Tavatar: Topology-Aware Gaussian Attribute Derivation for Animatable Human Avatars

Supplementary Material

7. Overview

This supplementary material provides additional implementation details, experimental settings, and comprehensive results that complement our main paper.

Algorithm 1 Analytical Face Gaussian Derivation

Input: Posed vertices $\mathbf{v}^p_{i,1}, \mathbf{v}^p_{i,2}, \mathbf{v}^p_{i,3} \in \mathbb{R}^3$
Hyperparameters: $\beta = 0.5, \epsilon = 10^{-3}$
Output: $\boldsymbol{\mu}^i_f \in \mathbb{R}^3, \mathbf{R}^i_f \in SO(3), \mathbf{s}^i_f \in \mathbb{R}^3$

// Compute edge lengths
$l_1 \leftarrow \|\mathbf{v}^p_{i,2} - \mathbf{v}^p_{i,3}\|$
$l_2 \leftarrow \|\mathbf{v}^p_{i,3} - \mathbf{v}^p_{i,1}\|$
$l_3 \leftarrow \|\mathbf{v}^p_{i,1} - \mathbf{v}^p_{i,2}\|$
// Position: incenter weighted by opposite edges
$\boldsymbol{\mu}^i_f \leftarrow \frac{l_1\mathbf{v}^p_{i,1} + l_2\mathbf{v}^p_{i,2} + l_3\mathbf{v}^p_{i,3}}{l_1+l_2+l_3}$
// Orientation: construct local coordinate frame
$\mathbf{n}^i_f \leftarrow \frac{(\mathbf{v}^p_{i,2}-\mathbf{v}^p_{i,1})\times(\mathbf{v}^p_{i,3}-\mathbf{v}^p_{i,1})}{\|(\mathbf{v}^p_{i,2}-\mathbf{v}^p_{i,1})\times(\mathbf{v}^p_{i,3}-\mathbf{v}^p_{i,1})\|}$
$\mathbf{t}^i_1 \leftarrow \frac{\mathbf{v}^p_{i,2}-\mathbf{v}^p_{i,1}}{\|\mathbf{v}^p_{i,2}-\mathbf{v}^p_{i,1}\|}$
$\mathbf{t}^i_2 \leftarrow \mathbf{n}^i_f \times \mathbf{t}^i_1$
$\mathbf{R}^i_f \leftarrow [\mathbf{t}^i_1, \mathbf{t}^i_2, \mathbf{n}^i_f]$
// Scale: tied to inradius for uniform coverage
$A_i \leftarrow \frac{1}{2}\|(\mathbf{v}^p_{i,2} - \mathbf{v}^p_{i,1}) \times (\mathbf{v}^p_{i,3} - \mathbf{v}^p_{i,1})\|$
$s_i \leftarrow (l_1 + l_2 + l_3)/2$
$r_i \leftarrow A_i/s_i$
$s^i_{f,x} \leftarrow \beta \cdot r_i, s^i_{f,y} \leftarrow \beta \cdot r_i, s^i_{f,z} \leftarrow \epsilon$
return $\boldsymbol{\mu}^i_f, \mathbf{R}^i_f, \mathbf{s}^i_f = (s^i_{f,x}, s^i_{f,y}, s^i_{f,z})$

8. Analytical Gaussian Derivation Details

We provide algorithmic details for analytically deriving Gaussian attributes from mesh geometry (Sec 4.2 of the main paper), where all geometric attributes ($\boldsymbol{\mu}$, $\mathbf{R}$, $\mathbf{s}$) are deterministically computed from local mesh topology.

Both algorithms output three geometric attributes for each Gaussian: (1) position $\boldsymbol{\mu} \in \mathbb{R}^3$, (2) rotation matrix $\mathbf{R} \in SO(3)$ encoding orientation via a local orthonormal frame, and (3) scale vector $\mathbf{s} \in \mathbb{R}^3$ with a small z-component (ϵ) for surface conformity.

8.1. Face Gaussian Derivation

Algorithm 1 derives Face Gaussian $\mathcal{G}^i_f$ for each face $f_i \in \mathbf{F}$ with vertices $\{\mathbf{v}^p_{i,1}, \mathbf{v}^p_{i,2}, \mathbf{v}^p_{i,3}\}$. The position $\boldsymbol{\mu}^i_f$ is placed at the weighted incenter, the rotation $\mathbf{R}^i_f$ aligns with the face's coordinate frame (tangents $\mathbf{t}^i_1$, $\mathbf{t}^i_2$ and normal $\mathbf{n}^i_f$), and the scale $\mathbf{s}^i_f$ is tied to the inradius r_i for uniform coverage.

Algorithm 2 Analytical Vertex Gaussian Derivation

Input: Posed vertex $\mathbf{v}^p_j \in \mathbb{R}^3$, adjacent faces $\mathcal{N}(j)$, one-ring neighbors $\mathcal{N}_1(j)$
Hyperparameters: $\gamma = 0.5, \epsilon = 10^{-3}$
Output: $\boldsymbol{\mu}^j_v \in \mathbb{R}^3, \mathbf{R}^j_v \in SO(3), \mathbf{s}^j_v \in \mathbb{R}^3$

// Position: anchor at vertex
$\boldsymbol{\mu}^j_v \leftarrow \mathbf{v}^p_j$
// Orientation: area-weighted normal from adjacent faces
$\mathbf{n}^j_v \leftarrow \frac{\sum_{f\in\mathcal{N}(j)} A_f\mathbf{n}_f}{\|\sum_{f\in\mathcal{N}(j)} A_f\mathbf{n}_f\|}$
// Construct stable orthonormal frame
$\mathbf{r} \leftarrow \arg\min_{\mathbf{u}\in\{\mathbf{e}_x,\mathbf{e}_y,\mathbf{e}_z\}} |\mathbf{u}^\top\mathbf{n}^j_v|$
$\mathbf{t}^j_1 \leftarrow \frac{\mathbf{r}\times\mathbf{n}^j_v}{\|\mathbf{r}\times\mathbf{n}^j_v\|}$
$\mathbf{t}^j_2 \leftarrow \mathbf{n}^j_v \times \mathbf{t}^j_1$
$\mathbf{R}^j_v \leftarrow [\mathbf{t}^j_1, \mathbf{t}^j_2, \mathbf{n}^j_v]$
// Scale: adapt to local mesh density
$s^j_{v,x} \leftarrow \gamma \cdot \min_{\mathbf{v}_k\in\mathcal{N}_1(j)} \|\mathbf{v}^p_k - \mathbf{v}^p_j\|$
$s^j_{v,y} \leftarrow s^j_{v,x}, s^j_{v,z} \leftarrow \epsilon$
return $\boldsymbol{\mu}^j_v, \mathbf{R}^j_v, \mathbf{s}^j_v = (s^j_{v,x}, s^j_{v,y}, s^j_{v,z})$

8.2. Vertex Gaussian Derivation

Algorithm 2 derives Vertex Gaussian $\mathcal{G}^j_v$ for each vertex $\mathbf{v}^p_j \in \mathbf{V}_p$. The position $\boldsymbol{\mu}^j_v$ is anchored at the vertex, the rotation $\mathbf{R}^j_v$ is computed from area-weighted adjacent face normals, and the scale $\mathbf{s}^j_v$ adapts to local mesh density via minimum edge length to one-ring neighbors.

9. Additional Results and Analysis

9.1. Computational Efficiency

Table A compares computational efficiency against IHuman on the male-3-casual subject (PeopleSnapshot dataset) using a single NVIDIA GeForce RTX 3090 GPU.

Our method achieves comparable training time (∼2 minutes) and higher rendering speed (50 FPS vs. 44 FPS) while using 62.5% fewer Gaussians (82,658 vs. 220,416), which directly translates to faster rasterization due to reduced computational overhead. Since geometric attributes ($\boldsymbol{\mu}$, $\mathbf{R}$, $\mathbf{s}$) are analytically derived rather than stored, the checkpoint size is reduced by 44.5% (32.36 MB vs. 58.26 MB), storing only appearance parameters (SH coefficients and shape encoder). The fixed Gaussian count (M Face +

Table A. Computational efficiency comparison on the male-3-casual subject from PeopleSnapshot dataset.

Method	# Gaussians	Train Time	Model Size	FPS
IHuman	220,416	∼2 min	58.26 MB	44
Ours	82,658	∼2 min	32.36 MB	50

Table B. Robustness to pose noise. We inject Gaussian noise $\mathcal{N}(0, \sigma)$ into SMPL pose parameters on PeopleSnapshot male-3-casual. Percentages indicate relative degradation from σ=0.

Noise (σ)	PSNR ↑		Normal($\times 10^{-3}$) ↓	
	Ours	IHuman	Ours	IHuman
0	28.92	26.66	1.88	2.04
0.005	27.02 (**-6.6%**)	23.75 (-10.9%)	2.01 (**+6.9%**)	2.53 (+24.0%)
0.01	25.72 (**-11.1%**)	22.54 (-15.5%)	2.47 (**+31.4%**)	3.13 (+53.4%)
0.02	23.81 (**-17.7%**)	20.69 (-22.4%)	3.88 (**+106.4%**)	4.56 (+138.2%)

N Vertex Gaussians) preserves topological correspondence without adaptive densification.

9.2. Robustness to Pose Noise

To evaluate the robustness of our topology-aware binding under inaccurate SMPL fitting, we inject Gaussian noise $\mathcal{N}(0, \sigma)$ into the pose parameters on the PeopleSnapshot male-3-casual subject and compare against IHuman. As shown in Table B, our method degrades more gracefully across all noise levels. For instance, at σ=0.01, our PSNR drops by 11.1% compared to 15.5% for IHuman, and our normal error increases by 31.4% versus 53.4%. This is because our analytically-derived Gaussians remain bound to the mesh surface and move coherently with the deformation, rather than drifting or floating as in methods with unconstrained Gaussian optimization.

9.3. Sensitivity to ER Weight

We ablate the equilateral regularization weight λ_t on the X-Avatar dataset. We follow IHuman's default settings for λ_n and λ_m, and vary only λ_t. As shown in Table C, λ_t=0.01 achieves the best trade-off across all metrics. Disabling the regularization (λ_t=0) leads to noticeable degradation in both rendering quality and geometric accuracy, confirming that equilateral regularization is essential for maintaining stable Gaussian binding under mesh deformation. Increasing λ_t beyond 0.01 over-constrains the mesh, slightly reducing reconstruction fidelity.

9.4. Additional Qualitative Results

Figures 8-10 present extended qualitative comparisons on the PeopleSnapshot and X-Avatar datasets. These results demonstrate our method's robustness on challenging out-of-distribution poses, validating the effectiveness of our topology-aware Gaussian representation for maintaining consistent appearance fidelity and geometric accuracy under extreme pose variations. The analytically-derived Gaussian attributes ensure stable deformation across diverse body configurations.

Table C. Sensitivity analysis of equilateral regularization weight λ_t on the X-Avatar dataset.

λ_t	PSNR ↑	SSIM ↑	Normal ↓
0	27.32	0.977	1.93
0.005	28.01	0.979	1.79
0.01 (Ours)	**28.13**	**0.981**	**1.74**
0.02	27.98	0.975	1.75

Figure 5. Qualitative comparison on novel out-of-distribution poses from the PeopleSnapshot dataset. *Please zoom in for detailed comparison.*

Figure 6. Qualitative comparison on male-3-casual subject performing Da-pose. *Please zoom in for detailed comparison.*

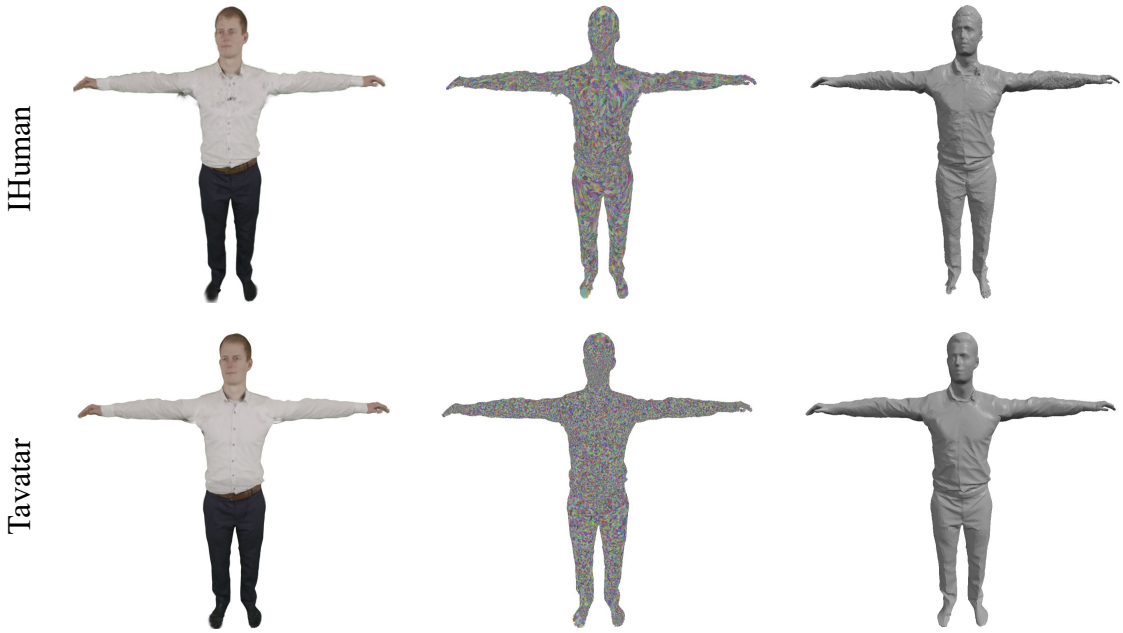

Figure 7. Qualitative comparison on male-4-casual subject in T-pose. *Please zoom in for detailed comparison.*

Figure 8. Qualitative comparison on female-3-casual subject performing Da-pose. *Please zoom in for detailed comparison.*

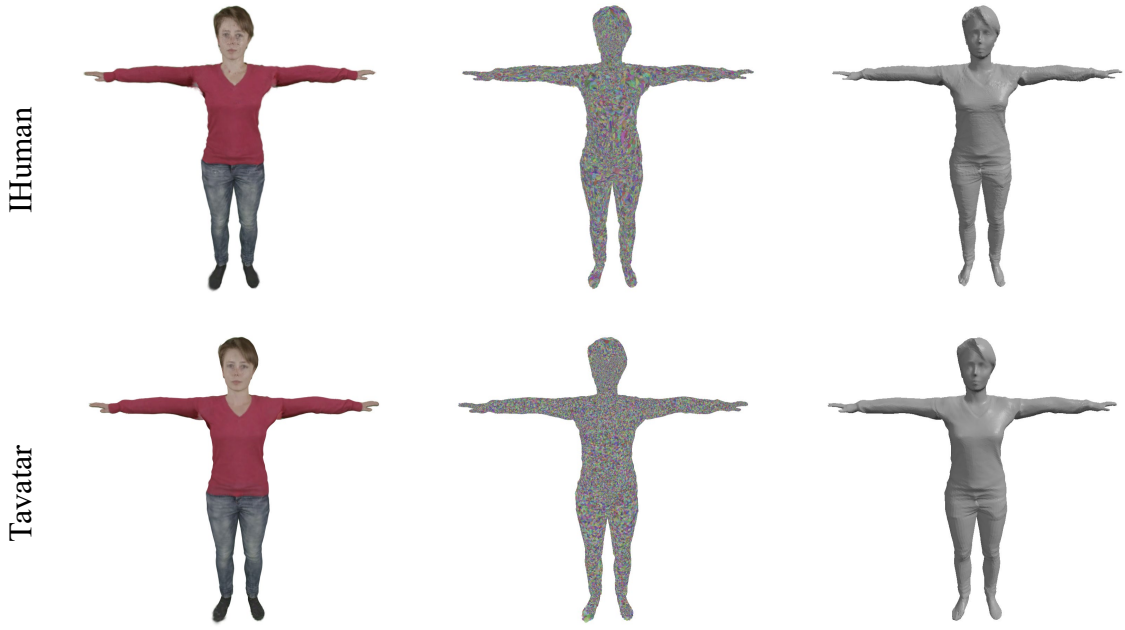

Figure 9. Qualitative comparison on female-4-casual subject in T-pose. *Please zoom in for detailed comparison.*

Figure 10. Additional qualitative comparison on the X-Avatar dataset. *Please zoom in for detailed comparison.*