \mathcal{C}} \|(\mathbf{F}_d^P)_i \mathbf{R}^P - (\mathbf{F}_d^Q)_j \mathbf{R}^Q\|_F$$

\mathcal{L}_s : maximize 'visual' feature similarity



$$\mathcal{L}_s = \mathbb{E}_{i \sim \mathcal{I}} \left[\log \left(1 + \sum_{j \in \mathcal{E}_p(i)} e^{\alpha_{ij}(d_{ij}^p - \Delta_p)} \cdot \sum_{k \in \mathcal{E}_n(i)} e^{\beta_{ik}(\Delta_n - d_{ik}^n)} \right) \right]$$

$$\alpha_{ij} = \gamma[d_{ij}^p - \Delta_p]_+, \quad \beta_{ik} = \gamma[\Delta_n - d_{ik}^n]_+ \quad d_{ij}^p = \|\hat{\mathbf{F}}_{s,i}^P - \hat{\mathbf{F}}_{s,j}^Q\|_2$$

\mathcal{L}_o : maximize 'occupancy' feature dissimilarity

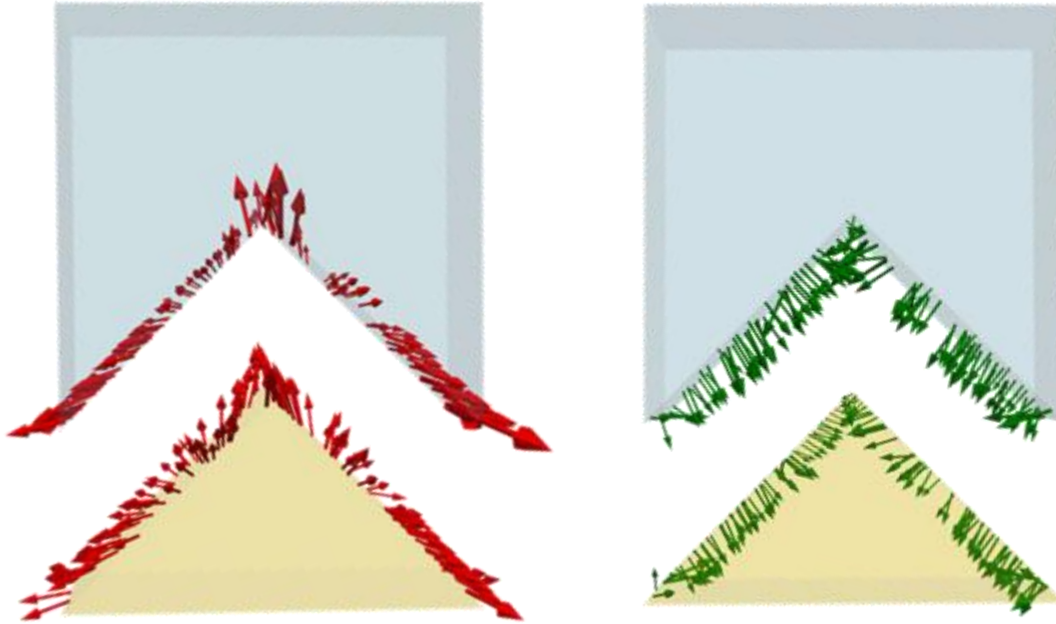


$$\mathcal{L}_o = \mathbb{E}_{i \sim \mathcal{I}} \left[\log \left(1 + \sum_{j \in \mathcal{E}_p(i)} e^{\alpha_{ij}(s_{ij}^p - \Delta_p)} \cdot \sum_{k \in \mathcal{E}_n(i)} e^{\beta_{ik}(\Delta_n - s_{ik}^n)} \right) \right]$$

$$\alpha_{ij} = \gamma[s_{ij}^p - \Delta_p]_+, \quad \beta_{ik} = \gamma[\Delta_n - s_{ik}^n]_+ \quad s_{ij}^p \approx \cos(\mathbf{F}_{o,i}^P, \mathbf{F}_{o,j}^Q)$$

Learned orientation analysis

Several notable patterns



$$\{\mathbf{x}_i\}_{i=1}^K$$

$$\{\mathbf{y}_i\}_{i=1}^K$$

\mathbf{x}_i : directed toward the center of the surface

\mathbf{x}_i : parallel to the 2D plane of the interface

\mathbf{y}_i : pointing in/outward on concave/convex

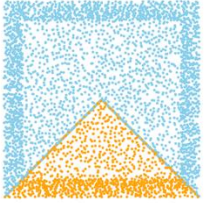
\mathbf{y}_i : magnitudes correlating with the degree of convexity/concavity

src and **trg** orientations are aligned in parallel
(enforced by objective)

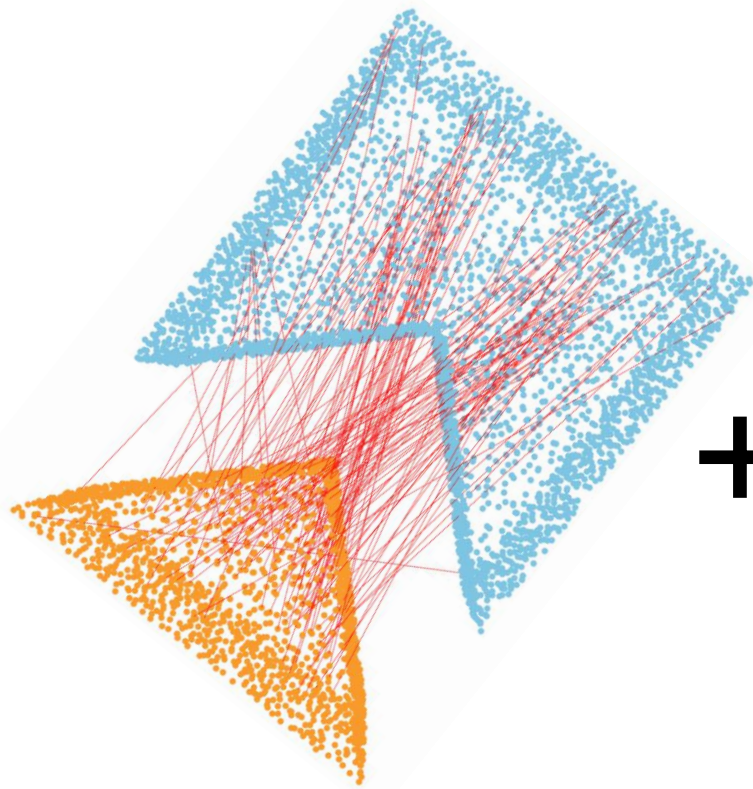
Learned orientations capture surface structures (convex, concave, curvature, 2D plane)
without any explicit supervisions dedicated to these aspects

Learned correlation analysis

Learning occupancy 'dissimilarity' effectively resolves local ambiguity

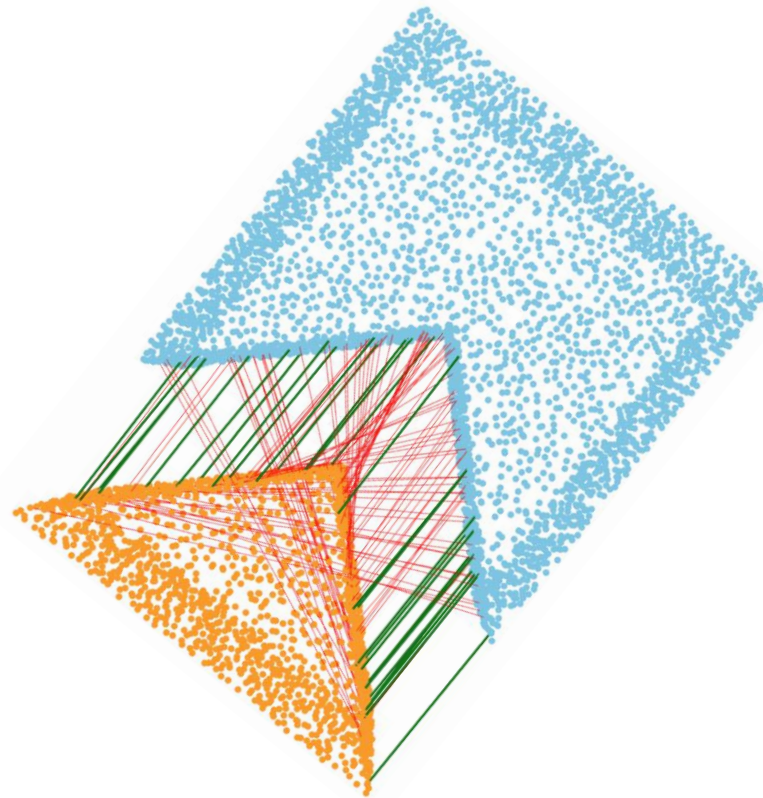


Assembly



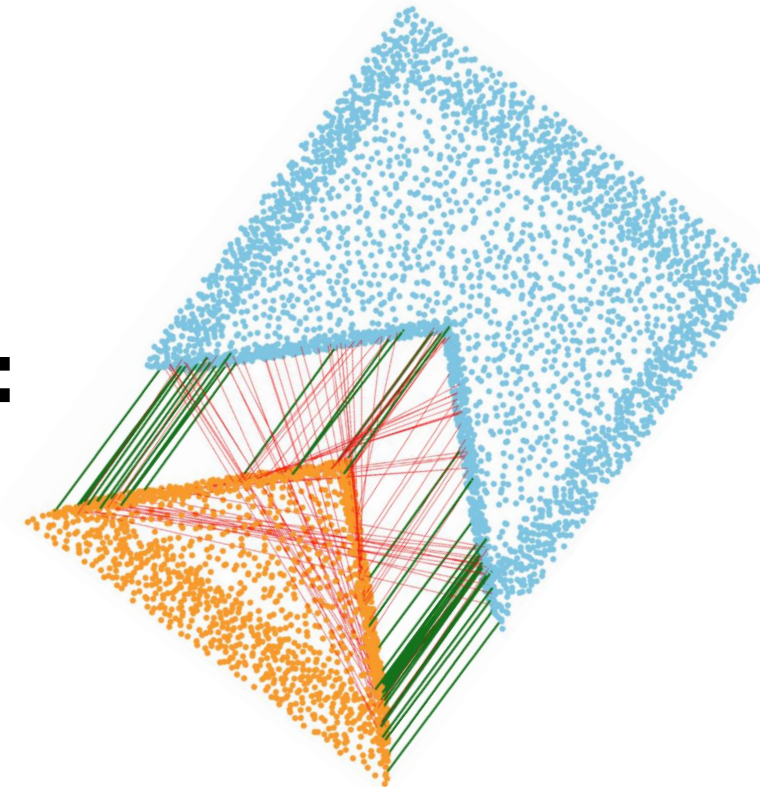
Shape only

+



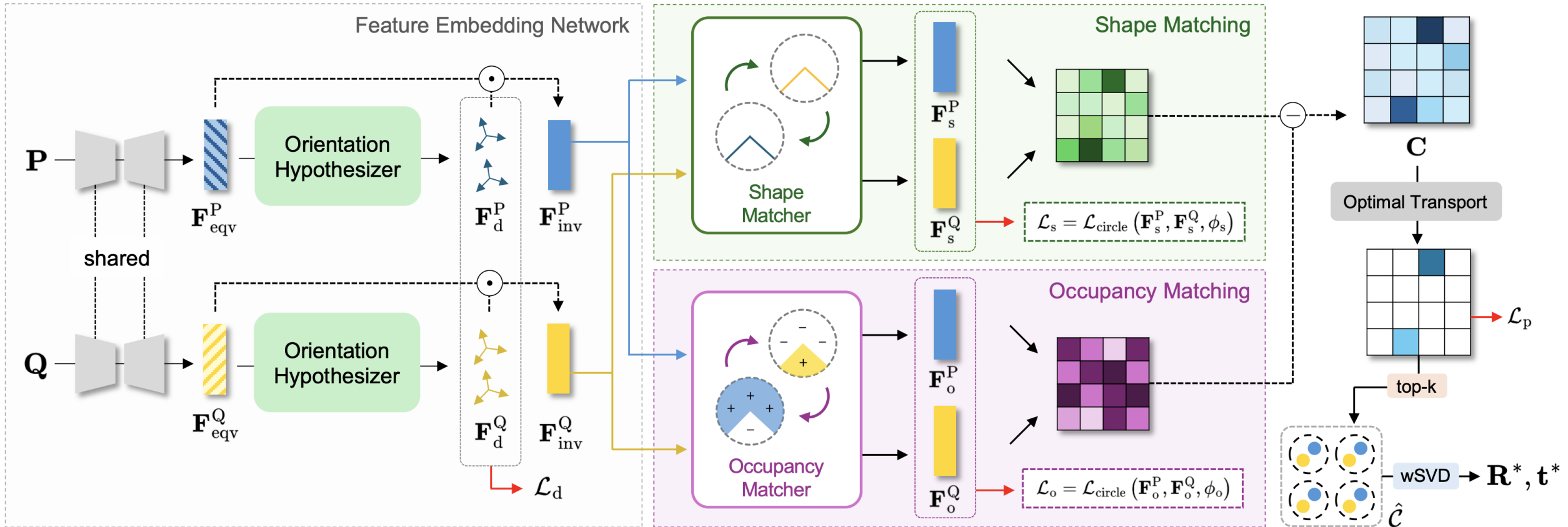
Occupancy only

=



Shape + Occupancy

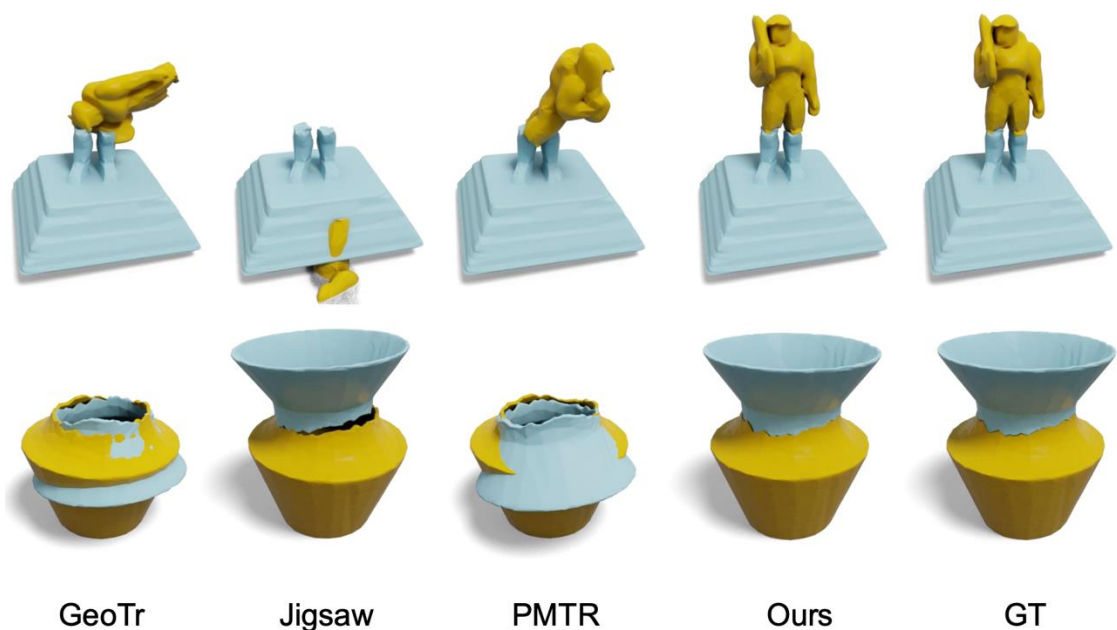
Network Architecture



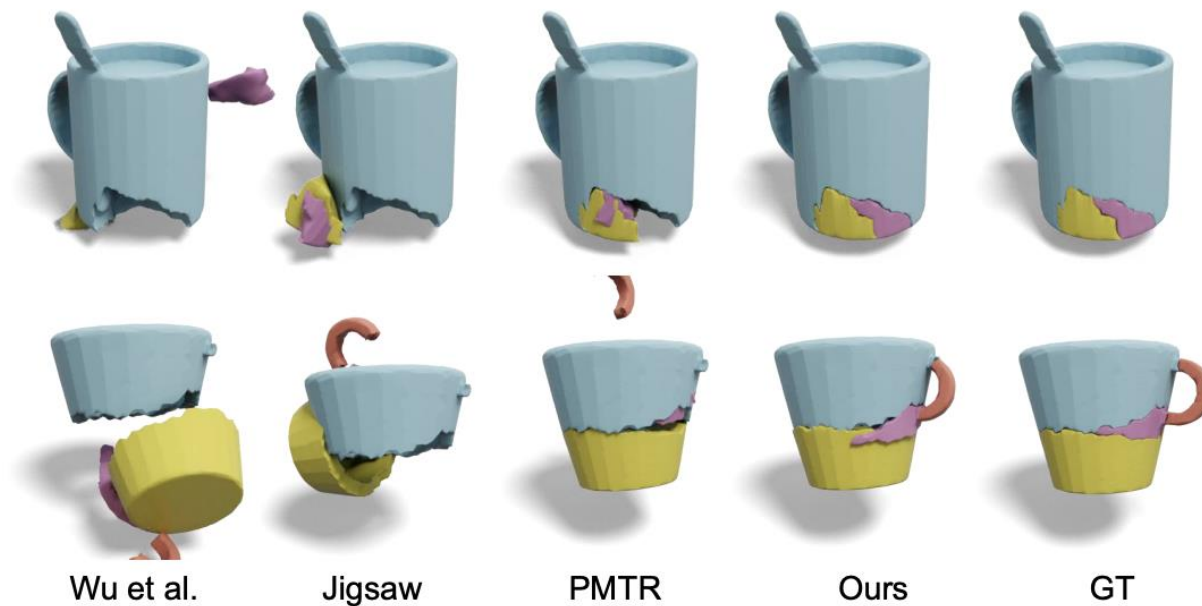
$$\mathcal{L} = \underbrace{\lambda_s \mathcal{L}_s + \lambda_o \mathcal{L}_o + \lambda_d \mathcal{L}_d}_{\text{Combinative matching objective}} + \mathcal{L}_p \leftarrow \text{Point matching loss}$$

Combinative matching objective

Qualitative Comparisons



Two-part assembly result



Multi-part assembly result

Comparison with SoTA

Method	CRD ↓ (10 ⁻²)	CD ↓ (10 ⁻³)	RMSE(R) ↓ (°)	RMSE(T) ↓ (10 ⁻²)
everyday				
NSM [5]	21.71	11.09	83.38	23.71
Wu et al. [48]	20.65	11.66	84.58	22.90
GeoTransformer [34]	0.61	0.51	22.81	7.28
Jigsaw [24]	5.48	1.34	38.73	2.73
PMTR [16]	<u>0.39</u>	<u>0.25</u>	<u>17.14</u>	5.53
CMNet (Ours)	0.28	0.17	12.88	<u>3.78</u>
artifact				
NSM [5]	19.44	6.33	83.22	21.41
Wu et al. [48]	19.17	7.97	85.04	20.90
GeoTransformer [34]	0.89	0.70	33.23	10.30
Jigsaw [24]	6.36	1.45	39.71	3.02
PMTR [16]	<u>0.60</u>	<u>0.42</u>	<u>23.28</u>	7.27
CMNet (Ours)	0.49	0.34	18.77	<u>5.57</u>

Pairwise Shape Assembly

Method	CRD ↓ (10 ⁻²)	CD ↓ (10 ⁻³)	RMSE(R) ↓ (°)	RMSE(T) ↓ (10 ⁻²)	PA _{CRD} ↑ (%)	PA _{CD} ↑ (%)
everyday						
Global [17, 37]	27.79	15.30	55.42	15.31	36.42	37.90
LSTM [47]	27.69	15.23	54.78	15.24	36.74	38.97
DGL [12]	27.90	13.23	55.76	15.33	36.99	39.70
Wu et al. [48]	28.18	19.70	54.98	15.59	35.66	36.28
Jigsaw [24]	14.13	11.82	41.12	11.74	52.48	60.26
PMTR [16]	<u>6.51</u>	<u>5.56</u>	<u>31.57</u>	<u>9.95</u>	<u>66.95</u>	<u>70.56</u>
CMNet (Ours)	5.18	3.65	27.11	8.13	73.88	77.88
artifact						
Global [17, 37]	26.42	14.92	54.41	14.48	36.67	36.97
LSTM [47]	28.15	14.61	53.59	15.49	36.67	37.25
DGL [12]	27.48	13.91	54.66	15.10	36.66	37.40
Wu et al. [48]	26.02	15.81	54.35	14.27	36.63	37.02
Jigsaw [24]	16.10	9.53	42.01	17.47	56.93	65.58
PMTR [16]	<u>5.67</u>	<u>4.33</u>	<u>31.58</u>	<u>10.08</u>	<u>66.96</u>	<u>71.61</u>
CMNet (Ours)	4.56	3.04	29.21	8.99	71.02	76.32

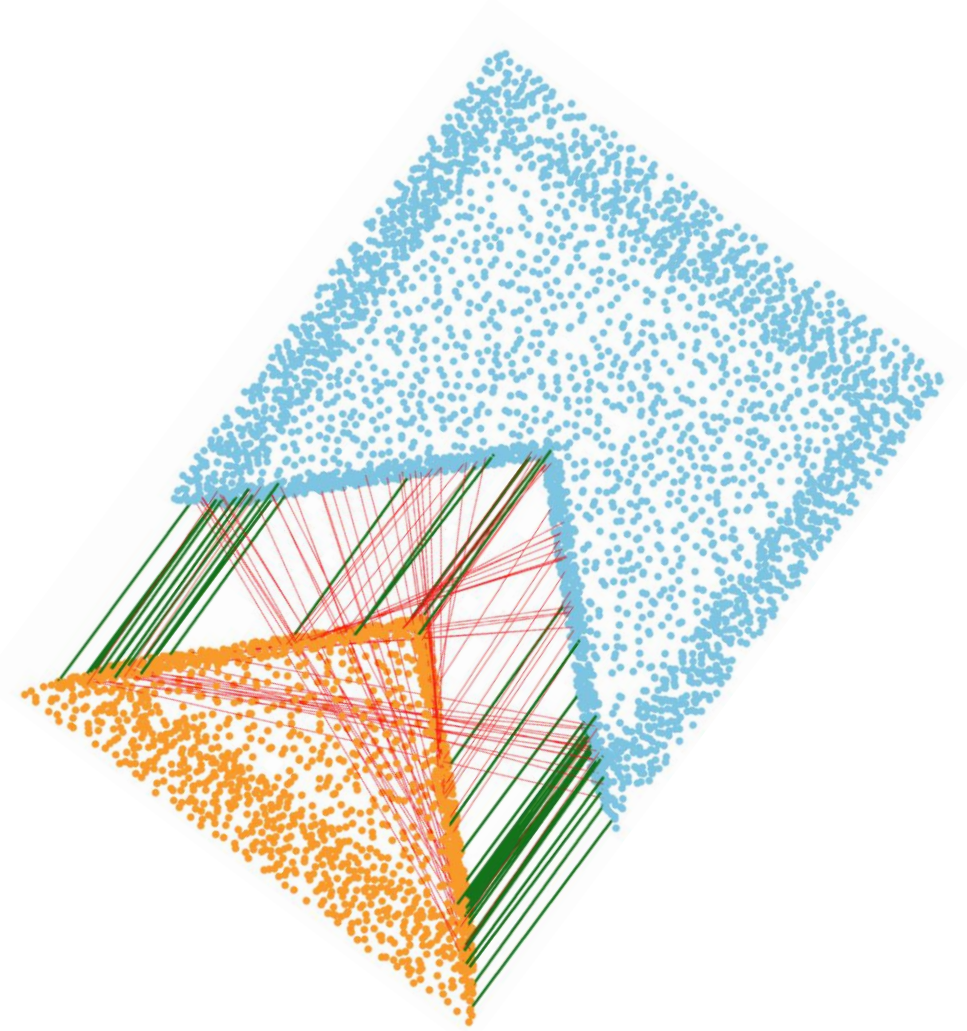
Multi-part Assembly (Volume-constrained ver.)

Method	RMSE(R) ↓ (°)	RMSE(T) ↓ (10 ⁻²)	PA _{CD} ↑ (%)	CD ↓ (10 ⁻³)
everyday				
Global [17, 37]	80.7	15.1	24.6	14.6
LSTM [47]	84.2	16.2	22.7	15.8
DGL [12]	79.4	15.0	31.0	14.3
Wu et al. [48]	79.3	16.9	8.41	28.5
DiffAssemble [36]	73.3	14.8	27.5	-
Jigsaw [24]	42.3	10.7	57.3	13.3
PuzzleFusion++ [46]	<u>38.1</u>	8.0	<u>71.0</u>	<u>6.0</u>
CMNet (Ours)	32.0	<u>9.6</u>	77.3	3.5
everyday → artifact				
Jigsaw [24]	52.4	22.2	45.6	14.3
PuzzleFusion++ [46]	52.1	13.9	49.6	14.5
CMNet (Ours)	46.0	<u>14.3</u>	52.6	9.8

Multi-part Assembly (Vanilla ver.)

Summary

- Assembly as a more realistic and challenging generative task is key to interacting AI with reality, e.g., manufacturing and design.
- Leveraging **geometric equivariance and invariance** in a careful design is crucial for addressing the tasks in a realistic environment.
- **Combinative matching**, which considers both shape and occupancy relations in matching, significantly improves the performance of geometric assembly and potentially extends to other relational tasks.



See you in Hawaii!

Combinative Matching for Geometric Shape Assembly

Nahyuk Lee^{1*} Juhong Min^{1,2*} Junhong Lee¹ Chunghyun Park¹ Minsu Cho^{1,3}

¹Pohang University of Science and Technology (POSTECH) ²Samsung Research America ³RLWRD

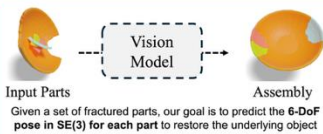
POSTECH
POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

SAMSUNG
RLWRD



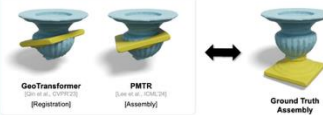
ICCV
OCT 19-23, 2025
HONOLULU
HAWAII

Geometric Shape Assembly?



Motivation and Overview

Traditional shape matching methods rely on **equative matching strategy**, which assumes that mating parts resemble each other at their surfaces.



Often falls short in geometric shape assembly, where parts are **not merely visually similar**, but are **structurally complementary**.

Then, what do we miss?

Complementary properties: Surfaces must not only look alike but also **fit by occupying opposite volumes**.

Main Contributions

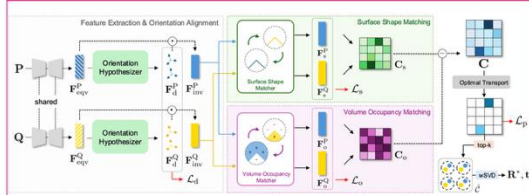
- Combinative Matching**: a new shape-matching methodology to combine interlocking parts for shape assembly.
- Combinative Matching Network**: a framework utilizing combinative matching, achieving SoTA on Breaking Bad.

Learning to interlock parts: Combinative Matching

We propose **Combinative Matching**, explicitly modeling both **identical surface shapes** and **opposite volume occupancy**, enabling robust combination of interlocking parts.

$$\mathcal{L}_s = \frac{1}{|\mathcal{C}|} \sum_{(i,j) \in \mathcal{C}} \|(\mathbf{F}_i^s, \mathbf{R}^s - (\mathbf{F}_j^s), \mathbf{R}^s)\|_F$$
$$\mathcal{L}_v = \sum_{i \in \mathcal{I}} \left[\log \left(1 + \sum_{j \in \mathcal{J}(i)} e^{\alpha_{ij}(\Delta_{ij}^v - \Delta_{ij}^s)} \right) + \sum_{j \in \mathcal{J}(i)} e^{\beta_{ij}(\Delta_{ij}^v - \Delta_{ij}^s)} \right]$$
$$\alpha_{ij} = \gamma \|\Delta_{ij}^v - \Delta_{ij}^s\|_2, \beta_{ij} = \gamma \|\Delta_{ij}^v - \Delta_{ij}^s\|_2, \Delta_{ij}^v = \|\mathbf{F}_i^v - \mathbf{F}_j^v\|_2$$
$$\mathcal{L}_o = \sum_{i \in \mathcal{I}} \left[\log \left(1 + \sum_{j \in \mathcal{J}(i)} e^{\alpha_{ij}(\Delta_{ij}^o - \Delta_{ij}^s)} \right) + \sum_{j \in \mathcal{J}(i)} e^{\beta_{ij}(\Delta_{ij}^o - \Delta_{ij}^s)} \right]$$
$$\alpha_{ij} = \gamma \|\Delta_{ij}^o - \Delta_{ij}^s\|_2, \beta_{ij} = \gamma \|\Delta_{ij}^o - \Delta_{ij}^s\|_2, \Delta_{ij}^o = \cos(\mathbf{F}_i^o, \mathbf{F}_j^o)$$

Combinative Matching Network (CMNet)



Step 1. Orientation Alignment

- Encode rotation-equivariant features (w/ VN-Layers)
- Predict per-point orientations
- Derive rotation-invariant embeddings via dot-product

Step 2. Combinative Matching

- Dual matching branches: shape & occupancy matching
- Learn to align both identical surface shape and opposite volume occupancy

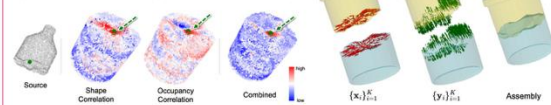
Step 3. Transformation Estimation

- Combine shape / occupancy correlations → Optimal Transport
- Estimate SE(3) pose via wSVD
- Further utilized as pairwise matcher for multi-part assembly

Experiments and Analysis

Learned Descriptor Analysis

(left: correlation distribution, right: orientation visualization)



By combining shape and occupancy information, the local ambiguity and match confidence uncertainty are resolved. Learned orientations capture complementarity between parts **without any explicit supervision**.

Ablation Studies on Combinative Matching

Occupancy Affinity	Orientation Loss (\mathcal{L}_o)	CRD \downarrow (10^{-3})	CD \downarrow (10^{-3})	RMSE(R) \downarrow ($^\circ$)	RMSE(T) \downarrow (10^{-3})	Equivariant Embedding	Shape Matching	Occupancy Matching	CRD \downarrow (10^{-3})	CD \downarrow (10^{-3})	RMSE(R) \downarrow ($^\circ$)	RMSE(T) \downarrow (10^{-3})
L2 dist	✓	0.42	0.31	14.88	4.31	✓	✓	✓	0.34	0.53	38.74	11.88
cosine	✓	0.31	0.23	13.77	4.44	✓	✓	✓	0.28	0.28	13.77	3.86
L2 dist	✓	0.38	0.30	13.29	3.81	✓	✓	✓	0.35	0.25	14.01	4.24
cosine	✓	0.38	0.17	12.88	3.78	✓	✓	✓	0.28	0.17	12.88	3.78

(a) Ablation study on combinative matching.

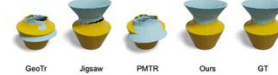
(b) Ablation study on model components.

Learning **complementary volume occupancy** under **orientation alignment** is crucial for accurate geometric shape assembly. Incorporating **joint shape and occupancy matching** under **SO(3)-equivariance** significantly enhances the assembly accuracy and robustness.

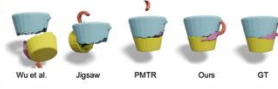
Improving SOTA for Geometric Shape Assembly

(top: pairwise assembly, bottom: multi-part assembly)

Method	CRD \downarrow (10^{-3})	CD \downarrow (10^{-3})	RMSE(R) \downarrow ($^\circ$)	RMSE(T) \downarrow (10^{-3})
everyday				
Wu et al. [44]	20.65	11.66	84.58	22.90
GeoTransformer [32]	0.61	0.51	22.81	7.28
Jigsaw [72]	5.48	1.34	38.73	2.73
PMTR [14]	0.39	0.25	17.14	5.53
CMNet (Ours)	0.28	0.17	12.88	3.78



Method	CRD \downarrow (10^{-3})	CD \downarrow (10^{-3})	RMSE(R) \downarrow ($^\circ$)	RMSE(T) \downarrow (10^{-3})	PA _{max} \uparrow (%)	PA ₅₀ \uparrow (%)
everyday						
Wu et al. [44]	28.18	19.70	54.18	15.59	35.66	36.78
Jigsaw [72]	14.13	11.82	41.12	11.74	52.48	60.26
PMTR [14]	6.51	5.56	31.57	16.65	70.36	70.36
CMNet (Ours)	5.18	3.68	27.11	8.13	73.88	77.88



OUR POSTER PRESENTATION

Project page



Tue 21 Oct, 2025
3:15 p.m. HST — 5:15 p.m. HST
@ Exhibit Hall I #885