1. Motivation

**Goal**: Learning 3D reconstruction from weak supervision of 2D masks

**Previous works**:

- **Full 3D supervision**\(^1\)\(^2\)\(^3\): 3D model is a very expensive label for practical use such as real image reconstruction.
- **2D mask supervision**\(^4\): Limited by visual hull. No concavity, symmetry, stability, etc.

**Proposed method**: Solving constrained optimization

\[
\begin{align*}
\min_x & \quad \text{ReprojectionError}(x) \\
\text{subject to} & \quad \text{Reconstruction } x \text{ to be a valid chair}
\end{align*}
\]

\(1\)

- **ReprojectionError** resembles **2D mask supervision** \(^4\) and Fig 1 (c)
- **The constraint** resembles Fig 1 (d)
- **Together learns correct 3D reconstruction**

\(2\)

2. Adversarial Constraint

1. Equation (1) can be re-written as

\[
\min_x \text{ReprojectionError}(x) - \frac{1}{t} \log g(x)
\]

using log barrier method where \(g(x) = 1\) iff reconstruction \(x\) is realistic and 0 otherwise

2. **Ideal discriminator** of GAN \(g^*(x)\), which outputs \(g^*(x) = 1\) iff reconstruction \(x\) is realistic, is analogous to the penalty function \(g(x)\)

3. Therefore, we can train \(g(x)\) as discriminator

\[
\min_g \mathbb{E}_{x \sim p} \log g(x) + \mathbb{E}_{\hat{x} \sim q} \log(1 - g(\hat{x}))
\]

\(3\)

3. Raytrace Pooling

- Bridge the gap between the target 3D reconstruction and **2D mask supervision** \(^4\) and Fig 1 (c)
- Takes reconstruction \(x\) and camera parameter of the ground-truth mask as input
- Renders mask of \(x\) through ray-voxel hit test
- Does not suffer from sampling artifacts as compared to \(^4\)

4. Experiments

1. **Ablation study on ShapeNet**: Our proposed method reconstructed a reasonable 3D shape from weak 2D supervision including concavity (red box in Fig 4). It is also worth noting that the adversarial constraint gives a noticeable performance boost especially on weak single-view supervision as shown in Figure 5

2. **Real image reconstruction**

3. **Representation analysis**

References


Acknowledgement

We acknowledge the support of Nvidia and Toyota (1186781-31-UDARO) to make this work possible.