

GeomHair: Reconstruction of Hair Strands from Colorless 3D Scans

Supplementary Material

Rachmadio Noval Lazuardi^{*1} Artem Sevastopolsky^{*1} Egor Zakharov²
Matthias Nießner¹ Vanessa Sklyarova^{3,2}

¹ Technical University of Munich ² ETH Zürich ³ Max Planck Institute for Intelligent Systems

1. Strands400 Dataset

Legal notice. All participants in our dataset signed an agreement form compliant with GDPR. Please note that GDPR compliance includes the right for every participant to request the timely deletion of their data, which we will enforce as part of the distribution process of our dataset.

Dataset statistics. In Fig. 1, we demonstrate the distribution of the age, stratified by gender, (left) and of ethnicity (right), reported by the participants in the Strands400 dataset. In the age histogram (left), the horizontal axis corresponds to the participants' age bins, and the vertical axis corresponds to the number of people in the bin. Only the votes of the participants who willingly disclosed that information were taken into account.

In the main paper, we visualize the t-SNE projection of VQA embeddings colored by hairstyle type. Here, we additionally show the projection colored by hair waviness in Fig. 7. Using the same BLIP embeddings [4] of LLaVA [5] answers and K-Means [2] clustering with 5 clusters, we observe clear groupings that separate straight, wavy, and curly hairstyles. The t-SNE locations are enhanced with LLaVA answers for the respective samples, and each second sample is shown to provide more space for the captions. Together with the hairstyle type visualization in the main paper, these plots confirm that the Strands400 dataset captures a broad spectrum of hairstyle attributes.

Capture setup. Our dataset consists of two parts – the subset of the latest version of NPHM that contains 383 scans, and additionally collected 17 scans with a setup similar to the NPHM setup. We processed 472 scans from NPHM dataset and categorize the reconstructed strands into three categories: good, improvable, and poor. This process yielded 322 high-quality samples. Additionally, we enhanced 61 samples from the second category by removing facial hair that had been inadvertently included

from hair segmentation errors, originally happened during preprocessing. Facial hair was removed by manually selecting the polygons in 2D, defining the hair regions to exclude. Finally, we collected 24 samples separately and exclude the 7 samples from the dataset. The additional scans collection setup consists of two handheld Artec Eva scanners, rotating over a 360° trajectory within ~ 3 seconds to capture a single person. Since our setup largely follows NPHM capture setup, we refer the reader to the NPHM paper for the remaining details regarding the capture setup [3]. The participants in this category were selected with an emphasis on hairstyles, more challenging for reconstruction (wavy, curly, etc.), to better align the overall distribution to the overall spectrum of hairstyles.

Representative samples. More samples from the Strands400 dataset are demonstrated in Fig. 2.

2. Technical Details

Questions employed in the calculation of the text prompt for the diffusion prior conditioning. We provide the list of questions for querying LLaVA model below.

1. *Describe in detail the bang/fringe of depicted hairstyle including its directionality, texture and coverage of face?*
2. *What is the overall hairstyle depicted in the image?*
3. *Does the depicted hairstyle longer than the shoulders or shorter than the shoulder?*
4. *Does the depicted hairstyle has short bang or long bang or no bang from frontal view?*
5. *Does the hairstyle has straight bang or Baby Bangs or Arched Bangs or Asymmetrical Bangs or Pin-Up Bangs or Choppy Bangs or curtain bang or side swept bang or no bang?*
6. *Are there any afro features in the hairstyle or no afro features?*
7. *Is the length of hairstyle shorter than middle of the neck or longer than middle of the neck?*

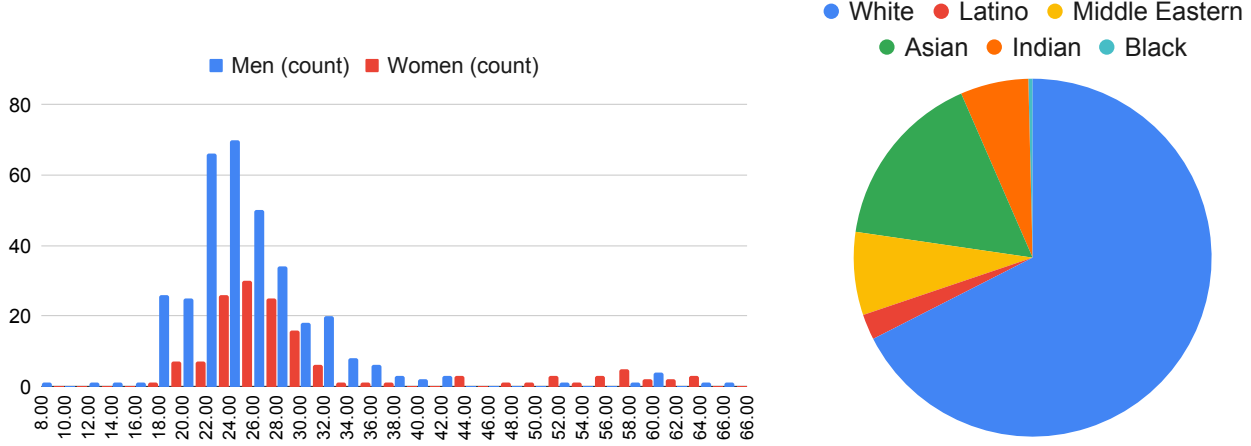


Figure 1. The distribution of the age, stratified by gender, (left) and of ethnicity (right), reported by the participants in the Strands400 dataset. Only the votes of the participants who willingly disclosed that information were taken into account.

8. What is the main geometry features of the depicted hairstyle?
9. What is the overall shape of the depicted hairstyle?
10. Is the hair short, medium, or long in terms of length?
11. What is the type of depicted hairstyle
12. What is the length of hairstyle relative to human body?
13. Describe the texture and pattern of hair in the image.
14. What is the texture of depicted hairstyle
15. Does the depicted hairstyle is straight or wavy or curly or kinky?
16. Can you describe the overall flow and directionality of strands?
17. Could you describe the bang of depicted hairstyle including its directionality and texture
18. Describe the main geometric features of the hairstyle depicted in the image
19. Is the length of hairstyle buzz cut, pixie, ear length, chin length, neck length, shoulder length, armpit length or mid-back length?
20. Describe actors with similar hairstyle type.
21. Does the hairstyle cover any parts of the face? Write which exactly parts.
22. In what ways is this hairstyle a blend or combination of other popular hairstyles?
23. Could you provide the most closest types of hairstyles from which this one could be blended?
24. How adaptable is this hairstyle for various occasions (casual, formal, athletic)?
25. How is this hairstyle perceived in different social or professional settings?
26. Are there historical figures who were iconic for wearing this hairstyle?

27. Could you describe the partition of this hairstyle if it is visible?

3D orientation extraction with crest lines. Following [7], we define crest lines on a surface S as the set of points where one of the principal curvatures reaches an extremum along its corresponding curvature direction. Mathematically, we can express this as:

$$e_{\max} = \frac{\partial k_{\max}}{\partial t_{\max}} = 0 \quad (\text{for convex crest lines}), \quad (1)$$

$$e_{\min} = \frac{\partial k_{\min}}{\partial t_{\min}} = 0 \quad (\text{for concave crest lines}), \quad (2)$$

where k_{\max} and k_{\min} are the maximum and minimum principal curvatures, with corresponding principal directions t_{\max} and t_{\min} . The associated extremality coefficients are e_{\max} and e_{\min} .

We identify crest lines on 3D hair scans using local cubic polynomial fitting at each mesh vertex, of the form:

$$h(x, y) = \frac{1}{2}(b_0x^2 + 2b_1xy + b_2y^2) + \frac{1}{6}(d_0x^3 + 3d_1x^2y + 3d_2xy^2 + d_3y^3)$$

After computing these values, crest lines are traced across the mesh by connecting points where the extremality coefficients vanish. To retain only the most salient hair features, we apply thresholding based on the cyclideness measure:

$$C = \sqrt{|e_{\max}|^2 + |e_{\min}|^2} \quad (3)$$

Let $C = \{c_1, \dots, c_N\}$ be a set of crest lines, where each $c_i = \{\mathbf{p}_i\}_{i=1}^{L_i}$ consists of L_i points, which may vary across

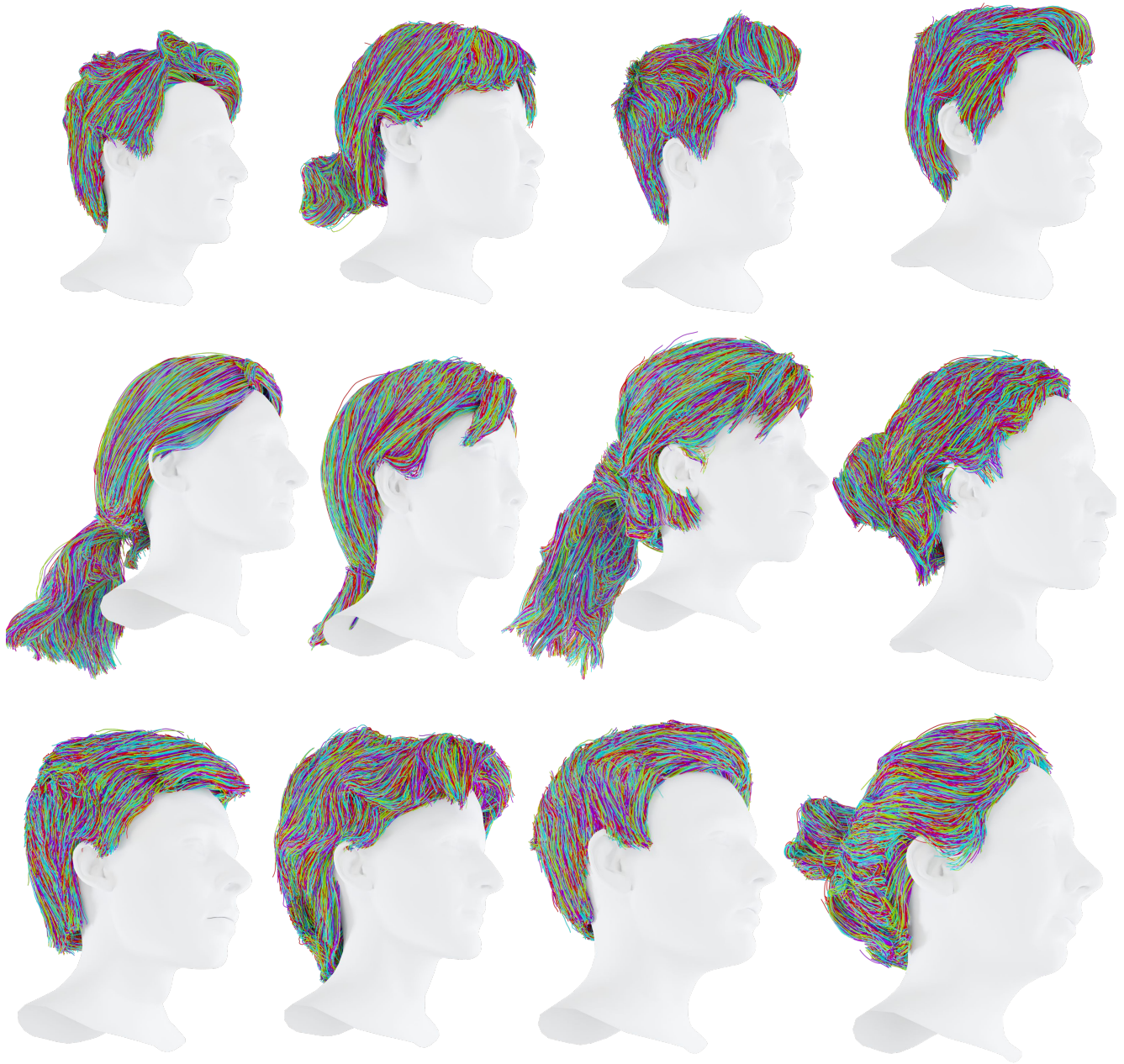


Figure 2. Sample strands reconstructions from Strands400 dataset.

crest lines. Each crest line is treated as a guide hair strand for computing local orientations. We estimate local curvature along each crest line c_i to guide adaptive window sizing. For each point \mathbf{p}_i , a local coordinate frame is computed via PCA [1] over its neighborhood. An adaptive smoothing step refines these frames, balancing noise reduction and preservation of directional variation. The resulting normalized orientations are used as supervision for strand-based reconstruction.

Gabor filter sensitivity to hair color. We demonstrate an experiment where the same hairstyle produces different extracted orientations when Gabor filters are applied to different hair colors. Fig. 3 shows that the orientation field obtained from lighter hair matches the straight hairstyle more closely, while the orientation field from darker hair appears noticeably noisier. This contrived example demonstrates that, even under decent capture conditions, hair reconstruction from multi-view RGB images can highly de-

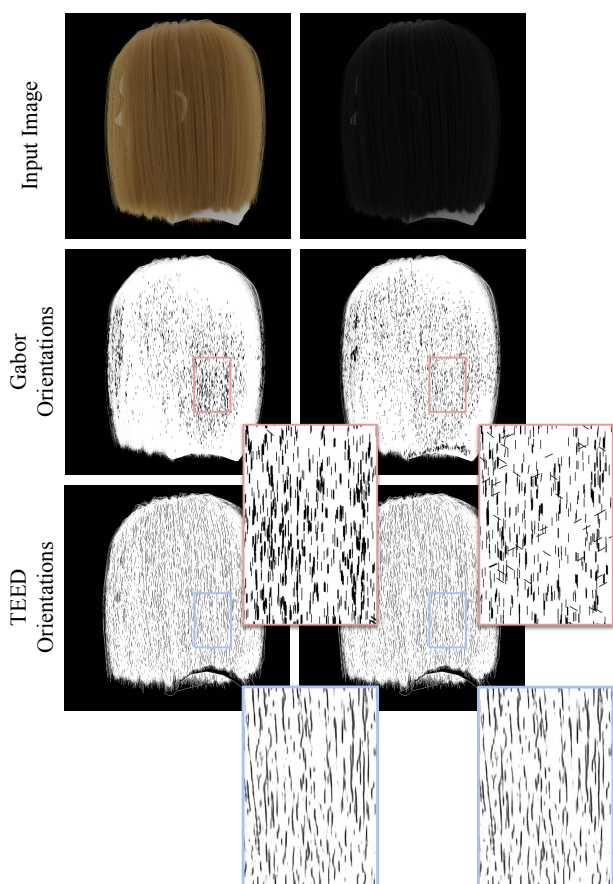


Figure 3. Commonly used hair orientation detector in RGB – Gabor filters – is sensitive to the hair color. The lighter color (left row) yields a better orientation than its darker (right row) counterpart. On the other hand, orientation detected from TEED is more robust to hair color. This example demonstrates that, even under decent capture conditions, hair reconstruction from multi-view RGB images can highly depend on features like hair color, unlike the typical reconstruction from geometry obtained from a structured-light 3D scanner.

pend on features like hair color, unlike the typical reconstruction from geometry obtained from a structured-light 3D scanner.

3. Results

Extended application. We provide additional results demonstrating our method’s versatility with head meshes from off-the-shelf generative models. Figure 4 presents extended reconstruction results using an image-to-mesh model, while Figure 5 shows results from a text-to-mesh model.

Limitations. While our method performs well on wavy and straight hairstyles, very curly scenes remain challeng-

ing because scans often fail to capture such high-frequency details, amplifying noise in the orientation estimators. As an example, we provide two samples where our method fails to reconstruct. Specifically, our method is less robust to curly and very short hairstyles, as illustrated in Fig. 6. One approach to address such limitation is by improving the hairstyle diffusion priors.

References

- [1] Hervé Abdi and Lynne J Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010. 3
- [2] Mohiuddin Ahmed, Raihan Seraj, and Syed Mohammed Shamsul Islam. The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, 9(8):1295, 2020. 1, 8
- [3] Simon Giebenhain, Tobias Kirschstein, Markos Georgopoulos, Martin Rünz, Lourdes Agapito, and Matthias Nießner. Learning neural parametric head models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21003–21012, 2023. 1
- [4] Junnan Li, Dongxu Li, Caiming Xiong, and Steven Hoi. Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation. In *ICML*, 2022. 1, 8
- [5] Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. Visual instruction tuning. In *NeurIPS*, 2023. 1, 8
- [6] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008. 8
- [7] Shin Yoshizawa, Alexander Belyaev, and Hans-Peter Seidel. Fast and robust detection of crest lines on meshes. In *Proceedings of the 2005 ACM symposium on Solid and physical modeling*, pages 227–232, 2005. 2

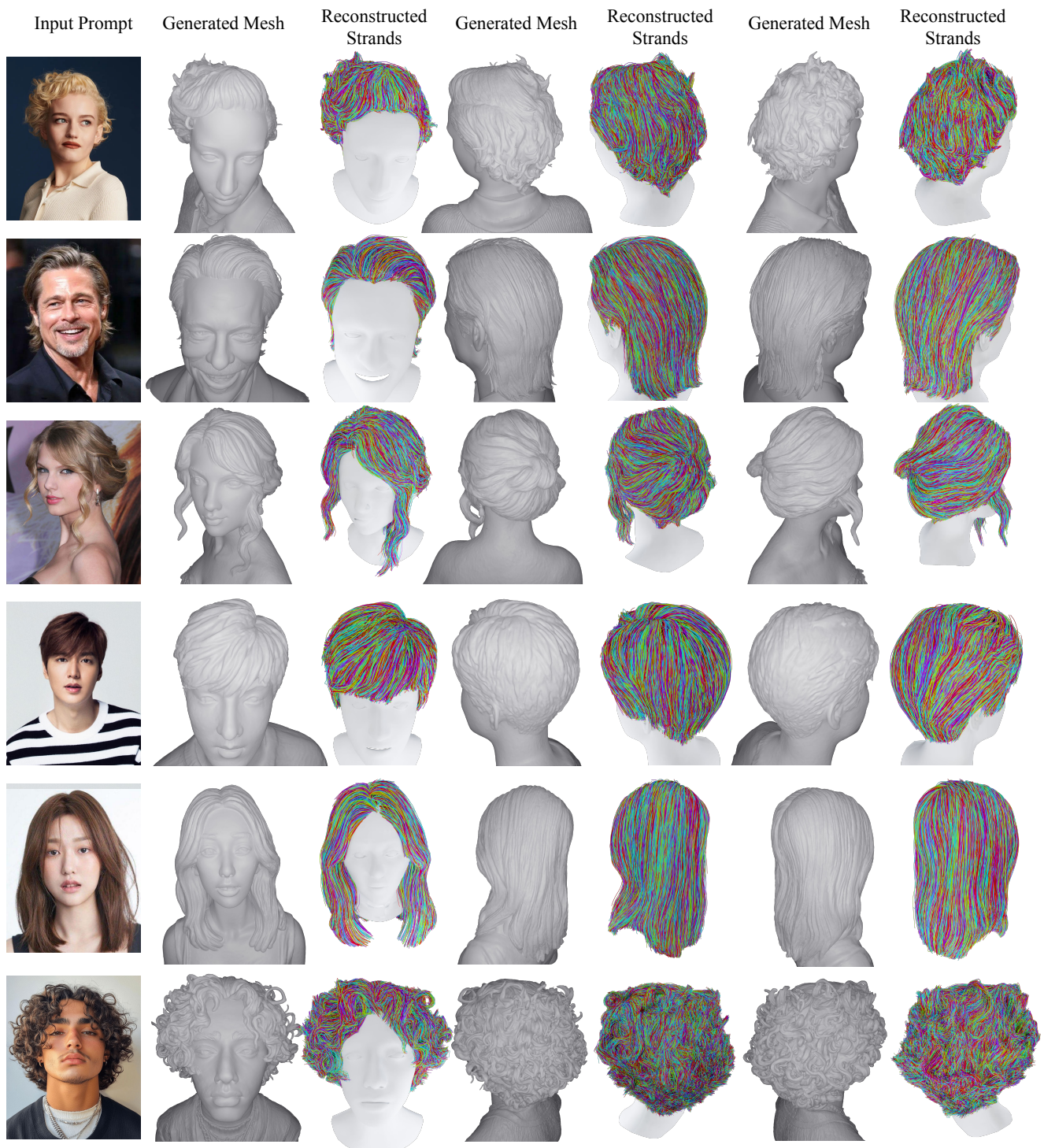


Figure 4. Reconstruction results with meshes from an off-the-shelf image-to-mesh model, demonstrating our method’s ability to generate diverse hairstyles from single images.

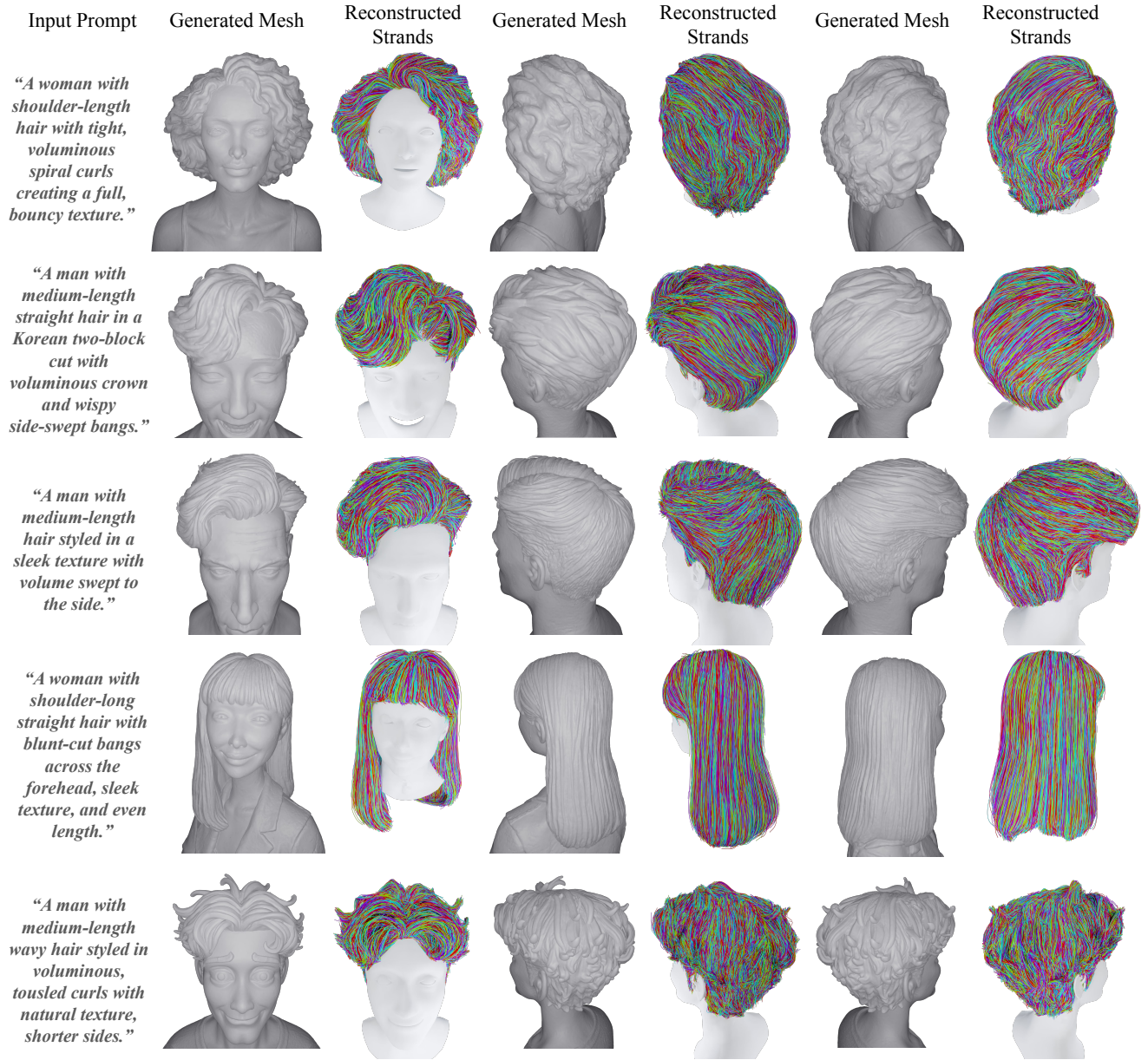


Figure 5. Reconstruction results with meshes from an off-the-shelf text-to-mesh model, demonstrating our method’s ability to generate diverse hairstyles from text descriptions.



Figure 6. Limitations of our method: reconstruction of very short and curly hairstyles. For very short hairstyles (two top rows), the hair mesh typically lacks sufficient curvature details. For curly hairstyles (two bottom rows), the crest lines algorithm fails to capture high-frequency details.

Strands-400: t-SNE of 768D Embeddings.

Captions are the LLaVA answers to:

Does the depicted hairstyle is straight or wavy or curly or kinky?

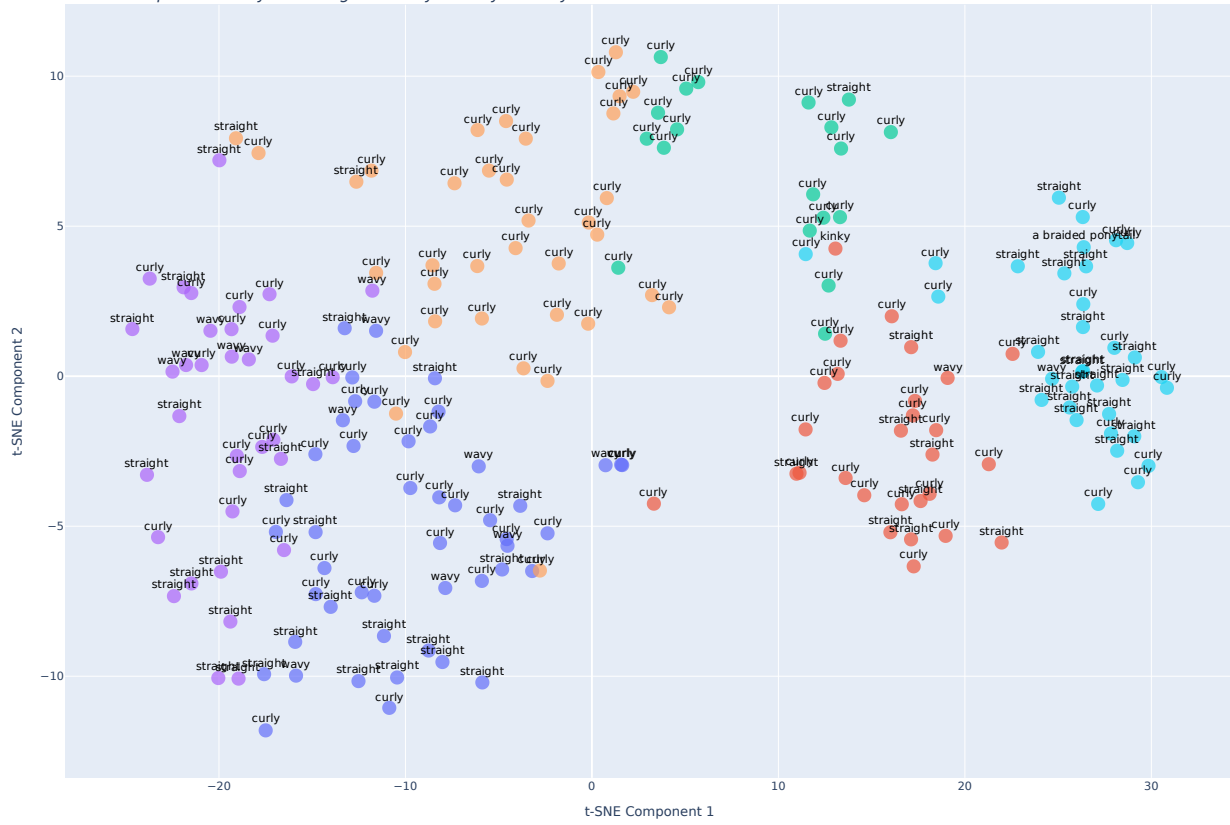


Figure 7. The distribution of hair waviness in the Strands400 dataset. The captions are collected from the answers of a VQA model (LLaVA [5]) after showing the rendered shading of the frontal and back views of the 3D scans in Strands400. The locations correspond to the t-SNE [6] over BLIP embeddings [4] of the LLaVA answers. The colors are calculated via K-Means [2].