



# GUAVA: Generalizable Upper Body 3D Gaussian Avatar

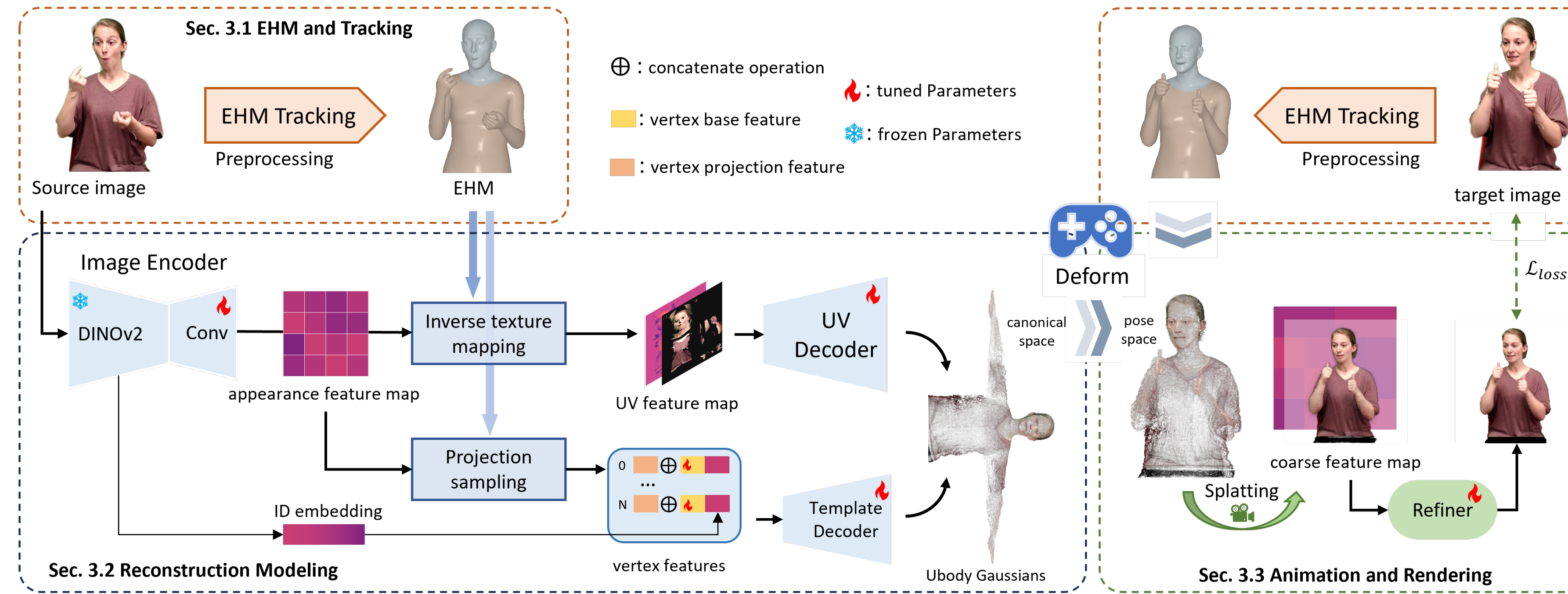


Dongbin Zhang, Yunfei Liu†, Lijian Lin, Ye Zhu, Yang Li, Minghan Qin, Yu Li‡, Haoqian Wang†  
Tsinghua University, International Digital Economy Academy

## Overview



## Pipeline



- We propose GUAVA, the first framework for generalizable upper-body 3D Gaussian avatar reconstruction from a single image.
- We introduce an expressive human template model with a corresponding upper-body tracking framework, providing an accurate prior for reconstruction.

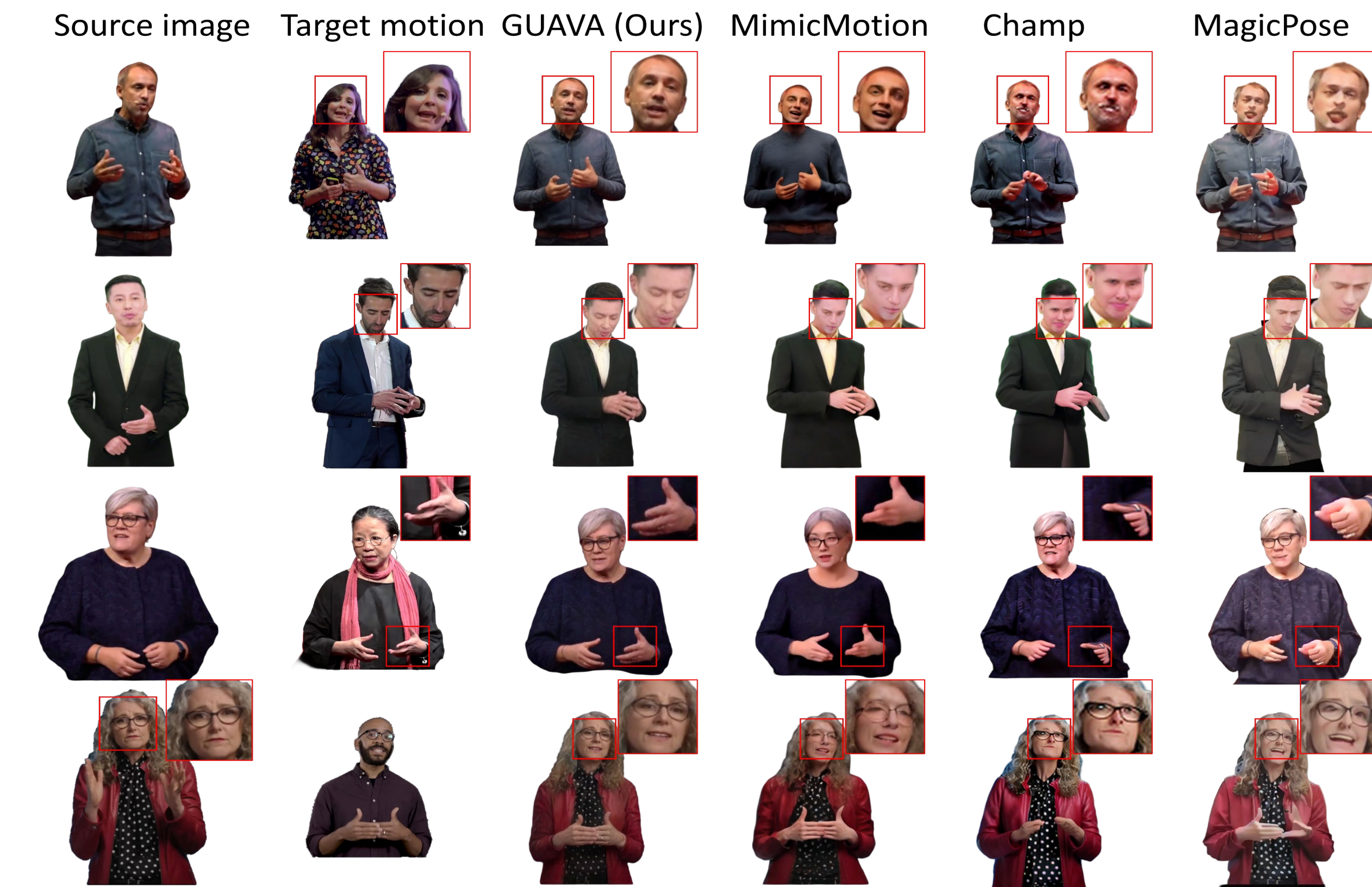
## Quantitative results on self-reenactment

	PSNR↑	L1↓	SSIM↑	LPIPS↓	FPS↑
GUAVA (Ours)	<b>25.87</b>	<b>0.0162</b>	<b>0.9000</b>	<b>0.0813</b>	<b>52.21</b>
MimicMotion	24.46	0.0200	0.8768	0.0879	0.21
Champ	22.01	0.0258	0.8643	0.1000	0.53
MagicPose	21.25	0.0333	0.8661	0.0913	0.12

	PSNR↑	L1↓	SSIM↑	LPIPS↓	Input	Time
GUAVA (Ours)	<b>25.70</b>	<b>0.0168</b>	<b>0.8976</b>	<b>0.0836</b>	single image	≈98ms
ExAvatar	24.09	0.0207	0.8783	0.1064	video	≈2.4 h
GaussianAvatar	23.62	0.0199	0.8780	0.1085	video	≈1.7 h
GART	24.46	0.0195	0.8780	0.1016	video	≈7 min

## Visual results on cross-reenactment



## Visual results on self-reenactment



For each single image with a tracked pose, GUAVA can reconstruct a 3D upper-body Gaussian avatar via feed-forward inference about **0.1s** ⚡, enabling real-time expressive animation at  $512 \times 512$  resolution.