CVPR JUNE 3-7, 2026 DENVER COLORADO

FlexAvatar

Learning Complete 3D Head Avatars with Partial Supervision

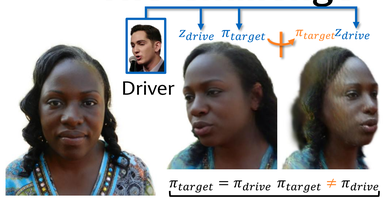
Tobias Kirschstein – Simon Giebenhain – Matthias Nießner

Project Page with videos and interactive 3D viewer

Highlights

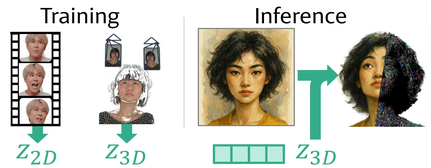
- Create a high-quality animatable **3D Head Avatar**
- **Input Flexibility**: Single-image, few-shot, monocular video, ...
- Heads are **full 360° viewable**
- **Bias Sinks** combine strengths of **2D datasets** (generalization) and **multi-view datasets** (completeness/quality)

The Challenge



- ✗ 3D data: **Not enough**
- ✗ 2D data: **Entangles** target viewpoint and expression => **Incomplete Heads**

The Solution



- Training: Learnable z_{3D}/z_{2D} tokens (bias sinks) indicate a **training sample's source dataset**
- Inference: Always feed in z_{3D} => inherit **completeness from 3D** and keep **generalization from 2D**

Complete 360° Head

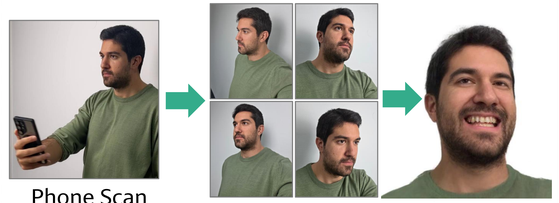
Effect of Fitting

3D Head Avatars from any Inputs

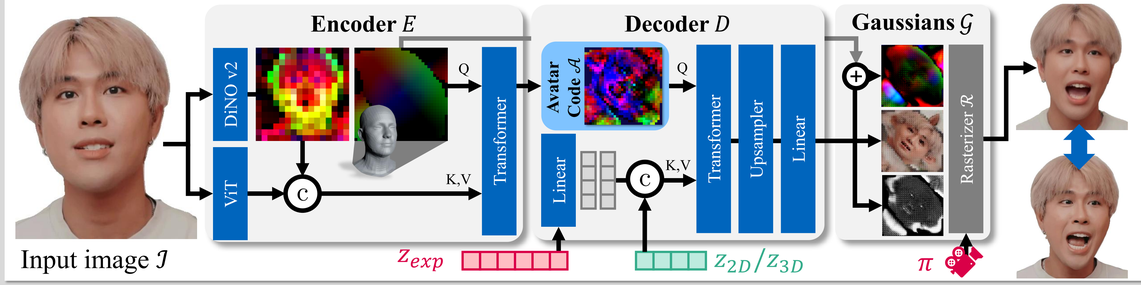
Portrait Image

Stylized Images

Phone Scans



Method



FlexAvatar Creation Steps

1. Run Encoder E on single image to get Avatar Code $\mathcal{A}$
2. Fit $\mathcal{A}$ with $z_{bias} = z_{3D}$ against all inputs
3. Animate via z_{exp} Render via π

Data & Bias Sinks Ablation

3D Data Requirement for Bias Sinks

Conclusion

- ✗ Simple joint 2D and 3D training leads to **data source separation**
- Capture desired properties in **learnable z_{3D} token**
- Bridge remaining gap via **fitting**