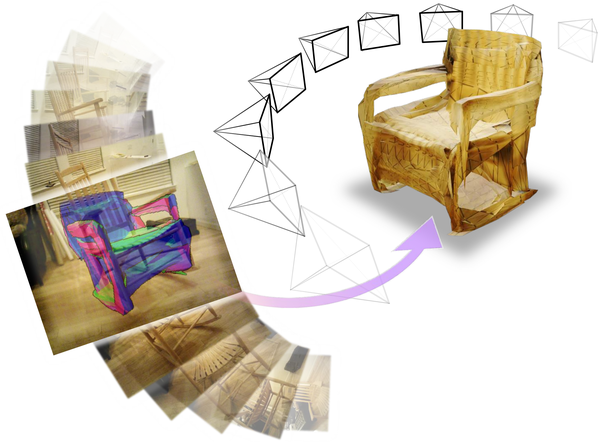
Photometric Mesh Optimization for Video-Aligned 3D Object Reconstruction

Chen-Hsuan Lin[12] Oliver Wang[2] Bryan C. Russell[2] Eli Shechtman[2] Vladimir G. Kim[2] Matthew Fisher[2] Simon Lucey[1]
[1]Carnegie Mellon University [2]Adobe Research

Code available!
https://github.com/chenhsuanlin/photometric-mesh-optim

Overview

We seek to recover **3D object meshes** aligned with the given object-centric **RGB video sequence**.



3D reconstruction, classic and modern

Multi-view geometric methods (SfM, MVS)

- ✅ Applicable to **generic** video sequences
- ❌ Recovers **point clouds**, no spatial structure
- ❌ **Lacks semantics** (which points are from the object?)

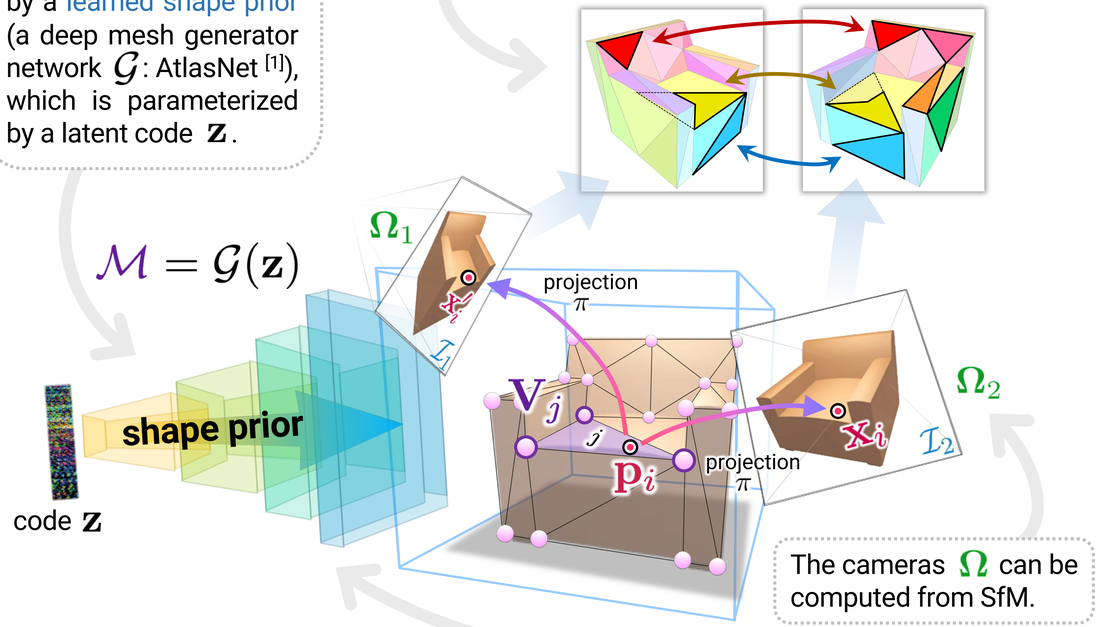
Data-driven priors (deep networks)

- ✅ Allows for **dense** 3D shape generation (**semantics**)
- ❌ Cannot generalize to **unseen sequences**

Approach

The mesh is constrained by a learned shape prior (a deep mesh generator network $\mathcal{G}$: AtlasNet [1]), which is parameterized by a latent code $\mathbf{z}$.

3D mesh optimization can be posed as a **2D piecewise image alignment** problem.



The cameras $\boldsymbol{\Omega}$ can be computed from SfM.

The photometric loss at $\mathbf{p}_i$ in triangle j is backpropagated to the latent code $\mathbf{z}$.

Photometric loss over mesh projections:

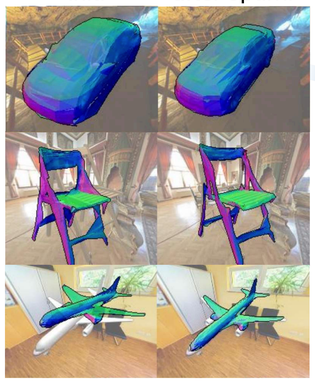
$$\sum_i \| \mathcal{I}_1(\mathbf{x}'_i) - \mathcal{I}_2(\mathbf{x}_i) \|_1$$

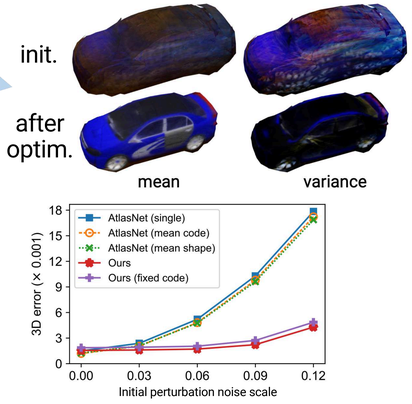
$$= \sum_j \sum_{i:\mathbf{p}_i \in \mathcal{P}_j} \| \mathcal{I}_1(\pi(\mathbf{p}_i(\mathbf{V}_j); \boldsymbol{\Omega}_1)) - \mathcal{I}_2(\pi(\mathbf{p}_i(\mathbf{V}_j); \boldsymbol{\Omega}_2)) \|_1$$

(mesh vertices; camera)

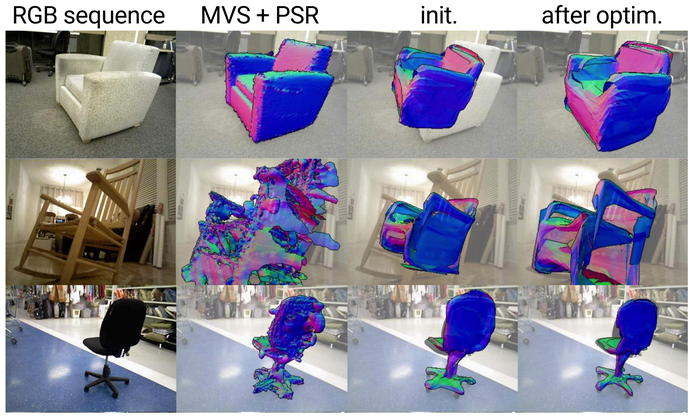
Results

ShapeNet + SUN360





Real-world videos



Check out our paper and code for more details and discussions!

[1] Groueix et al. "AtlasNet: A Papier-Mâché Approach to Learning 3D Surface Generation." CVPR 2018