

Capturing and Synthesis of Digital Humans

Wojciech Zielonka



Personal Website

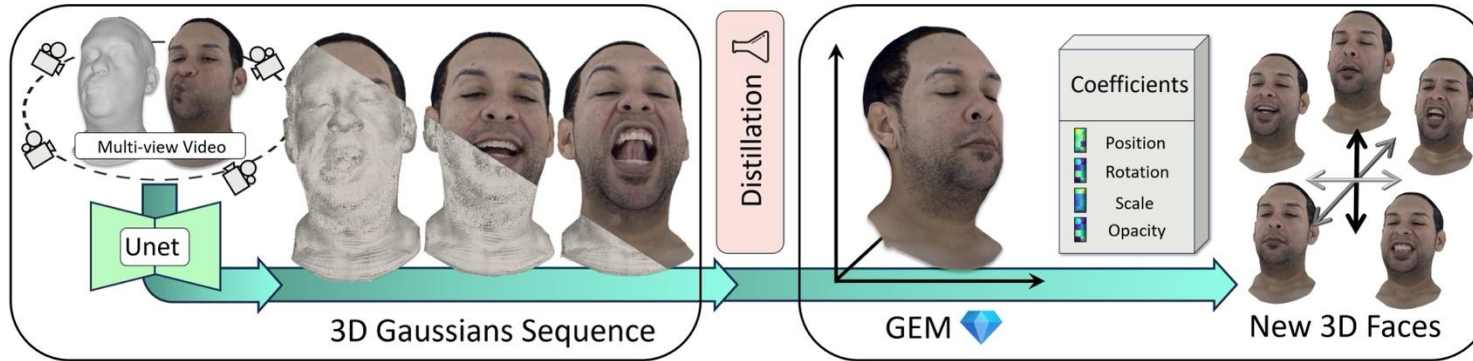


MAX PLANCK INSTITUTE
FOR INTELLIGENT SYSTEMS



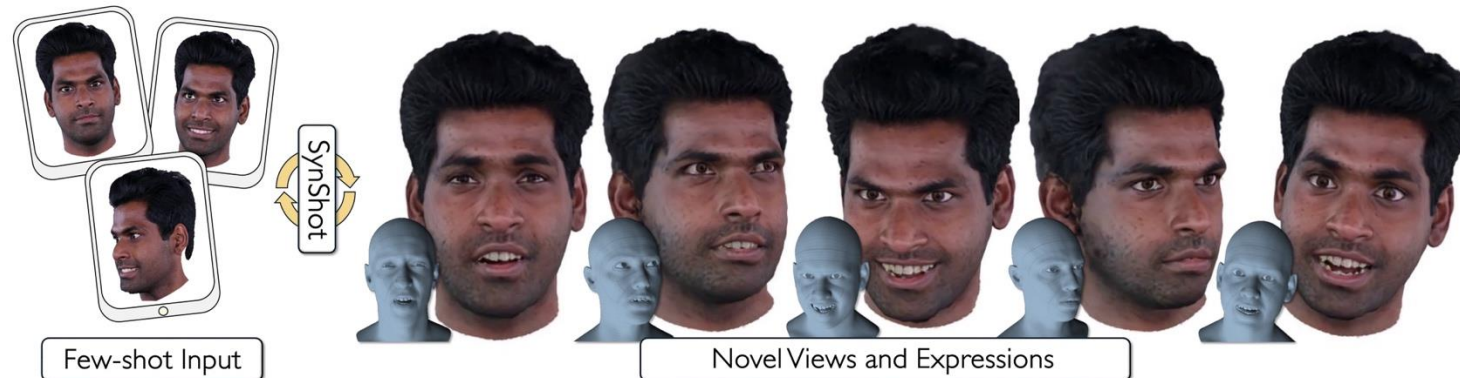
TECHNISCHE
UNIVERSITÄT
DARMSTADT

Capturing and Synthesis of 3D Digital Humans



GEM - Gaussian Eigen Models for Human Heads

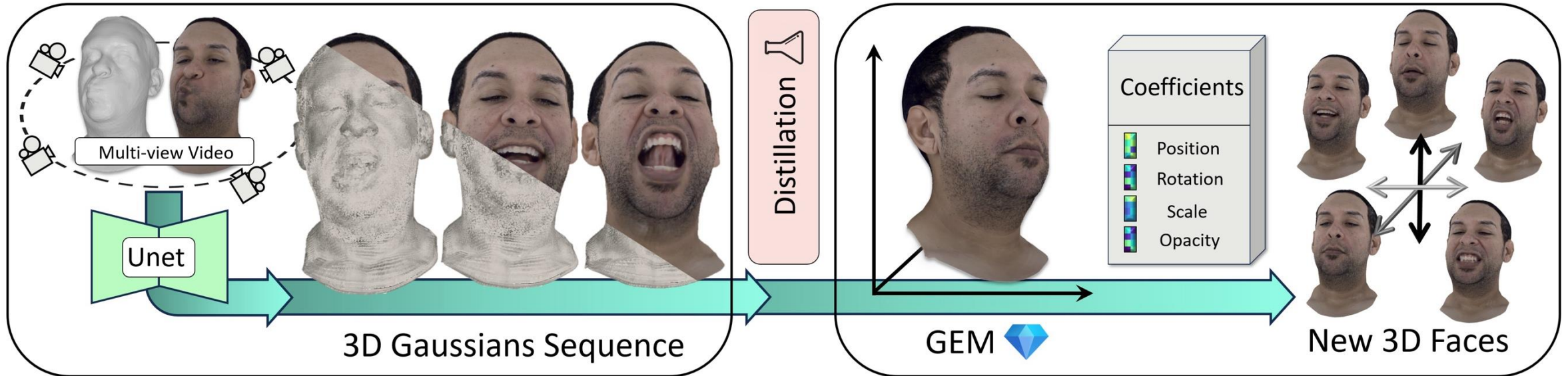
Wojciech Zielonka, Timo Bolkart, Thabo Beeler, Justus Thies



SynShot - Synthetic Prior for Few-Shot Drivable Head Avatar Inversion

Wojciech Zielonka, Stephan J. Garbin, Alexandros Lattas,
George Kopanas, Paulo Gotardo, Thabo Beeler, Justus Thies, Timo Bolkart

Capturing and Synthesis of 3D Digital Humans



GEM - Gaussian Eigen Models for Human Heads

Wojciech Zielonka, Timo Bolkart, Thabo Beeler, Justus Thies

GEM - Motivation

Place A



GEM  Avatar

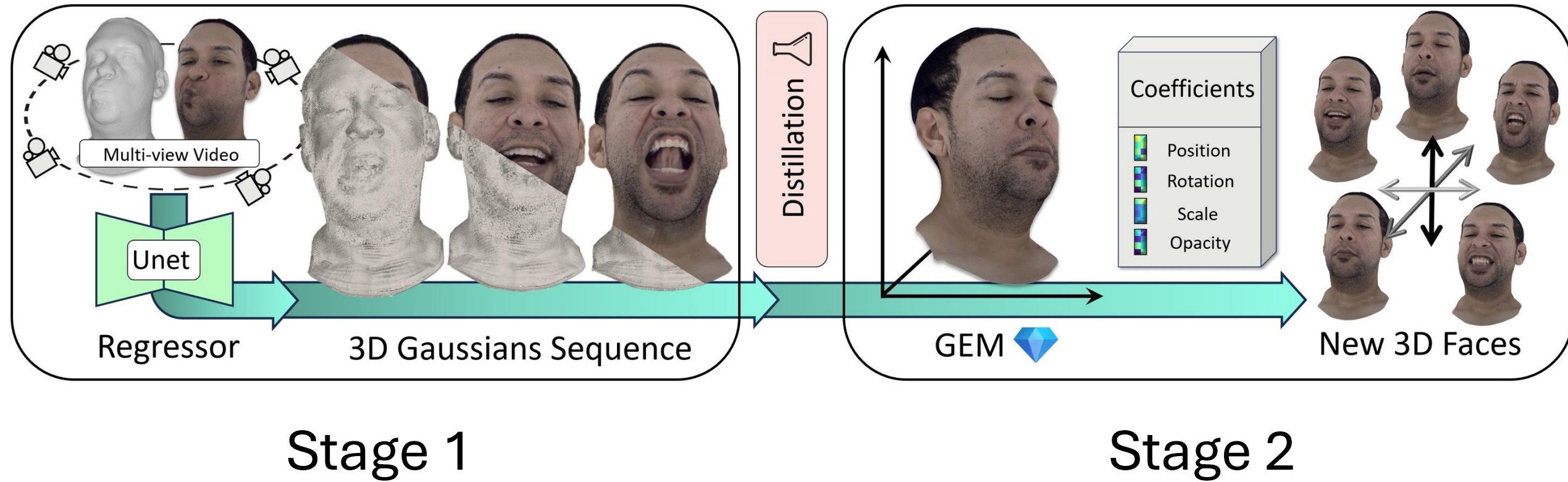
Place B



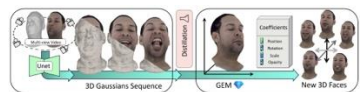
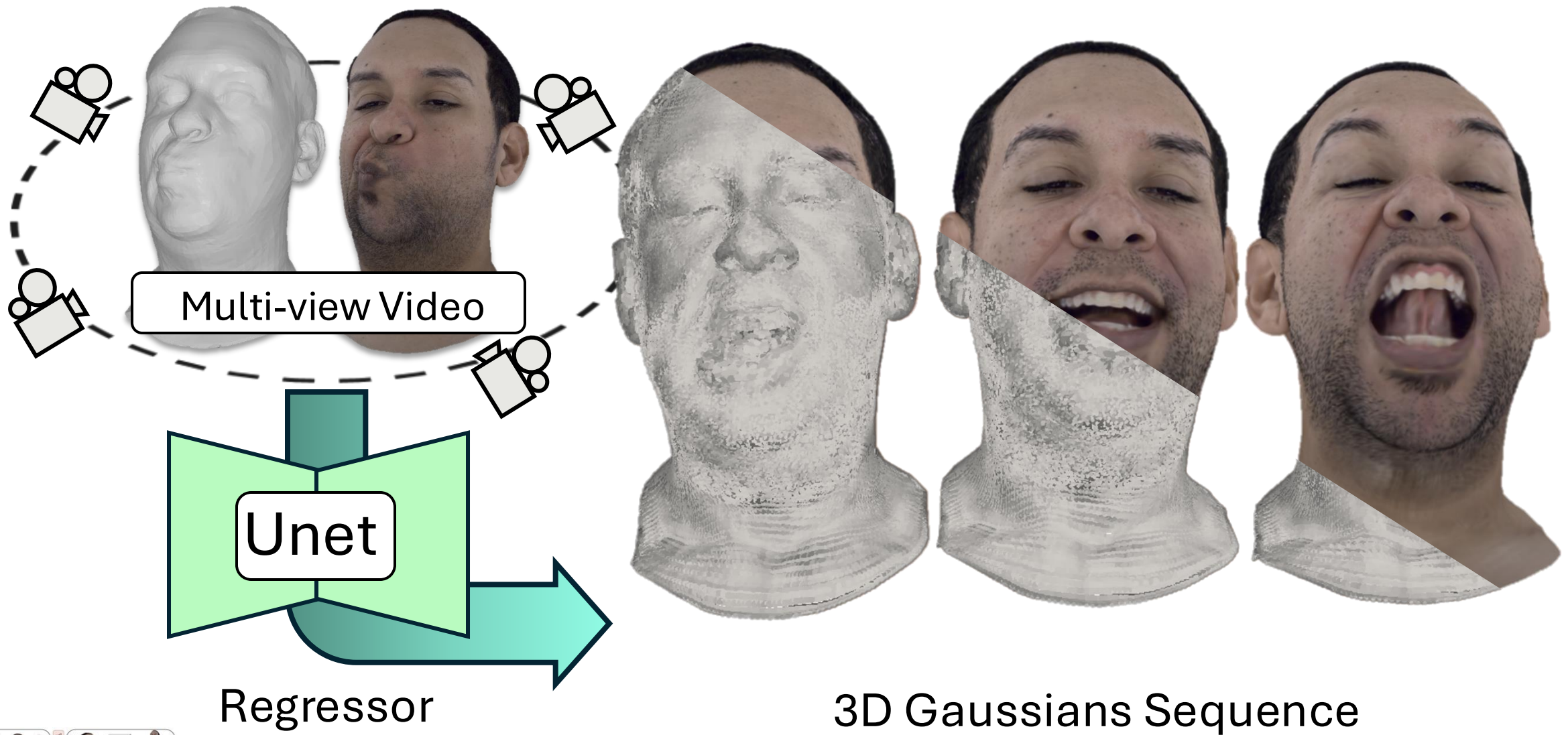
GEM  Avatar




Method

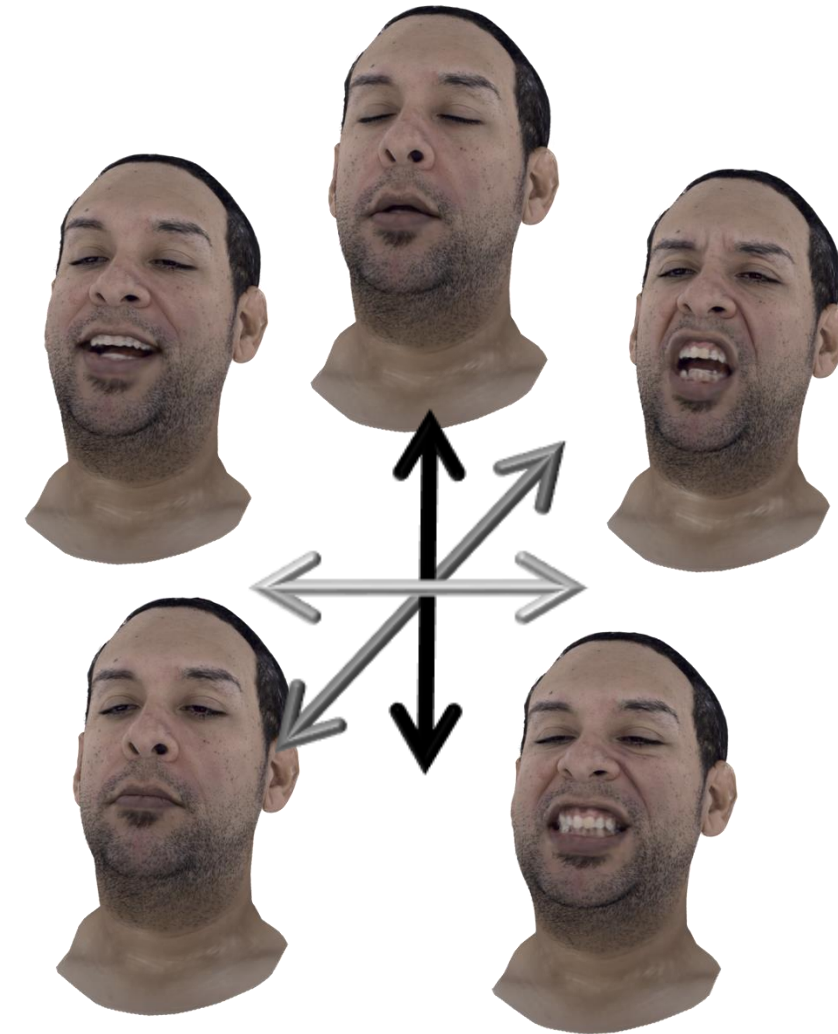
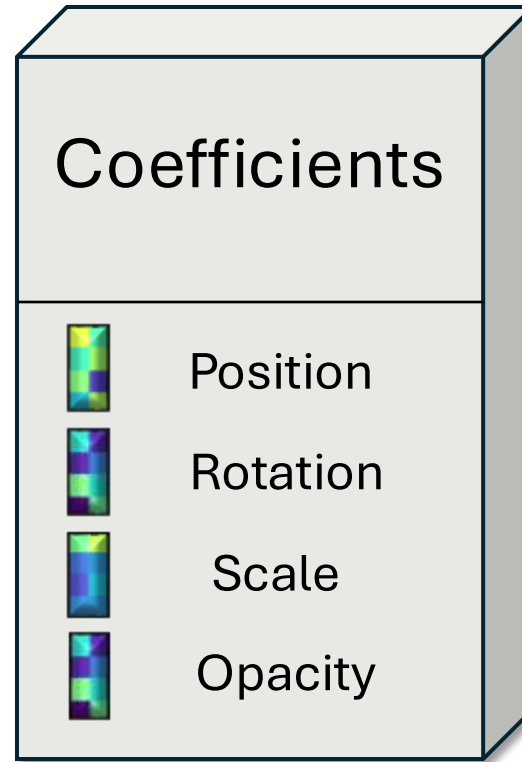


Method (Stage 1)



Method (Stage 2)

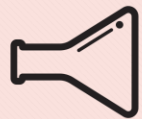
Distillation 



New 3D Faces

GEM 





Distillation

Ensamble of eigen basis



GEOMETRY COMP 00 = -2.6



Position



OPACITY COMP 00 = -2.6



Opacity



ROTATION COMP 00 = -2.6



Rotation



SCALES COMP 01 = -2.6



Scale

Space traversal $[-3\sigma, 3\sigma]$



Faces



Results – Novel Expressions and Views



Ground Truth

Ours  GEM

Ours Net 

AG 

GA 

INSTA 

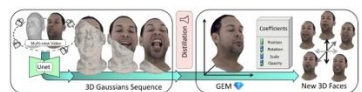
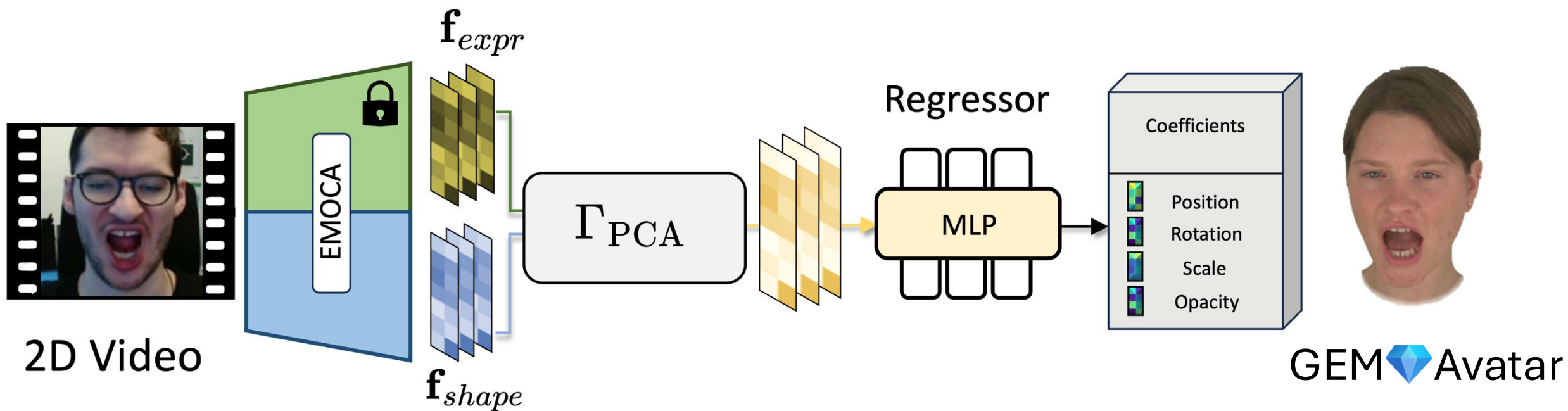
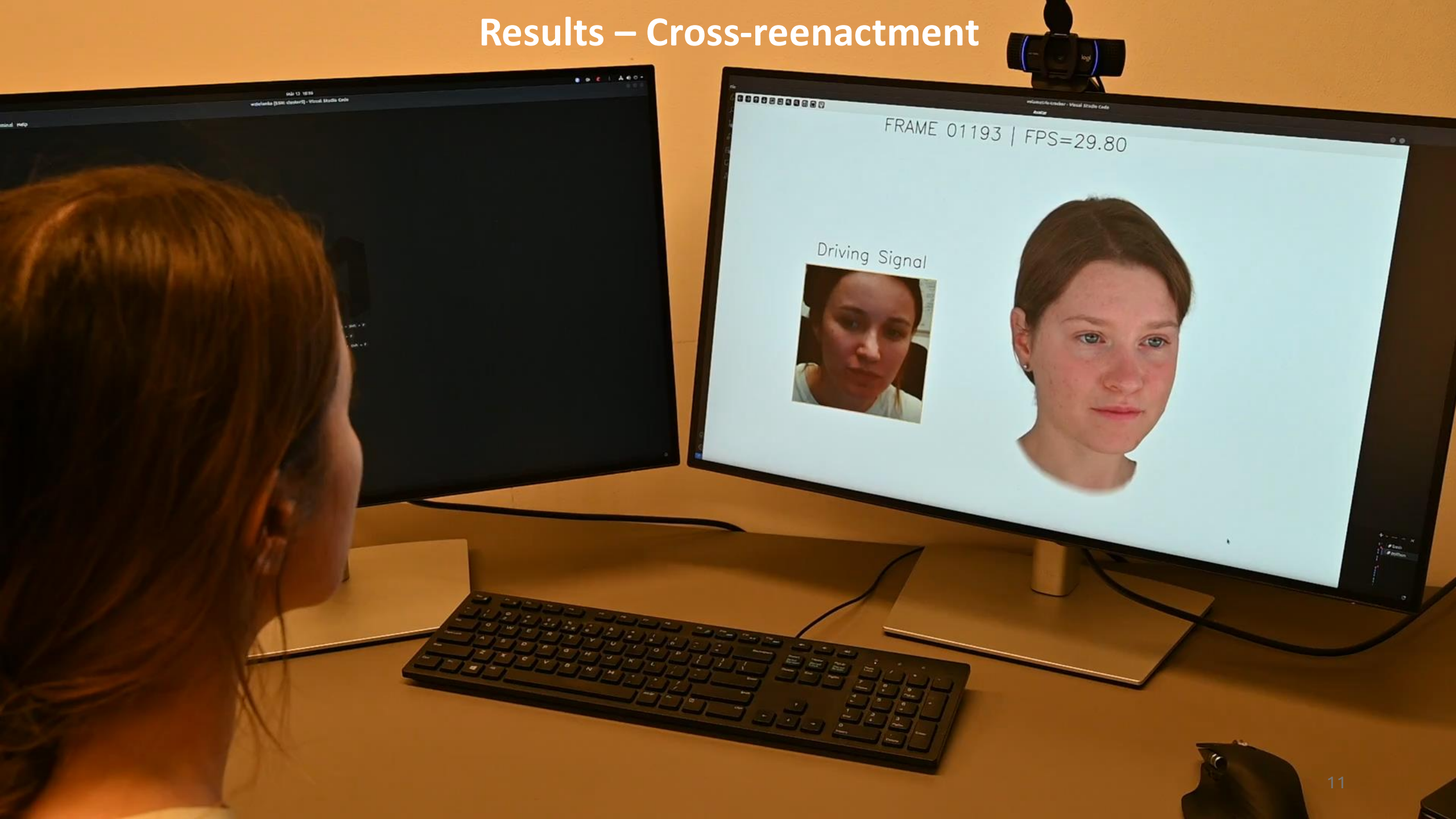


Image-base Cross-reenactment

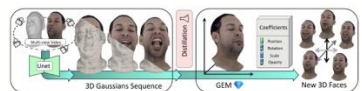
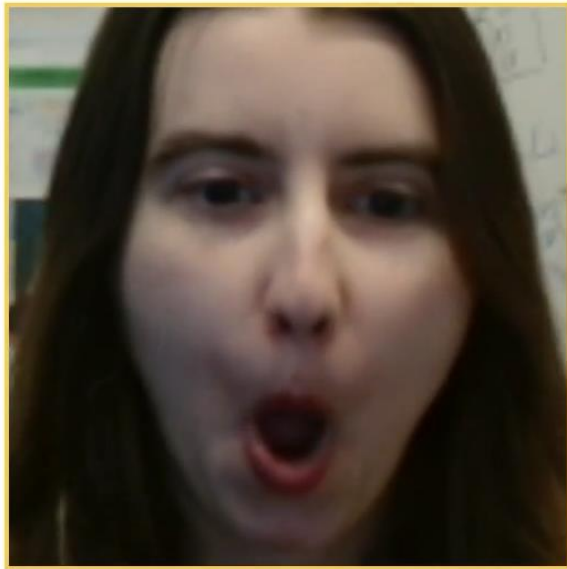


Results – Cross-reenactment



FRAME 03669 | FPS=29.60

Driving Signal



FRAME 00279 | FPS=30.15

Driving Signal



Localized PCA



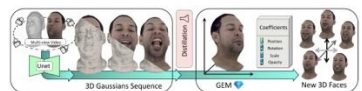
Source





GEM




Localized GEM



Summary



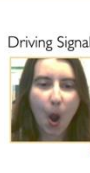
TECHNISCHE
UNIVERSITÄT
DARMSTADT



MAX PLANCK INSTITUTE
FOR INTELLIGENT SYSTEMS



Live Demo

GEM Avatars





GEM: Gaussian Eigen Models for Human Heads

Wojciech Zielonka, Timo Bolkart, Thabo Beeler, Justus Thies

Motivation



CNN Networks²



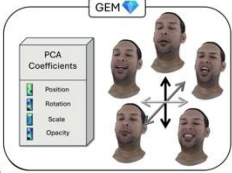

3DGS¹

Challenges

Neural head avatars trade off detail for efficiency: lightweight models lack realism, while high-quality ones are too resource-intensive for commodity devices like VR glasses.

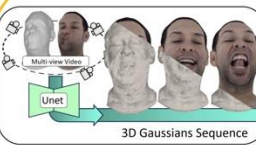
Solution

To address this gap, we introduce Gaussian Eigen Models, a compact, single-layer representation distilled from high-quality CNNs, enabling fast face synthesis via simple dot products.

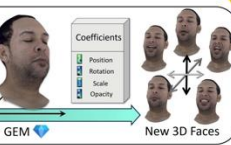



GEM Avatars

Method




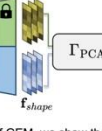
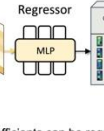
Stage 1



Stage 2







Based on our Gaussian regressor (Stage 1), we synthesize a dataset that is distilled into GEM (Stage 2).

Application









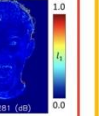

As an application of GEM, we show that the coefficients can be regressed from a single image, enabling real-time facial animation and cross-reenactment.

Results










Both our CNN and GEM perform better on novel views, especially around the mouth and wrinkles.



Compression error depending on the number of used principal components in GEM.

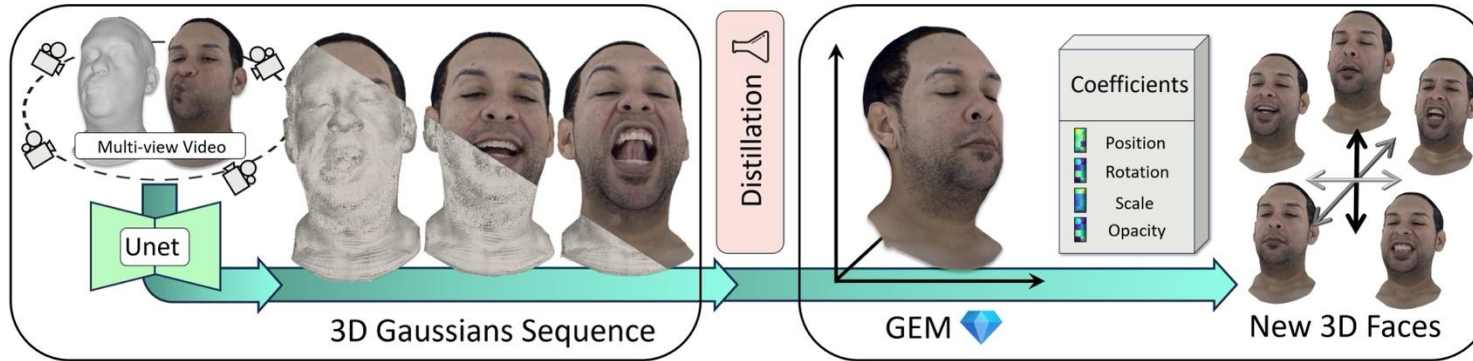

One of the applications of our GEM is real-time (cross-)reenactment by utilizing image-based regressor.

¹⁾ Kertt et al. 3D Gaussian Splatting for Real-Time Radiance Field Rendering
²⁾ Li et al. Animatable Gaussians: Learning Pose-dependent Gaussian Maps for High-Fidelity Human Avatar Modeling
³⁾ Chen et al. GaussianAvatars: Photorealistic Head Avatars with Rigged 3D Gaussians
⁴⁾ Zielonka et al. Instant Volumetric Head Avatars



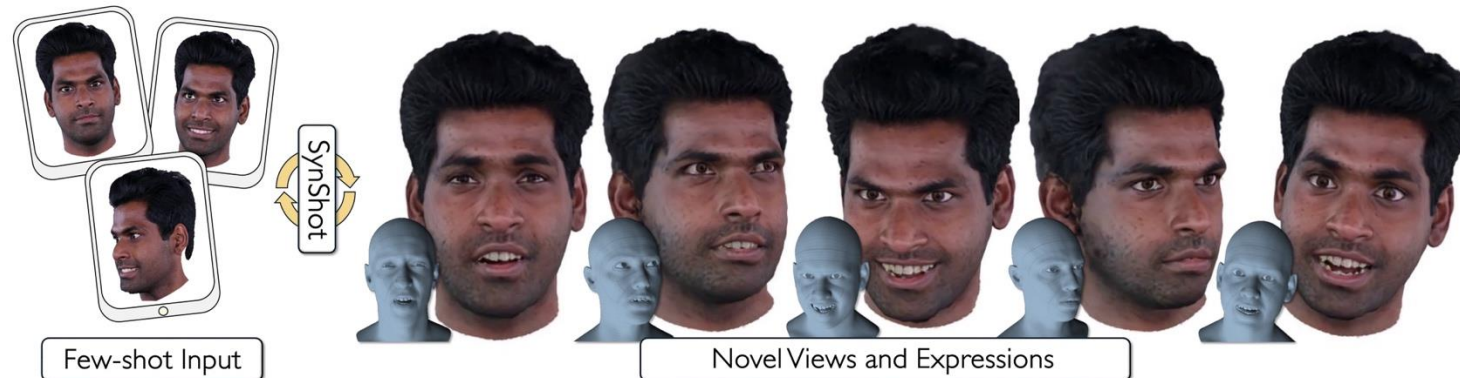
Saturday, June 14
 17:00 - 19:00 Poster Session 4 & Exhibit Hall (Hall D)
 Poster 7

Capturing and Synthesis of 3D Digital Humans



GEM - Gaussian Eigen Models for Human Heads

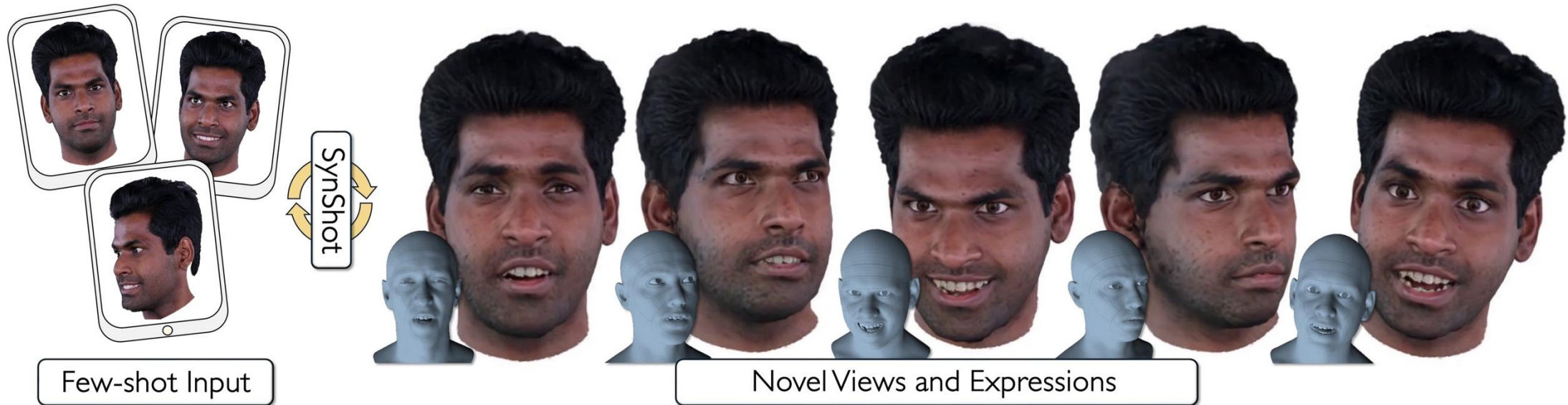
Wojciech Zielonka, Timo Bolkart, Thabo Beeler, Justus Thies



SynShot - Synthetic Prior for Few-Shot Drivable Head Avatar Inversion

Wojciech Zielonka, Stephan J. Garbin, Alexandros Lattas,
George Kopanas, Paulo Gotardo, Thabo Beeler, Justus Thies, Timo Bolkart

Capturing and Synthesis of 3D Digital Humans



SynShot - Synthetic Prior for Few-Shot Drivable Head Avatar Inversion

Wojciech Zielonka, Stephan J. Garbin, Alexandros Lattas,
George Kopanas, Paulo Gotardo, Thabo Beeler, Justus Thies, Timo Bolkart

SynShot – Motivation

Few-shot Input



Novel Views and Expressions



Motivation



Source



INSTA⁴



Splatting Avatar⁵



Flash Avatar⁶



4) Zielonka *et al.* Instant Volumetric Head Avatars

5) Xiang *et al.* FlashAvatar: High-fidelity Head Avatar with Efficient Gaussian Embedding

6) Zhijing *et al.* SplattingAvatar: Realistic Real-Time Human Avatars with Mesh-Embedded Gaussian Splatting

Motivation



Source



InvertAvata
 r^7



Portrait4D⁶



Next3D⁸

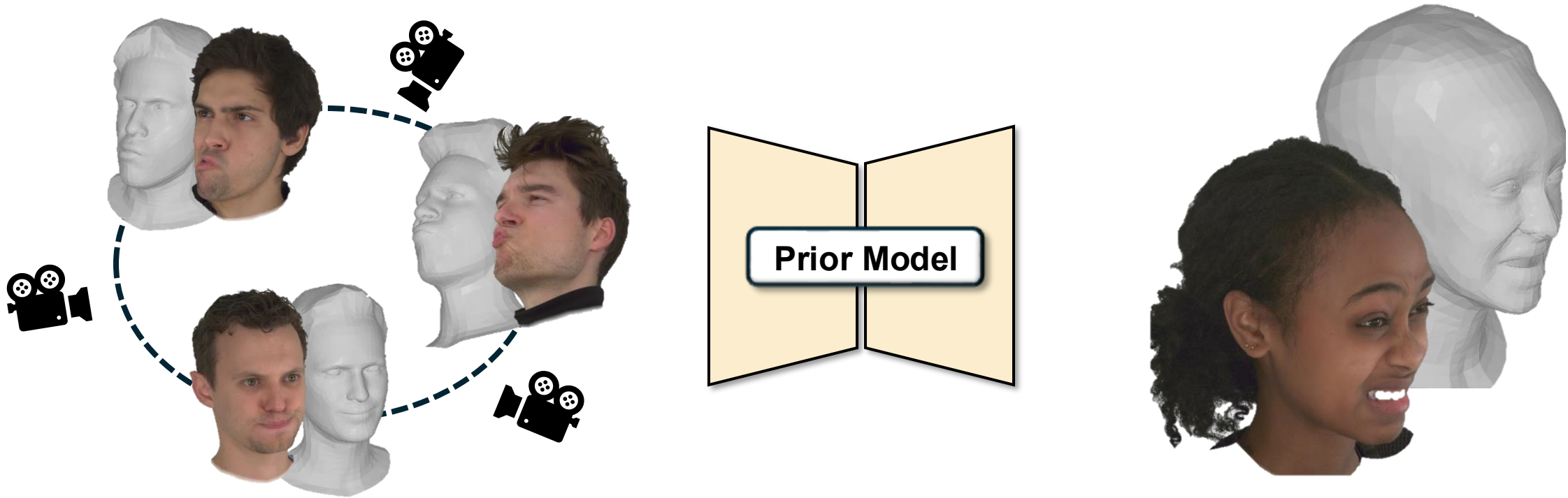


7) Zhao *et al.* InvertAvatar: Incremental GAN Inversion for Generalized Head Avatars

8) Deng *et al.* Portrait4D: Learning One-Shot 4D Head Avatar Synthesis using Synthetic Data

9) Sun *et al.* Next3D: Generative Neural Texture Rasterization for 3D-Aware Head Avatars

Motivation (Solution)



Use multi-view dataset with tracked meshes to build a prior model used for inversion and driving the avatars



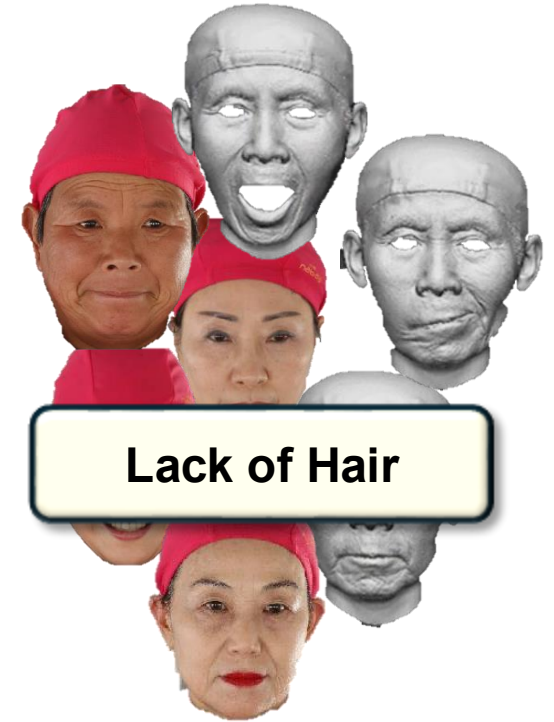
Motivation



Nersemble¹



Multiface²



FaceScape³



1) Kirschstein *et al.* NeRSemble: Multi-View Radiance Field Reconstruction of Human Heads

2) Wu *et al.* Multiface: A Dataset for Neural Face Rendering

3) Zhu *et al.* FaceScape: 3D Facial Dataset and Benchmark for Single-View 3D Face Reconstruction

Motivation



The **General Data Protection Regulation (GDPR)** is an EU law that protects individuals' personal data and privacy, enforced since May 25, 2018.

What does it mean for **Digital Humans research**:

1. Dataset derivatives must be frequently deleted e.g., each 30 days.
2. Trained models the same, periodically removed.



Nersemble¹

Multiface²

FaceScape³



SynShot - Synthetic Prior for Few-Shot Drivable Head Avatar Inversion

1) Kirschstein *et al.* NeRSemble: Multi-View Radiance Field Reconstruction of Human Heads

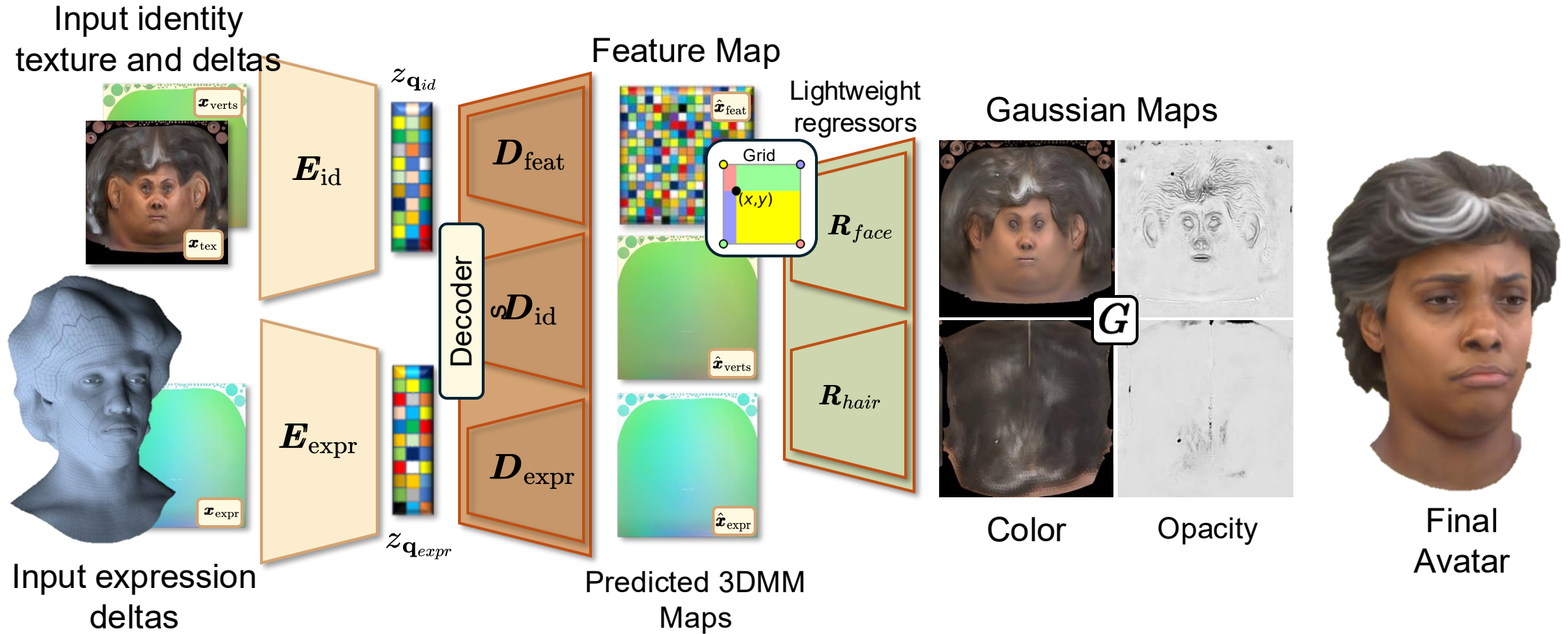
2) Wu *et al.* Multiface: A Dataset for Neural Face Rendering

3) Zhu *et al.* FaceScape: 3D Facial Dataset and Benchmark for Single-View 3D Face Reconstruction

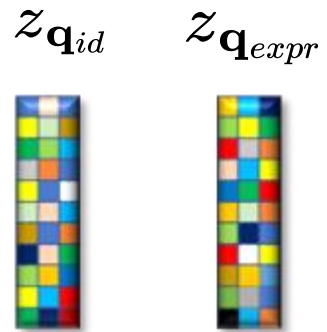
Motivation – Synthetics



Method



Method – Latent Space Interpolation



Method – Inversion



Few-shot Input



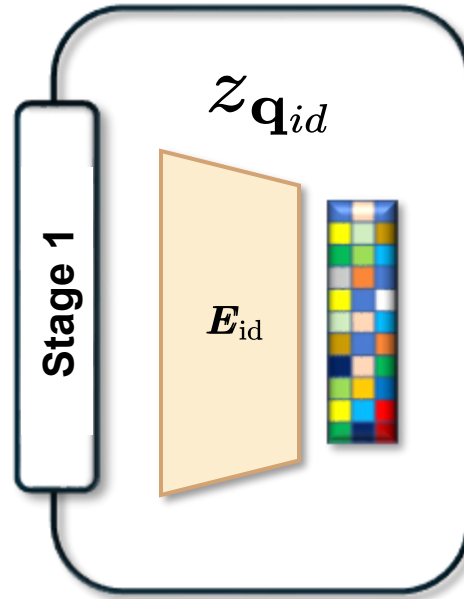
Inverted Avatar



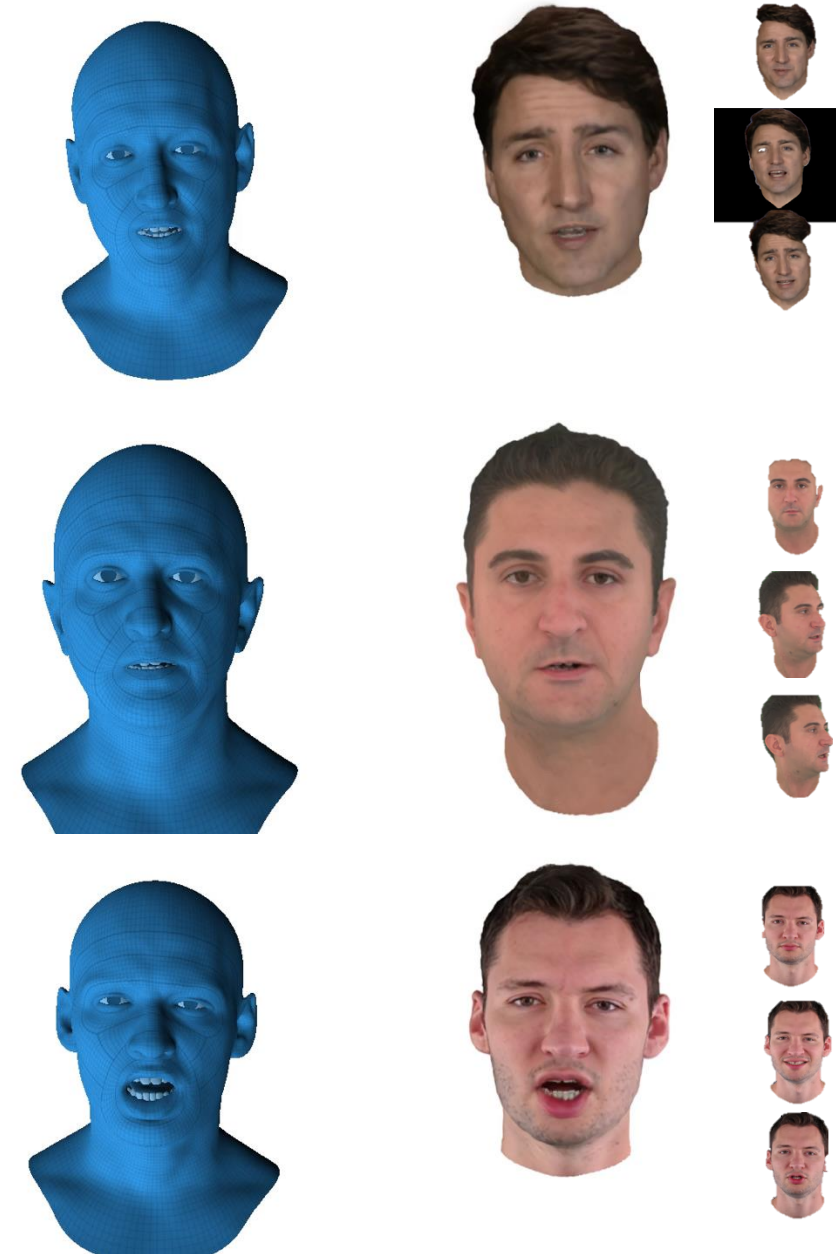
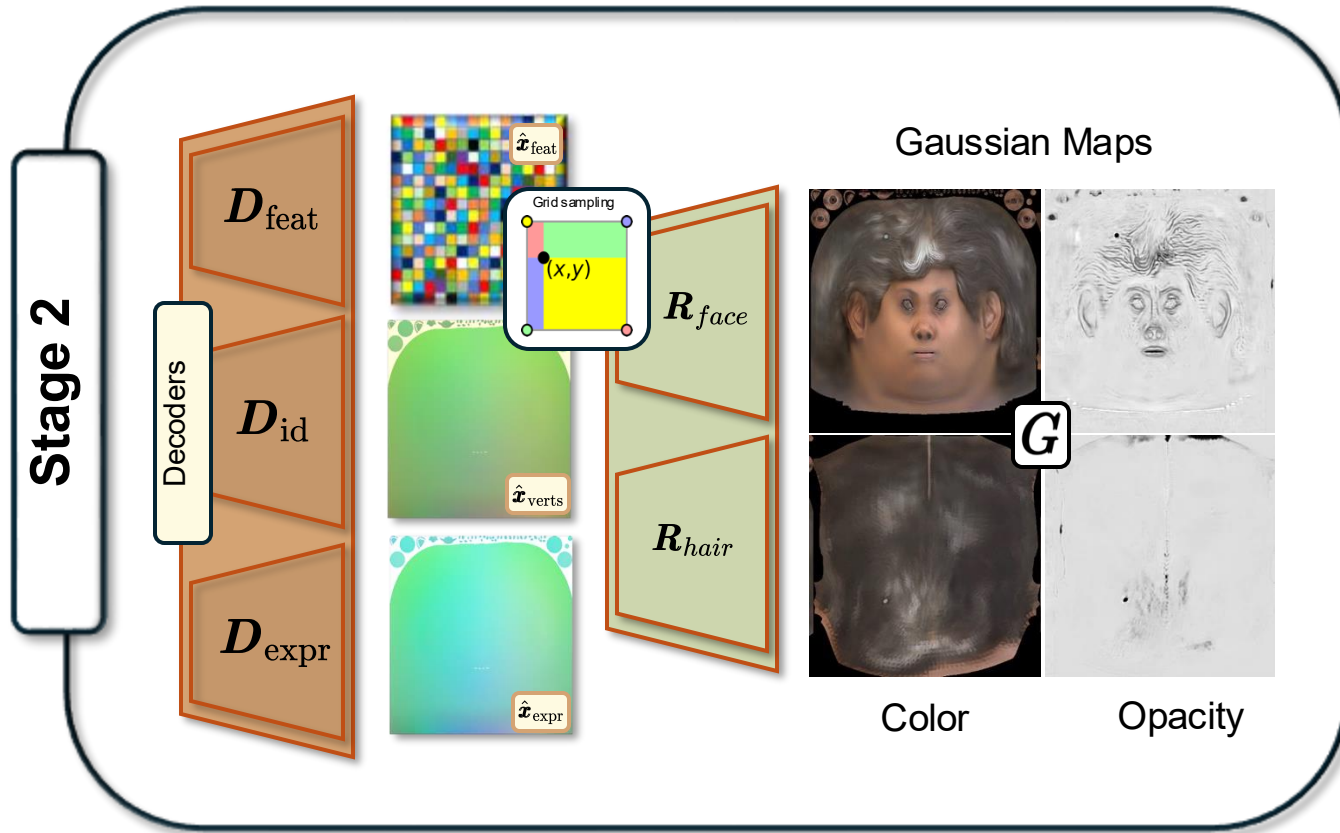
Input



Inversion – Stage 1



Inversion – Stage 2



Results – Personalized-baselines



Source



Ours



INSTA⁴



**Splatting
Avatar⁵**



Flash Avatar⁶



4) Zielonka *et al.* Instant Volumetric Head Avatars

5) Xiang *et al.* FlashAvatar: High-fidelity Head Avatar with Efficient Gaussian Embedding

6) Zhijing *et al.* SplattingAvatar: Realistic Real-Time Human Avatars with Mesh-Embedded Gaussian Splatting

Results – GAN-baselines



Source



Ours



InvertAvata
 r^7



Portrait4D⁶



Next3D⁸



7) Zhao *et al.* InvertAvatar: Incremental GAN Inversion for Generalized Head Avatars

8) Deng *et al.* Portrait4D: Learning One-Shot 4D Head Avatar Synthesis using Synthetic Data

9) Sun *et al.* Next3D: Generative Neural Texture Rasterization for 3D-Aware Head Avatars

Results



Summary

TECHNISCHE
UNIVERSITÄT
DARMSTADT

MAX PLANCK INSTITUTE
FOR INTELLIGENT SYSTEMS

Few-shot Input

Novel Views and Expressions

CVPR Nashville
JUNE 11-15, 2025

SynShot: Synthetic Prior for Few-Shot Drivable Head Avatar Inversion

Wojciech Zielonka, Stephan J. Garbin, Alexandros Lattas, George Kopanas, Paulo Gotardo, Thabo Beeler, Justus Thies, Timo Bolkart

Motivation

Challenges

- I. Controllable 3D generative models need diverse image-mesh pairs, which are scarce.
- II. Real data use is limited by privacy laws (e.g., GDPR).
- III. Monocular avatar models lack strong priors and often overfit, limiting generalization.

Solution

To address those challenges, we propose a method that learns a prior from a large synthetic dataset of diverse heads. Given a few input images, it fine-tunes this prior to generate a photorealistic avatar that generalizes to new expressions and views.

Synthetic Prior

The prior network uses a VQ-VAE generative framework to predict feature maps, which are later bilinearly sampled and used as conditioning inputs for per-part (face, hair) regressors of Gaussian primitives.

Our latent space for expressions and identity is represented by a smooth manifold.

Prior Fine-tuned Prior Fine-tuned

To bridge the sim-to-real gap, we designed an optimization-based refinement technique for our prior network, like pivotal tuning.

Few-shot Inversion

GAN-based baselines

Input Ground Truth Ours Invert Avatar⁷ Portrait4D⁶ Next3D⁸

SynShot outperform existing GAN-based methods especially in novel-view consistency.

Monocular baselines

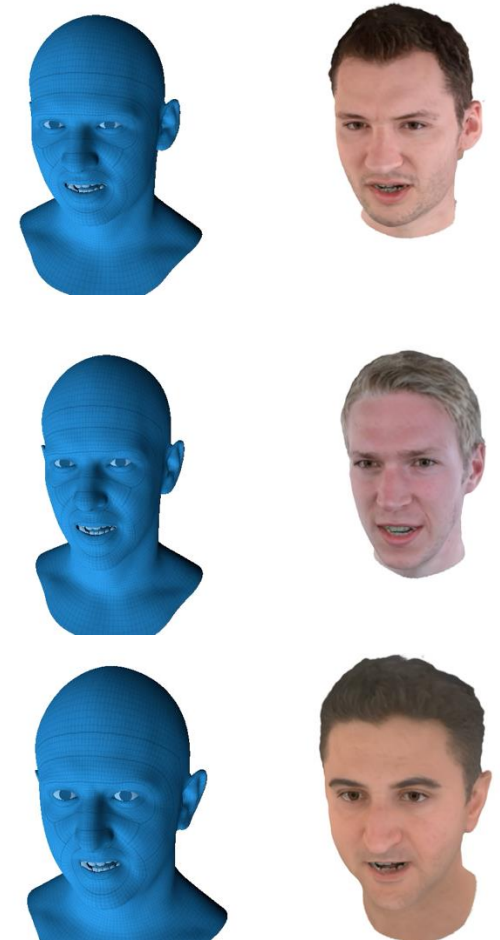
Driving Ours INSTA⁴ Splatting Avatar⁵ Flash Avatar⁶

Synthetic prior significantly boosts cross-reenactment over monocular methods.

Driving Prior Fine-tuned Input Prior Fine-tuned

We can bridge the gap between synthetics and real data.

1) Kirschstein et al. NerfScribe: Multi-View Radiance Field Reconstruction of Human Heads
2) Wu et al. Multiface: A Dataset for Neural Face Rendering
3) Zhu et al. FaceScape: 3D Facial Dataset and Benchmark for Single-View 3D Face Reconstruction
4) Zielonka et al. Instant Volumetric Head Avatars
5) Xiang et al. FlashAvatar: High-Quality Head Avatar with Efficient Gaussian Embedding
6) Zhang et al. SplattingAvatar: Realistic Real-Time Human Avatars with Mesh-Embedded Gaussian Splatting
7) Zhou et al. InvertAvatar: Incremental GAN Inversion for Generalized Head Avatars
8) Deng et al. Portrait4D: Learning One-Shot 4D Head Avatar Synthesis using Synthetic Data
9) Sun et al. Next3D: Generative Neural Texture Representation for 3D-Avatar Head Avatars



Saturday, June 14

10:30 - 12:30 Poster Session 3 & Exhibit Hall (ExHall D)

Poster 8

Capturing and Synthesis of Digital Humans

Wojciech Zielonka



Personal Website



MAX PLANCK INSTITUTE
FOR INTELLIGENT SYSTEMS



**TECHNISCHE
UNIVERSITÄT
DARMSTADT**