

Generating Part-Aware Editable 3D Shapes without 3D Supervision

Konstantinos Tertikas^{1,3} Despoina Paschalidou² Boxiao Pan² JJ Park² Mika Uy² Ioannis Emiris^{3,1} Yannis Avrithis⁴ Leonidas Guibas²

> ¹National and Kapodistrian University of Athens ²Stanford University ³Athena RC Greece ⁴Institute of Advanced Research in Artificial Intelligence



Motivation

enable local control but require explicit 3D supervision and only produce textureless objects.

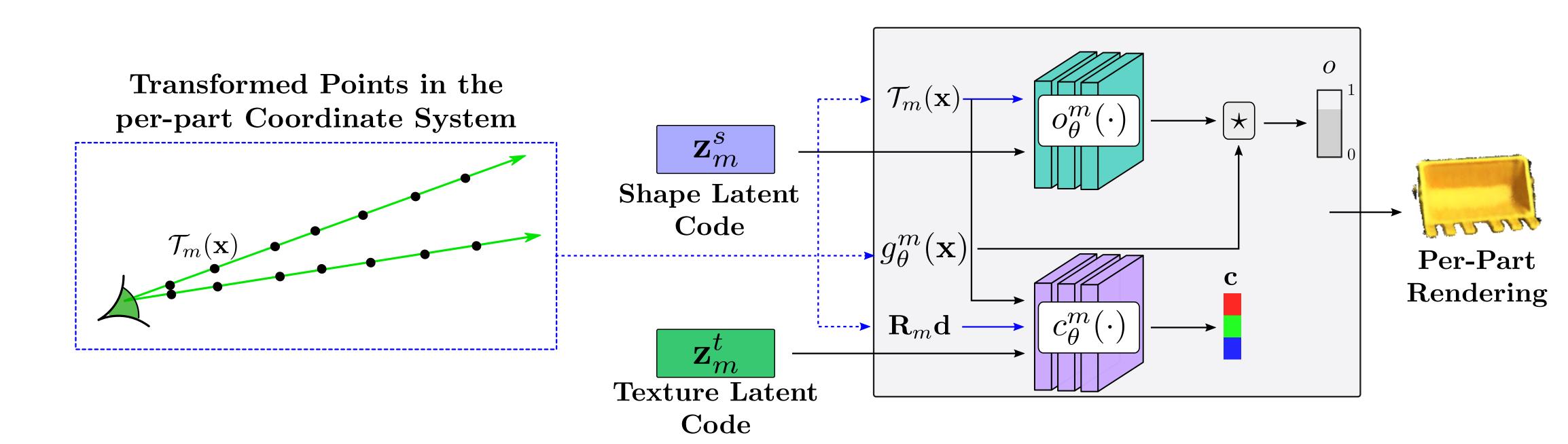
	Representation	Supervision	Parts	Shape Editing	Texture Editing	Mixing
GET3D	Mesh	2D	X	X	X	X
GRAF			X	×	X	X
Pi-GAN	Neural Field	2D	X	X	X	X
EG3D			X	X	X	X
DualSDF	Implicit	3D	✓		X	X
SPAGHETTI					X	
PartNeRF	Neural Field	2D				

Contributions:

- First part-aware generative model for editable 3D shapes.
- Does not require explicit 3D supervision.
- Enables new editing capabilities.

Part Parametrization

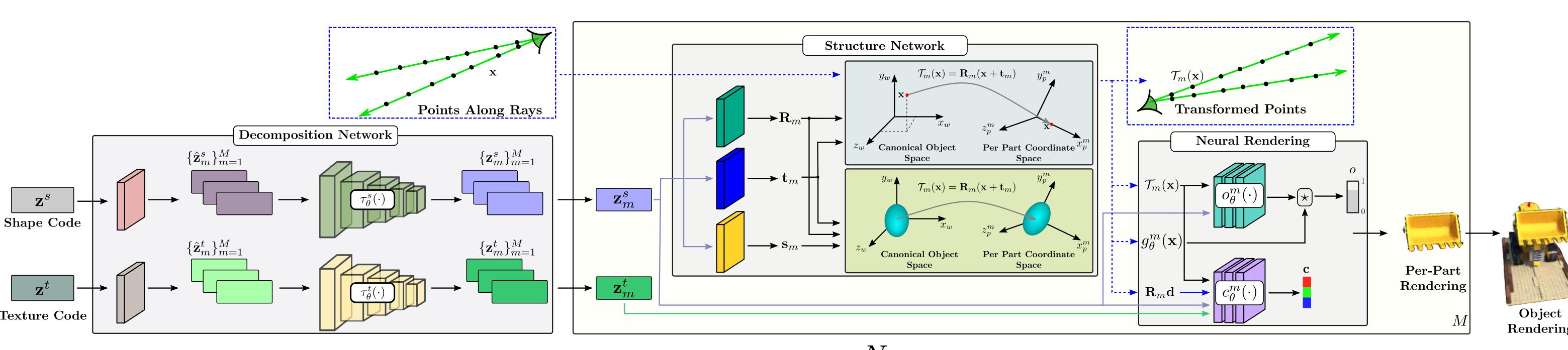
Each part is parametrized as a locally defined NeRF, implemented using a **color** $c_{\theta}^{m}(\cdot)$ and an **occupancy** network $o_{\theta}^{m}(\cdot)$ that map a 3D point $\mathbf{x} \in \mathbb{R}^3$, a viewing direction $\mathbf{d} \in \mathbb{S}^2$, a shape \mathbf{z}_m^s and a texture code \mathbf{z}_m^t into a **color** $\mathbf{c} \in \mathbb{R}^3$ and an **occupancy** value $o \in [0, 1]$.



To be able to modify the shape, size and appearance of specific parts of the object independently, we transform the geometric inputs of each NeRF in its local coordinate system $T_m(\mathbf{x})$.

Our Method

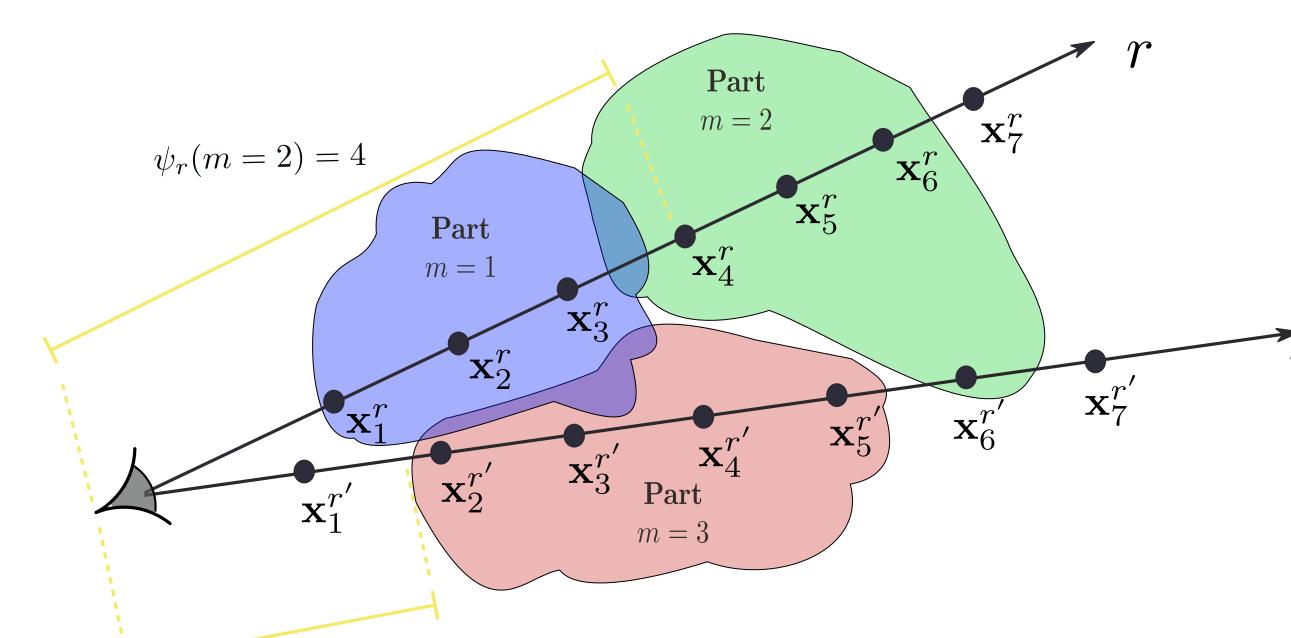
NeRF-based generative models generate high quality 3D objects Given a collection of posed images accompanied by object masks, we formulate our, with texture but lack local control. Part-based generative models—generative model as an auto-decoder that represents 3D objects as M NeRFs.



The **per-part rendering equation** is
$$\hat{C}_m(r) = \sum_{i=1}^N h_{\theta}^m(\mathbf{x}_i^r) \prod_{j < i} (1 - h_{\theta}^m(\mathbf{x}_i^r)) c_{\theta}^m(\mathbf{x}_i^r, \mathbf{d}^r),$$

where $h_{\theta}^{m}(\mathbf{x}_{i}^{r})$ is the occupancy and $c_{\theta}^{m}(\mathbf{x}_{i}^{r}, \mathbf{d}^{r})$ its color at point \mathbf{x}_{i}^{r} .

To ensure distinct, manipulable parts, we introduce a hard assignment between rays and parts, namely the color of each ray is determined by a single NeRF.



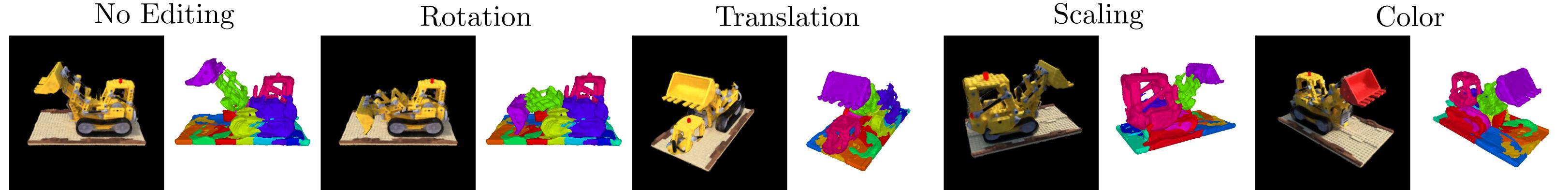
The set of rays \mathcal{R}_m associated with the m-th part is the set of rays that first intersect with it:

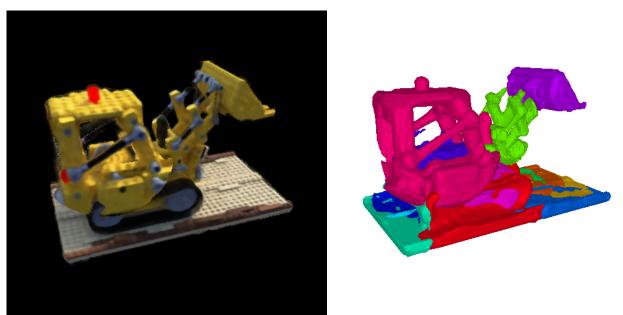
$$\mathcal{R}_m = \left\{ r \in \mathcal{R} : m = \underset{k \in \{0...M\}}{\operatorname{argmin}} \psi_r(k) \right\}.$$

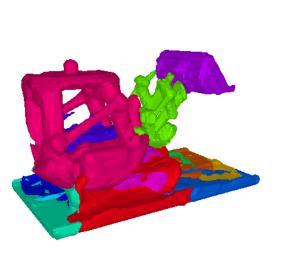
The **per-object rendering equation** using
$$M$$
 NeRFs is $\hat{C}(r) = \sum_{m=1}^{M} \mathbf{1}_{r \in \mathcal{R}_m} \hat{C}_m(r)$.

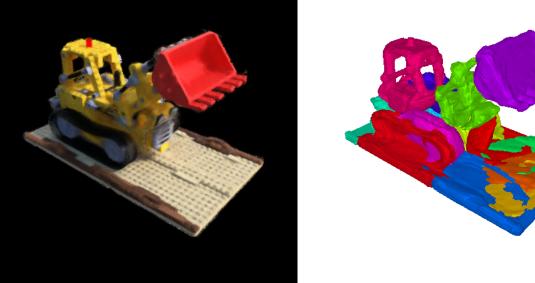
Scene-Specific Editing Results

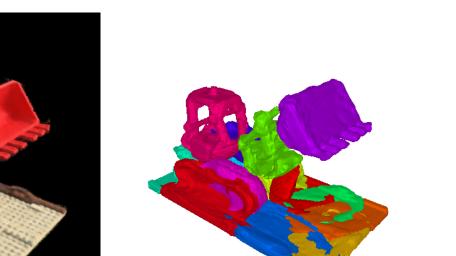
During all edits only a specific part of the object changes, while the rest do not.





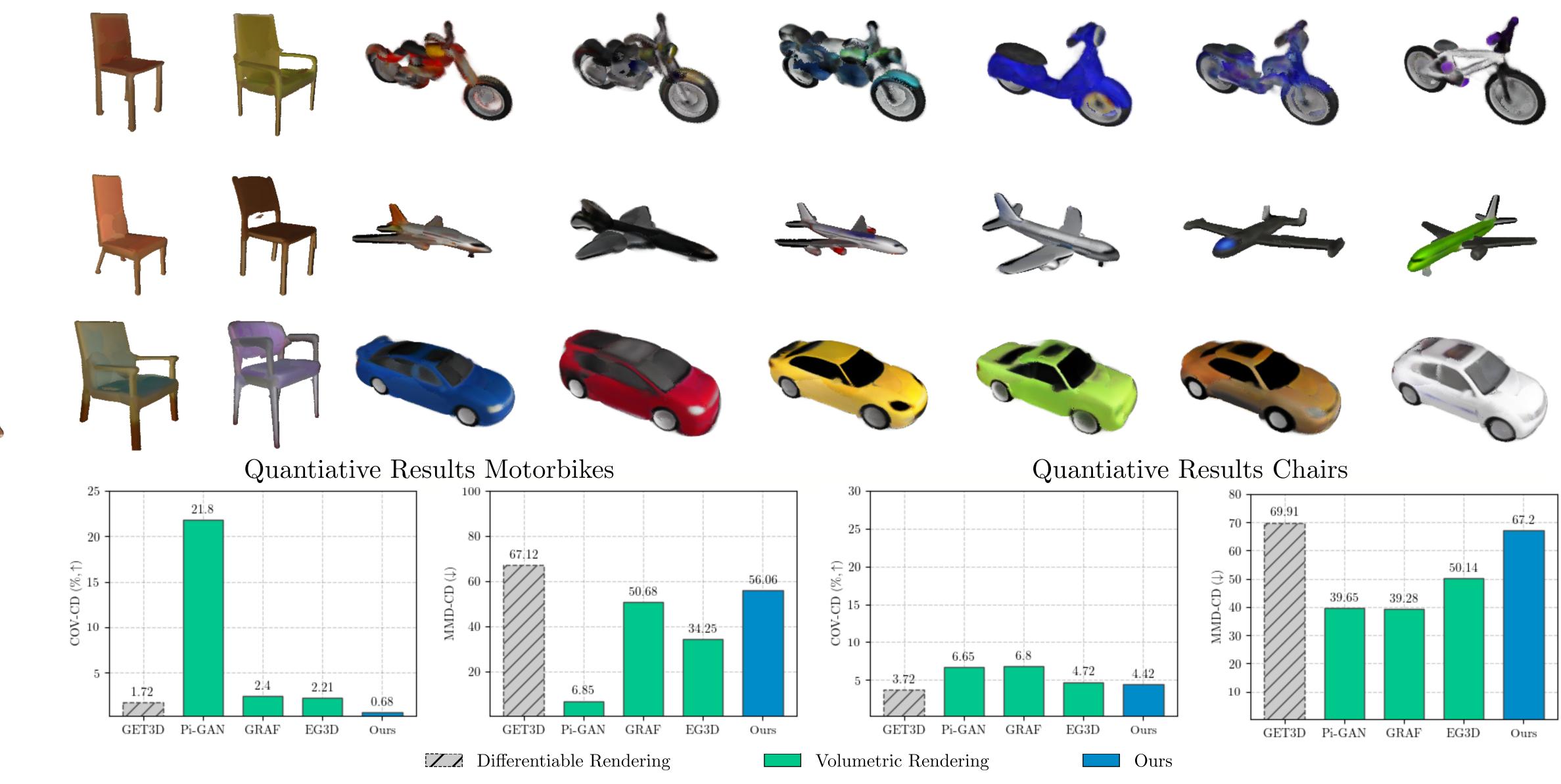




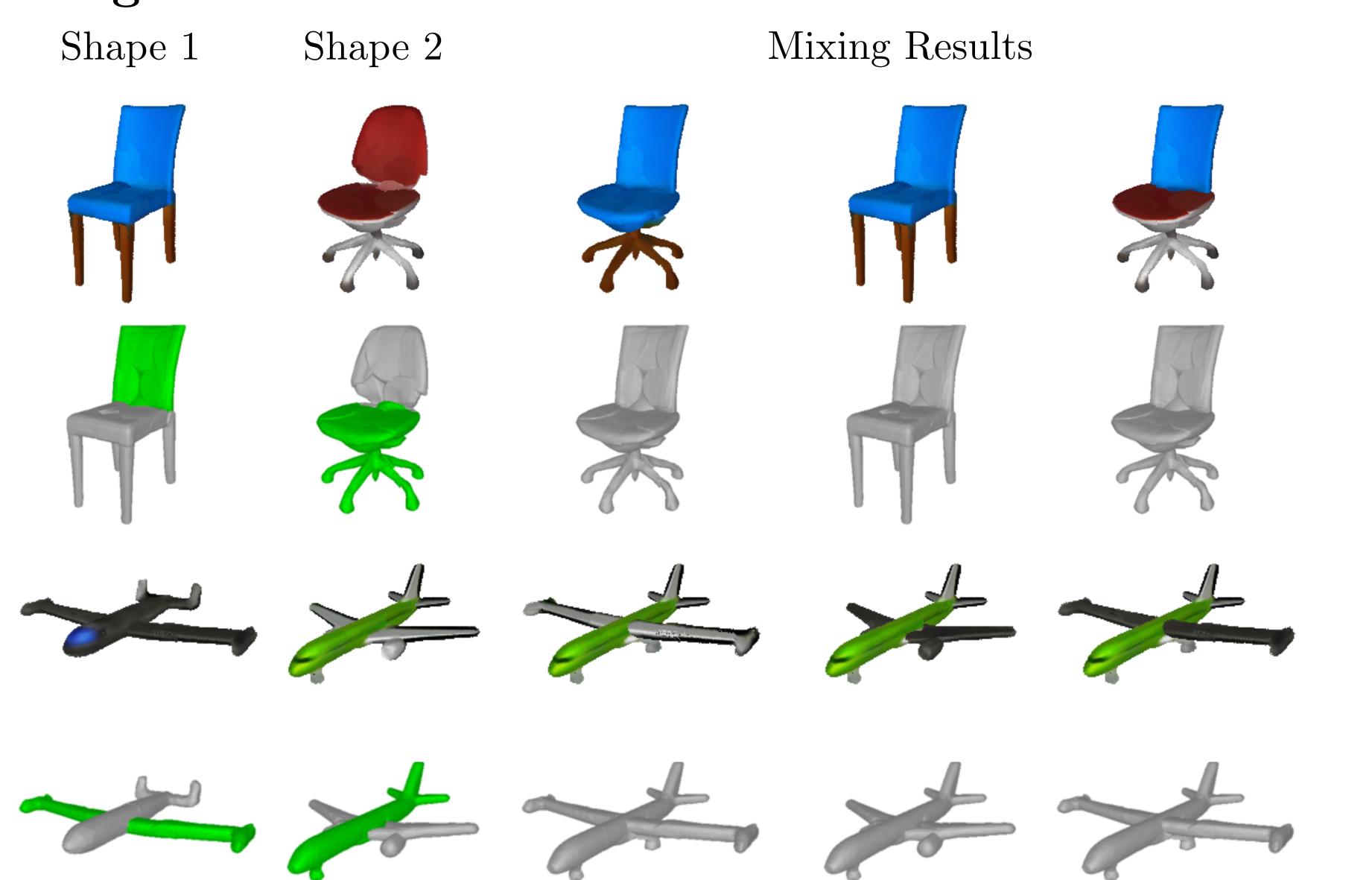


Shape Generation and Editing





Shape Mixing:



Shape Interpolations:







