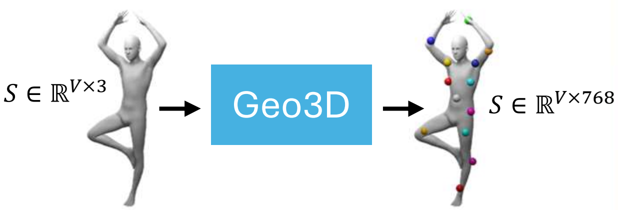
GeoAware3D: Enhancing Semantic Features for 3D Shape Decoration

Marco Börner Jed Guzelkabaagac

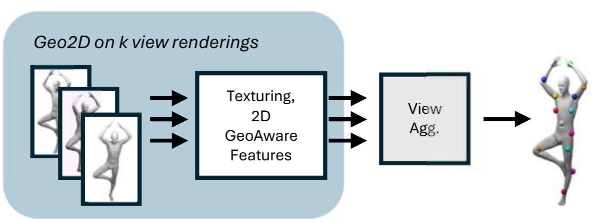
Zero-shot Shape Decoration:

- Want to capture geometric and semantic info
- Foundational 2D models capture semantics well
- Untextured meshes not suitable for SD & DINO – must add texture for descriptive features
- Geometric Awareness: remembers left & right



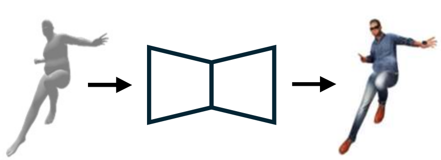
Projective Geometry:

- Render k views in 2D
- Get semantics using foundational models
- Unproject and view aggregate

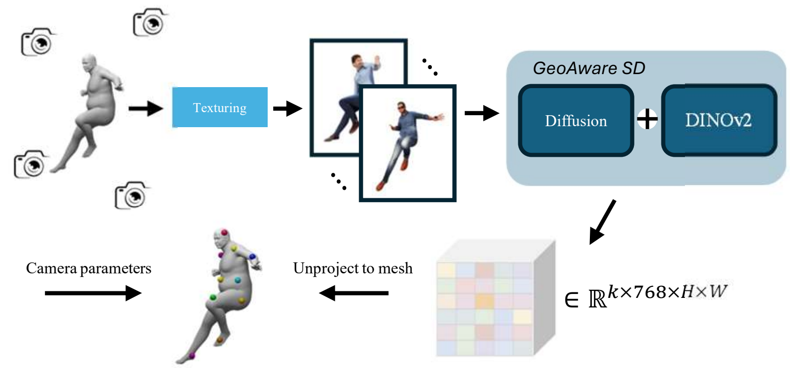


Texturing 2D Images:

- Renderings are textured first to encode more semantic information
- Image features are robust to inconsistent texturing from different views [1]



Pipeline:



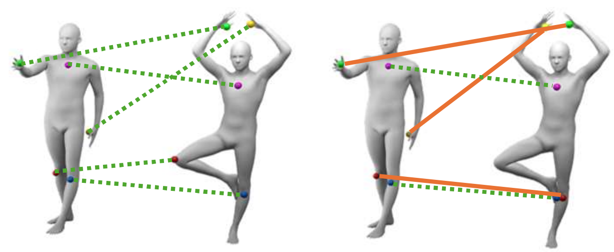
- Our method is more lightweight than DIFF3F [1], requiring over 4x less compute time and less VRAM

Method → Benchmark ↓	DIFF3F (zero-shot)	SE-ORNet (trained)	GeoAware3D (zero-shot)
SHREC'19 Accuracy (%)	**26.41**	21.41	23.42
Runtime (min) / mesh	4.42	?	**1.02**

Geometry-Aware Correspondence:

- SD + DINO are natural complements [3] for capturing semantics. However, struggle to encode consistent geometric info across different views
- GeoAware SC [2] addresses this issue, leading to improved understanding, particularly in humans and animals

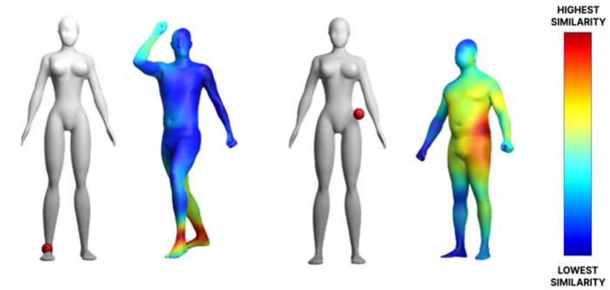
GeoAware vs SD+DINO Example:



GeoAware Features SD+DINO Features

Semantic Understanding:

- Our method encodes semantics of the meshes, even with challenging symmetries



Dense Correspondence:

- Robust to pose variation and extends to some non-isometric shapes

[1] DIFF3F

Method →	DIFF3F	SE-ORNet	GeoAware
Zero-shot?	✓	×	✓
Class agnostic?	✓	×	✓
Low dim features? (VRAM efficient)	×	✓	✓

References:
[1] Dutt, Muralikrishnan, et al. Diffusion 3D Features (Diff3F). CVPR '24
[2] Zhang, Hermann, et al. Telling Left from Right: Identifying Geometry-Aware Semantic Correspondence. CVPR '24
[3] Zhang, Hermann, et al. A Tale of Two Features: Stable Diffusion Complements DINO for Zero-Shot Semantic Correspondence. NeurIPS '23