PanoHair: Detailed Hair Strand Synthesis on Volumetric Heads

Shashikant Verma Shanmuganathan Raman

CVIG Lab, Indian Institute of Technology Gandhinagar, India

Abstract

Achieving realistic hair strand synthesis is essential for creating lifelike digital humans, but producing high-fidelity hair strand geometry remains a significant challenge. Existing methods require a complex setup for data acquisition, involving multi-view images captured in constrained studio environments. Additionally, these methods have longer hair volume estimation and strand synthesis times, which hinder efficiency. We introduce PanoHair, a model that estimates head geometry as signed distance fields using knowledge distillation from a pre-trained generative teacher model for head synthesis. Our approach enables the prediction of semantic segmentation masks and 3D orientations specifically for the hair region of the estimated geometry. Our method is generative and can generate diverse hairstyles with latent space manipulations. For real images, our approach involves an inversion process to infer latent codes and produces visually appealing hair strands, offering a streamlined alternative to complex multi-view data acquisition setups. Given the latent code, PanoHair generates a clean manifold mesh for the hair region in under 5 seconds, along with semantic and orientation maps, marking a significant improvement over existing methods, as demonstrated in our experiments.

Contributions

Our contributions are as follows:

- We introduce PanoHair, a generative framework that distills knowledge from PanoHead to model human heads as SDFs while predicting semantic labels and 3D orientations.
- We propose a multi-view orientation projection loss to ensure view-consistent 3D orientation learning and mitigate ambiguities from single-view 2D projections.
- We employ a fast and robust hair volume extraction method using boolean operations, leveraging the structured and bounded nature of the learned implicit representation.

PanoHair Network Architecture

We leverage knowledge distillation from PanoHead [1], which predicts radiance field densities, guiding our SDF-based surface modeling. Our network also predicts binary segmentation labels on the SDF iso-surface to identify hair regions. Supervision for segmentation and orientation comes from Gabor maps and segmentation masks rendered from PanoHead.

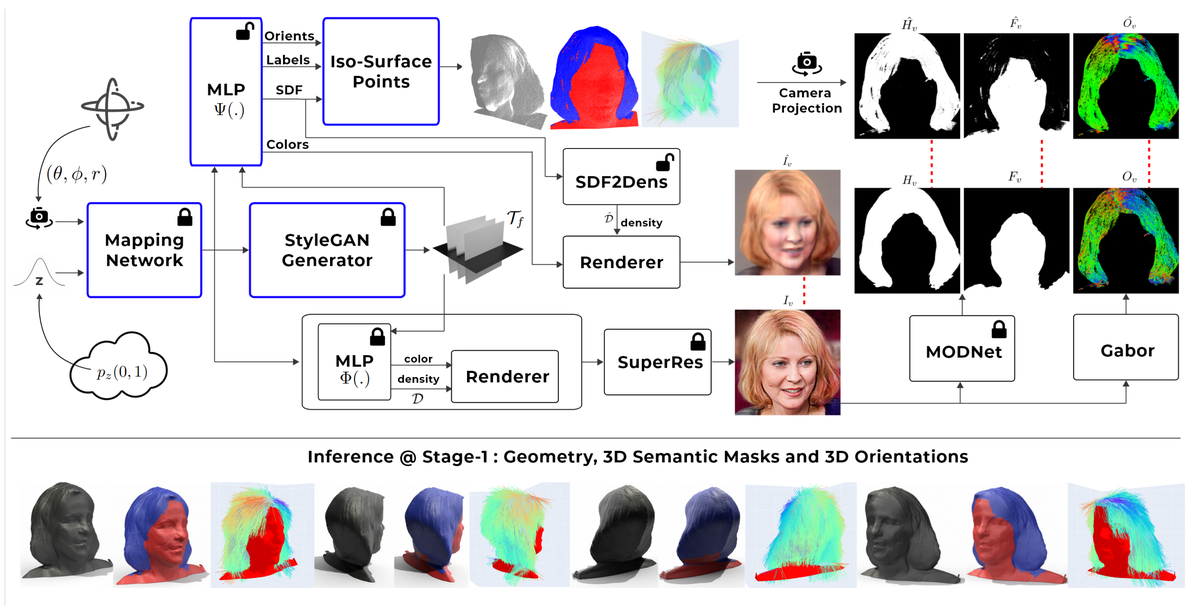


Figure 1. **PanoHair Overview.** Given latent code z and camera parameters $C_{\text{cam}} = (\theta, \phi, r)$, we first obtain tri-grid features $\mathcal{T}_f$ using the pre-trained PanoHead [1]. A set of MLPs $\Psi(.)$ then predicts semantic labels, 3D orientations, and SDF values. Training is supervised by distilling knowledge from PanoHair image renderings, Gabor orientations, and semantic segmentation. During inference, only the Blue-boxed components are required. The bottom row illustrates PanoHair's outputs—geometry, semantic labels, and 3D orientations.

Loss Function

We train PanoHair using a combination of reconstruction, semantic mask, density, and orientation constraints. The orientation loss $\mathcal{L}_{\text{orient}}$ enforces tangential orientation to the surface and multi-view consistency across latent codes. Full formulations and weighting schemes are provided in the supplementary material.

$$\mathcal{L} = \underbrace{\lambda_1 \cdot \delta(t) \cdot \|\mathcal{D} - \hat{\mathcal{D}}\|_1}_{\text{Density Loss}} + \underbrace{\lambda_2 \cdot \|I^+ - I\|^2}_{\text{Image Reconstruction Loss}} + \underbrace{\lambda_3 \cdot \text{BCE}(I_M, \hat{M}_{2D})}_{\text{Semantic Loss}} + \mathcal{L}_{\text{orient}} \quad (1)$$

$$\mathcal{L}_{\text{orient}} = \lambda_4 \mathcal{L}_{\text{tan}} + \lambda_5 \mathcal{L}_{\text{proj}} \quad (2)$$

$$\mathcal{L}_{\text{tan}} = \frac{1}{N} \sum_{i=1}^{N} |o_S^i \cdot n_S^i| \quad \text{where} \quad n_S = \frac{\nabla \mathcal{S}}{||\nabla \mathcal{S}||} \quad (3)$$

$$\mathcal{L}_{\text{proj}} = \sum_{i=0}^{k-1} \min \left(\left| {}^0\hat{\mathcal{O}}_{2D} - {}^iI_O \right|, \left| {}^0\hat{\mathcal{O}}_{2D} - {}^iI_O - \pi \right|, \left| {}^0\hat{\mathcal{O}}_{2D} - {}^iI_O + \pi \right| \right) \quad (4)$$

Hair Volume Estimation

By selecting points classified as hair, we obtain the hair's outer surface and 3D orientations, as shown in Figure 2.

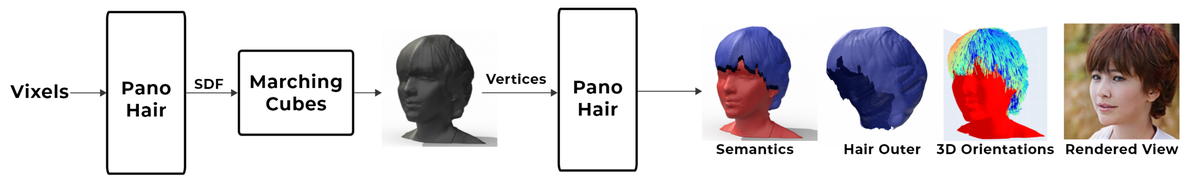


Figure 2. Inferencing hair's outer surface and orientations. A $512 \times 512 \times 512$ volumetric grid is sampled within $[-0.75, 0.75]$ to estimate SDF values. The geometry is extracted using marching cubes, and semantic labels and 3D orientations are computed for the mesh vertices.

We compute the volume by performing Boolean subtraction between the fitted FLAME model and empirically estimated universal fit.

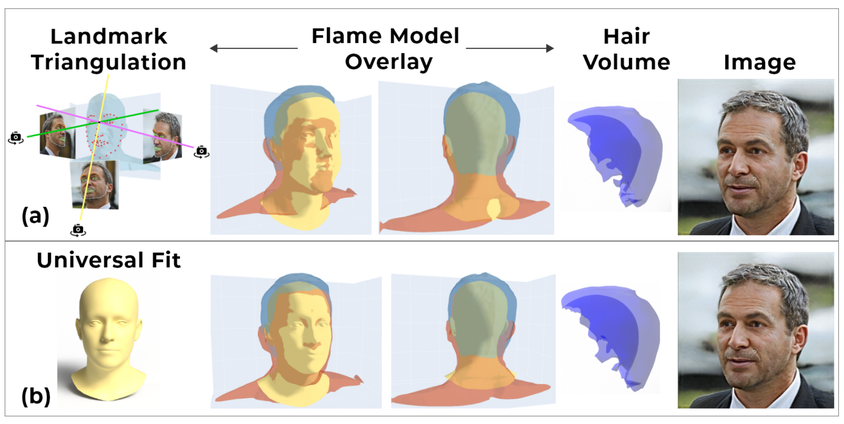


Figure 3. Estimating hair volume. (a) FLAME model fitted to estimated landmarks. (b) Empirically computed universal fit, leveraging the structured representation of volumetric faces within a unit volume by design during training. Subsequently, we employ Boolean subtraction between the fitted FLAME model and the extracted mesh to compute volume.

Hair Strand Growing

We generate hair strands from the localized scalp region of the fitted FLAME model and grow them using orientation constraints predicted by PanoHair at the outermost shell.

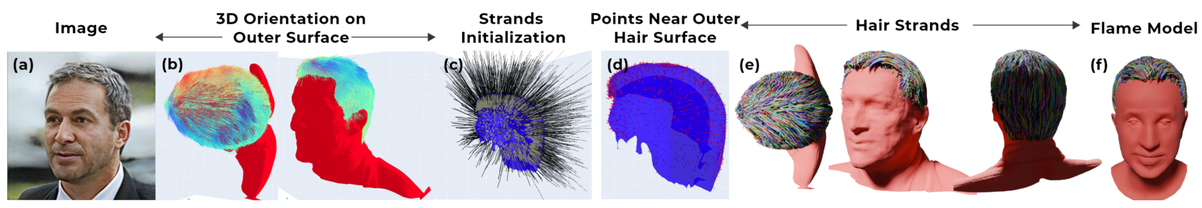
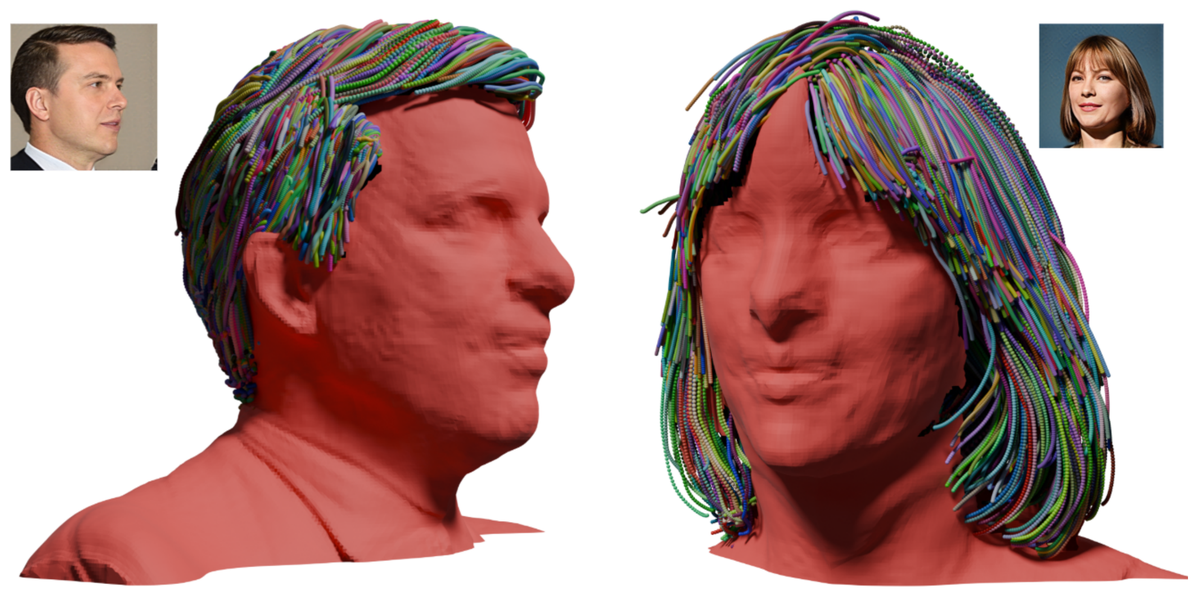


Figure 4. Strand-based hair growth. Given a latent code, (a) shows the generated image, and (b) visualizes 3D orientations on the outer hair surface. Strands are initialized using shape textures from [2]. Unlike their rasterization-based optimization, we directly optimize boundary strand points (d) to align with predicted orientations (b), producing the final hair (e). (f) shows the FLAME model with strands rooted at the scalp.

Results



Model	Input	Hair Seg. (IoU)	Orient. Error at Hair Region ($\mathcal{L}_{\text{proj}}$)	Hair Volume Est. time	Strand Reconstruction time
NeuralHaircut	Multi-View	0.815	0.391	10-12 Hrs	15-18 Hrs
HairStep	Single View	0.682	0.526	~3 sec	~5 mins.
PanoHair (Ours)	Latent code	**0.846**	**0.368**	~5 sec	4-5 Hrs.

Table 1. Quantitative comparison and the total runtime on an NVIDIA RTX-4090 GPU for a single head.

References

[1] Sizhe An, Hongyi Xu, Yichun Shi, Guoxian Song, Umit Y. Ogras, and Linjie Luo. Panohead: Geometry-aware 3d full-head synthesis in 360deg. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 20950–20959, June 2023.

[2] Vanessa Sklyarova, Jenya Chelishev, Andreea Dogaru, Igor Medvedev, Victor Lempitsky, and Egor Zakharov. Neural haircut: Prior-guided strand-based hair reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 19762–19773, 2023.

shashikant.verma@iitgn.ac.in