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

CVPR2023 Highlight

Yujian Zheng, Zirong Jin, Moran Li, Haibin Huang,

Chongyang Ma, Shuguang Cui and Xiaoguang Han*

千家智联网络研究院

香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

kwai

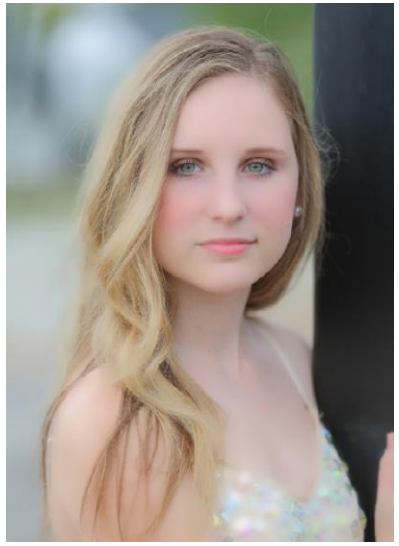
Outline

- Background
- HairStep
- Experiment
- Conclusion

Outline

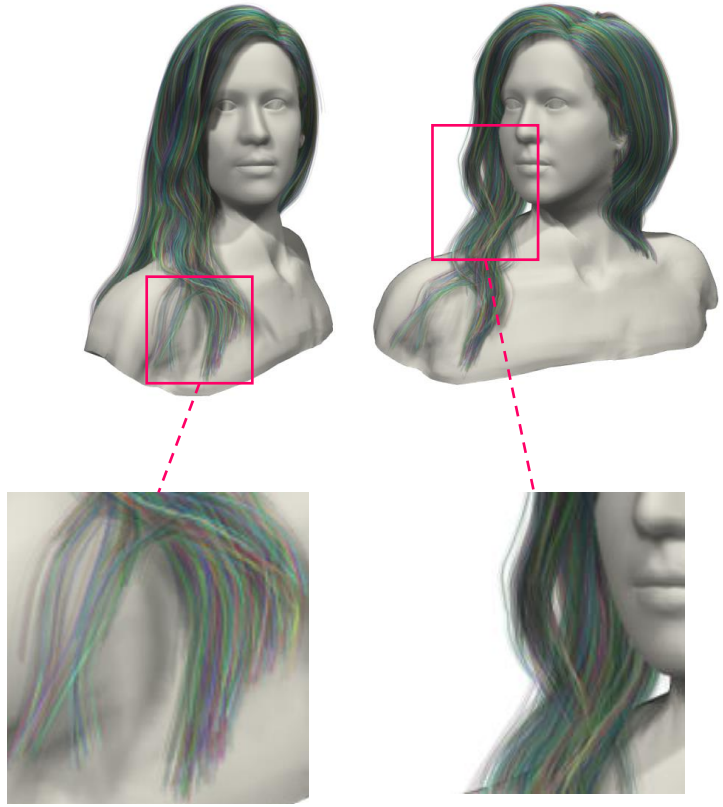
- Background
  - Single-view 3D Hair Modeling
  - Previous Approaches
  - Domain Gap
- HairStep
- Experiment
- Conclusion

Single-view 3D Hair Modeling



Input Image

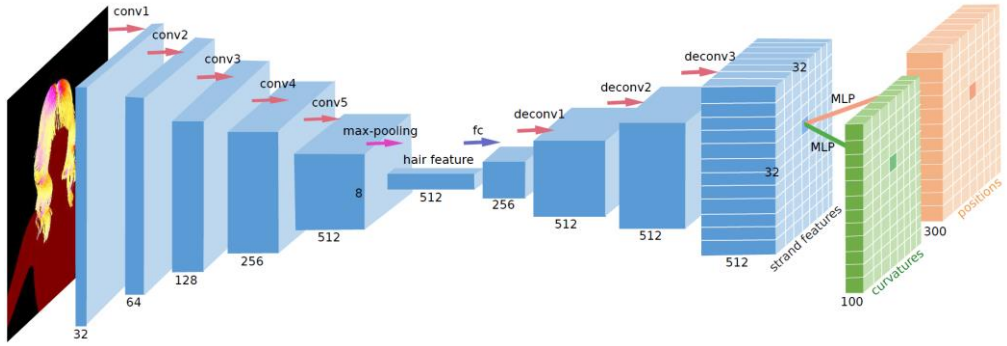
3D Hair Reconstruction



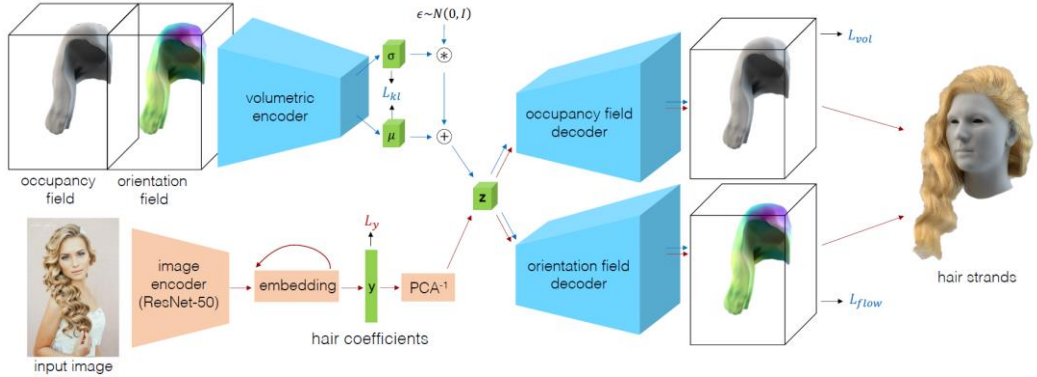
Strand-level 3D Hair

Chai et al. 2016

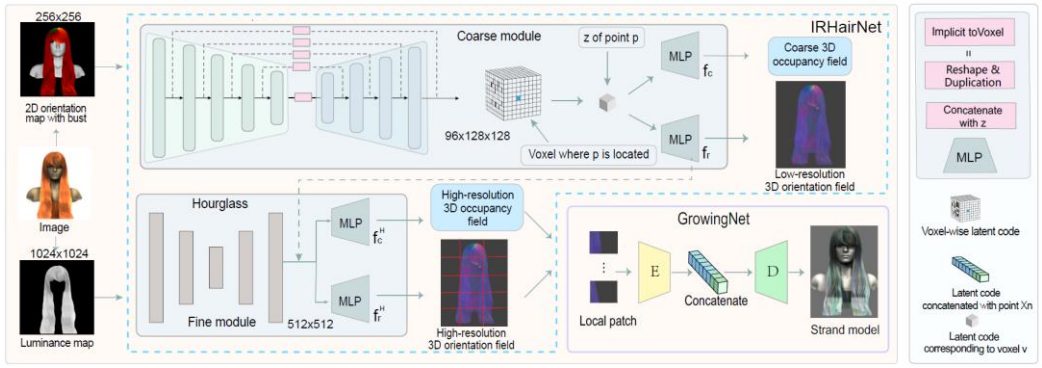
Previous Approaches



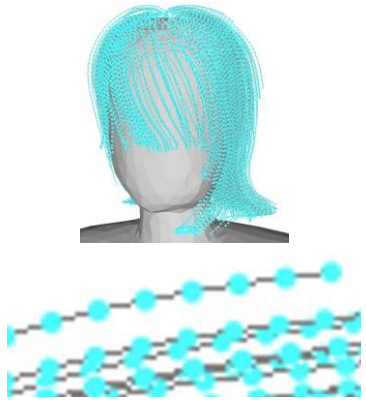
Zhou et al. 2018



Saito et al. 2018

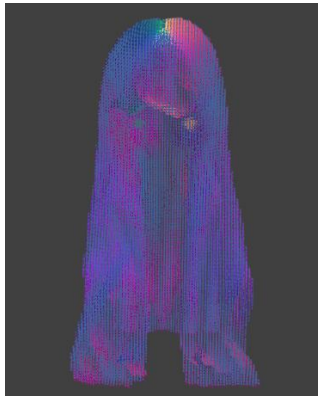


Wu et al. 2022



Explicit point sequences

Volumetric orientation field



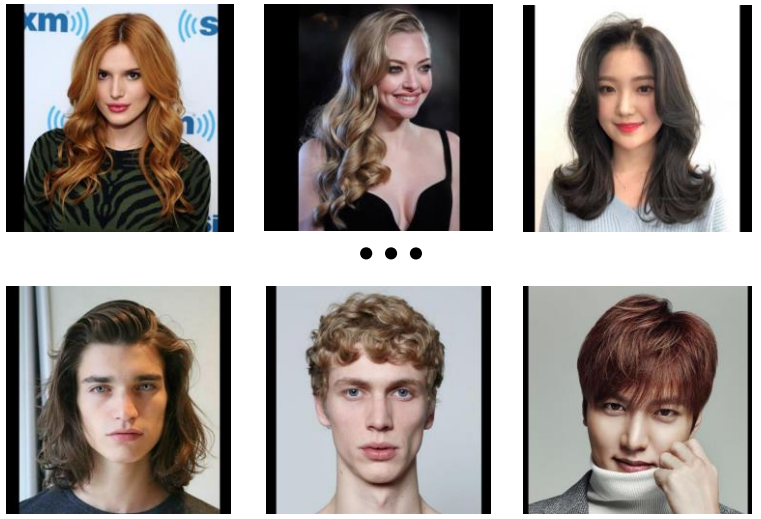
Implicit orientation field

Domain Gap

Trained on Synthetic Data

Domain Gap

Failure Cases



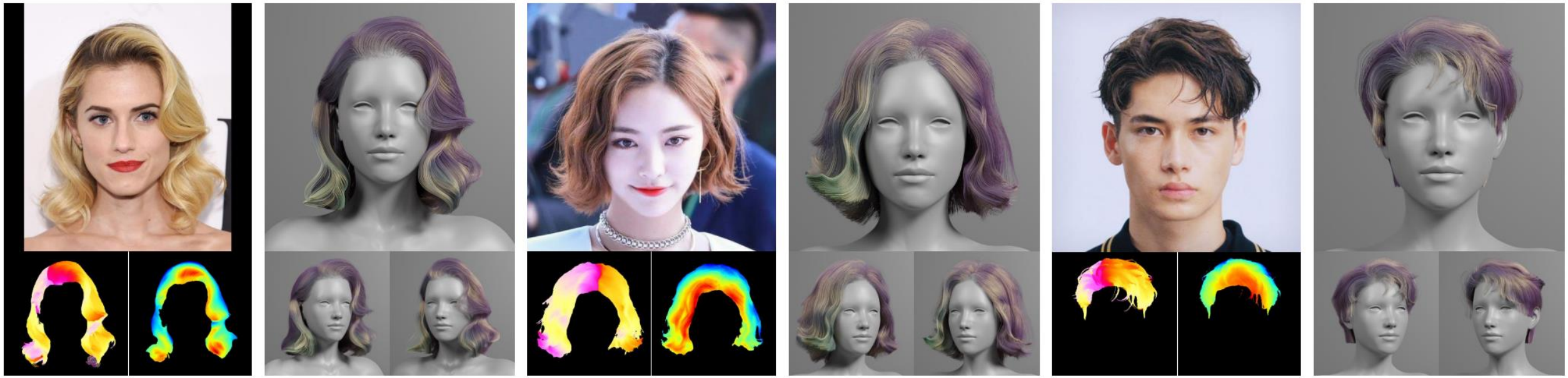
Tested on Real Images

Outline

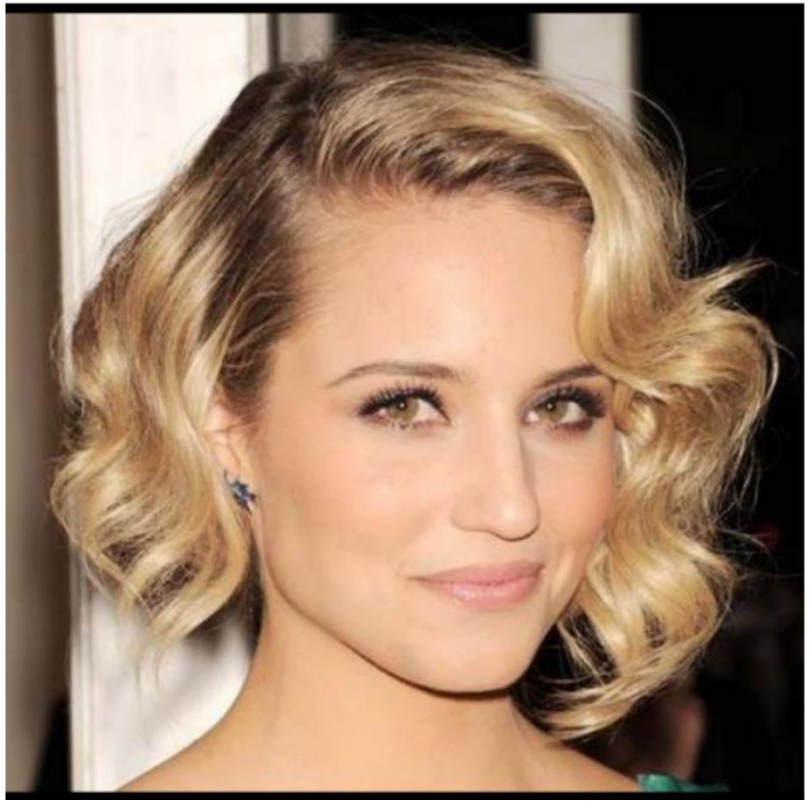
- Background
- HairStep
  - Motivation & Overview
  - Extraction of HairStep
  - Single-view Hair Reconstruction
- Experiment
- Conclusion

HairStep

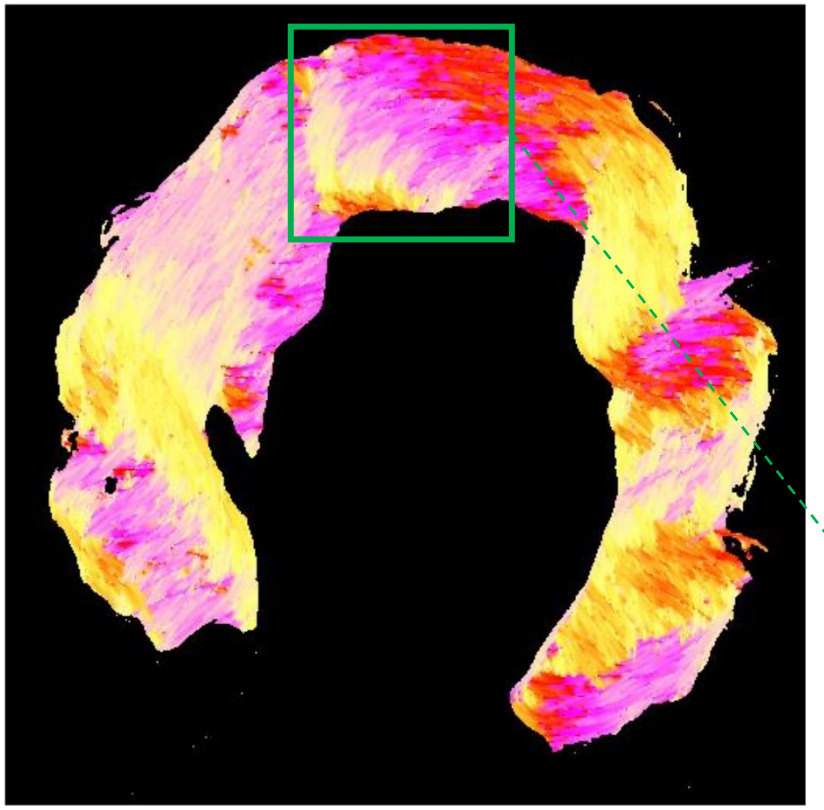
HairStep: Transfer Synthetic to Real Using Strand and Depth Maps for Single-View 3D Hair Modeling



Motivation

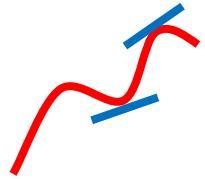


Image

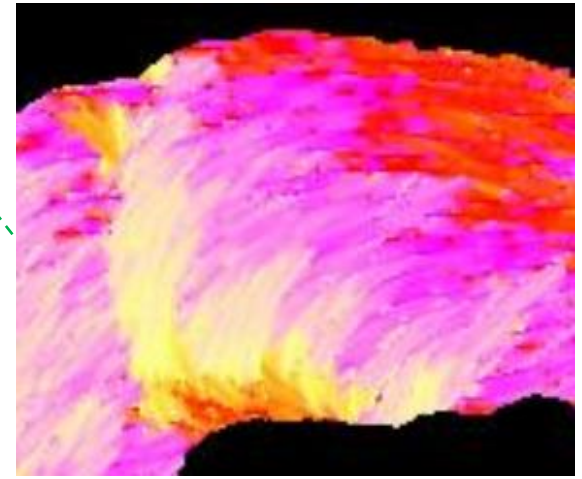


Orientation Map
from Gabor Filters

Un-directed Lines

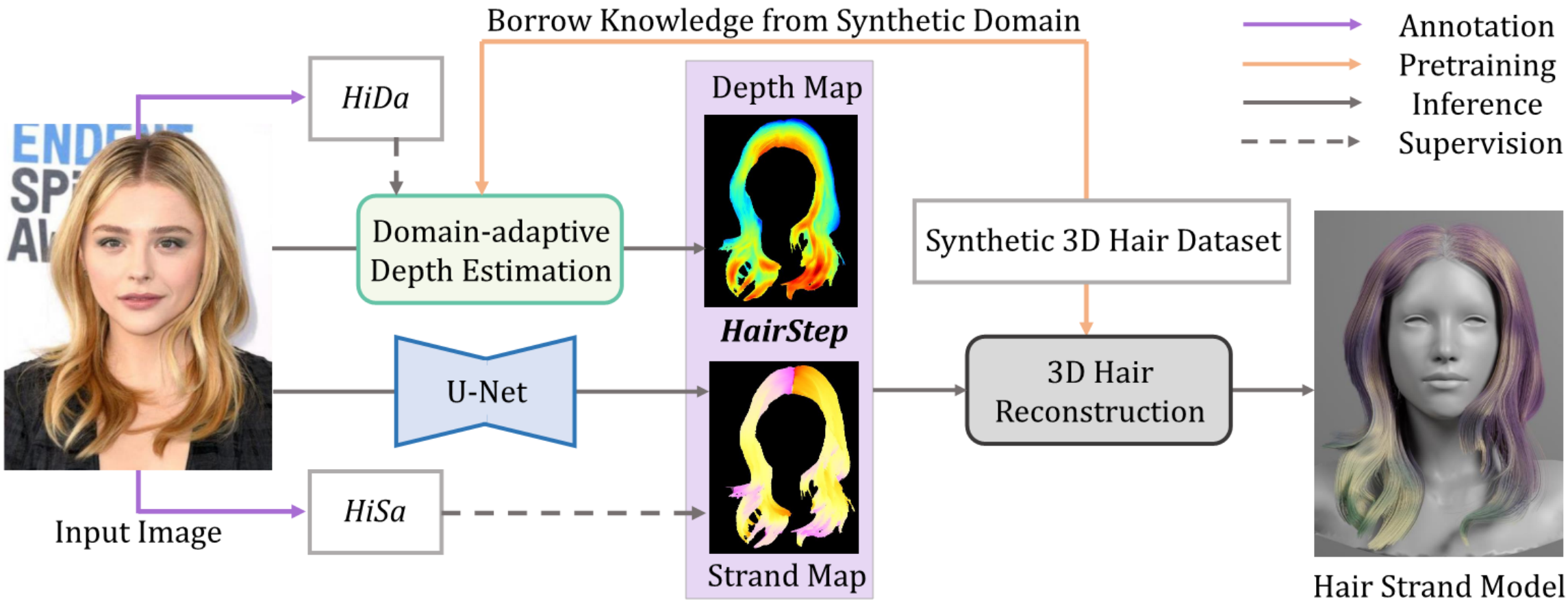


[0, 180] Only



Ambiguous & Noisy!

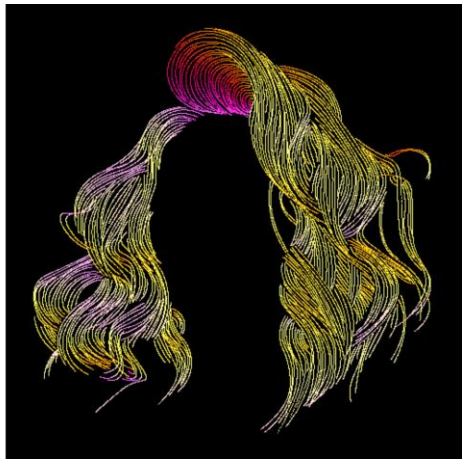
Overview


Borrow Knowledge from Synthetic Domain
Annotation
Pretraining
Inference
Supervision
HiDa
Depth Map
Domain-adaptive Depth Estimation
HairStep
U-Net
Synthetic 3D Hair Dataset
3D Hair Reconstruction
HiSa
Input Image
Strand Map
Hair Strand Model

HiSa Dataset & Strand Map

Image

Strokes

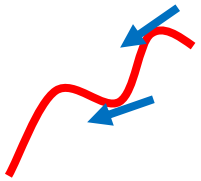


Stroke Map

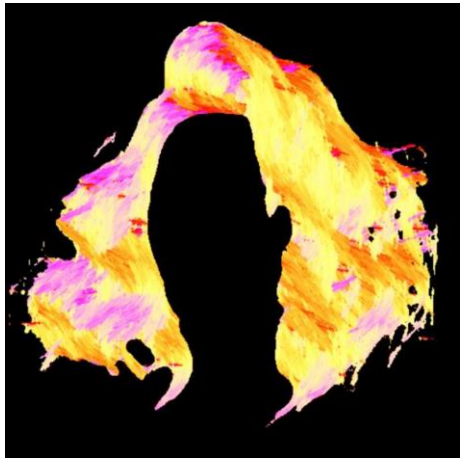
300 vector curves
from hair roots to hair ends

Strand Map

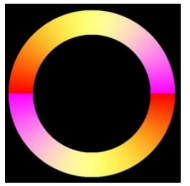
[0, 360]



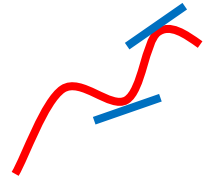
Directed Lines



Orientation Map



[0, 180]



Un-directed Lines

Extraction of Strand Map

Image

U-Net

Predicted Strand Map

Image Loss

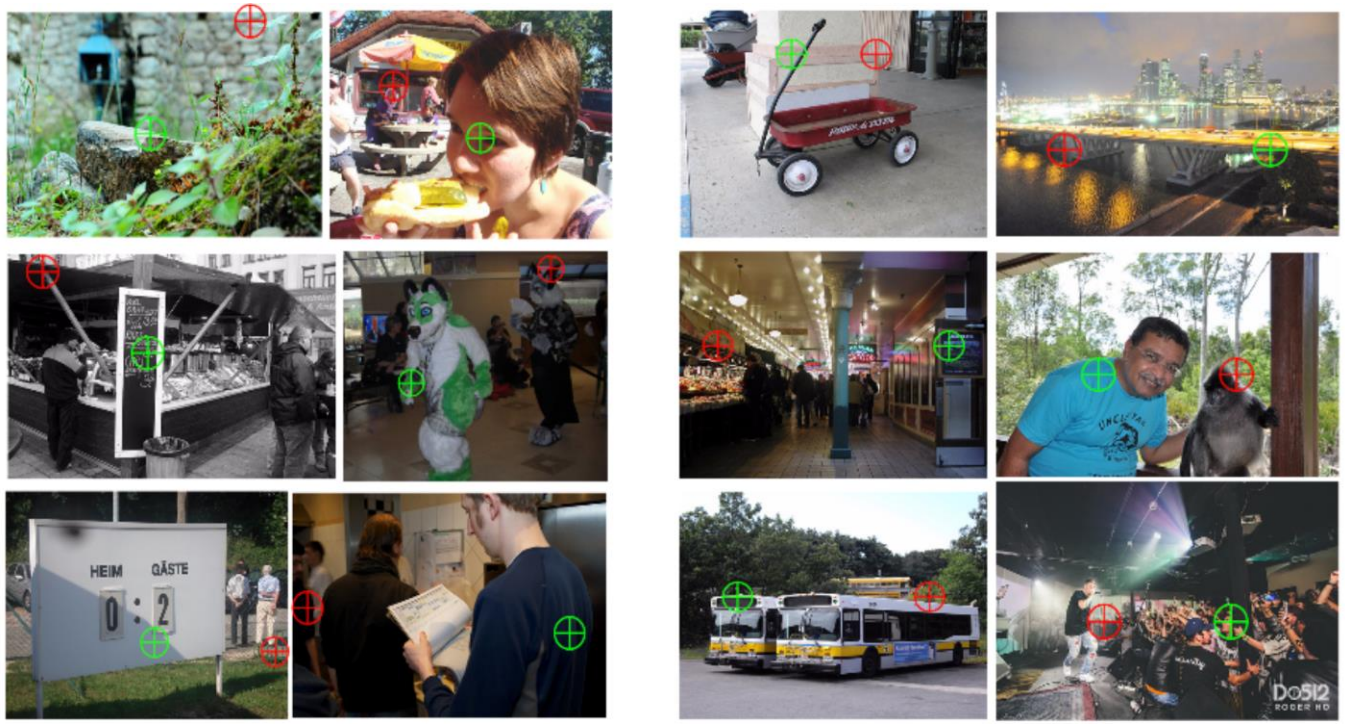
GT Strand Map

Pixel-wise L_1 Loss
+
Perceptual Loss

$$L_{strand} = \frac{1}{C \cdot \sum \mathbf{M}} \left\| \hat{\mathbf{O}} - \mathbf{O} \right\|_1 + \alpha \cdot \frac{1}{W_j H_j C_j} \left\| \phi_j(\hat{\mathbf{O}}) - \phi_j(\mathbf{O}) \right\|_2^2$$

HiDa Dataset & Relative Depth

How to **annotate depth** of hair?



Depth-in-the-wild
Relative depth

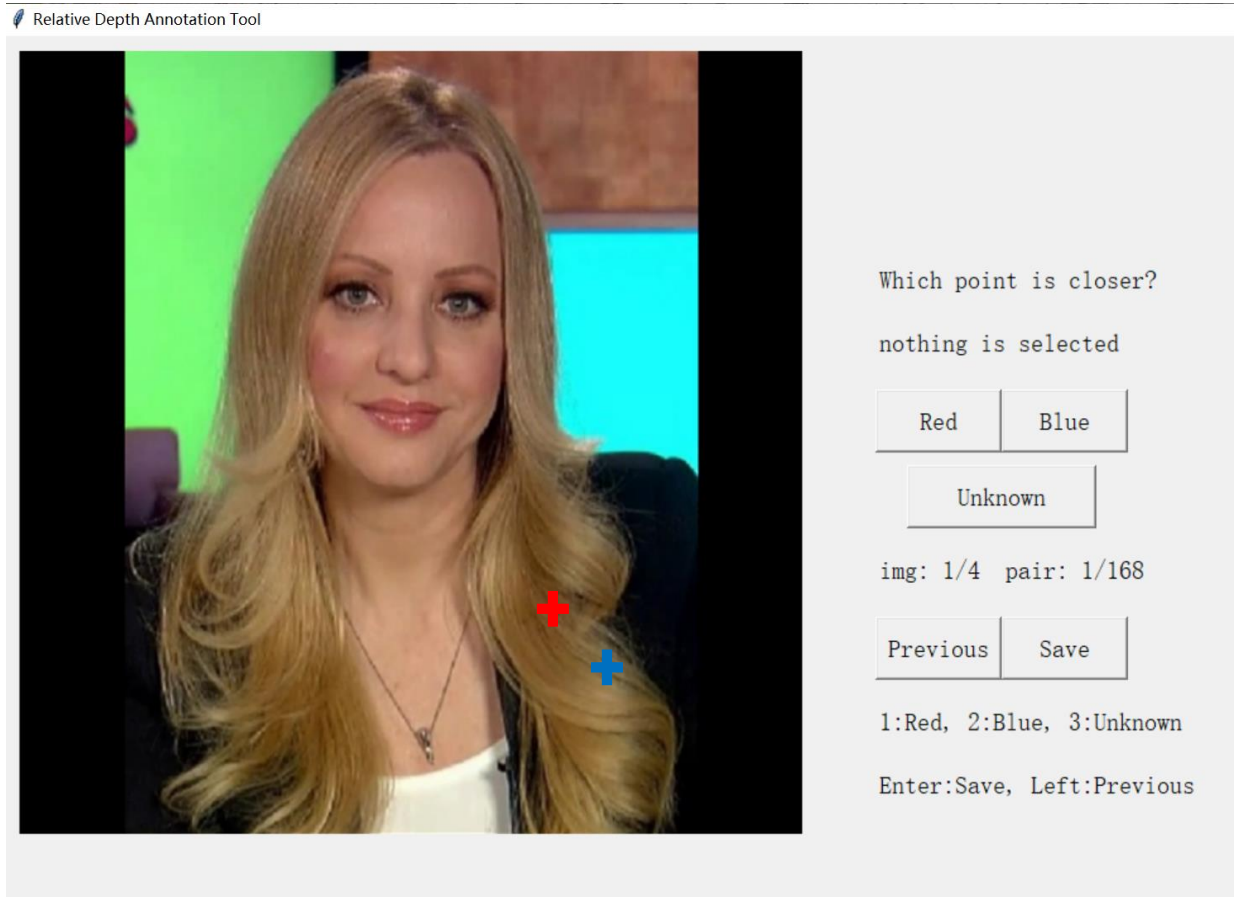
HiDa Dataset & Relative Depth

Sample Pairs in Neighbors of Super-pixels

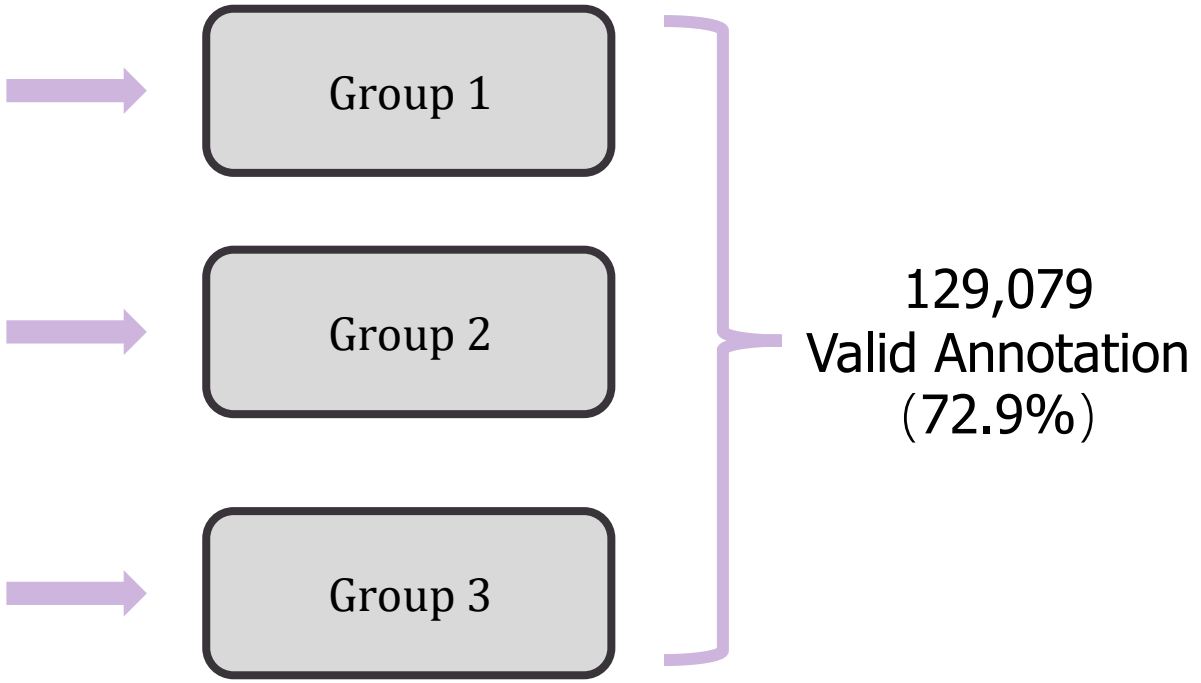
1250 Images

177,074 Pixel Pairs

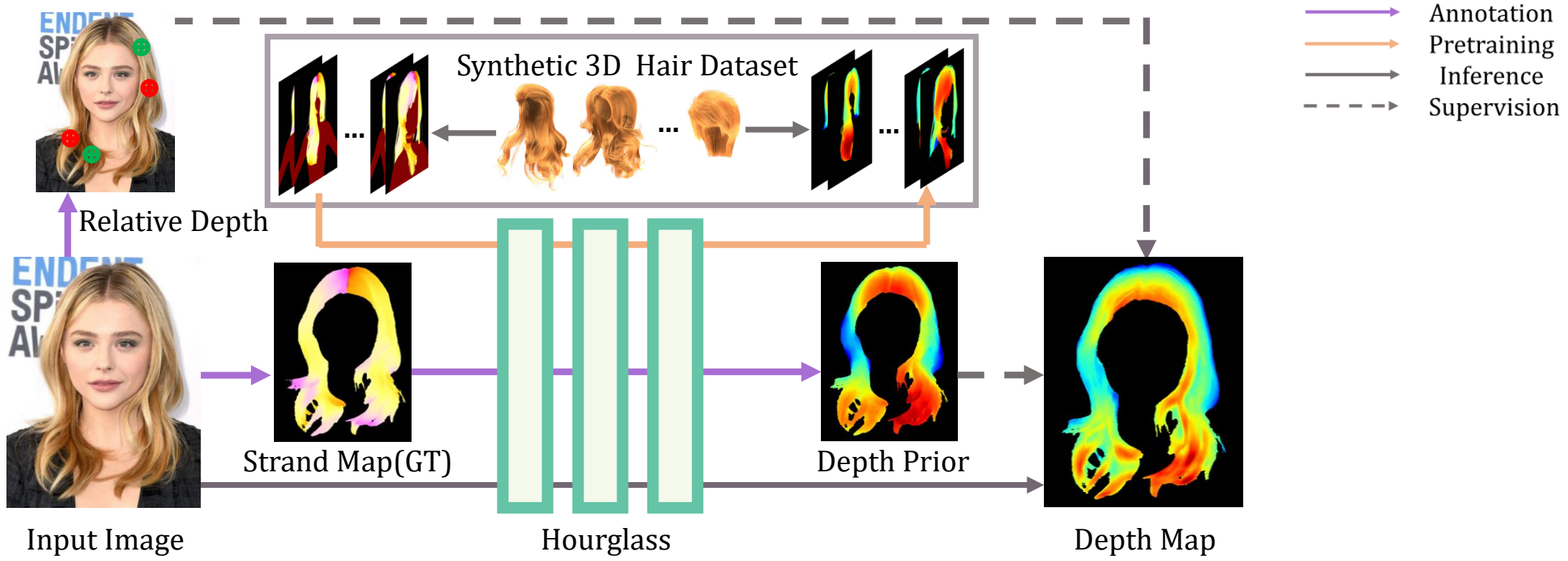
HiDa Dataset & Relative Depth



Annotation Tool



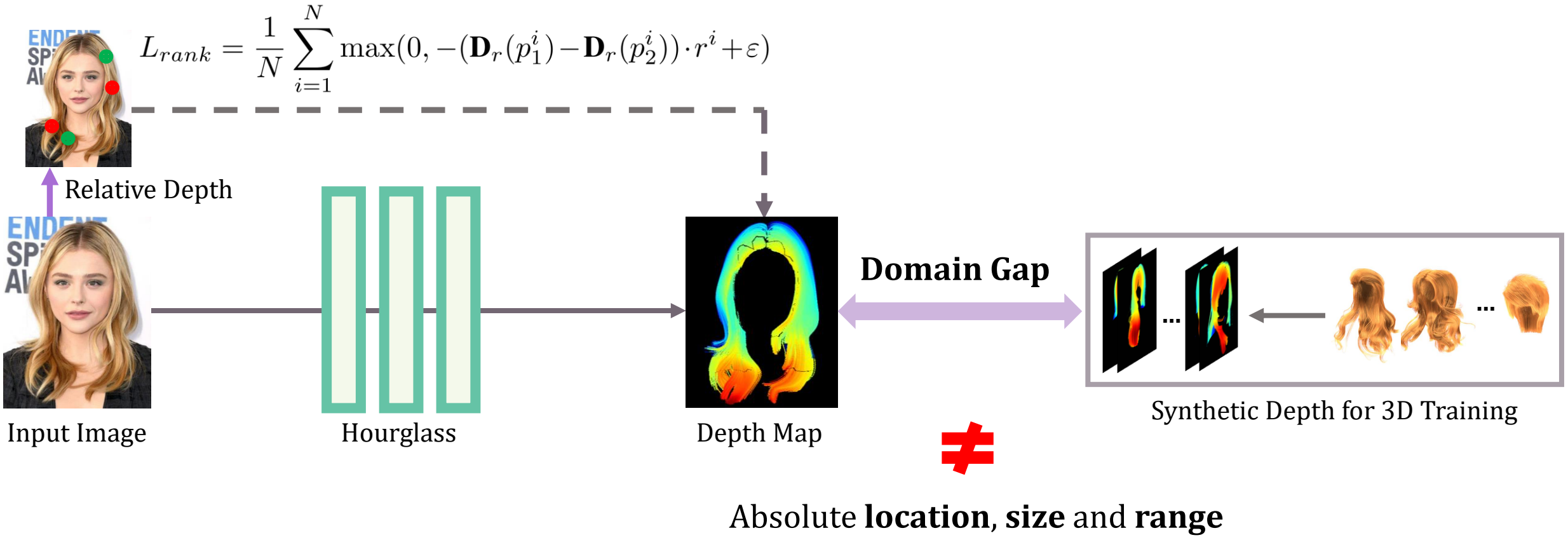
Domain-Adaptive Depth Estimation



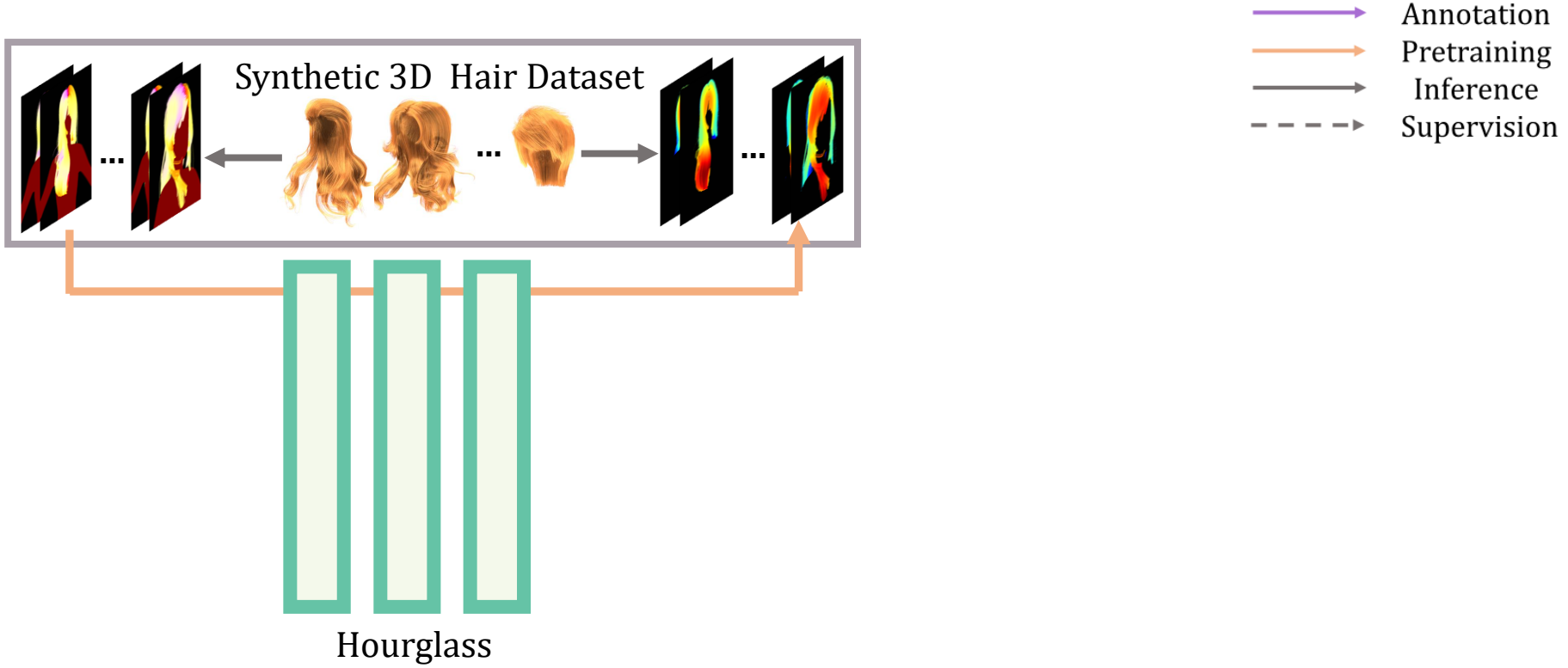
$$L_{rank} = \frac{1}{N}\sum_{i=1}^{N}\max(0, -(\mathbf{D}_r(p_1^i) - \mathbf{D}_r(p_2^i))\cdot r^i + \varepsilon)$$

$$L_{depth} = \beta \cdot \left\|\mathbf{D}_r - \bar{\mathbf{D}}\right\|_1 + L_{rank}$$

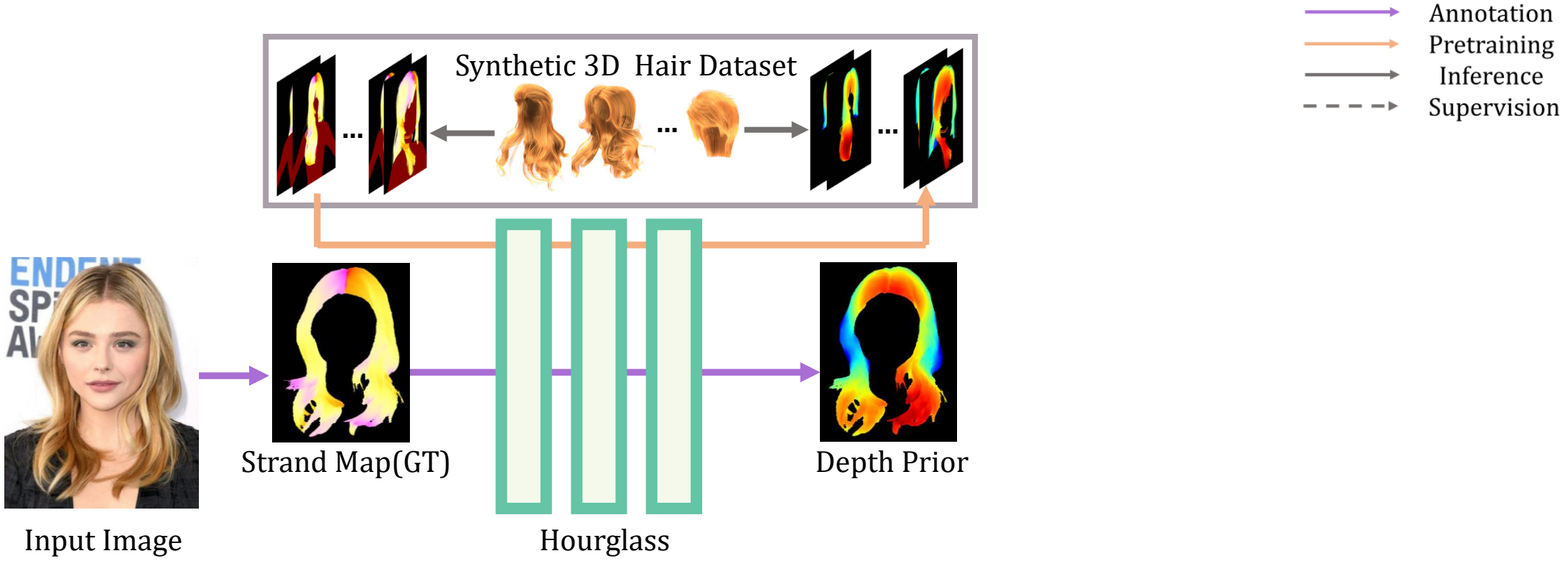
Domain-Adaptive Depth Estimation



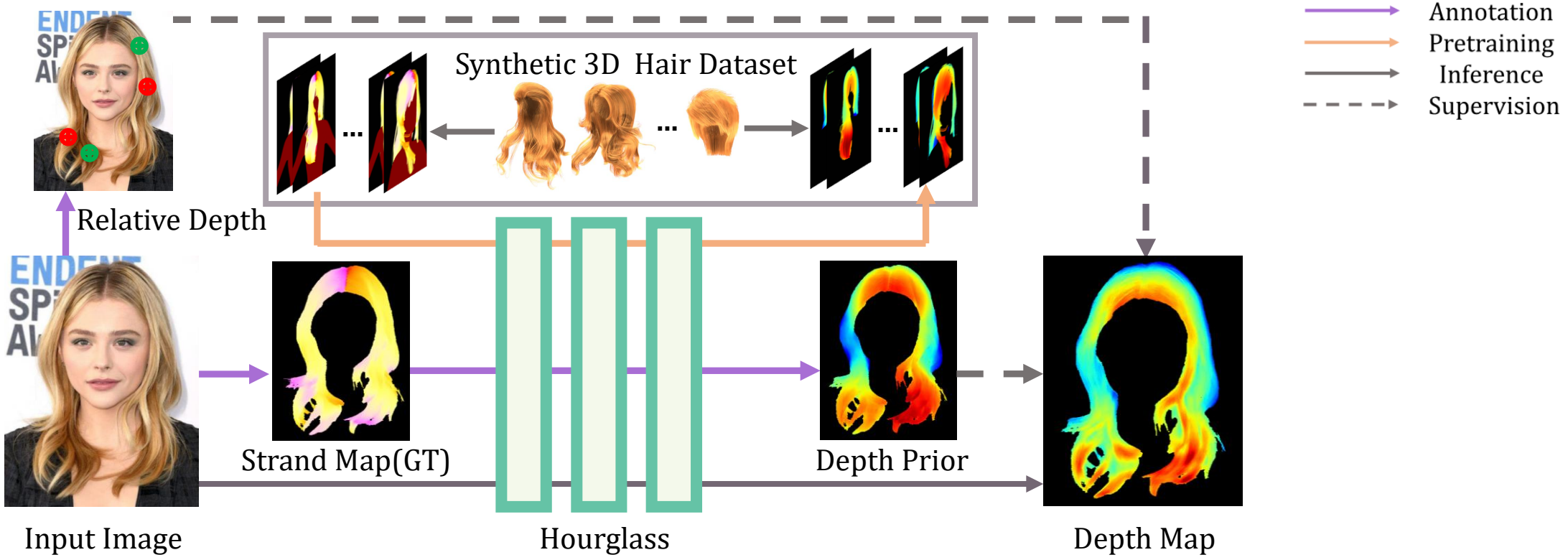
Domain-Adaptive Depth Estimation



Domain-Adaptive Depth Estimation



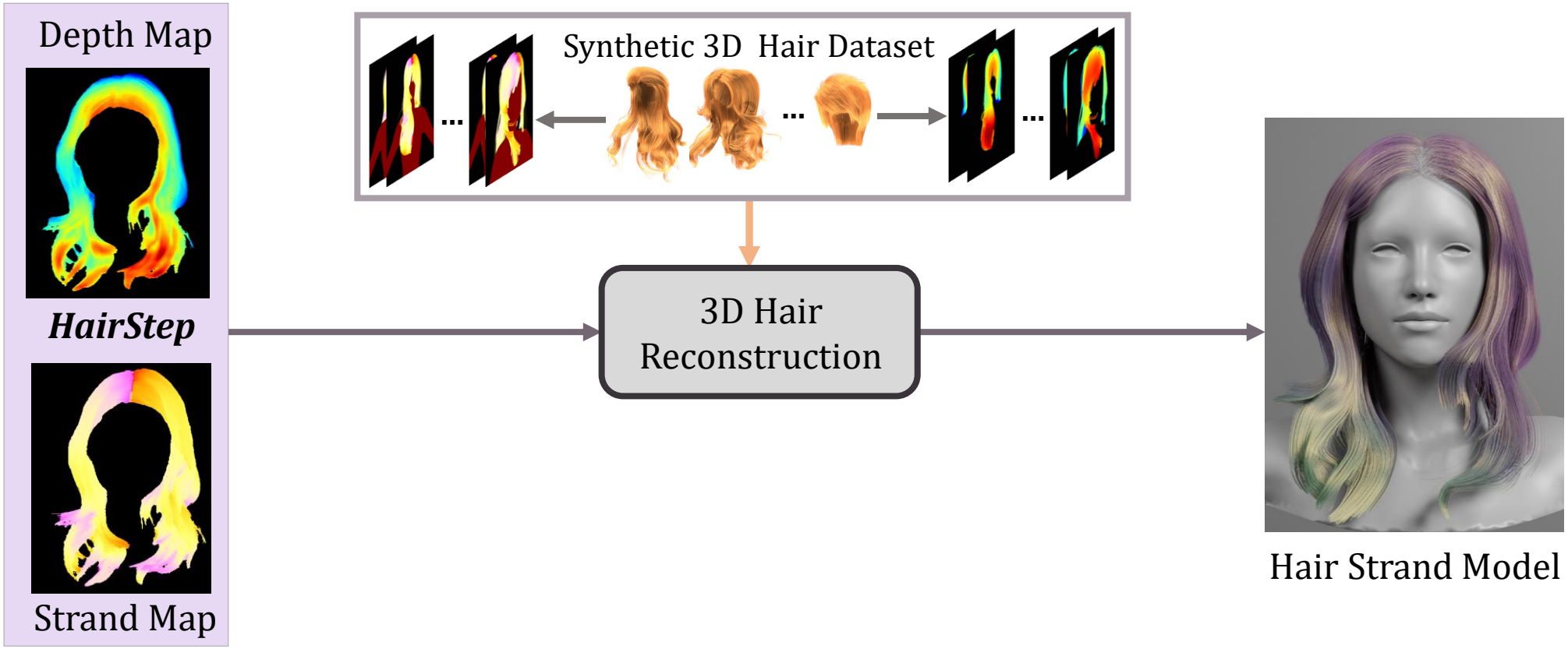
Domain-Adaptive Depth Estimation



$$L_{rank} = \frac{1}{N}\sum_{i=1}^{N}\max(0, -(\mathbf{D}_r(p_1^i) - \mathbf{D}_r(p_2^i))\cdot r^i + \varepsilon)$$

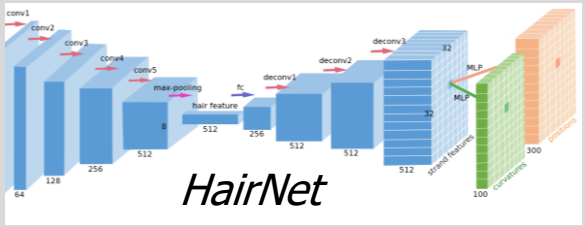
$$L_{depth} = \beta \cdot \left\|\mathbf{D}_r - \bar{\mathbf{D}}\right\|_1 + L_{rank}$$

Single-View 3D Hair Modeling

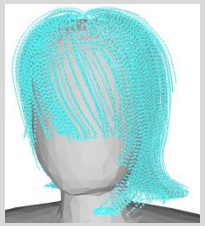


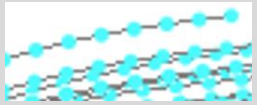
Single-View 3D Hair Modeling

HairStep

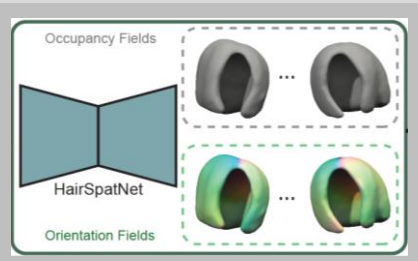


HairNet

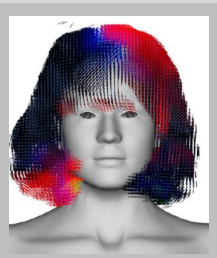




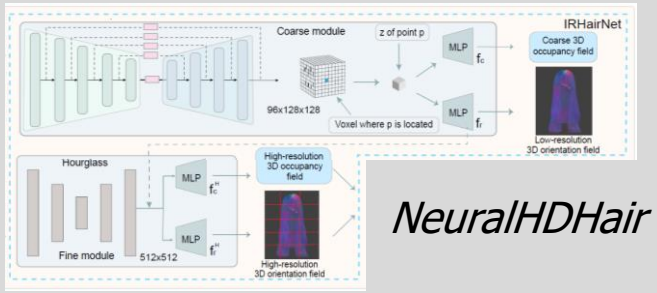
Explicit
point sequences



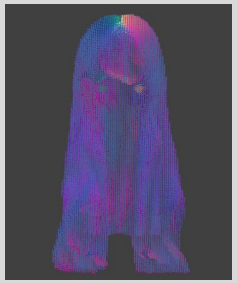
DynamicHair (Single-view)



Volumetric orientation field



NeuralHDHair



Implicit orientation field

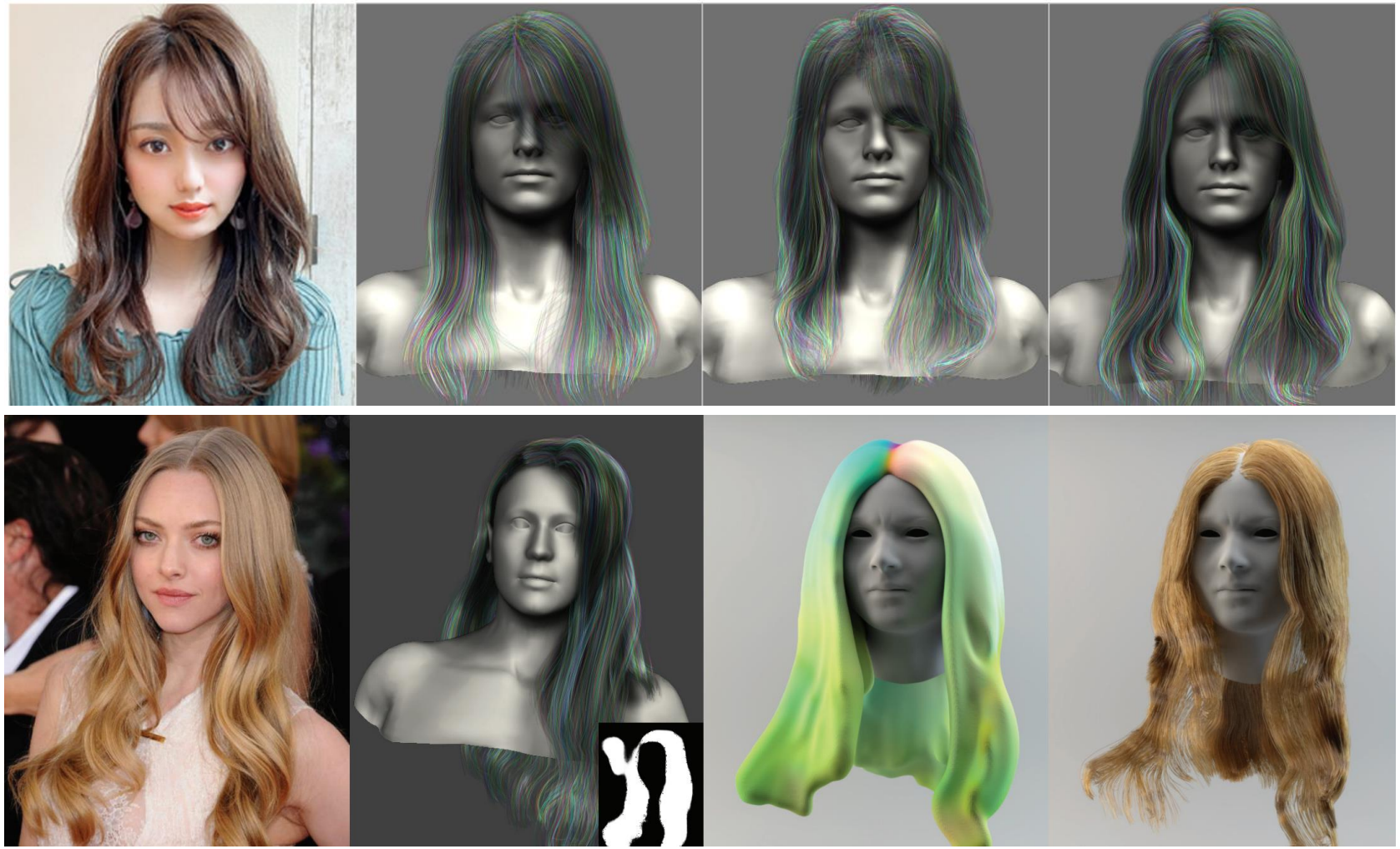
3D Hair Model

3D Hair Reconstruction

Outline

- Background
- HairStep
- **Experiment**
  - Fair Metrics
  - Evaluation on HairStep Extraction
  - Comparisons of 3D Reconstruction
  - Ablation Study
- Conclusion

Fair Metrics



Wu et al. 2022

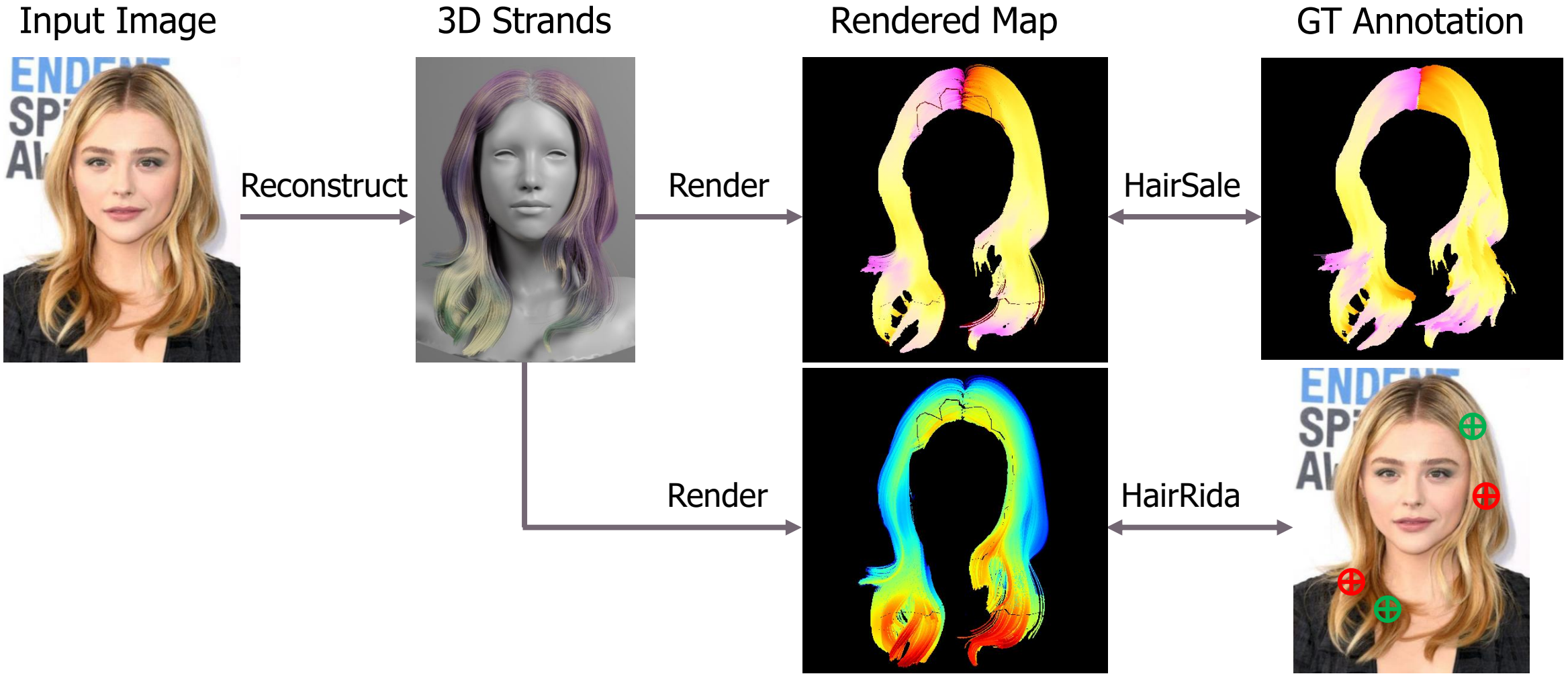
Only Comparisons of

the **Visual Quality**

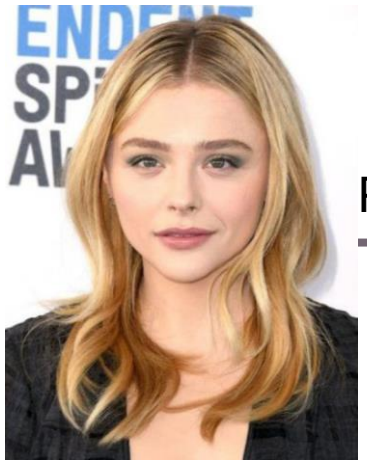
in Existing Methods

Saito et al. 2018

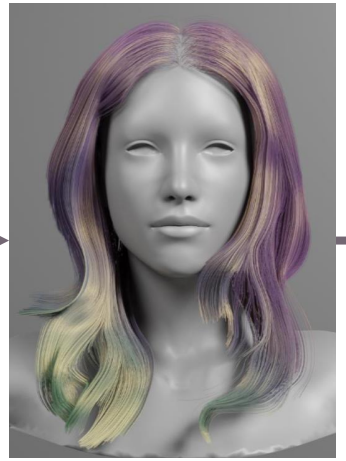
Fair Metrics



HairSale

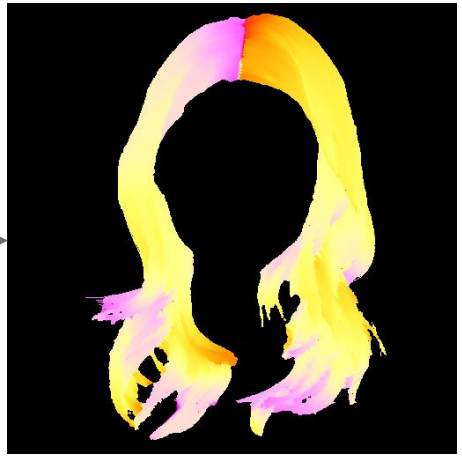


Reconstruct



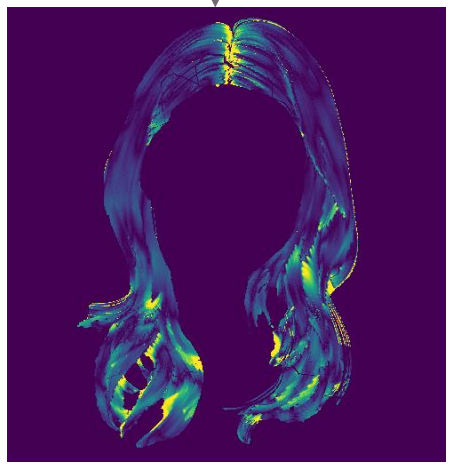
Render

HairSale

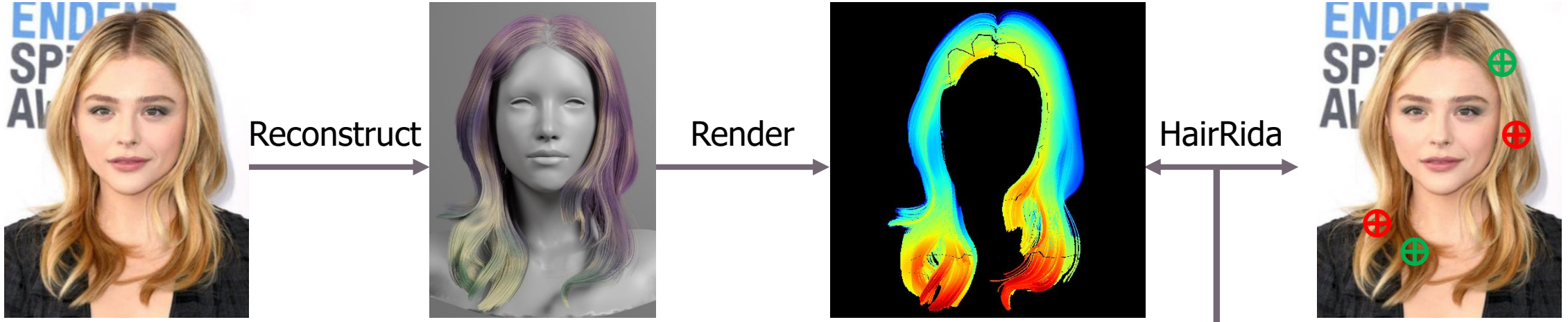


$$HairSale = \frac{1}{K} \sum_{\mathbf{x}^i \in U} \arccos(\mathcal{V}(\mathbf{O}_r(\mathbf{x}^i)) \cdot \mathcal{V}(\mathbf{O}_{gt}(\mathbf{x}^i)))$$

Mean Angle Error of Growing Direction



HairRida



$$HairRida = \frac{1}{Q}\sum_{i=1}^{Q}\max(0, r^i \cdot \mathrm{sign}(\mathbf{D}_r(p_1^i) - \mathbf{D}_r(p_2^i)))$$

The Relative Depth Accuracy

Evaluation on the Extraction of Strand Map

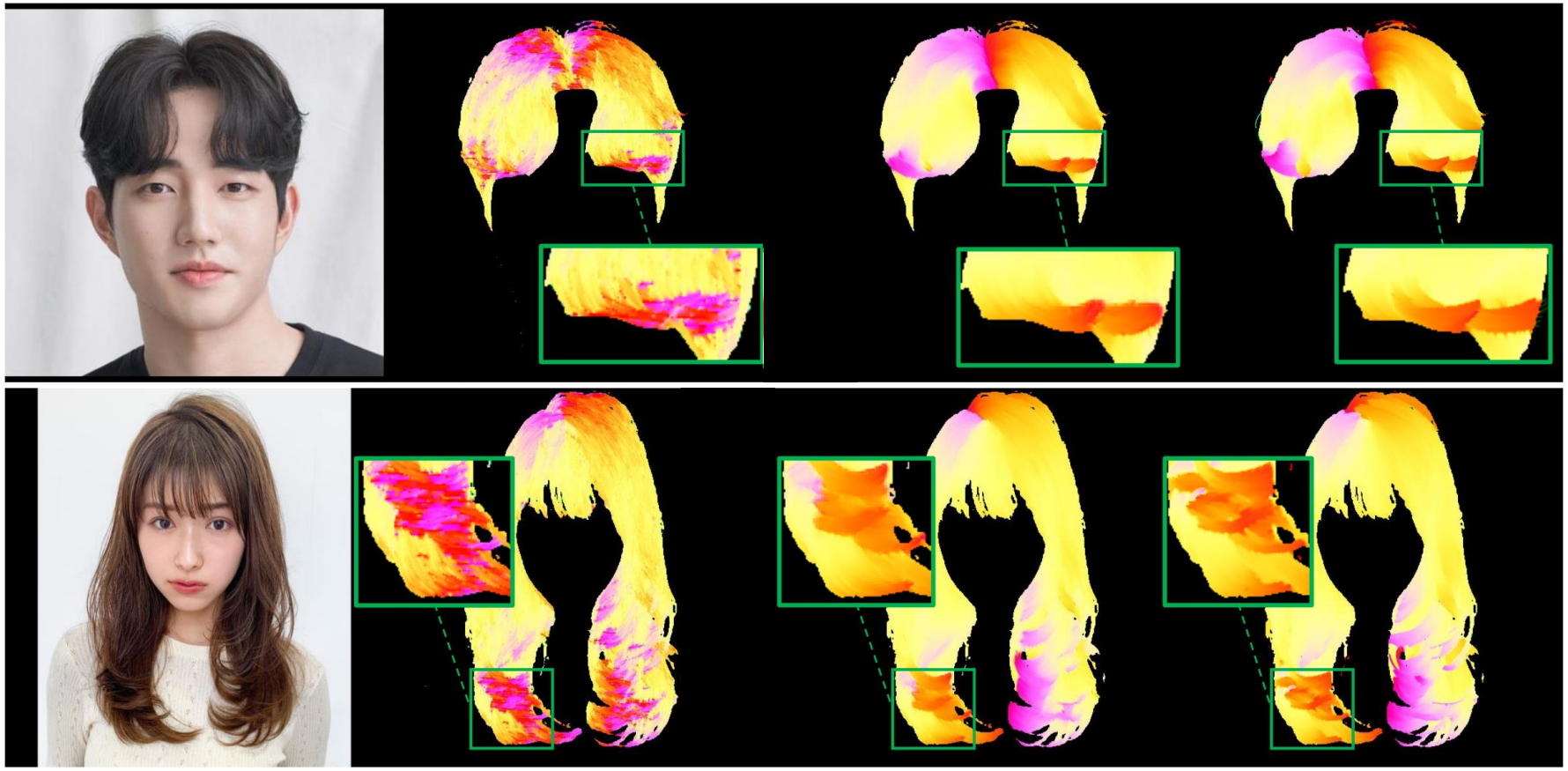
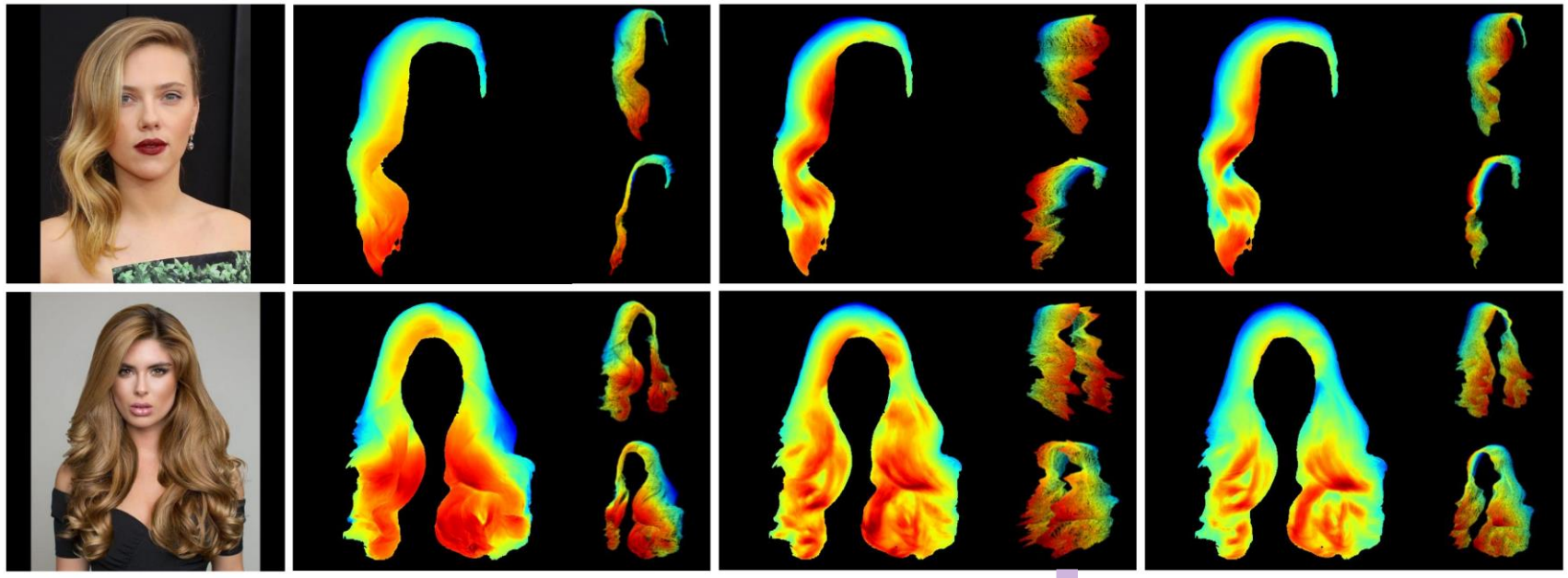


Image | Orientation Map (Gabor Filter) | Strand Map (Predicted) | Strand Map (GT)

	HairSale↑ (Undirected)
Gabor Filters	18.4
Ours	**14.2**

22.8% Better

Evaluation on Depth Estimation



Input Image $Depth_{pseudo}$ $Depth_{weak}$ $Depth_{DA}$

Flat Geometry

Serration-like Artifacts

Balance the **local details of depth** and the **similarity of global shape** to the synthetic prior

	HairRida↑	L1↓
$Depth_{pseudo}$	80.47%	/
$Depth_{weak}$	85.17%	0.2470
$Depth_{DA}$	**85.20%**	**0.1768**

Visual Comparisons with Existing Methods

Visual Comparisons with Existing Methods

Quantitative Comparisons on Representation

On Synthetic Data

Method	Orien. err. ↓	Occ. acc. ↑
NeuralHDHair* (Orientation map)	0.1324	82.59%
NeuralHDHair* (Strand map)	0.0722 (-41.7%)	84.18%
NeuralHDHair* (HairStep)	0.0658 (-50.3%)	86.77%
DynamicHair (Orientation map)	0.1352	78.19%
DynamicHair (Strand map)	0.1185 (-12.4%)	79.62%
DynamicHair (HairStep)	0.1174 (-13.2%)	79.78%
HairNet (Orientation map)	0.02349	/
HairNet (Strand map)	0.02206 (-6.1%)	/
HairNet (HairStep)	0.02184 (-7.0%)	/

On Real Data

Method	IoU ↑	*HairSale* ↓	*HairRida* ↑
NeuralHDHair* (Orientation map)	77.56%	19.6	70.67%
NeuralHDHair* (Strand map)	77.6%	**16** (-18.4%)	72.37%
NeuralHDHair* (HairStep)	77.22%	16.36 (-16.5%)	**76.79%**
DynamicHair (Orientation map)	56.39%	32.66	74.08%
DynamicHair (Strand map)	59.51%	26.53 (-18.8%)	73.42%
DynamicHair (HairStep)	59.14%	27.51 (-15.8%)	73.58%
HairNet (Orientation map)	57.15%	31.97	75.65%
HairNet (Strand map)	57.48%	28.6 (-10.5%)	74.81%
HairNet (HairStep)	57.01%	27.68 (-13.4%)	74.97%

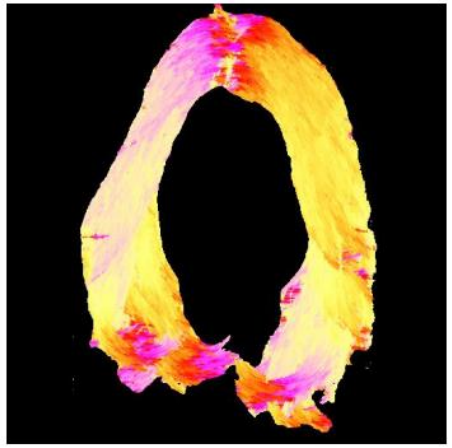
HairStep **Benefits all Methods**

On Both **Synthetic & Real** Data

Visual Comparisons on Representation (NeuralHDHair*)

Image

From Orientation Map

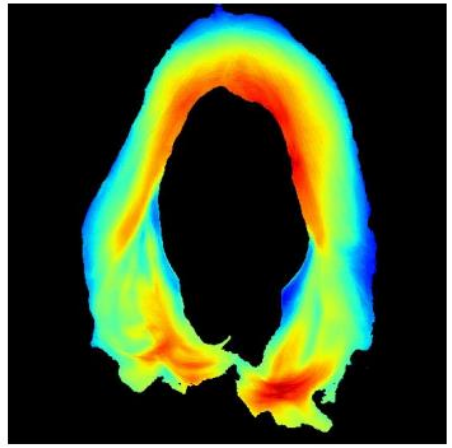


Orientation Map

From Strand Map

Strand Map

From HairStep

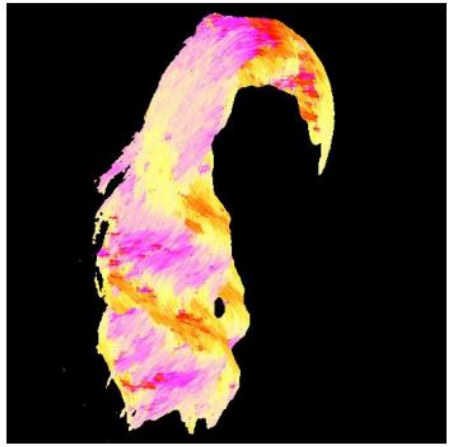


Depth Map

Visual Comparisons on Representation (NeuralHDHair*)

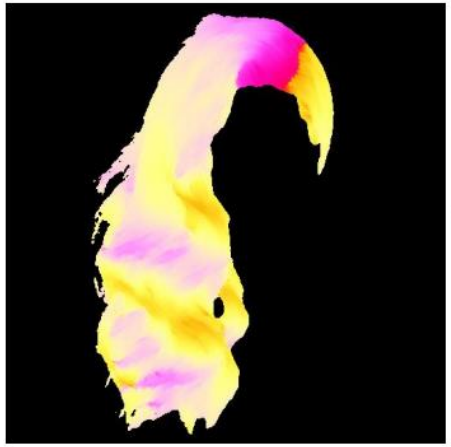
Image

From Orientation Map



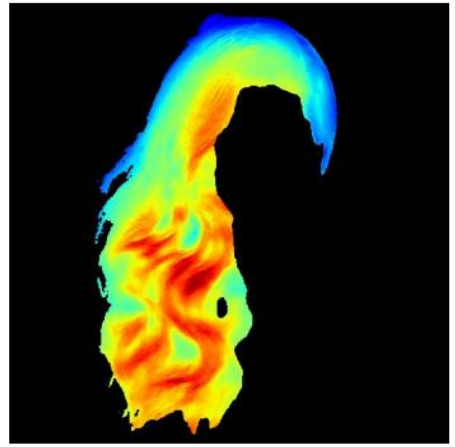
Orientation Map

From Strand Map



Strand Map

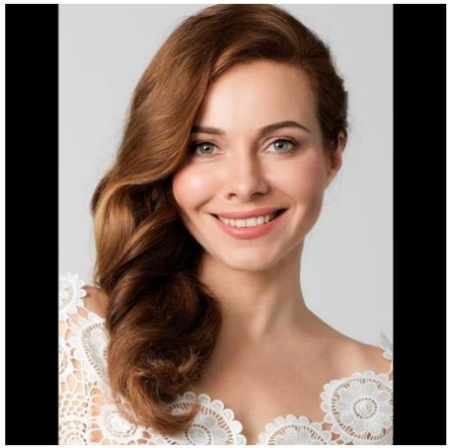
From HairStep



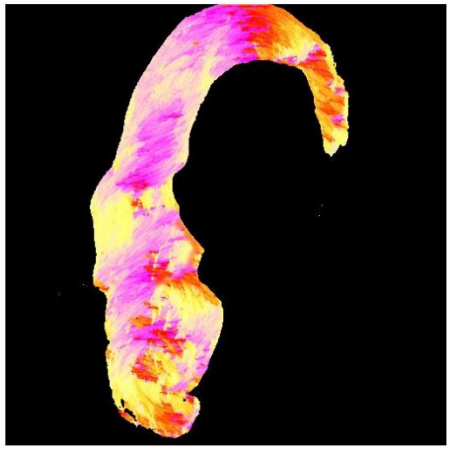
Depth Map

Visual Comparisons on Representation (NeuralHDHair*)

Image

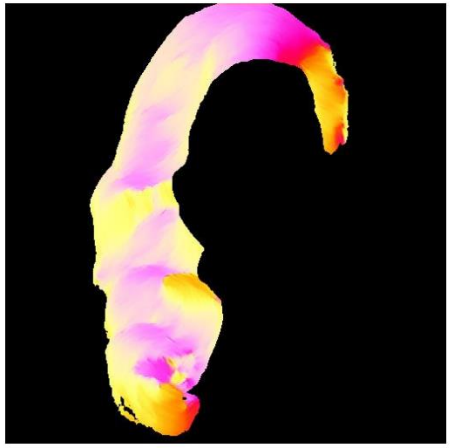


From Orientation Map



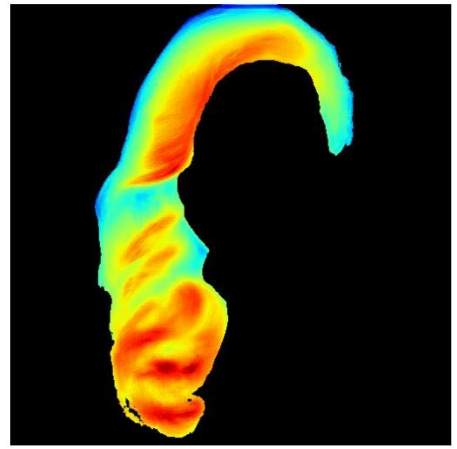
Orientation Map

From Strand Map



Strand Map

From HairStep

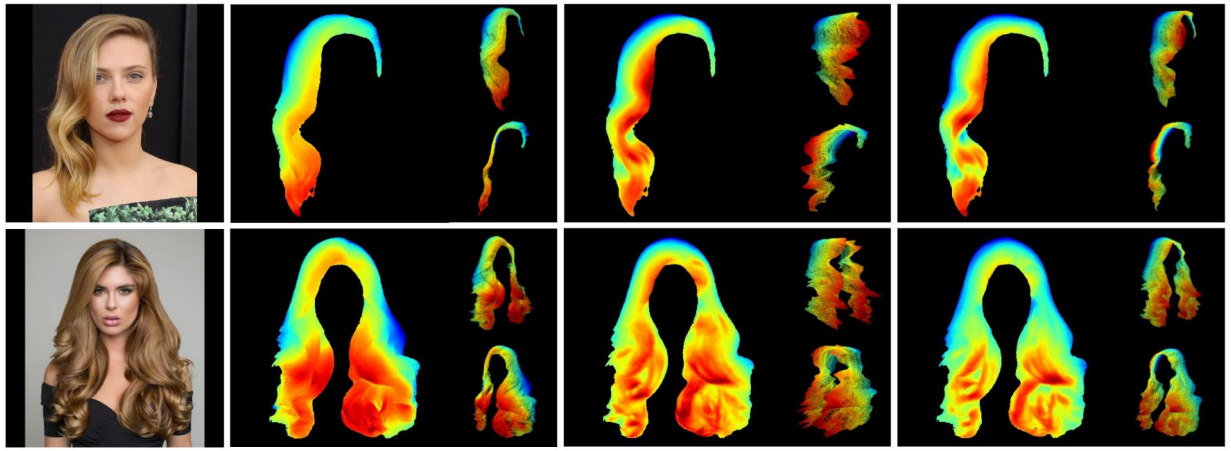


Depth Map

Visual Comparisons on Representation (Multi-view)

Comparisons with Orientation Map

Ablation Study



Input Image | $Depth_{pseudo}$ | $Depth_{weak}$ | $Depth_{DA}$

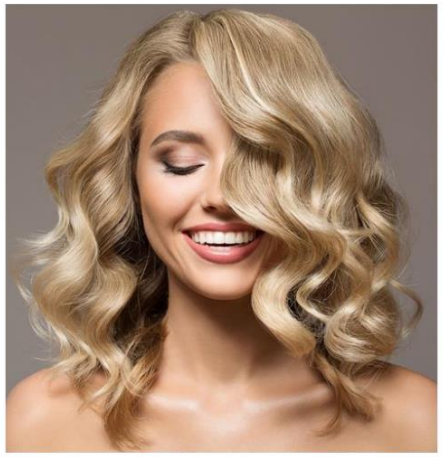
- C_0: strand map + $Depth_{pseudo}$.
- C_1: strand map + $Depth_{weak}$.
- $Full$: strand map + $Depth_{DA}$.

Method	IoU ↑	*HairSale* ↓	*HairRida* ↑
C_0	77.75%	16.03 (-18.2%)	73.57%
C_1	77.11%	16.54 (-15.6%)	75.8%
$Full$	77.22%	16.36 (-16.5%)	**76.79%**

C_1 gives **worse alignment of hair growth**.

C_0 suffers from the **flat geometry of depth**.

Ablation Study



HairRida

55.88%

HairRida

75.00%

HairRida

78.26%

Image

C_0

C_0

Full

Ablation Study

HairRida

78.26%

HairRida

80.49%

HairRida

89.80%

Image

C_0

C_0

Full

Outline

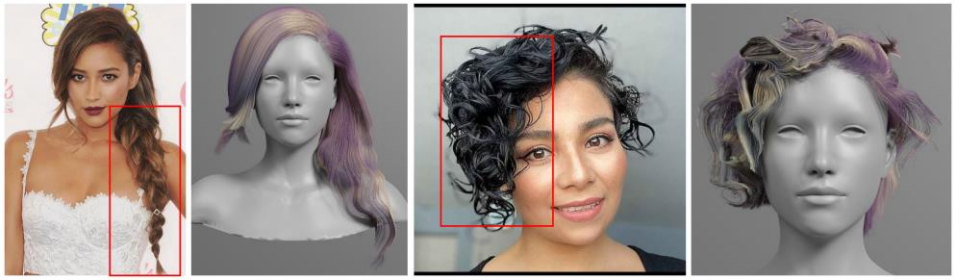
- Background
- HairStep
- Experiment
- Conclusion

Conclusion

- A novel **representation** HairStep
- Two **datasets** HiSa & HiDa for future research
- Fair **metrics** for single-view 3D hair modeling

Limitation

- No update for current pipeline of 3D hair modeling
- Fails on hairstyles with braid and complex curly pattern



Reference

[1] Chai, M., Shao, T., Wu, H., Weng, Y., & Zhou, K. (2016). Autohair: Fully automatic hair modeling from a single image. ACM Transactions on Graphics, 35(4).

[2] Zhou, Y., Hu, L., Xing, J., Chen, W., Kung, H. W., Tong, X., & Li, H. (2018). Hairnet: Single-view hair reconstruction using convolutional neural networks. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 235-251).

[3] Saito, S., Hu, L., Ma, C., Ibayashi, H., Luo, L., & Li, H. (2018). 3D hair synthesis using volumetric variational autoencoders. ACM Transactions on Graphics (TOG), 37(6), 1-12.

[4] Wu, K., Ye, Y., Yang, L., Fu, H., Zhou, K., & Zheng, Y. (2022). 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 (pp. 1526-1535).

[5] Paris, S., Briceno, H. M., & Sillion, F. X. (2004). Capture of hair geometry from multiple images. ACM transactions on graphics (TOG), 23(3), 712-719.

[6] Chen, W., Fu, Z., Yang, D., & Deng, J. (2016). Single-image depth perception in the wild. Advances in neural information processing systems, 29.

[7] Yang, L., Shi, Z., Zheng, Y., & Zhou, K. (2019). Dynamic hair modeling from monocular videos using deep neural networks. ACM Transactions on Graphics (TOG), 38(6), 1-12.

The End

Thank You!