GuideFlow3D: Optimization-Guided Rectified Flow For Appearance Transfer

Sayan Deb Sarkar Sinisa Stekovic Vincent Lepetit Iro Armeni

Stanford University · Gradient Spaces · École des Ponts ParisTech

Project Page

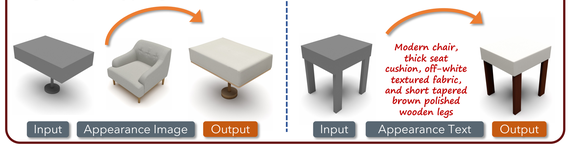
NEURAL INFORMATION PROCESSING SYSTEMS

Why 3D Appearance Transfer

Input Base 3D shape + appearance cue from image, mesh or text.

Output Geometry-preserving 3D model restyled with appearance.

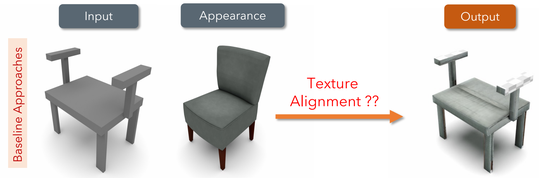
Goal Accelerate stylized 3D asset creation for gaming, AR/VR and digital prototyping.



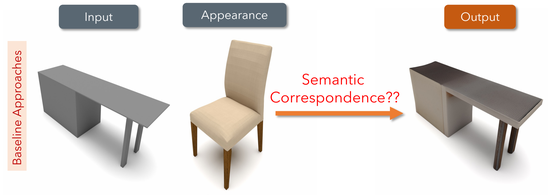
Key Challenges

Why is Appearance Transfer **Really Hard?**

- **Geometric irregularity** and absence of part-aware grounding disrupt texture alignment and structural consistency.



- **Large semantic gaps across categories** break matching correspondences causing style leakage and textures that fail to align with object geometry.

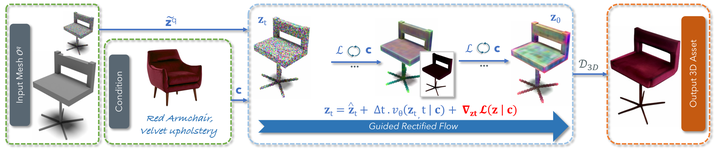


- 3D appearance transfer cannot be treated as a texture mapping or multi-view rendering problem.

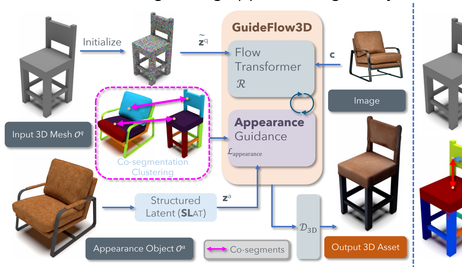
Guided Flow For Appearance Transfer

Our Approach

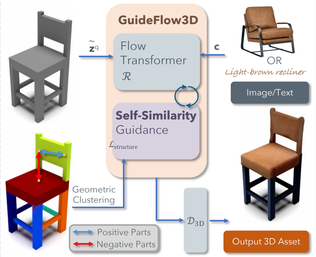
Differentiable Guidance During Sampling



- Training-free framework, interleaving rectified flow sampling and semantic + geometric prior as guidance objective.
- **Part-Aware Guidance:** Matches input and appearance latents through part-based co-segmentation and encourage semantically consistent style transfer.
- **Self-Similarity Guidance:** Promotes local consistency using geometric clustering without homogenizing appearance globally.

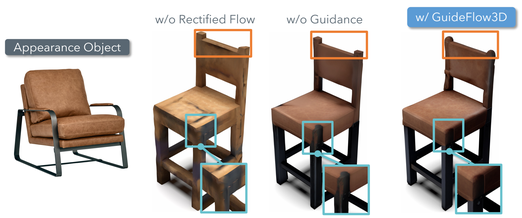


Part-Aware Guidance



Self-Similarity Guidance

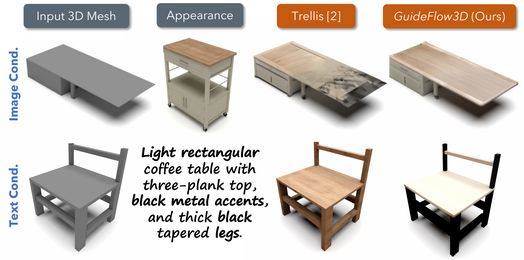
How does guiding structured latents help?



- No flow: unrealistic, distorted outputs.
- No guidance: style transfer fails.
- ✓ **w/ GuideFlow3D:** preserve global realism and fine-grained control.

Experimental Results

Furniture-To-Furniture Transfer



- Preserves geometric detail and delivers coherent transfer.
- Interprets cross-category textual cues with generalizability.

In-The-Wild Transfer

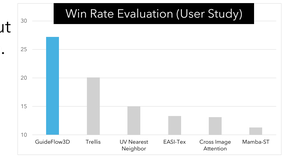
Robustly transfers appearance and generalizes to unseen shape categories and object styles, reinforcing the adaptability.

Our method shows superior texture fidelity and structural preservation.

Methods	Ranking metrics: Fidelity ↓ (Intra-Category)	Clarity ↓ (Intra-Category)	Adaptation ↓ (Intra-Category)	Fidelity ↓ (Inter-Category)	Clarity ↓ (Inter-Category)	Adaptation ↓ (Inter-Category)
Simple-Complex						
UV Nearest Neighbor	4.12	3.84	4.43	4.06	3.51	4.17
MambaST [8]	4.94	3.55	4.42	4.87	3.57	4.38
Cross Image Attention [2]	3.56	3.48	3.47	3.54	3.55	3.52
EasiTex [52]	3.18	4.30	4.18	3.25	4.21	4.10
Trellis [72]	2.51	2.58	2.61	2.64	2.85	2.76
GuideFlow3D (Ours)	**1.89**	**2.41**	**2.28**	**1.99**	**2.75**	**2.45**

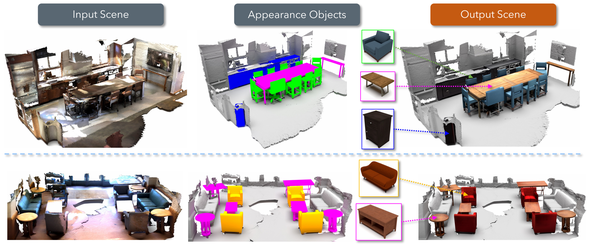
Failure Case

Interpreting abstract semantics without ambiguity remains an open challenge.



3D Scene Editing

Seamlessly stylizes objects while preserving their geometry and spatial layout for interactive context-aware 3D scene restyling.



Key Takeaways

- Novel framework for 3D appearance transfer that applies differentiable guidance to a pretrained rectified flow model.
- Training-free approach, generalizable to different appearance representations.

[1] Xiang et. al, Structured 3D Latents For Scalable and Versatile 3D Generation, *CVPR 2025*,. [2] Perla et. al, EASI-Tex: Edge-Aware Mesh Texturing from Single Image, *SIGGRAPH 2024*.
[3] Alaluf et. al, Cross-Image Attention for Zero-Shot Appearance Transfer, *SIGGRAPH 2024*. [4] Botti et. al, Mamba-ST: State Space Model for Efficient Style Transfer, *WACV 2025*.