3D Geometric Shape Assembly via Efficient Point Cloud Matching

POSTECH
POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

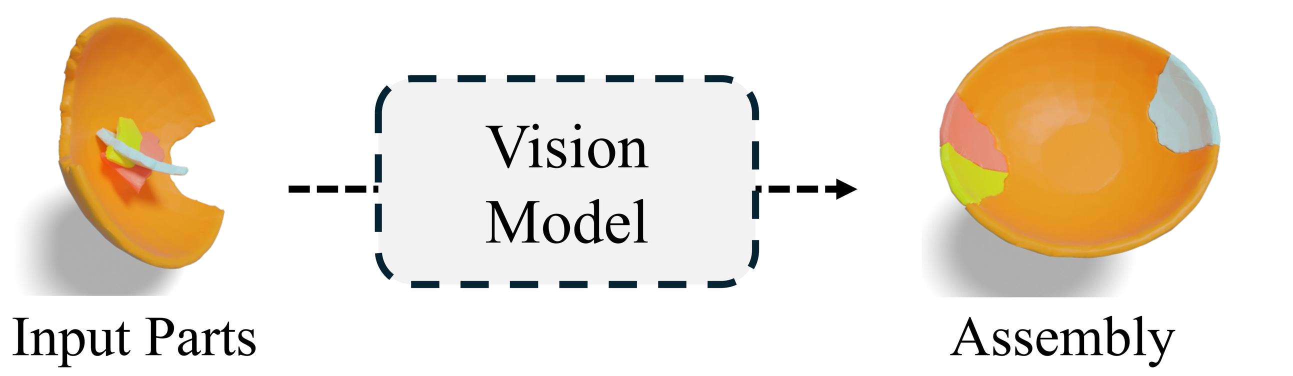
Nahyuk Lee* [1] Juhong Min* [1] Junha Lee [1] Seungwook Kim [1] Kanghee Lee [2] Jaesik Park [2] Minsu Cho [1]

[1] Pohang University of Science and Technology (POSTECH) [2] Seoul National University

Project page

ICML International Conference On Machine Learning

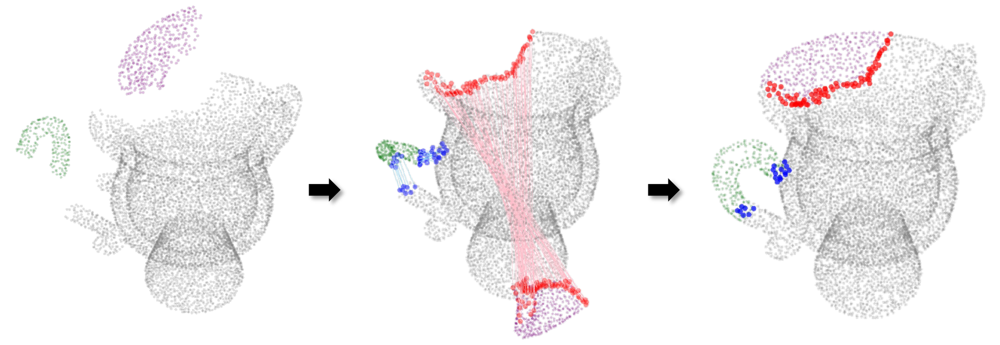
Geometric Shape Assembly?



Given a set of fractured parts, our goal is to predict the **6-DoF pose in SE(3) for each part** to restore the underlying object.

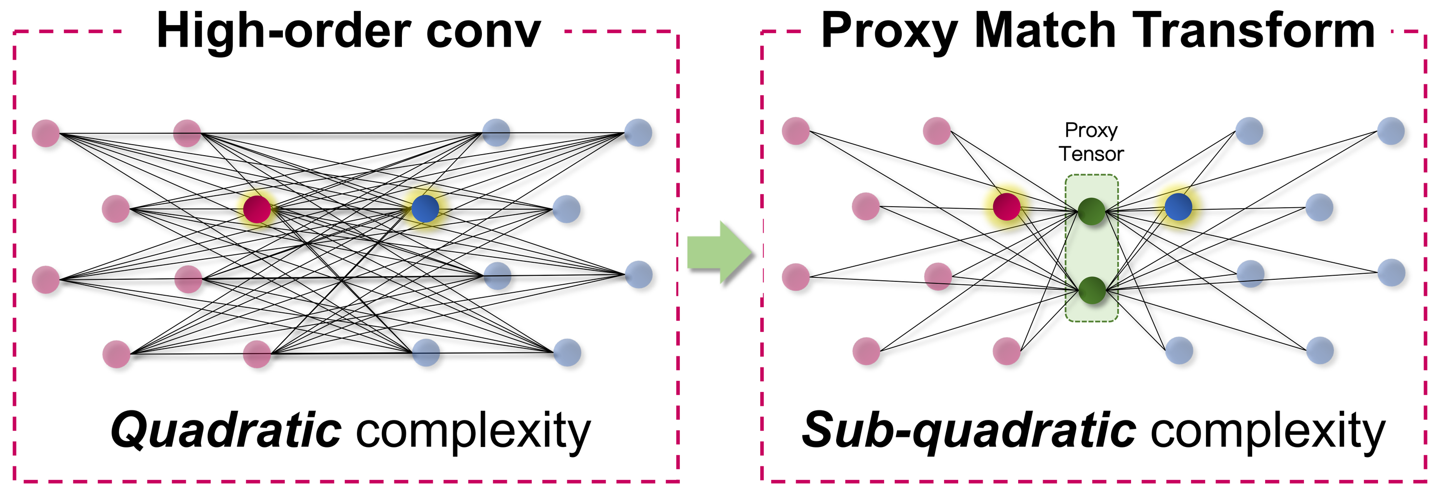
Motivation and Overview

★ 1 Assembly requires to find ***reliable*** correspondences.



➡ Use ***high-order conv***: a trending approach for correlation analysis! [NCNet, HSNet, etc.]

★ 2 However, high-order conv has ***quadratic*** complexity.



➡ Propose ***PMT: effective approximation of existing high-order conv*** with only sub-quadratic complexity!

Main Contributions

- **Proxy Match Transform (PMT):** a novel high-order feature transform layer with only sub-quadratic complexity.
- **Proxy Match TransformeR (PMTR):** a coarse-to-fine framework with a series of PMTs, achieving SoTA on Breaking Bad.

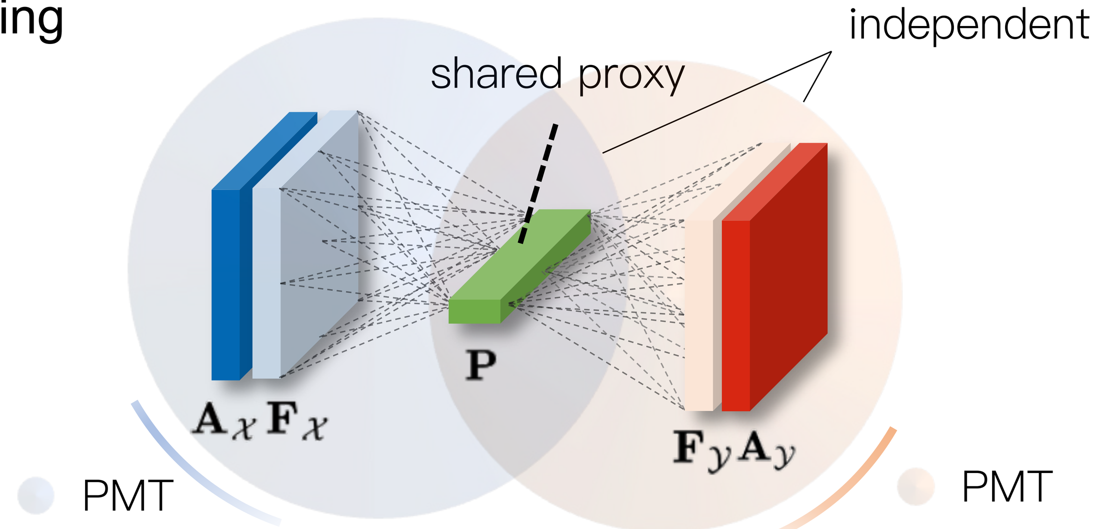
Proxy Match Transform: Efficient High-order Feature Transform

We propose ***low-complexity high-order*** feature transform layer, Proxy Match Transform (PMT), for analyzing **6D correlation** between two 3D point clouds.

$$\mathrm{PMT}(\mathbf{F}_{\mathcal{X}}) := \sum_{h\in[N_h]} \mathbf{A}_{\mathcal{X}}^{(h)}\mathbf{F}_{\mathcal{X}}\mathbf{P}^{(h)\top}w_{\mathcal{X}}^{(h)}$$

shared

$$\mathrm{PMT}(\mathbf{F}_{\mathcal{Y}}) := \sum_{h\in[N_h]} \mathbf{A}_{\mathcal{Y}}^{(h)}\mathbf{F}_{\mathcal{Y}}\mathbf{P}^{(h)\top}w_{\mathcal{Y}}^{(h)}$$



Information exchange by sharing a proxy tensor!

Dot product of PMTs can ***approximate high-order conv.*** with two constraints (complete proof in paper):

$$(\mathrm{PMT}(\mathbf{F}_{\mathcal{X}}) \cdot \mathrm{PMT}(\mathbf{F}_{\mathcal{Y}})^\top)_{(\mathbf{x},\mathbf{y})} \approx \mathrm{Conv}(\mathbf{F}_{\mathcal{X}}, \mathbf{F}_{\mathcal{Y}})_{(\mathbf{x},\mathbf{y})}$$

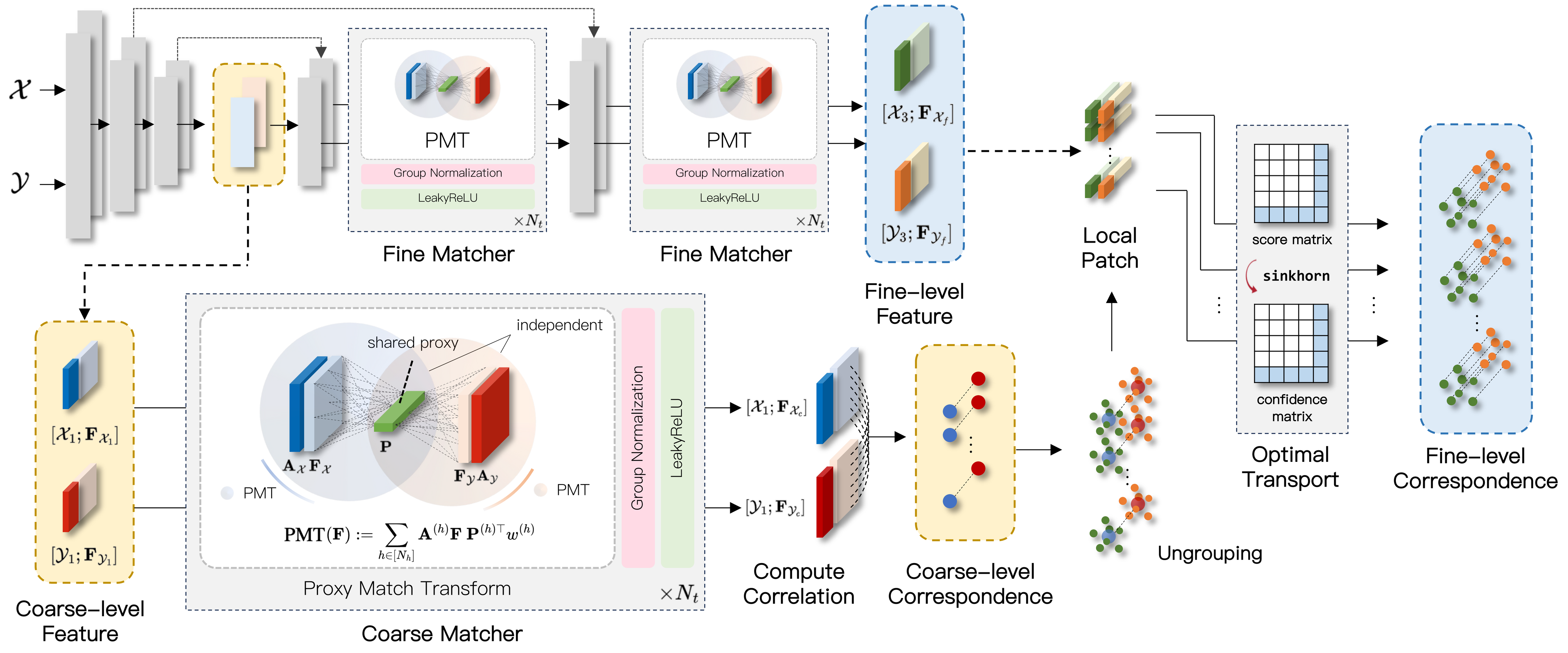
➡

$$\mathcal{L}_{\text{orth}} = \sum_{(i,j)\in[N_h]^2} \delta(i,j)(\mathbf{P}^{(i)\top}\mathbf{P}^{(j)} - \mathbf{I}_{D_{\text{emb}}})$$

$$\mathcal{L}_{\text{zero}} = \sum_{(i,j)\in[N_h]^2} (1-\delta(i,j))\mathbf{P}^{(i)\top}\mathbf{P}^{(j)}$$

☺ Effectively express high-order convolution with **sub-quadratic** complexity!

Proxy Match TransformeR (PMTR) for Geometric Assembly



Step 1. Coarse-level Matching
PMT processes the coarse feature pair, evaluating ***potential*** local correspondences between the point features.

Step 2. Fine-level Matching
Two PMT matchers refine the high-resolution feature pairs, facilitating the capture of more ***detailed*** fine-level geometric feature correlations.

Step 3. Transformation Estimation
Given fine-level correspondences, we utilize Optimal Transport to estimate the relative transformation between a pair of input parts.

Experiments and Analysis

Improving SOTA for Geometric Assembly

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

We evaluate our PMTR on the tasks of (1) pairwise assembly and (2) multi-part assembly. Our results outperform state-of-the-art baselines while remaining highly efficient.

Method	CRD ↓ (10^{-2})	CD ↓ (10^{-3})	RMSE (R) ↓ (°)	RMSE (T) ↓ (10^{-2})
	everyday			
GeoTransformer (2022)	0.61	0.51	22.81	7.28
Jigsaw (2023)	5.48	1.34	38.73	**2.73**
PMTR (Ours)	**0.39**	**0.25**	**17.14**	5.53

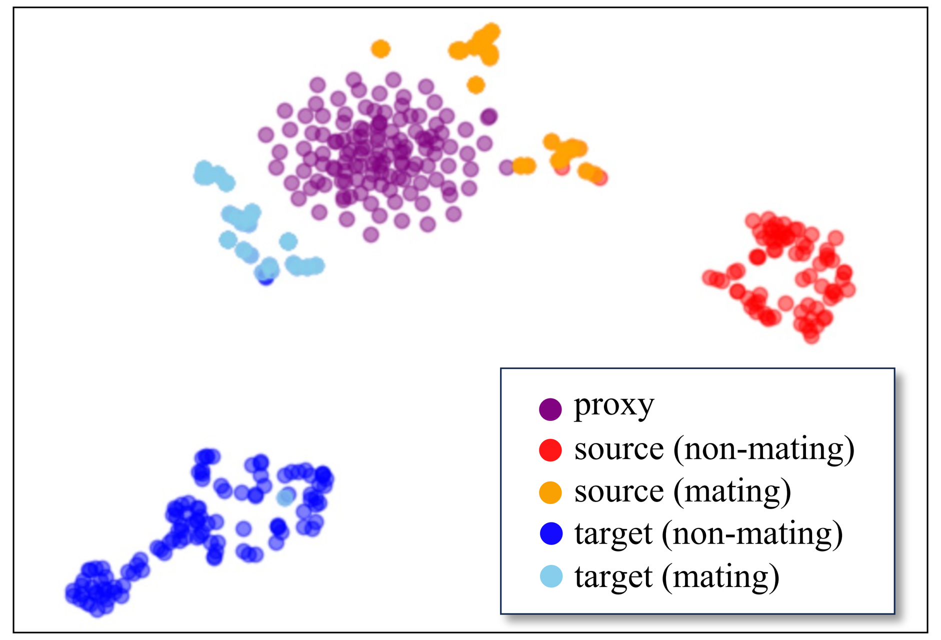
Method	CRD ↓ (10^{-2})	CD ↓ (10^{-3})	RMSE (R) ↓ (°)	RMSE (T) ↓ (10^{-2})	PA_{CRD} ↑ (%)	PA_{CD} ↑ (%)
	everyday					
Wu et al. (2023b)	28.18	19.70	54.98	15.59	35.66	36.28
Jigsaw (2023)	14.13	11.82	41.12	11.74	52.48	60.26
PMTR (Ours)	**6.51**	**5.56**	**31.57**	**9.95**	**66.95**	**70.56**

Qualitative Comparison

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

t-SNE Visualization of Proxy Tensor

The visualization indicates that point features on each mating surface are closer to the proxy, implying **strong correlations** in the feature space.



Ablation Analysis

proxy	shared proxy	CRD ↓ (10^{-2})	CD ↓ (10^{-3})	RMSE (R) ↓ (°)	RMSE (T) ↓ (10^{-2})
✗	✗	0.53	0.47	21.04	6.93
✓	✗	0.44	0.31	18.66	5.97
✓	✓	**0.39**	**0.25**	**17.14**	**5.53**

By **sharing a proxy tensor** in each PMT, two ***independent*** feature transforms exchange relevant information for effective matching.

Coarse-level Matcher	Fine-level Matcher	CRD ↓ (10^{-2})	CD ↓ (10^{-3})	RMSE (R) ↓ (°)	RMSE (T) ↓ (10^{-2})
PMT	None	0.53	0.43	20.70	6.63
	Linear	0.47	0.37	17.55	5.68
	MLP	0.49	0.38	17.35	5.69
	HDC	Out of memory error			
	GeoTr	Out of memory error			
	PMT	**0.39**	**0.25**	**17.14**	**5.53**

PMT layer at fine-level yields the best assembly accuracy while incurring **lower computational complexity** compared to other baselines.

Efficiency of Proxy Match Transform (PMT)

Method	Coarse-level Matcher	Fine-level Matcher	FLOPS ↓ (G)	# Param. ↓ (K)	Mem. train ↓ (GB)	Mem. test ↓ (GB)	Train time ↓ (ms)	Inference time ↓ (ms)
GeoTransformer (2022)	GeoTr	None	9.67	926.85	6.96	3.10	8.93	8.04
PMTR (Coarse-only)	PMT	None	**0.45**	**273.85**	**2.12**	**0.28**	**4.06**	**3.23**
PMTR (Coarse + Fine)	PMT	PMT	0.78	296.15	3.78	0.88	5.35	3.75

PMT is approximately **x21.5** more efficient in FLOPS with **x3.4** less number of parameters and **x3.28 / x11.07** less memory required for training / inference phases compared to the previous SoTA matching module, *e.g.*, GeoTr.