Monocular Identity-Conditioned Facial Reflectance Reconstruction

CVPR 2024

Xingyu Ren, Jiankang Deng*, Yuhao Cheng, Jia Guo, Chao Ma*,
Yichao Yan, Wenhan Zhu, Xiaokang Yang

Shanghai Jiao Tong University

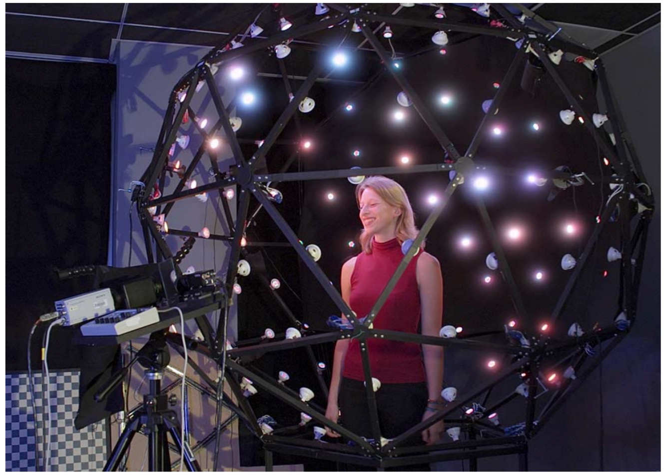
饮 水 思 源 • 爱 国 荣 校

Overview | ID2Reflectance

Overview | ID2Reflectance

Industrial pipeline | Capture System


Lightstage 1


Lightstage 3

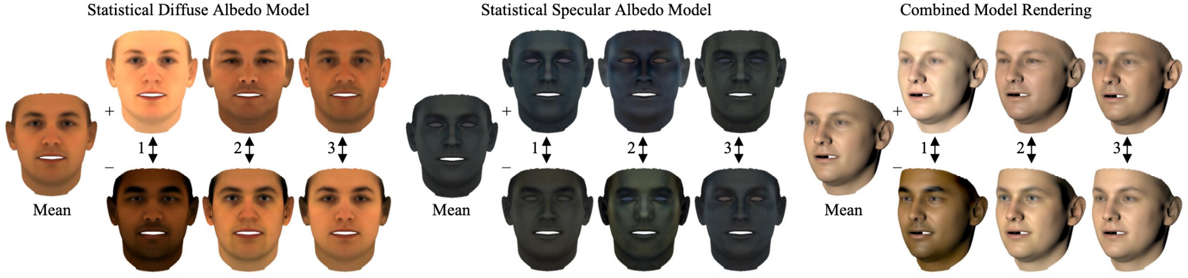

Lightstage 2


Lightstage 5

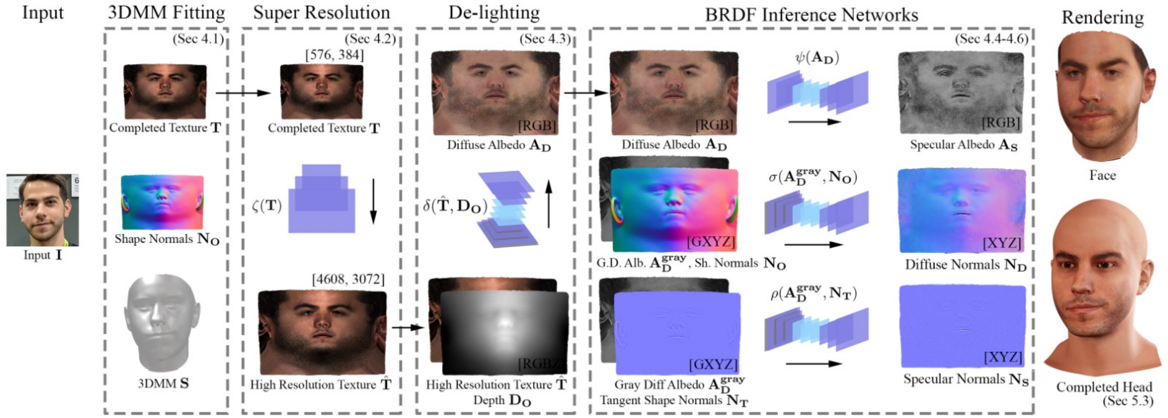

Our LightStage System

Complicated and High Cost

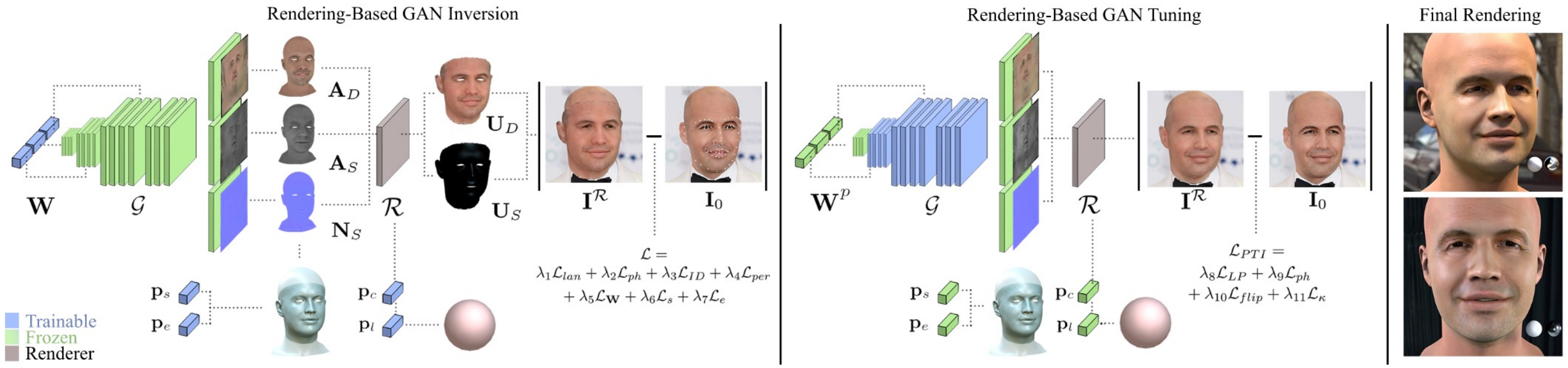
Related Works



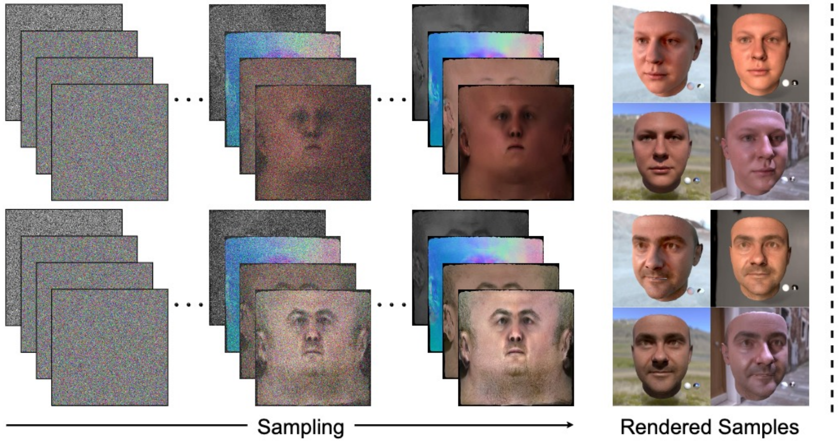
Facial Albedo Morphable, Smith et al. CVPR 2020



AvatarMe, Lattas et al. CVPR 2020

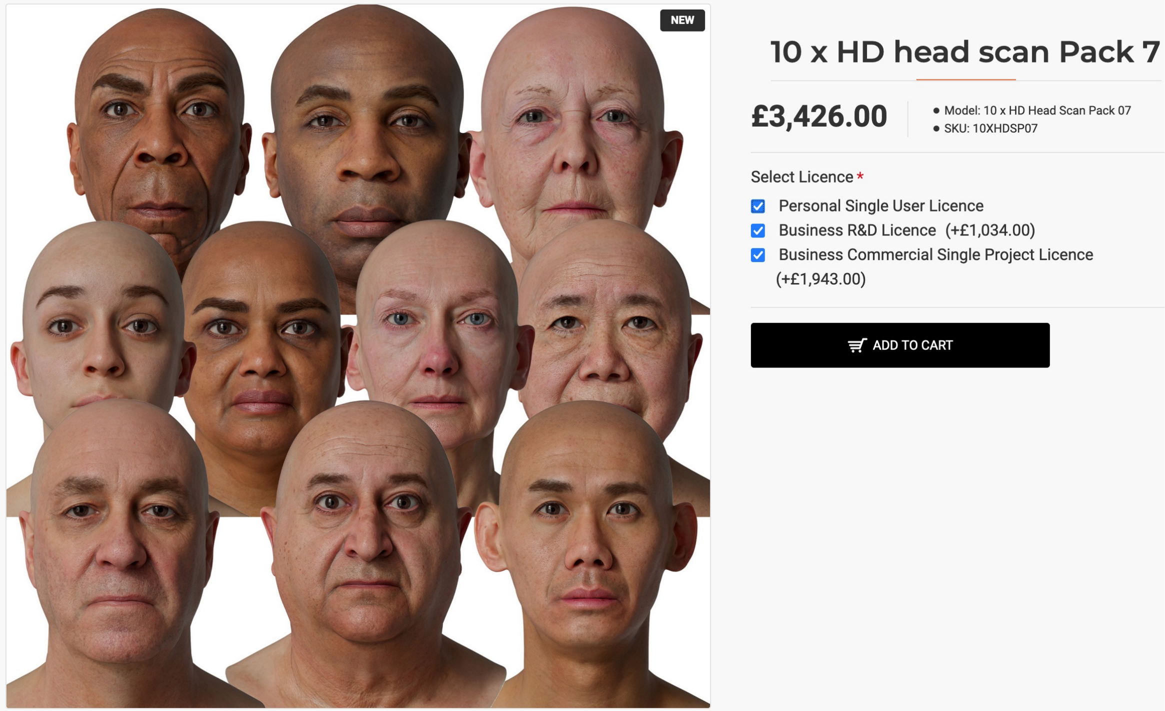


FitMe, Lattas et al. CVPR 2023



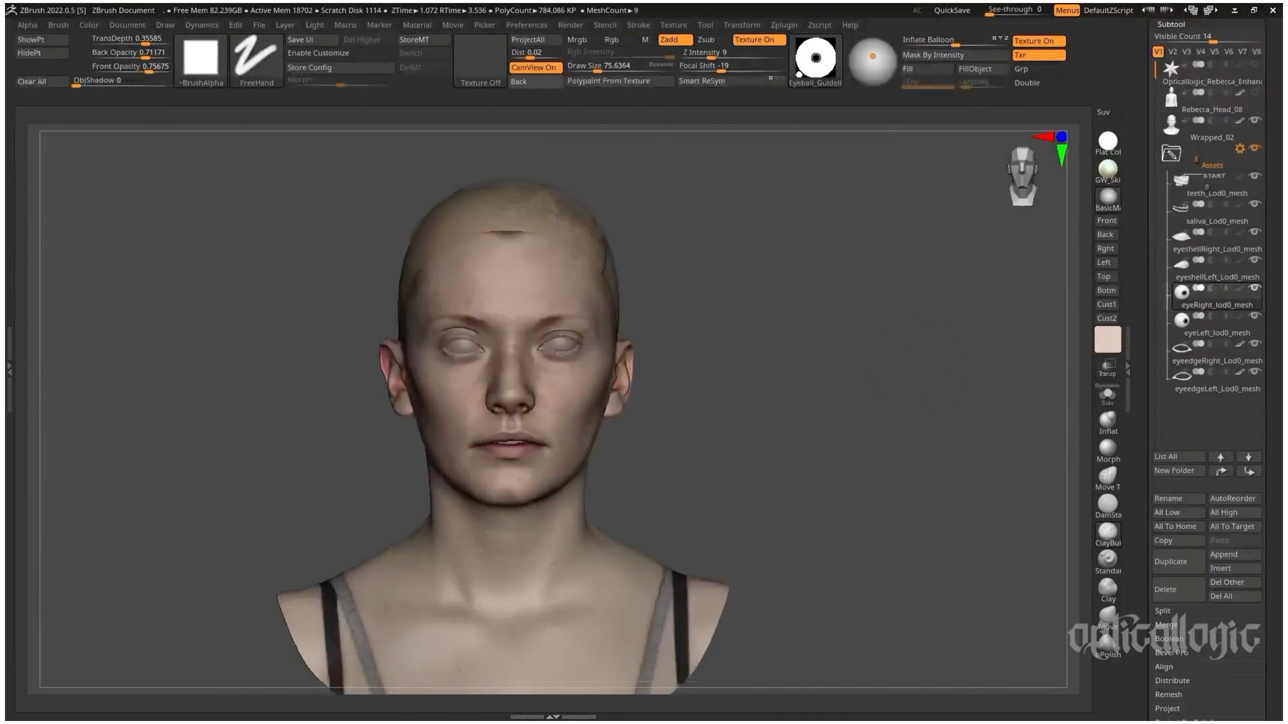
Relightify, Papantoniou et al. ICCV 2023

Challenge



Limited high-quality reflectance data available for purchase

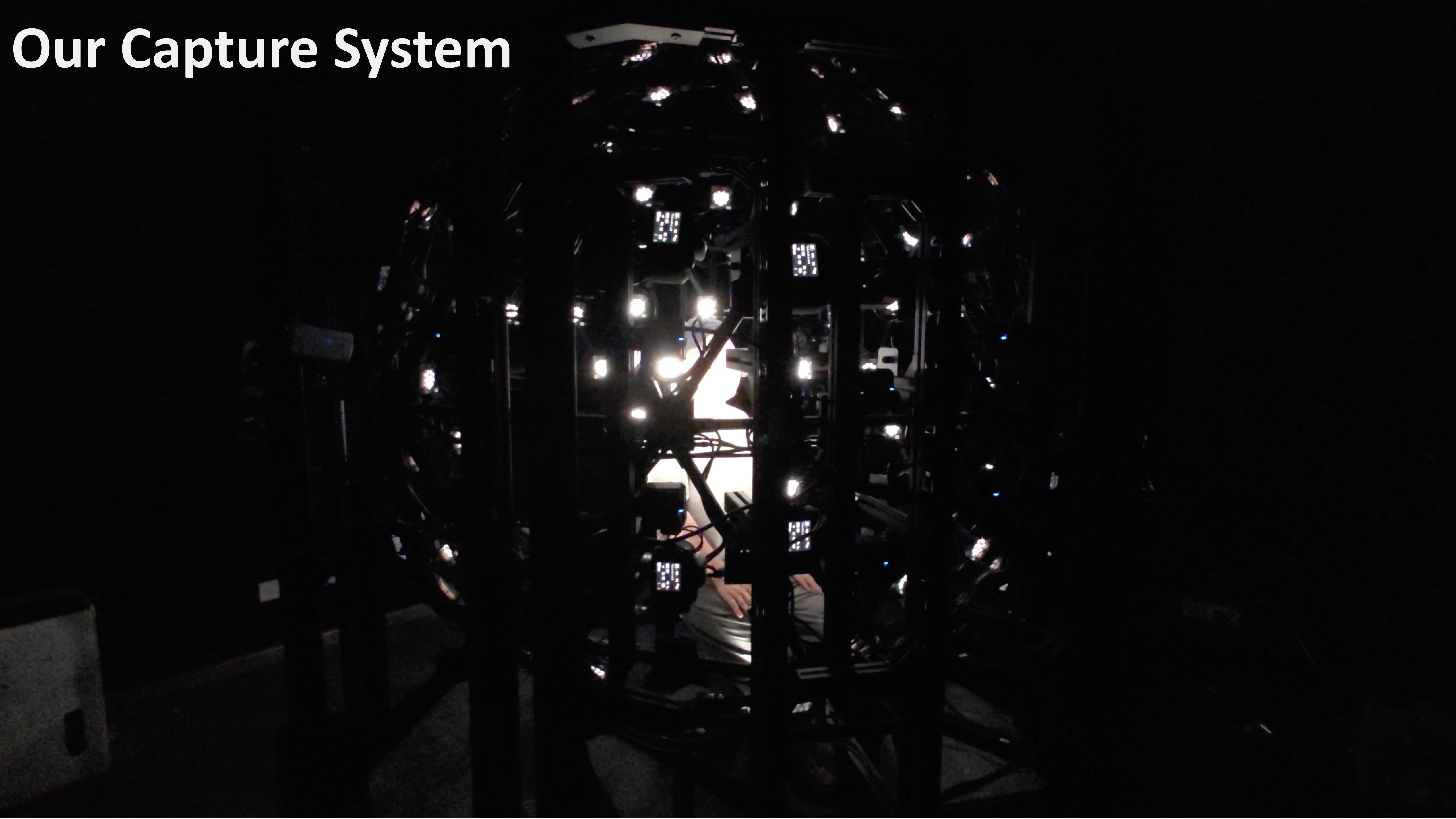
High post-processing costs



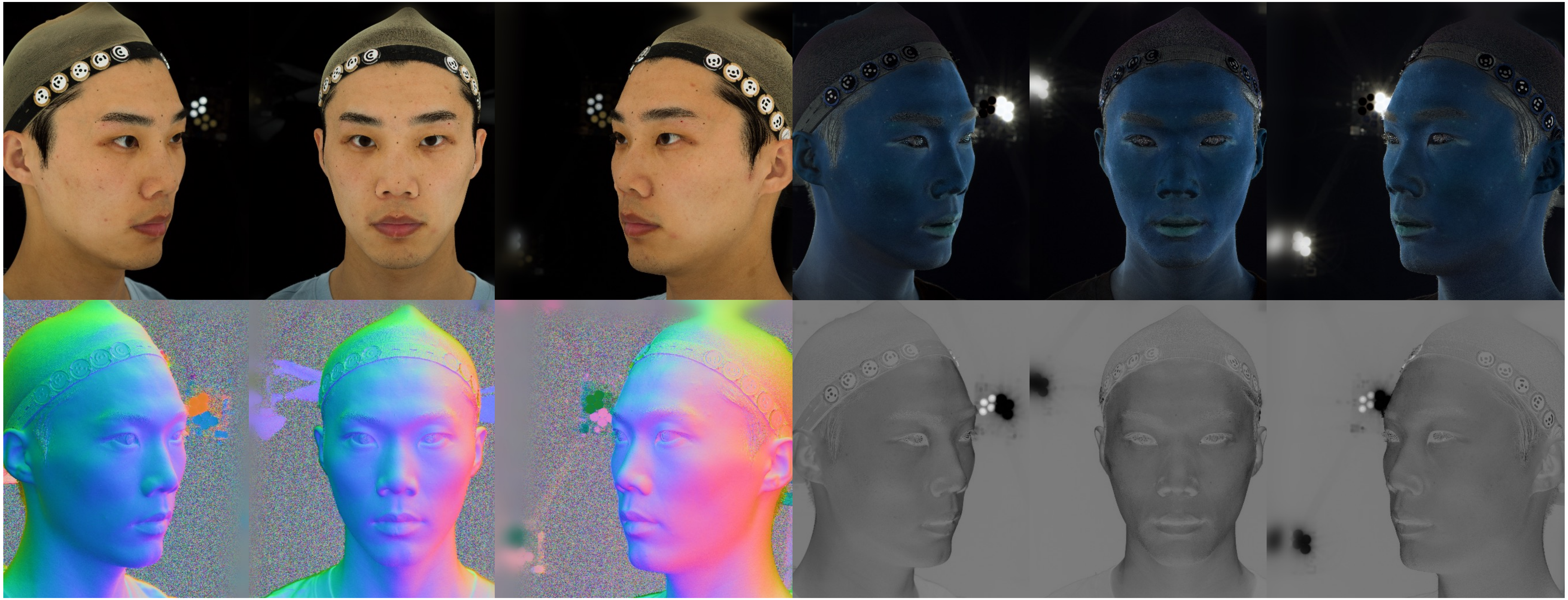
Motivation

How to achieve facial reflectance reconstruction for a single image with limited captured raw data?

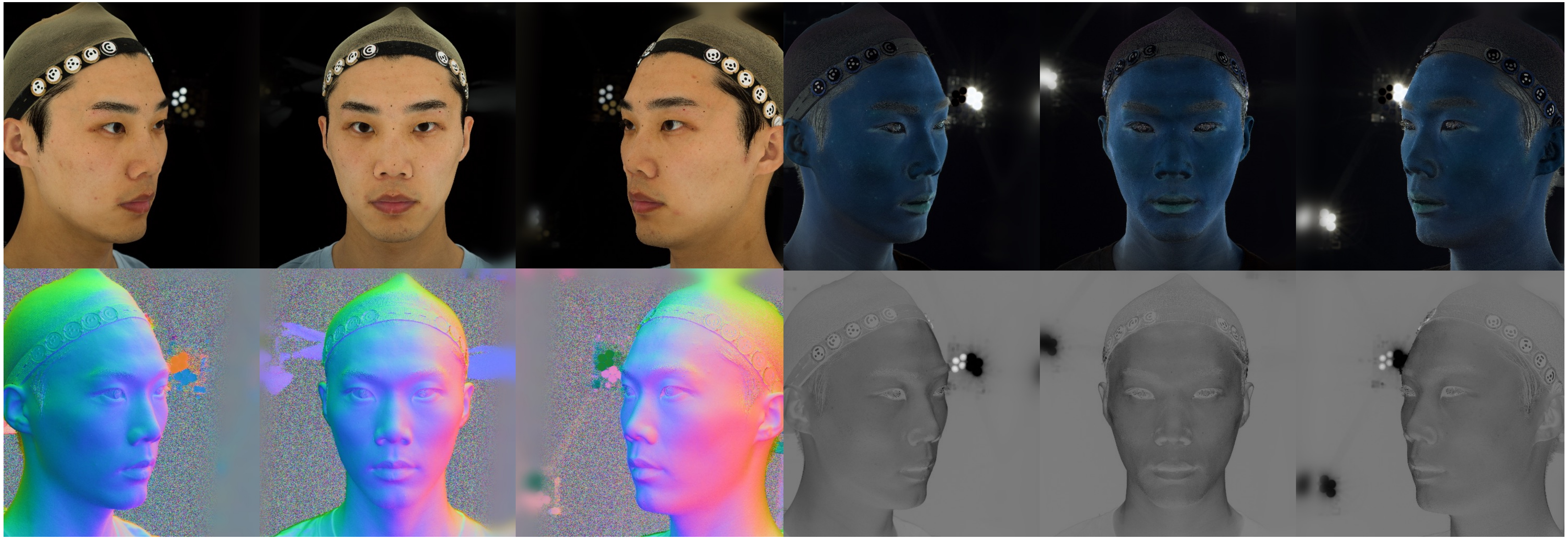
Our Capture System



Motivation

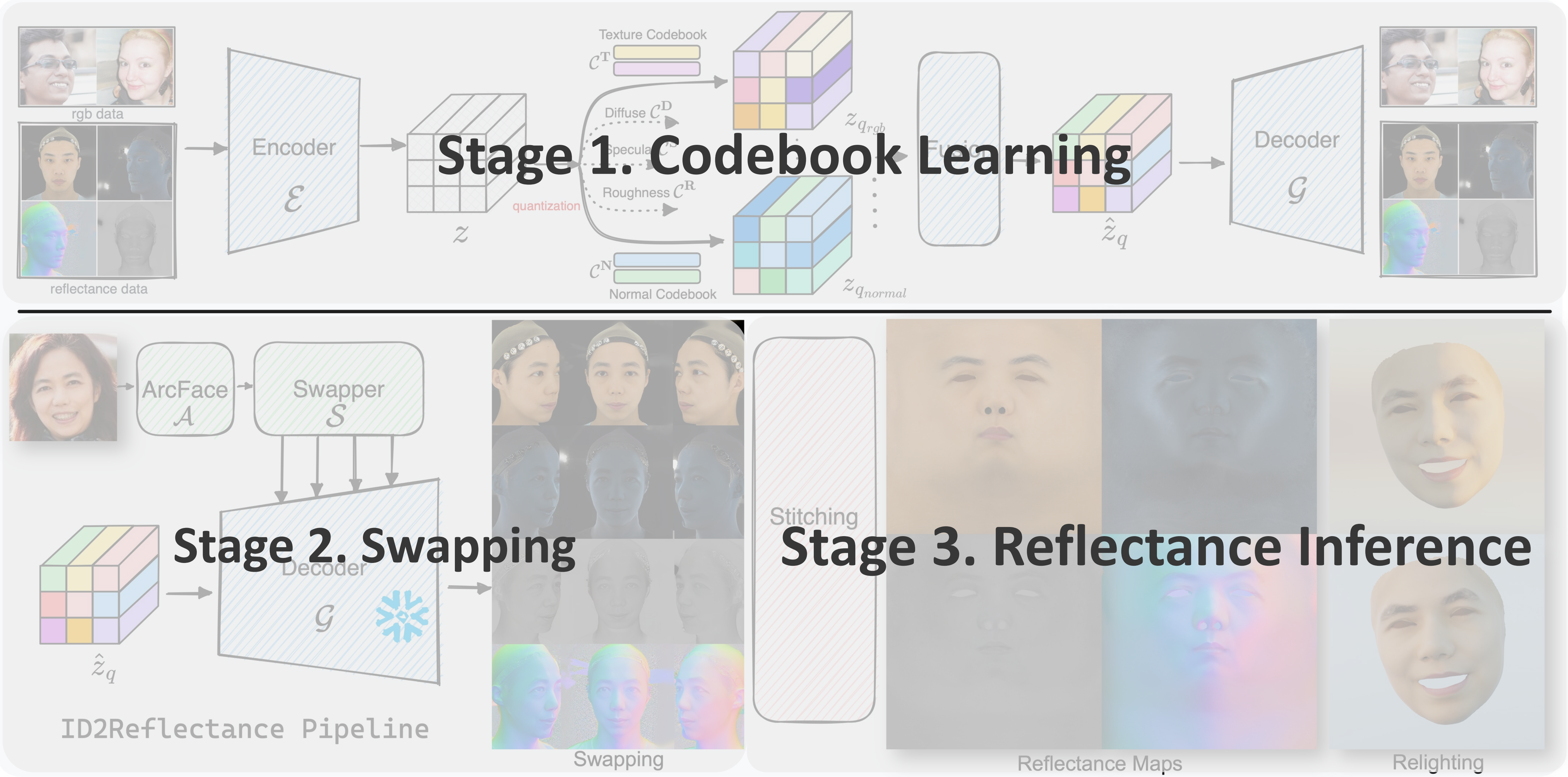


Motivation



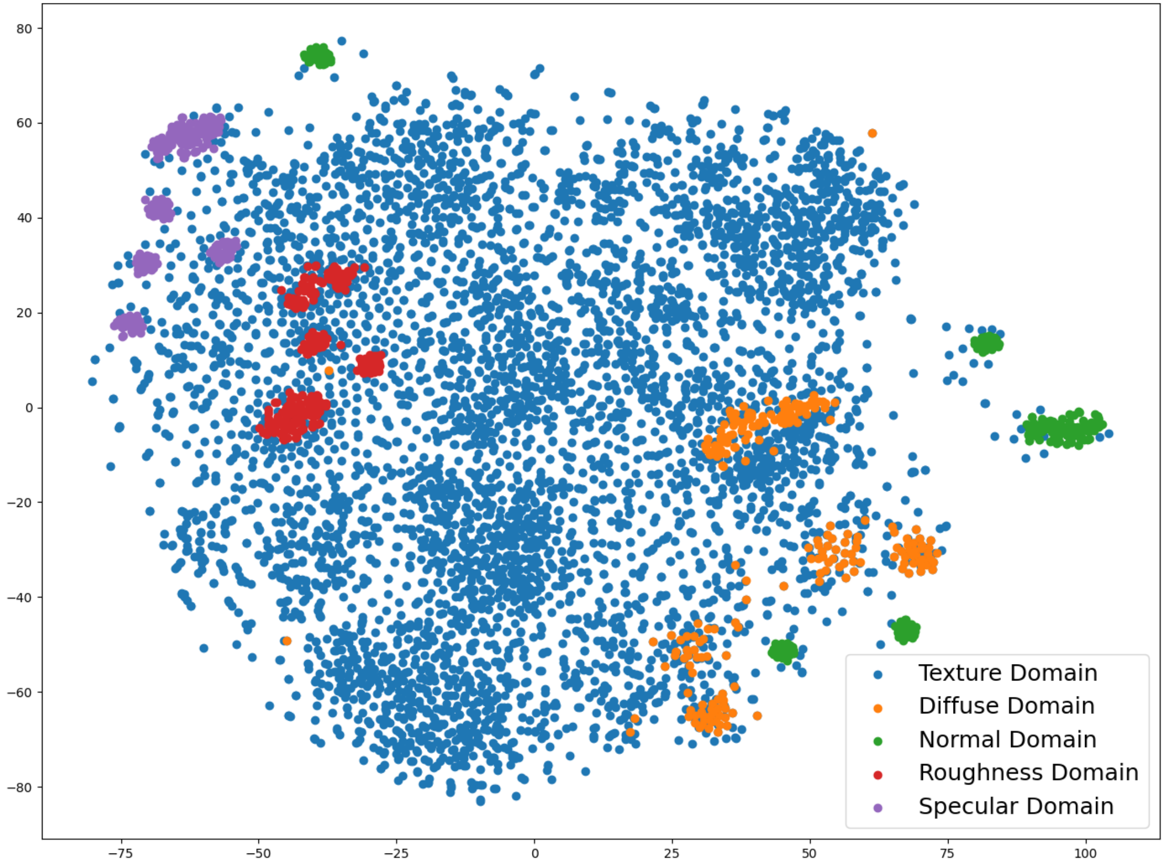
Insight: Model the facial structure (lots of RGB data) and the appearance (limited reflectance data) separately.

Method | Pipeline

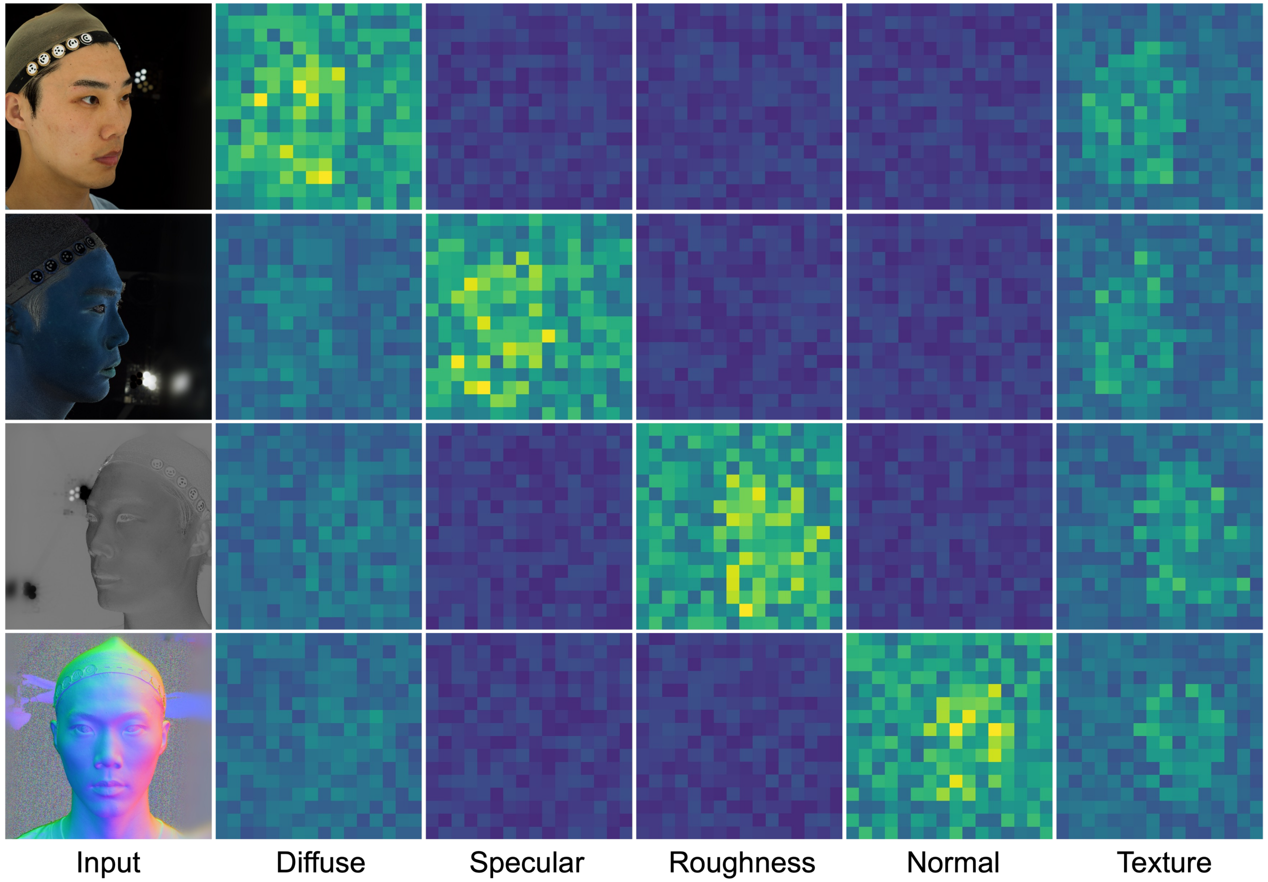


Method | Codebook Learning

- **Stage 1:** Train a shared codebook by high quality facial RBG and reflectance data.
- **Stage 2:** Train a multi-domain codebook to further improve facial reflectance reconstruction.



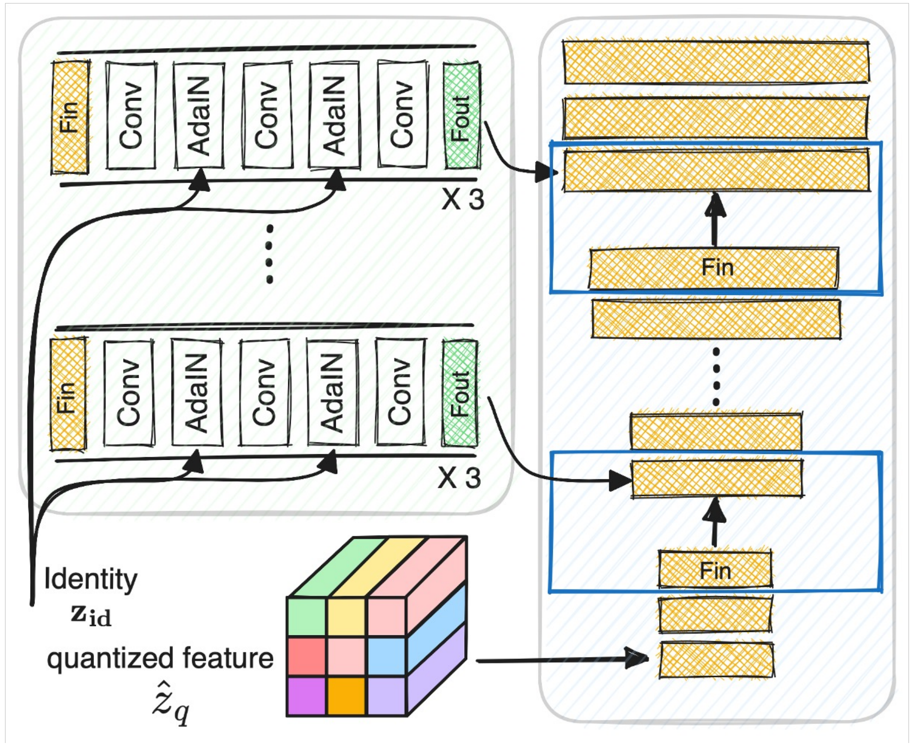
t-SNE distribution of latent feature z_q for reflectance data and RGB data.



Visualization of codebook fusion weights.

Method | Identity Swapping

- **Codebook-based identity swapping: Once** identity injection module trained on RGB data, which can be automatically migrated to the facial reflectance domain.

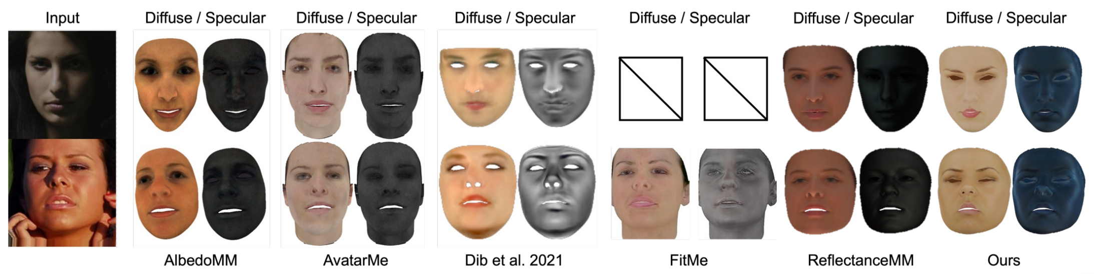


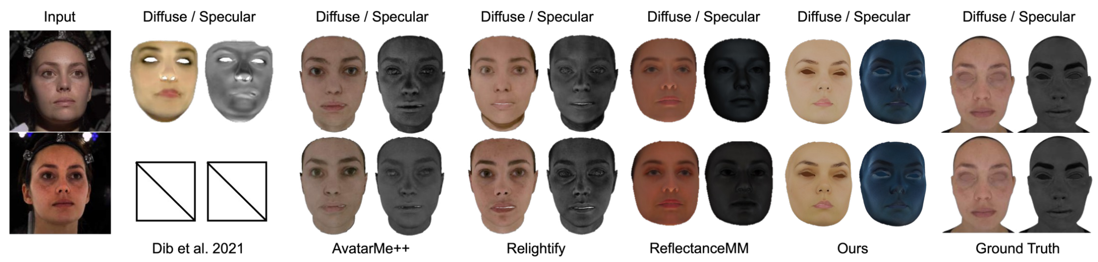
Detailed architecture of our swapper module.

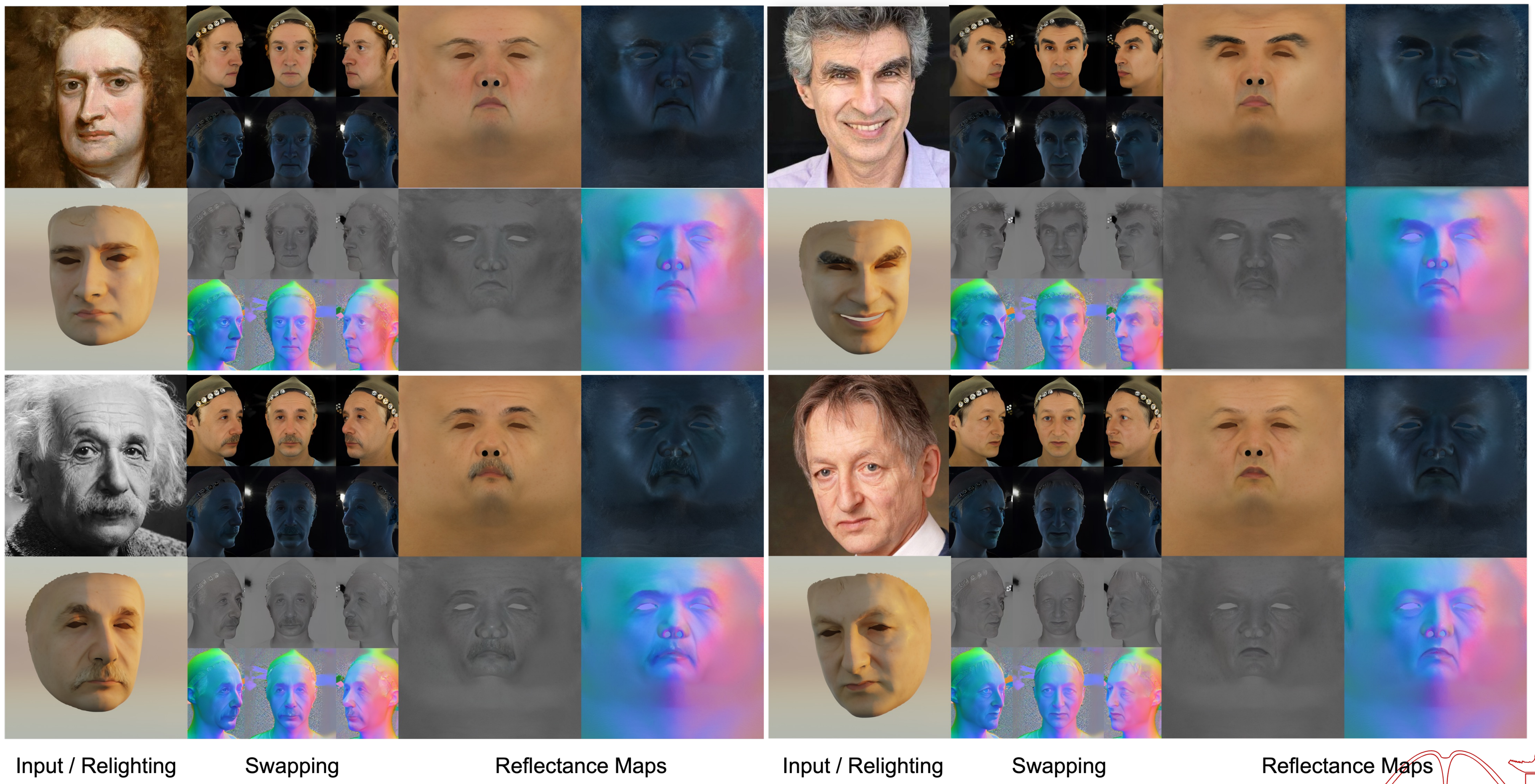
Comparison of swapper module under different configurations. F16 means that the source identity is injected from an upsample layer with a feature size of 16**x**16.

Configuration	ID-Retrival ↑	pose ↓
F16	N/A	N/A
F16 + F32	0.894	0.0143
F16 + F32 + F64	0.941	0.0132
F16 + F32 + F64 + F128 + F256	0.933	**0.0128**
F16 + F32 + F64 + F128 (Ours)	**0.965**	0.0129

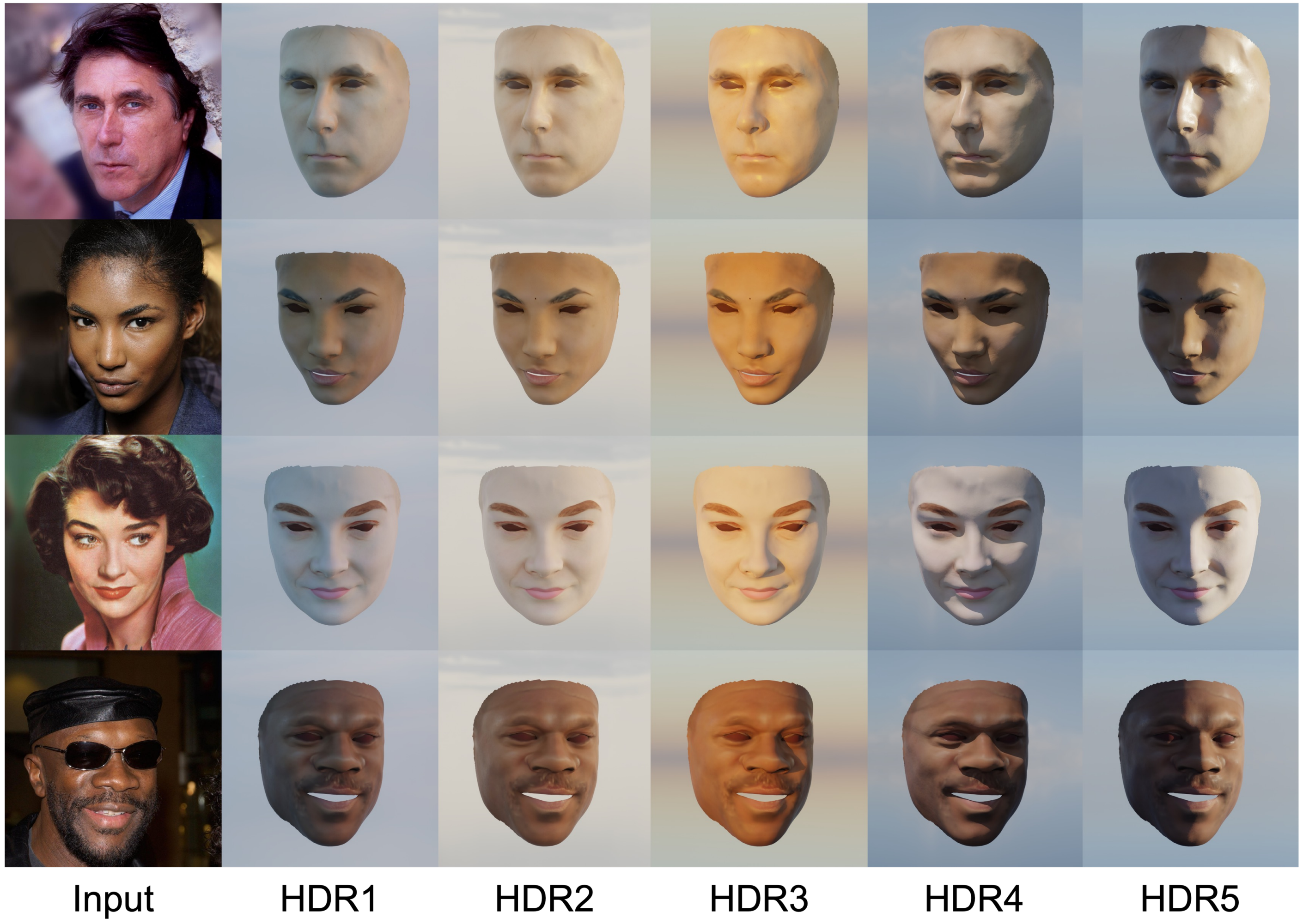
Experiments | Facial Reflectance Reconstruction

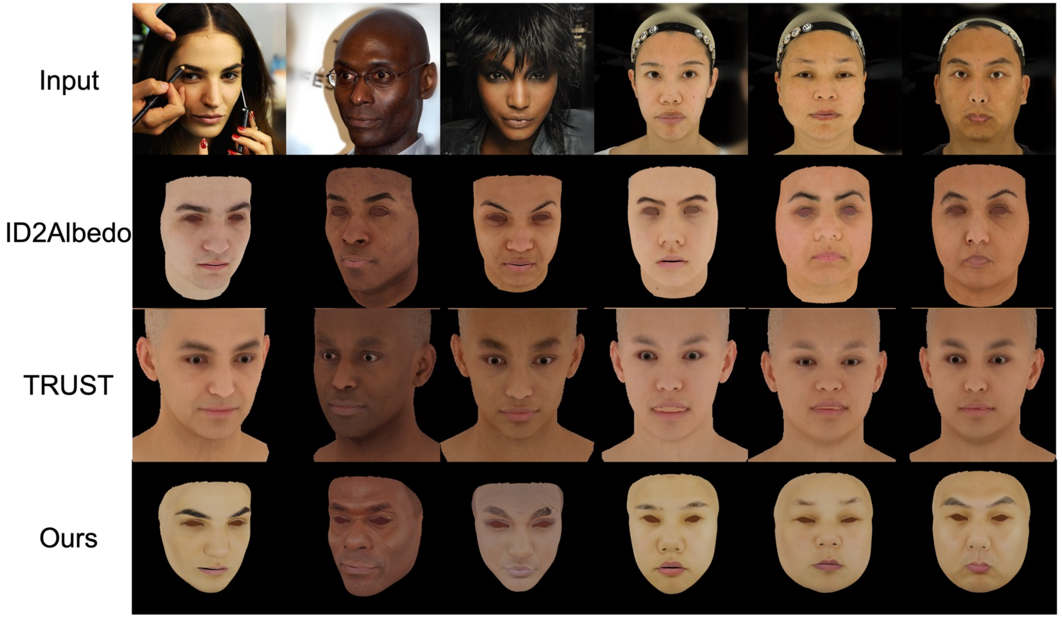






Experiments | Albedo Reconstruction & Relighting





Comparisons of our method with previous methods on albedo estimation

Methods	PSNR↑	SSIM↑	LPIPS↓	ID↑
TRUST [17]	21.63	0.852	0.2014	0.478
ID2Albedo [54]	23.72	0.884	0.1549	0.532
Ours	**28.47**	**0.923**	**0.1248**	**0.735**

SJTU

Experiments | Relighting results via hard inputs

Input HDR1 HDR2 HDR3 HDR4 HDR5 HDR6 HDR7 HDR8 HDR9 HDR10

Ablation Studies | Reflectance Codebooks

Comparison to state-of-the-arts on the FAIR benchmark

Method	Avg. ITA ↓	Bias ↓	Score ↓	MAE ↓	ITA per skin type ↓					
					I	II	III	IV	V	VI
Deep3D [13]	22.57	24.44	47.02	27.98	**8.92**	**9.08**	8.15	10.90	28.48	69.90
GANFIT [20]	62.29	31.81	94.11	63.31	94.80	87.83	76.25	65.05	38.24	11.59
MGCNet [57]	21.41	17.58	38.99	25.17	19.98	12.76	8.53	**9.21**	22.66	55.34
DECA [18]	28.74	29.24	57.98	38.17	9.34	11.66	11.58	16.69	39.10	84.06
INORig [2]	27.68	28.18	55.86	33.20	23.25	11.88	**4.86**	9.75	35.78	80.54
CEST [68]	35.18	12.14	47.32	29.92	50.98	38.77	29.22	23.62	21.92	46.57
TRUST [17]	13.87	**2.79**	**16.67**	**18.41**	11.90	11.87	11.20	13.92	**16.15**	18.21
ID2Albedo [54]	**12.07**	4.91	16.98	23.33	18.30	9.13	5.83	9.46	19.09	**10.59**
Ours	14.21	4.22	18.43	22.02	12.91	13.11	9.68	10.22	17.72	21.63

Comparison of ID2Reflectance framework under different training data.

Numbers	Diffuse	Specular	Roughness	Normal
30 subjects	25.09	24.12	24.78	23.55
60 subjects	28.63	27.58	28.22	27.05
90 subjects	30.54	29.84	30.41	29.22
115 subjects	31.62	30.96	31.59	30.32

Comparison of ID2Reflectance framework under different configurations. (1) joint codebook v.s. multi-domain codebooks for reflectance reconstruction, and (2) fixed swapping template v.s. closest swapping template for identity-conditioned reflectance prediction.

Configs	Diffuse	Specular	Roughness	Normal
Joint codebook	24.87	20.95	21.31	20.56
Multi-domain codebooks	31.62	30.96	31.59	30.32
Fixed Template	25.26	26.44	29.56	25.77
Closest Template	28.47	26.68	30.32	26.83

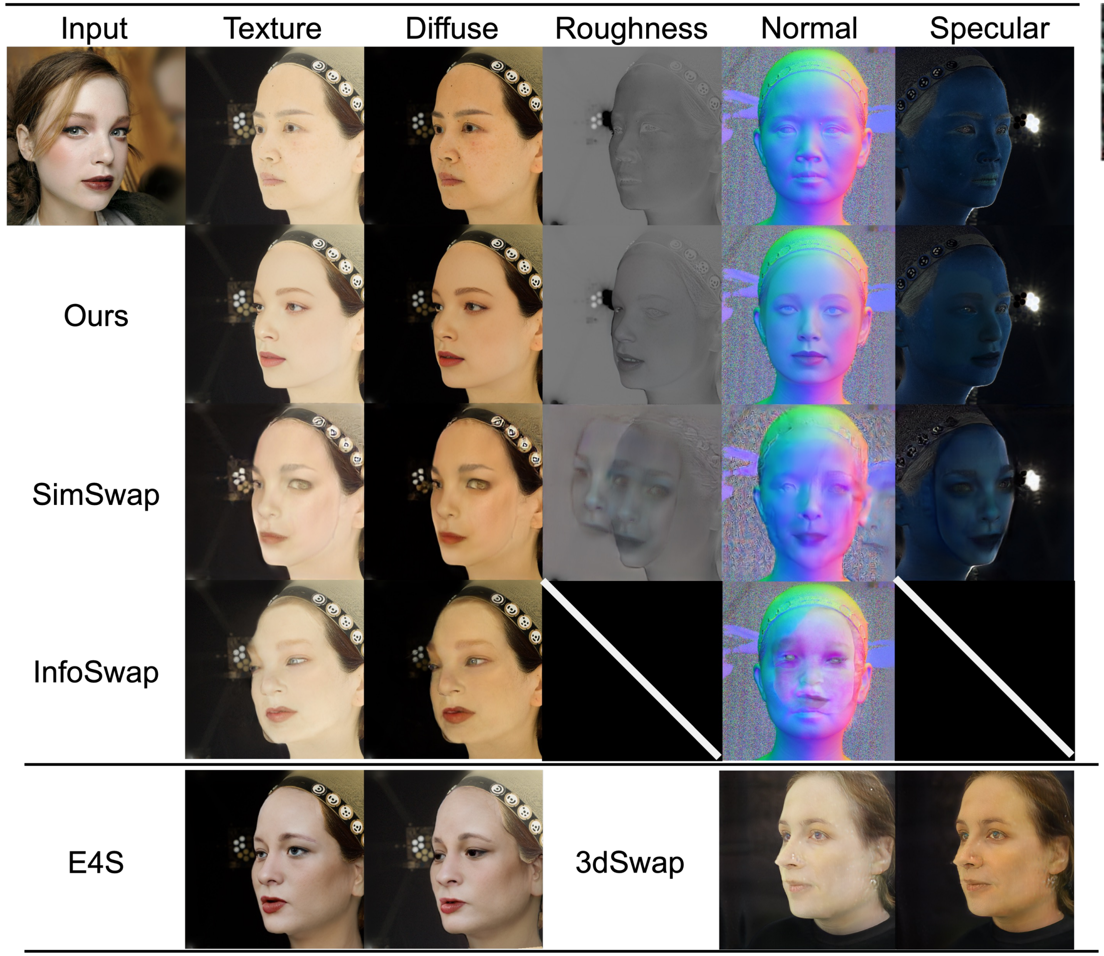
Reconstruction comparison of using joint and multi-domain codebooks. Inputs are the same faces from PBR domains.

Input

Joint codebook

Multi-domain codebooks

Ablation Studies | Face Swapping



Cross-domain swapping comparison with other methods.

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

人工智能研究院
Artificial Intelligence Institute

Thank you

饮 水 思 源　爱 国 荣 校