Motivation / Challenges

Q: "Can we learn high-resolution geometry with an arbitrary topology without 3D supervision?"

Need Effective 3D Data Representation

- Voxel: +Topology, -Fidelity
- Point cloud: +Topology, -Fidelity
- Mesh: -Topology, +Fidelity
- Occupancy field: ++Fidelity

How to Differentiably Render Implicit Function Efficiently?

- Explicit shape representations
- Differentiable rendering
- Projection (Rasterization)
- Implicit representations
- Field probing (Ray tracing)

Method

Our Framework: Single-view 3D Object Reconstruction using Implicit Surface

Efficient Sampling-based 2D Supervision

Geometric Regularization on Implicit Surface

\[
\mathcal{L}_{geo} = \frac{1}{N_p} \sum_{j=1}^{N_p} \sum_{i=1}^{6} W(\phi(p_j)) \frac{\|\mathbf{n}(p_j) - \mathbf{n}(q^i_j)\|_F}{\sum_{i=1}^{6} W(\phi(q^i_j))}
\]

Ablation Study

Results (ShapeNet)

<table>
<thead>
<tr>
<th>Category</th>
<th>Airplane</th>
<th>Bench</th>
<th>Table</th>
<th>Car</th>
<th>Chair</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTN [4]</td>
<td>0.5554</td>
<td>0.4875</td>
<td>0.4938</td>
<td>0.7123</td>
<td>0.4494</td>
<td>0.5399</td>
</tr>
<tr>
<td>NMR [1]</td>
<td>0.6172</td>
<td>0.4998</td>
<td>0.4829</td>
<td>0.7095</td>
<td>0.4990</td>
<td>0.5617</td>
</tr>
<tr>
<td>SoftRas [2]</td>
<td>0.6419</td>
<td>0.5080</td>
<td>0.4487</td>
<td>0.7697</td>
<td>0.5270</td>
<td>0.5789</td>
</tr>
<tr>
<td>Ours</td>
<td>0.6530</td>
<td>0.5360</td>
<td>0.5250</td>
<td>0.7820</td>
<td>0.5540</td>
<td>0.6100</td>
</tr>
</tbody>
</table>

Input images | Ground Truth | SoftRas (Mesh) | Ours (Implicit field) | Input images | Ground Truth | SoftRas (Mesh) | Ours (Implicit field) | Input images | Ground Truth | PTN (Voxel) | DFC (Point clouds) | N3MR (Mesh) | SoftRas (Mesh) | Ours (Implicit occupancy field)