

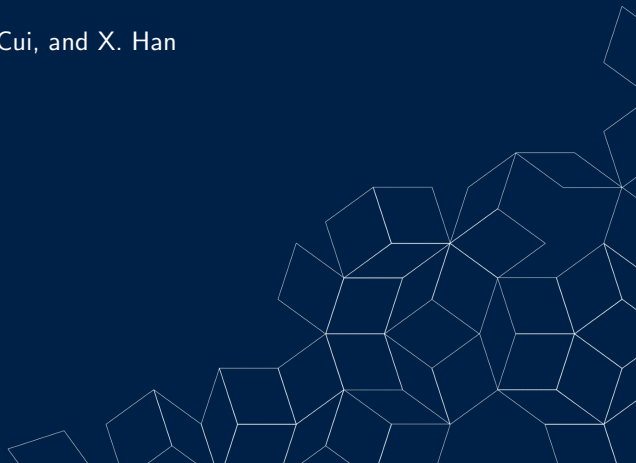
# HairStep: Transfer Synthetic to Real Using Strand and Depth Maps for Single-View 3D Hair Modeling

Y. Zheng, Z. Jin, M. Li, H. Huang, C. Ma, S. Cui, and X. Han

Computer Vision and Pattern Recognition, 2023

ALIREZA HEIDARI

December 2024



# Outline

---

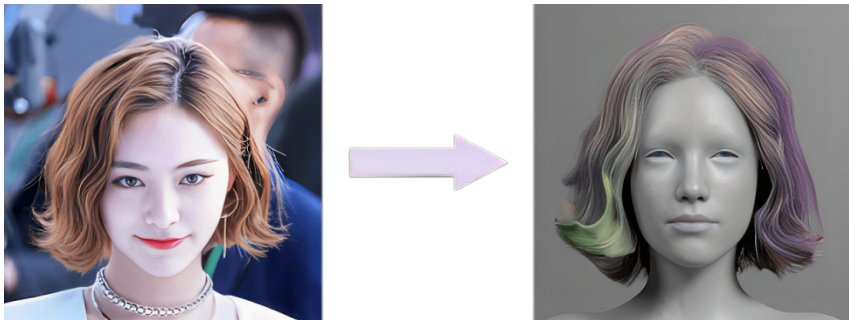
1. Objective
2. Related Works
3. Motivation
4. Contribution
5. Proposed Method
6. Conclusion

# Objective



# 3D Hair Reconstruction from a Single Image

---



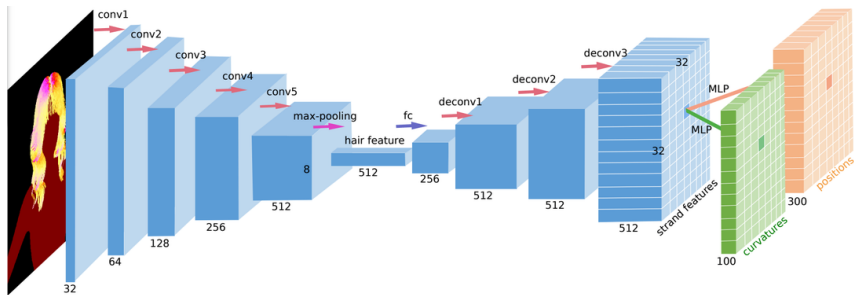
Single-view 3D Hair Modeling

# Related Works



# Related Works

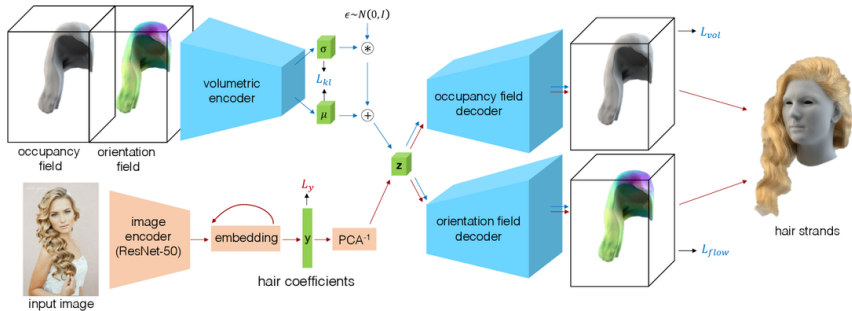
## *Single-View Hair Reconstruction using Convolutional Neural Networks*



Zhou et al. 2018[1]

# Related Works

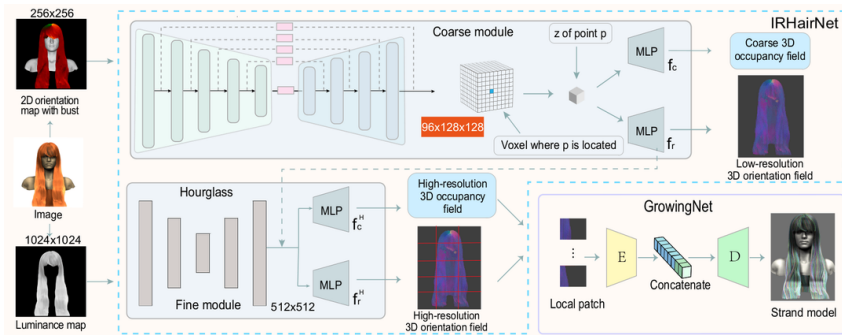
## 3D Hair Synthesis using Volumetric Variational Autoencoders



Saito et al. 2018[2]

# Related Works

## *NeuralHDHair: Automatic High-fidelity Hair Modeling from a Single Image Using Implicit Neural Representations*



Wu et al. 2022[3]



# Motivation



# Motivation

---

**Main Challenge:** Using synthetic data as a prior for real-world 3D hair modeling introduces a domain gap.

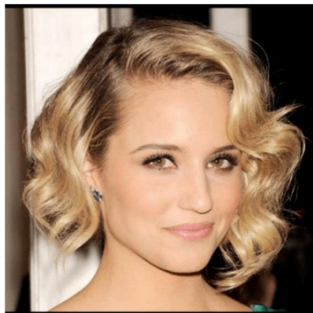
**Existing Solutions:**

- ▶ Utilize undirected 2D orientation maps as an intermediate representation between the input image and the 3D hair model.

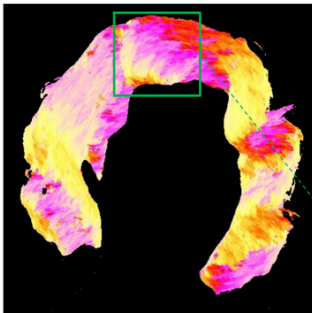
**Limitations:**

- ▶ Ambiguous directionality: Loses 3D cues from the image.
- ▶ Reliance on image filters: Adds noise and inaccuracies.

# Undirected Orientation Maps Challenges



Image



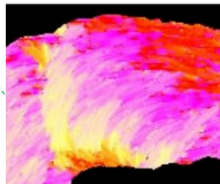
Orientation Map  
from Gabor Filters



Un-directed Lines



$[0, 180]$  Only



**Ambiguous & Noisy!**

Example of a 2D orientation map used in existing solutions.

# Contribution



# Contribution

---

- ▶ Proposed *HairStep*, a novel intermediate representation combining strand maps and depth maps for 3D hair reconstruction.
- ▶ Developed a weakly-supervised domain adaptation method for depth estimation using synthetic priors and real-world sparse annotations.
- ▶ Created *HiSa* (strand annotations) and *HiDa* (relative depth annotations) datasets for 1,250 real portrait images.
- ▶ Introduced new metrics, *HairSale* (strand alignment error) and *HairRida* (relative depth accuracy), for quantitative evaluation of 3D hair modeling.
- ▶ Achieved state-of-the-art performance in single-view 3D hair modeling.

# Proposed Method



# HairStep Representation

---

*HairStep* is defined as  $\mathbf{H} = \{\mathbf{O}, \mathbf{D}\}$  for each input image  $\mathbf{I} \in \mathbb{R}^{W \times H \times 3}$ , where:

- ▶  $\mathbf{O} \in \mathbb{R}^{W \times H \times 3}$  is the **Strand Map**.
- ▶  $\mathbf{D} \in \mathbb{R}^{W \times H \times 1}$  is the **Depth Map**.

The **Strand Map**  $\mathbf{O} \in \mathbb{R}^{W \times H \times 3}$  is defined at each pixel  $x$  as:

$$\mathbf{O}(x) = (\mathbf{M}(x), \frac{\mathbf{O}_{2D}(x)}{2} + 0.5), \quad (1)$$

- ▶  $\mathbf{M}(x) \in \{0, 1\}$  is the hair mask indicating hair regions (1) and background (0).
- ▶  $\mathbf{O}_{2D}(x) \in \mathbb{R}^2$  is the unit vector of 2D hair-growth orientation at pixel  $x$ .

The **Depth Map**  $\mathbf{D} \in \mathbb{R}^{W \times H \times 1}$  defines relative depth differences among hair strands.

- ▶ Each pixel  $\mathbf{D}(x) \in [0, 1]$ :
  - ▶  $\mathbf{D}(x) = 0$ : Farthest from the camera (background or distant strands).
  - ▶  $\mathbf{D}(x) = 1$ : Closest to the camera.

# Method Overview

---

The pipeline consists of three main components:

## 1. Strand Map Extraction and Prediction

- ▶ Extract strand maps from real images using the *HiSa* dataset.
- ▶ Train a network to predict strand maps from input images.

## 2. Domain-Adaptive Depth Estimation

- ▶ Estimate relative depth from real images using the *HiDa* dataset.
- ▶ Employ domain adaptation techniques to refine depth estimation.

## 3. 3D Hair Reconstruction

- ▶ Reconstruct 3D hair strands from the predicted strand and depth maps.
- ▶ Utilize implicit fields for volumetric hair representation.



# Conclusion



# Conclusion

---

## Contributions:

- ▶ Proposed *HairStep*, a novel intermediate representation combining strand and depth maps.
- ▶ Collected new datasets *HiSa* and *HiDa* with annotated real images.
- ▶ Proposed two quantitative metrics: *HairSale* and *HairRida*.
- ▶ Achieved state-of-the-art performance in single-view 3D hair modeling.

## Limitations:

- ▶ Manual annotation is time-consuming, limiting scalability.
- ▶ Generalization to unseen hairstyles and diverse real-world conditions needs further investigation.

# References I

---

- [1] Y. Zhou, L. Hu, J. Xing, *et al.*, “Single-view hair reconstruction using convolutional neural networks,” in *European Conference on Computer Vision*, 2018. [Online]. Available: <https://api.semanticscholar.org/CorpusID:49666680>.
- [2] S. Saito, L. Hu, C. Ma, H. Ibayashi, L. Luo, and H. Li, “3d hair synthesis using volumetric variational autoencoders,” *ACM Transactions on Graphics (TOG)*, vol. 37, pp. 1–12, 2018. [Online]. Available: <https://api.semanticscholar.org/CorpusID:54101192>.
- [3] K. Wu, Y. Ye, L. Yang, H. Fu, K. Zhou, and Y. Zheng, “Neuralhdhair: Automatic high-fidelity hair modeling from a single image using implicit neural representations,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 1526–1535.
- [4] Y. Zheng, Z. Jin, M. Li, *et al.*, “Hairstep: Transfer synthetic to real using strand and depth maps for single-view 3d hair modeling,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 12 726–12 735.