Leveraging SE(3) Equivariance for Learning 3D Geometric Shape Assembly

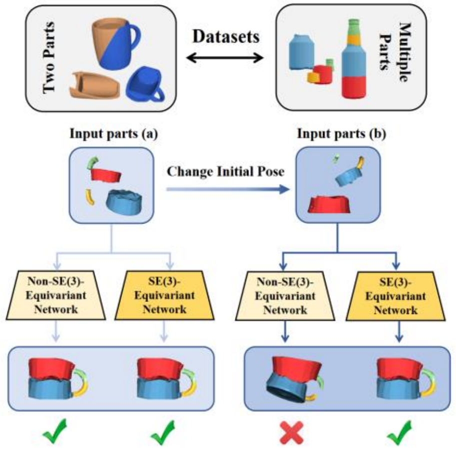
Ruihai Wu*, Chenrui Tie*, Yushi Du, Yan Shen, Hao Dong

PEKING UNIVERSITY

CVPR SEATTLE, WA JUNE 17-21, 2024

Geometric Shape Assembly

Geometric Shape Assembly aims to assemble different fractured parts into a whole shape. We propose to leverage **SE(3) Equivariance** for learning Geometric Shape Assembly, which disentangles poses and shapes of fractured parts, and performs better than networks without SE(3)-equivariant representations.



SE(3) equivariance & invariance

SE(3), which represents the Special Euclidean Group in three dimensions, characterizes the rigid body motion in 3D space, including 3D translation and rotation.

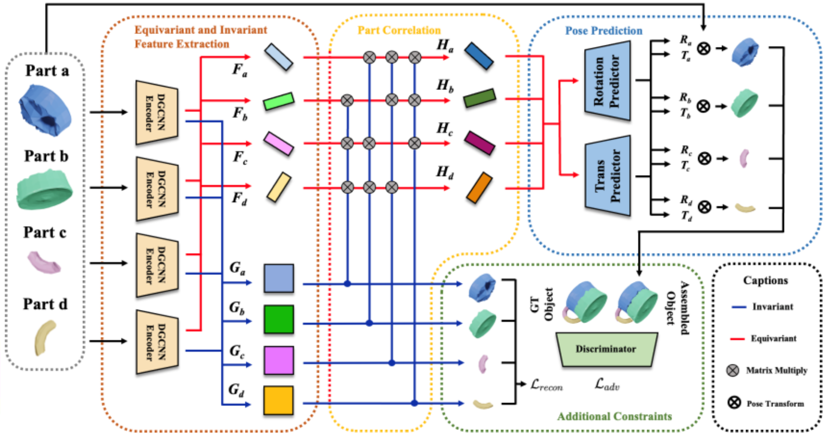
Given an input point cloud $\mathcal{P} \in \mathbb{R}^{n\times 3}$, for any rotation matrix $R \in \mathbb{R}^{3\times 3}$ and translation vector $T \in \mathbb{R}^{3}$. A SE(3) equivariant encoder $\mathcal{E}_{equiv}$ means

$$\mathcal{E}_{equiv}(\mathcal{P}R + T) = \mathcal{E}_{equiv}(\mathcal{P})R + T$$

And a SE(3) invariant encoder $\mathcal{E}_{inv}$ means

$$\mathcal{E}_{inv}(\mathcal{P}R + T) = \mathcal{E}_{inv}(\mathcal{P})$$

Method



Taking as input the point cloud of each part i, our framework first outputs the equivariant representation F_i and invariant representation G_i, computes the correlation between part i and each part j using the matrix multiplication of F_i and G_i, and thus gets each part's equivariant representation H_i with part correlations. H_i is equivariant with input part i, and invariant with input part j

The rotation decoder and the translation decoder respectively take H and decode the rotation and translation of each part. Additional constraints such as adversarial training and canonical point cloud reconstruction using G further improves the performance of our method.

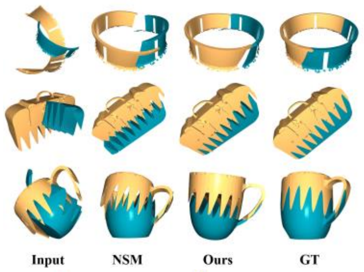
Project Page

Ruihai Wu

Chenrui Tie

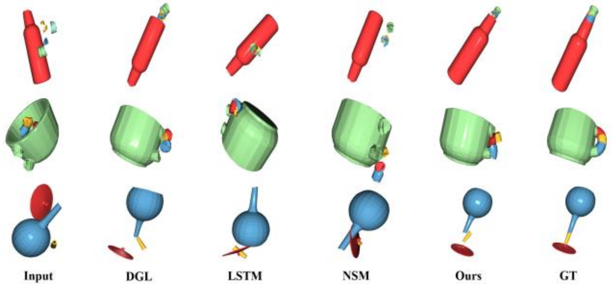
Hao Dong

Experiments



We test the performance of out method on two datasets

- Geometric Shape Mating dataset for two-part geometric shape assembly
- Breaking Bad dataset for multi-part geometric shape assembly



We compare our method with SOTA part assembly methods.

Conclusion

We propose to leverage SE(3)-equivariant representations that disentangle shapes and poses to facilitate the task. Our method leverages SE(3) equivariance in part representations considering part correlations, by learning both SE(3)-equivariant and -invariant part representations and aggregating them into SE(3)-equivariant representations. To the best of our knowledge, we are the first to explore leveraging SE(3) equivariance on multiple objects in related fields. Experiments demonstrate the effectiveness of our method.

Acknowledgements

National Natural Science Foundation of China (No. 62136001).